To Buyers of Mahler's Guide to Advanced Ratemaking

Howard C. Mahler, FCAS, MAAA hmahler@mac.com

This study guide is split into two volumes: Sections 1 to 9, and Sections 10 and following.

In the electronic version use the bookmarks / table of contents in the Navigation Panel in order to help you find what you want.
You may find it helpful to print out selected portions, such as the Table of Contents.

Information in bold is more important to pass your exam.
Sections and material in italics is less likely to be needed to directly answer exam questions, and should be skipped on the first time through.

I have doubled underlined highly recommended questions to do on your first pass through the material, underlined recommended questions to do on your second pass, and starred additional questions to do on a third pass through the material.[1] No questions were labeled from the 2011 exam or later, in order to allow you to use them as practice exams.

Added to the Exam 8 syllabus for Fall 2025:

Penalized Regression & Lasso Credibility, by Thomas Holmes, and Mattia Casotto.
CAS Monograph #13.

Actuarial Standard of Practice **(ASOP) #25** "Credibility Procedure Applicable to Accident and Health, Group Term Life, and Property/Casualty Coverages".

The reading on the NCCI Retrospective Rating Plan may be updated.
The version in the 2026 Study Kit is not available at the time of publication.
Check my webpage for any updates.

The CAS is not releasing exams starting with Fall 2020.

The CAS used computer based testing for the Fall 2020 Exam,
and expects to do so going forward. Be sure to check the CAS webpage for information.

My solutions to questions are intended to be model solutions.[2] Often they are more detailed and contain more explanation than would be needed in order to get full credit. This was done in order to give you a clearer and better understanding of the subject material.[3]

After you have done one of the released exams, be sure to look closely at the CAS Examiner's Report. See the sample solutions in the Examiner's Reports, and read the comments of the examiners.

In the case of verbal questions, do not concentrate on grammar or complete sentences. Feel free to list or outline your ideas. The selected use of abbreviations can save some time.

---

[1] Obviously feel free to do whatever questions you want. This is just a guide for those who find it helpful.
[2] In some cases, I even quote a reading word for word. This does not mean you need to be able to do so!
[3] Sometimes much of what I say is directed at those students who did not answer a question correctly and need to learn more. In any case, you want to know as much as possible to help answer the question that will be on your exam, as opposed to whatever they happened to have asked on some past exam.

In any case, remember that spending considerable additional time to increase 80% credit to 100% credit on a single question is usually not a good use of your limited exam time. You can come back later to a question, if you have the time.

As stated in the CAS Syllabus: “The model response to the typical essay question is brief, less than one-half of a written page. Be concise — candidates do not need to answer in complete sentences when a well-composed outline format is more appropriate. Candidates should not waste time on obscure details. They should show that they have learned the relevant material and that they understand it. They should state the obvious, if it is part of the answer.”

Also read “The Importance of Adverbs on Exams,” in which the Exam Committee notes the difference between: briefly discuss, discuss, and fully discuss.[4]
“Brief descriptions, discussions, etc., are worth 1/4 point.
(Unmodified) discussions or descriptions are worth 1/2 point.
Full descriptions or discussions are worth at least 1 point.
Please look carefully for these word choices and point values on all CAS upper-level exams.
Most importantly, answer the question in accordance with the amount of information being asked.”

The CAS is gradually moving towards an integrative testing framework.
Integrative Questions (IQs) will require candidates to understand multiple facets of the syllabus material and concepts in addressing complex business problems in a single exam question. IQs will differ from a typical exam question in three significant ways.

1. An IQ will be worth more points. One IQ could be worth 10-15% of the total exam.
2. Each IQ will require candidates to draw from multiple syllabus learning objectives in order to answer the question.
3. IQs will test at a higher average Bloom’s Taxonomy level than a standard exam question.

To assist candidates with preparing to answer an IQ, t**he CAS released sample IQs and responses**.[5] It should be noted that while the samples were constructed in parallel with the IQ that will appear on the exam, they may not be structured in the same manner nor cover the same learning objectives as the exam question. It is advised that candidates use the samples to validate preparation and identify potential areas for improvement after completing the majority of their study, rather than using them during their initial study as one might use text book exercises.

Exam 8 featured one IQ on the Fall 2017 exam,
and two each on the Fall 2018 and Fall 2019 exams.

It is expected that Exams 7, 8, and 9 will continue to include IQs in future sittings, and the number of IQs that will appear on the exams will gradually increase over time. At the same time, there will be fewer exam questions overall to account for the presence of IQs in order to avoid any increase in the time length of the exam. There will be no change to the normal grading process, as described in the Syllabus, for IQs.

---

[4] http://www.casact.org/admissions/index.cfm?fa=adverbs
[5] See the CAS webpage for updated information.

In March 2022 the CAS added New Questions to the CBT Sample Exams 5-9.[6]
Sample questions 12 and 13 are from the unreleased Fall 2021 Exam 8.[7]
I have included them in my Sections 1 and 18. .

My study guide includes question written by me, and some by Sholom Feldblum.[8] In addition, the former exam questions are arranged in chronological order. The more recent exam questions are on average more similar to what you will be asked on your exam, than are less recent questions.

Note that In some cases, numerical values shown in one of my spreadsheets are unrounded, while the corresponding value in my text may be rounded.

It is important that you **do problems when learning a subject and then some more problems a few weeks later.**
As you get closer to the exam, the portion of time spent doing problems should increase.

There are two manners in which you should be doing problems. First you can do problems in order to learn the material. Take as long on each problem as you need to fully understand the concepts and the solution. Reread the relevant syllabus material. Carefully go over the solution to see if you really know what to do. Think about what would happen if one or more aspects of the question were revised. This manner of doing problems should be gradually replaced by the following manner as you get closer to the exam.

The second manner is to do a series of problems under exam conditions, with the items you will have when you take the exam. Take in advance a number of points to try based on the time available. For example, if you have an uninterrupted hour, then one might try 60 / 4 = 15 points of problems. Do problems as you would on an exam in any order, skipping some and coming back to some, until you run out of time. Leave time to double check your work.

It is important that you develop the skill of quickly and clearly writing down what you know. Many of you will benefit by giving some of your solutions to questions to someone else to "grade".[9] They should give you feedback on whether they were able to follow what you did.[10] They should point out where you wrote more than was necessary or not enough.

Read the "Hints on Study and Exam Techniques" in the CAS Syllabus.

The CAS has posted a pdf on Bloom's Taxonomy of question writing.
You might want to look at it.
http://www.casact.org/admissions/syllabus/Blooms-Taxonomy.pdf

---

[6] https://abe-prd-1.pvue2.com/st2/driver/startDelivery?sessionUUID=972271b1-7c06-4e8f-8a4b-499d4e047cd0
[7] Also in the Sample Questions are Fall 2019 questions 17 and 19, which are in my study guide.
[8] I thank Sholom Feldblum for the kind permission to use his material. Any mistakes are my responsibility.
[9] Someone else taking this exam or who has just passed this exam would be a good choice.
[10] On average you get less credit on essay questions when graded by someone else then when you self-grade.

I thank Sholom Feldblum for the kind permission to use his material.
Finally, thanks to the many past students who have helped me to improve these study guides.

Sold separately are my seminar style slides. They are electronic.

Feel free to send me any questions or suggestions: hmahler@mac.com
Please send me any suspected errors by Email.
(Please specify as carefully as possible the page and Exam number.)

I will post a list of errata on my webpage: www.howardmahler.com/Teaching

CAS Post Exam Summaries:

"In light of the discontinuation of Examiners' Reports in 2020, the CAS has recognized the need to fill the void left in candidates' understanding of effective study strategies and overall exam performance. To bridge this knowledge gap, we are introducing the new Post Exam Summary crafted by the Syllabus and Examination Working Group. This resource is designed to provide candidates with insightful observations on candidates' exam performance, coupled with expert recommendations for improvement. The Post Exam Summary comprises a general summary section that applies universally to all constructed response exams, followed by individual sections for each of the exams administered during the last sitting. In the future, we look forward to expanding on this format and continuing to enhance this summary."

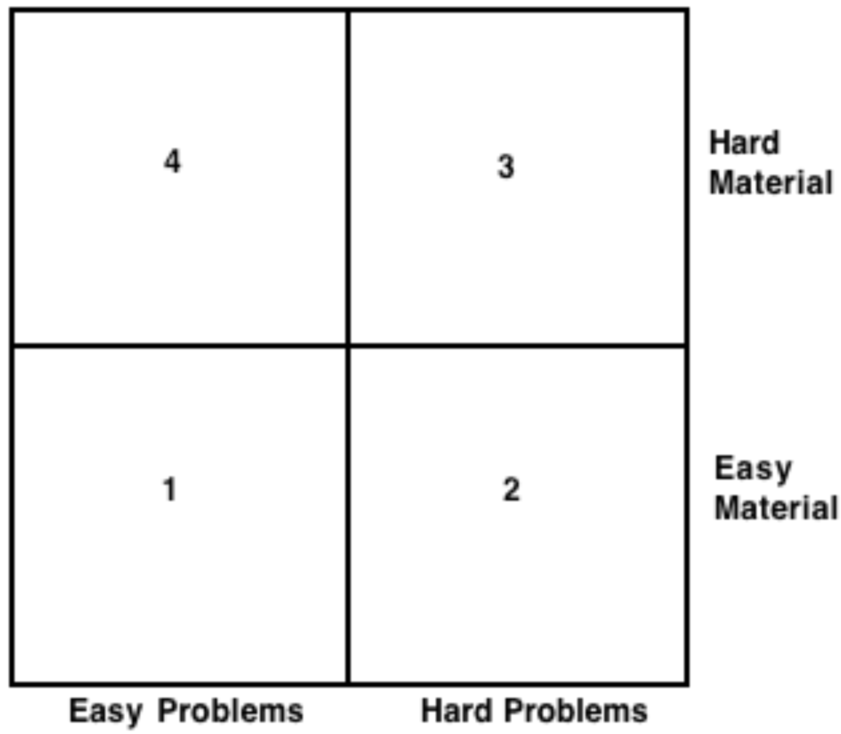
The CAS should be releasing such Post Exam Summaries after each exam sitting.
To the extent you might find them useful, be sure to check them out.[11]

Preparing for a CAS Exam--what to do with hard material
by Dr. J. Eric Brosius, FCAS

The syllabus for a typical CAS exam includes both easy and hard material. Many students learn the easy material well, but adopt less-than-optimal strategies for learning the hard material. Some spend a lot of time trying to understand syllabus readings that are nearly incomprehensible. Others ignore the more difficult readings altogether. Neither approach is a good idea, not if you hope to pass! I will suggest a better way to approach these readings. Your goal in studying is not to understand the material in general but to be able to answer the questions. Do not study the syllabus readings in a vacuum; consider also what types of questions are likely to be asked. Each exam contains both easy problems and hard problems.

[11] They might be particularly useful after taking an exam and unfortunately failing.

We can divide the problems into four categories based on the difficulty of the material and the difficulty of the problem, as follows:



Box 1 contains easy problems on easy material. These are easy to answer; unfortunately, there are not enough of them!

Box 2 contains hard problems on easy material. You can prepare for these by practicing problems from old tests and other sources of sample problems.

Box 3 contains hard problems on hard material. Few students can afford to spend the time required to answer all of these. Fortunately, the Examination Committee does not ask many of these question: even if they understand the reading well enough to do so, there isn't much point in a question that no one can answer. Be prepared to skip Box 3 problems if necessary.

Box 4 contains easy problems on hard material. These problems can supply the extra points you need to change a "5" into a “6". They appear often, because the Examination Committee tends to ask easy questions about hard readings. When a reading is technically difficult, and especially if it was recently added to the syllabus, even the simplest question poses a challenge. Study these readings with an eye to answering the obvious questions. It is a shame not to get points for a question that could have been answered if only you had read the first paragraph of the reading.

Plan for your exam in such a way that you **focus on Box 2 and Box 4**. Prepare for Box 2 questions by studying the easy material in detail, and by doing many sample problems. Prepare for Box 4 questions by outlining the high points of the material, and by trying to guess, alone or with other students, what questions on this material might appear on the exam.

Use whatever order to go through the material that works best for you.
Here is a schedule that may work for some people.
Modify it to meet your own needs.
In any case, leave plenty of time to go back and review material.

A 14 week Study Schedule for Exam 8:

Week	Sections of Study Guide
1	1-2
2-3	3
4	4-5
5-6	6-8
7-8	9-10
9	11-12
10	13-14
11	15-16
12	17-18
13	19-21
14	22

Since 2011, the points on exam questions are similar to the present. Going back a few more years further in time, a 5 point exam question might only be worth 3 points today.[12]

Exam 8	Points	Number of Questions	Integrated Questions	Average % of Exam per Integrated Question
2011	59	25		
2012	54.75	23		
2013	57.5	25		
2014	60.25	25		
2015	59.5	23		
2016	53.25	21		
2017	53.75	20	1	15.8%
2018	52	17	2	17.8%
2019	52.5	19	2	10.5%

The CAS has stopped releasing pass marks:

Exam 8	Pass Mark	Percent of Available Points	95th Percentile	75th Percentile
2011	43.75	74.15%	47.38	43.00
2012	37.75	68.95%	44.25	39.75
2013	40.75	70.87%	47.50	43.63
2014	37.50	62.24%	44.50	40.63
2015	40.75	68.49%	48.50	43.13
2016	37.25	69.95%	42.88	38.88
2017	37.5	69.77%	43.00	39.50
2018	33.75	64.91%	47.38	43.00
2019	37	70.48%	42.88	38.50

[12] For my problems, it depends on when I wrote them.
My older ones are probably more like the older exam questions as far as points go.
I am sorry that my study guides are not more consistent with respect to “points”.
The CAS stopped releasing exams with the Fall 2020 exam.

Exam 8	Exams Taken	Passed	Raw Pass Ratio	Effective Pass Ratio
2011	418	93	22.2%	23.9%
2012	519	218	42.0%	43.7%
2013	592	283	47.8%	49.3%
2014	729	350	48.0%	50.2%
2015	771	313	40.60%	42.18%
2016	791	301	38.05%	40.13%
2017	945	376	39.8%	41.7%
2018	953	314	32.9%	35.1%
2019	1080	376	34.8%	37.0%
2020	228	86	37.7%	41.1%
S2021	174	66	37.9%	42.0%
F2021	900	316	35.1%	38.5%
2022	870	329	37.8%	39.8%
2023	875	422	48.2%	51.1%
2024	898	359	40.0%	42.1%
2025	614	278	45.3%	47.0%

One measure of the difficulty of an exam is the ratio of the 75th percentile to the available points:

Exam 8	Points	75th Percentile	Ratio
2011	59	43.00	72.9%
2012	54.75	39.75	72.6%
2013	57.5	43.63	75.9%
2014	60.25	40.63	67.4%
2015	59.5	43.13	72.5%
2016	53.25	38.88	73.0%
2017	53.75	39.50	73.5%
2018	52	35.00	67.3%
2019	52.5	38.50	73.3%

The lower the ratio of the 75th percentile to the available points, the harder the exam.

Mahler’s Guide to

Advanced Ratemaking

CAS Exam 8

prepared by
Howard C. Mahler, FCAS

Study Aid 2026-**8**

Howard Mahler
hmahler@mac.com
www.howardmahler.com/Teaching

Mahler's Guide to Advanced Ratemaking

Information in bold or sections whose title is in bold are more important for passing the exam. Larger bold type indicates it is extremely important. Information presented in italics (including subsections whose titles are in italics) should rarely be needed to directly answer exam questions and should be skipped on first reading. It is provided to aid the reader's overall understanding of the subject, and to be useful in practical applications.

I have doubled underlined highly recommended questions to do on your first pass through the material, underlined recommended questions to do on your second pass, and starred additional questions to do on a third pass through the material.[1] No questions were labeled from the 2011 exam or later, in order to allow you to use them as practice exams.

Solutions to problems are at the end of each section.[2]

[1] Obviously feel free to do whatever questions you want. This is just a guide for those who find it helpful.
[2] Note that problems include both some written by me and some from past exams. The latter are copyright by the Casualty Actuarial Society and are reproduced here solely to aid students in studying for exams. The solutions and comments are solely the responsibility of the author; the CAS bears no responsibility for their accuracy. While some of the comments may seem critical of certain questions, this is intended solely to aid you in studying and in no way is intended as a criticism of the many volunteers who work extremely long and hard to produce quality exams. There are also some past exam questions copyright by the Society of Actuaries.

Volume	Section #	Pages	Section Name
one	1	9-110	Mahler, An Example of Credibility and Shifting Risk Parameters
one	2	111-213	Bailey & Simon, Credibility of a Single Car
one	3	214-665	Goldburd, Khare and Tevet, Generalized Linear Models
one	4	666-697	ASOP 12: Risk Classification
one	5	698-781	Couret & Venter, Class Frequency Vectors
one	6	782-904	Experience Rating
one	7	905-987	NCCI Experience Rating Plan
one	8	988-1091	ISO Experience Rating Plan
one	9	1092-1188	Frequency and Loss Distributions
two	10	1189-1513	Bahnemann, Distributions for Actuaries
two	11	1514-1638	Lee Diagrams, Loss Distributions
two	12	1639-1778	Retrospective Rating
two	13	1779-1876	Table M Construction
two	14	1877-1935	NCCI Retrospective Rating
two	15	1936-2027	Table L
two	16	2028-2126	Lee Diagrams, Retrospective Rating
two	17	2127-2156	Limited Table M
two	18	2157-2188	Other Loss Sensitive Plans
two	19	2189-2307	Pricing Large Dollar Deductible Policies
two	20	2308-2325	Concluding Remarks, Individual Risk Rating
two	21	2326-2334	ASOP 25: Credibility Procedures
two	22	2335-2459	Holmes & Casotto, Lasso Credibility

Past Exam Questions by Section

Sec.		1995 Exam 9	1996 Exam 9	1997 Exam 9	1998 Exam 9
1	Mahler, Shifting Risk Parameters	10, 31	20	44, 45, 46	13, 14, 25
2	Bailey & Simon, Cred. Single Car	6, 30, 32	50	19	26
3	Goldburd, Khare and Tevet, GLMs				
4	ASOP 12: Risk Classification			18	15, 22
5	Couret & Venter, Class Freq.				
6	Experience Rating	20, 40, 42	4, 27, 28c&d	31a, 32	18, 37b, 38, 39
7	NCCI Experience Rating Plan	16, 41	24, 25	10, 34	17, 20, 36
8	ISO Experience Rating Plan	17	1, 21, 22, 23	9, 33	41
9	Frequency and Loss Distributions				
10	Bahnemann, Distrib. for Actuaries	11, 33, 35	36, 38, 41, 42	13, 36a, 40a	30a, 31, 33, 34
11	Lee Diagrams, Loss Distributions		39	37	29
12	Retrospective Rating	21, 22, 24, 44, 46, 47	29, 31, 32, 34	1, 27	4, 44c, 42, 47
13	Table M Construction	45	10	22, 23	
14	NCCI Retro. Rating				46
15	Table L	25	30, 35		43
16	Lee Diagrams, Retro. Rating	50		4, 26	
17	Limited Table M				
18	Other Loss Sensitive Plans				
19	Pricing LDD Policies				
20	Conclud. Remarks, Indiv. Risk Rat.				

Some questions are based on more than one syllabus reading, particularly on recent exams.[3]
In any case, sometimes it is unclear what is the best section in which to put a question.
In those cases, I have made one of the possible reasonable choices of where to put a question.

[3] Integrated questions involve several different syllabus readings.

Sec.		1999 Exam 9	2000 Exam 9	2001 Exam 9	2002 Exam 9
1	Mahler, Shifting Risk Parameters	48	34	1	
2	Bailey & Simon, Cred. Single Car	1	32	2, 22	47
3	Goldburd, Khare and Tevet, GLMs				
4	ASOP 12: Risk Classification	2, 43b			48
5	Couret & Venter, Class Freq.				
6	Experience Rating	12, 13, 31	1, 4, 40		
7	NCCI Experience Rating Plan	28	17, 42	25	33
8	ISO Experience Rating Plan	30	2	27	11, 12, 34
9	Frequency and Loss Distributions				
10	Bahnemann, Distrib. for Actuaries	35, 38, 40, 41	39	11, 35, 37c	41, 42
11	Lee Diagrams, Loss Distributions	34, 39	37		43
12	Retrospective Rating	5, 6, 9, 21, 22, 23, 25	5, 6, 44	8, 9, 10, 31, 32, 34	14, 15, 16, 35, 40
13	Table M Construction		19, 48	30	36
14	NCCI Retro. Rating				
15	Table L	26	45		38, 39
16	Lee Diagrams, Retro. Rating				17
17	Limited Table M				
18	Other Loss Sensitive Plans				
19	Pricing LDD Policies	42	38		1
20	Conclud. Remarks, Indiv. Risk Rat.				

Sec.		2003 Exam 9	2004 Exam 9	2005 Exam 9	2006 Exam 9
1	Mahler, Shifting Risk Parameters	21	3	2	
2	Bailey & Simon, Cred. Single Car	22	2	3	2
3	Goldburd, Khare and Tevet, GLMs	25			5
4	ASOP 12: Risk Classification		23		
5	Couret & Venter, Class Freq.				
6	Experience Rating	2, 6, 26, 28	15, 16, 39	26	23, 27
7	NCCI Experience Rating Plan	27		24, 27	24
8	ISO Experience Rating Plan	3, 4, 5	14, 41	28	28
9	Frequency and Loss Distributions				
10	Bahnemann, Distrib. for Actuaries	13, 37, 38, 43	5, 6, 19 25, 26	6, 7, 10 23a, 35	6, 8
11	Lee Diagrams, Loss Distributions				
12	Retrospective Rating	7, 10, 31, 32, 33	18, 20, 45, 47	31, 32	30, 32, 35
13	Table M Construction		43	8	9
14	NCCI Retro. Rating				
15	Table L	30	44		7
16	Lee Diagrams, Retro. Rating	8, 9, 29	4, 17	33	29, 34
17	Limited Table M				
18	Other Loss Sensitive Plans				
19	Pricing LDD Policies	35	46, 48	34, 36	31, 33, 36
20	Conclud. Remarks, Indiv. Risk Rat.				

Sec.		2007 Exam 9	2008 Exam 9	2009 Exam 9	2010 Exam 9
1	Mahler, Shifting Risk Parameters	6			
2	Bailey & Simon, Cred. Single Car	2	5	4	5
3	Goldburd, Khare and Tevet, GLMs	4a	3	3	3
4	ASOP 12: Risk Classification				
5	Couret & Venter, Class Freq.				
6	Experience Rating	26	23	20	23
7	NCCI Experience Rating Plan	25, 28	25	21	20
8	ISO Experience Rating Plan	27	24	22	21
9	Frequency and Loss Distributions				
10	Bahnemann, Distrib. for Actuaries	7, 8, 10	26, 27	17, 18, 26	17, 26
11	Lee Diagrams, Loss Distributions			24	
12	Retrospective Rating	32, 35	36	28, 30, 31	27, 29
13	Table M Construction	30, 34	28		
14	NCCI Retro. Rating				
15	Table L		32, 33	32	
16	Lee Diagrams, Retro. Rating	31	29		25, 31
17	Limited Table M				
18	Other Loss Sensitive Plans				
19	Pricing LDD Policies	33, 36	30, 31	29a	28
20	Conclud. Remarks, Indiv. Risk Rat.			27	24

Sec.		2011 Exam 8	2012 Exam 8	2013 Exam 8	2014 Exam 8
1	Mahler, Shifting Risk Parameters		3		
2	Bailey & Simon, Cred. Single Car	1	6		5
3	Goldburd, Khare and Tevet, GLMs	*3*	2, 4	2	3
4	ASOP 12: Risk Classification				
5	Couret & Venter, Class Freq.	2	5	3	1, 4
6	Experience Rating	15, 16b&c	11, 16a&c	9, 10b	9, 11
7	NCCI Experience Rating Plan	12	13		10
8	ISO Experience Rating Plan	14	14	8	8
9	Frequency and Loss Distributions				
10	Bahnemann, Distrib. for Actuaries	10, 17	15	6	7
11	Lee Diagrams, Loss Distributions	11	22		6
12	Retrospective Rating	20, 21, 25	19, 23	14	17
13	Table M Construction			12	13
14	NCCI Retro. Rating				
15	Table L		18	13	
16	Lee Diagrams, Retro. Rating	22	21	15	12, 18
17	Limited Table M				
18	Other Loss Sensitive Plans				
19	Pricing LDD Policies	18, 19	20	16, 19	16, 19
20	Conclud. Remarks, Indiv. Risk Rat.	23			

Added for the 2011 Exam: Couret & Venter,
For the 2016 exam, Goldburd, M.; Khare, A.; and Tevet, D., "Generalized Linear Models for Insurance Rating," replaced Anderson, D.; Feldblum, S; Modlin, C; Schirmacher, D.; Schirmacher, E.; and Thandi, N., "A Practitioner's Guide to Generalized Linear Models"

Questions from Robertson no longer on the syllabus: 2011 Q.4, 2012 Q.1, 2013 Q.4, 2014 Q.2, 2015 Q.6, 2016 Q.2, 2017 Q.2, 2019 Q.4.
Questions from Clark Reinsurance Pricing no longer on the syllabus of this exam: 2011 Q.7&8, 2012 Q.7&10, 2013 Q.21&23&25, 2014 Q.20&21&22&23&25, 2015 Q.21&23, 2016 Q.20, 2017 Q.19, 2018 Q.15, 2019 Q.17&18.
Questions from Bernegger no longer on the syllabus of this exam: 2011 Q.9, 2012 Q.8, 2013 Q.20&22, 2015 Q.20, 2016 Q.21, 2017 Q.18.
Questions from Grossi & Kunreuther Catastrophes no longer on the syllabus of this exam: 2011 Q.5&6, 2012 Q.9, 2013 Q.24, 2014 Q.24, 2015 Q.22, 2016 Q.18&19, 2017 Q.20, 2018 Q.16&17, 2019 Q.19..

Sec.		2015 Exam 8	2016 Exam 8	2017 Exam 8	2018 Exam 8	2019 Exam 8
1	Mahler, Shifting Risk Parameters	4				
2	Bailey & Simon, Cred. Single Car	1	1	3	3	3
3	Goldburd, Khare and Tevet, GLMs	3	4, 5, 6, 7	4, 5, 6	5, 6, 7	2, 5, 6
4	ASOP 12: Risk Classification		3		4	
5	Couret & Venter, Class Freq.	5				1
6	Experience Rating	10, 11, 12	11	11	9, 10	9, 10, 11
7	NCCI Experience Rating Plan		9, 10			
8	ISO Experience Rating Plan	9		9, 10	11	
9	Frequency and Loss Distributions				1*	
10	Bahnemann, Distrib. for Actuaries	8a		7, 8, 14	8, 13	13
11	Lee Diagrams, Loss Distributions	7		12		
12	Retrospective Rating	15, 16, 17		13, 15, 17	14	14, 15*
13	Table M Construction		12	16		7*, 12
14	NCCI Retro. Rating					
15	Table L		14			
16	Lee Diagrams, Retro. Rating		13			
17	Limited Table M					16
18	Other Loss Sensitive Plans			1*		
19	Pricing LDD Policies	13, 14, 18, 19	15, 16		2*, 12	8
20	Conclud. Remarks, Indiv. Risk Rat.					

ASOP No. 12 Risk Classification was added to the syllabus for 2017.
It replaced American Academy of Actuaries "Risk Classification Statement of Principles."
For the 2017 exam, many previous readings were replaced by:
a CAS Study Note "Individual Risk Rating," by Fisher, McTaggart, Petker, and Pettingell,
and a CAS Monograph "Loss Distributions for Actuaries," by Bahnemann.

For the 2025 Exam were added: ASOP #25 and Holmes and Casotto.

Integrated questions (which cover material in more than one section) are marked with a star.
The 2017 Exam 8 Sample Integrative Question is in Section 18.

The CAS stopped releasing exams, starting with the 2020 Exam.

CAS Sample Q.11 (from the Fall 2021 Exam 8) is in my Section 1.
CAS Sample Q.6 (from the Fall 2021 Exam 8) is in my Section 18.

Section 1, Mahler, Shifting Risk Parameters[1]

Errata for "An Example of Credibility and Shifting Risk Parameters" by Howard C. Mahler:[2]

Page 286, first sentence:[3]
τ is distributed on the range [-1, 1].
If the actual correlation, $\rho = 0$, then τ is symmetrically distributed on the range [-1, 1].

Page 297, fourth line:[4]

$$\text{Cov}[X_i, X_j] = \begin{cases} \ell(|j-i|)\,\zeta^2 & i \neq j \\ \zeta^2 + \delta^2 & i = j \end{cases}$$

Uses of Credibility:

This paper is principally thinking of two uses of credibility in insurance.

In classification ratemaking, one credibility weights the indicated relativity for a class with that for overall.[5] The indicated relativity for a class is: $\frac{\text{expected pure premium for the class}}{\text{overall expected pure premium}}$.

In individual risk rating, one credibility weights the indicated relativity for an individual risk with a relativity of one.[6]
The indicated relativity for the individual is: $\frac{\text{expected pure premium for the individual}}{\text{expected pure premium for its class}}$.[7]

---

[1] "An Example of Credibility and Shifting Risk Parameters", by Howard C. Mahler, PCAS 1990.
Candidates will not be tested on the Appendices.
CAS Domains/Tasks A2, A4.
[2] Not official.
[3] In Appendix B, not on the syllabus.
[4] In Appendix D, not on the syllabus.
[5] The overall or state relativity is one.
[6] Its class in this case is the bigger item and for this purpose has a relativity of one.
[7] As discussed elsewhere in the syllabus, individual risk rating plans can be more complicated than this simplification.

<u>Shifting Risk Parameters</u>:

Shifting risk parameters: The parameters defining the risk process for an individual insured are not constant over time. There are (a series of perhaps small) permanent changes to the insured's initial risk process as one looks over several years.[8]

For example, a private passenger automobile insured's risk parameters might shift if a major new road were opened in his locality or if he changed the location to which he commutes to work.

In another example, the private passenger automobile insurance experience of a town relative to the rest of the state, in other words the town's relativity, could shift as that town becomes more densely populated.

In yet another example, the procedures and machines used to manufacture widgets change over time. This could result in changes over time in the expected pure premium and therefore the relativity for the Widget Manufacturing Class for Workers Compensation Insurance.

For insurance situations, risk parameters are never totally constant over decades. However, depending on the length of the time period considered and the particular data, the magnitude of the shifts can be large or small.

If risk parameters shift significantly over time, this will significantly effect the optimal credibility to assign to years of past data in order to predict the future.

<u>The Baseball Paradigm</u>:

In Mahler's "An Example of Credibility and Shifting Risk Parameters," the author evaluates various estimates for baseball teams' future losing percentages using historical losing percentages Mahler discusses the impact of shifting parameters over time in this context.

Mahler combines a substantive actuarial topic, the effect of shifting risk parameters on optimal credibility values, with an excellent baseball analogy.
Mahler seeks optimal credibility values, primarily for experience rating but also for class ratemaking, reserving, and other actuarial topics. Section 11 of the paper explains the covariance structure and provides the formulas for estimating optimal credibility values. But many readers of the paper have trouble digesting the theory. The baseball analogy is an excellent means of explaining the intuition.

This is an analogy with the characteristics needed, but without the problems of insurance data.

---

[8] Taken from page 456 of "Credibility With Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty," by Howard C. Mahler. PCAS 1998, <u>not</u> on the syllabus.

1. Insurance applications of credibility are complex, since different size risks have different degrees of partial credibility. The baseball teams all play the same number of games; they are the same size, so there is no need for partial credibilities.
2. Insurance is complicated by loss development. There is no loss development in baseball; when the season is over, we know the won-loss record.
3. An insurance portfolio changes over time, as new insureds are added and as old insureds leave. Mahler has the same baseball teams for 60 years.

The analogy of the baseball example to an insurance industry situation:

losing percentage of baseball team. ⇔ loss ratio of an insured (or class).

losing percentage of team compared to average.
⇔ loss ratio of an insured compared to average. ⇔ relativity of a class.

predicting future losing percentage of a team.
⇔ experience rating an insured. ⇔ determining new class relativity.

Advantages of the Baseball Data:[9]

1. Over a very extended period of time there is a constant set of risks (teams).
In insurance, there are generally insureds who leave the data base and new ones that enter.

2. The loss data over this extended period of time are readily available, accurate and final.
In insurance, the loss data are sometimes hard to compile or obtain and are subject to possible reporting errors and loss development.

3. Each of the teams in each year plays roughly the same number of games.
Thus the loss experience is generated by risks of roughly equal "size."
Thus, in this example, one need not consider the dependence of credibility on size of risk.

Sampling Error:

The use of credibility mitigates distortions caused by sampling error. Part of sampling error is the inability to get accurate readings because the measuring instruments are too crude. We don't get accurate estimates of incurred losses until years after the accident, because we can not observe future court decisions. Mahler wants to avoid this topic, so that he can focus on shifting risk parameters over time. Therefore, Mahler analyses a data set, baseball won-loss records, that mitigates sampling error problems.

[9] See Section 3.1 of the paper.

Team Differences and the Binomial Test:[10] [11]

Mahler demonstrate that baseball losing percentages have the characteristics that are relevant for credibility studies. If all insureds were the same, there would be no use for experience rating. Using the so-called Binomial Test, Mahler shows that the losing percentages of the various teams are not random; there are better teams and worse teams.

If the experience for each team were drawn from the same probability distribution, the results for each team would be much more similar.

A Binomial distribution with a 50% chance of losing, for 9000 games, has a variance of:
9000 (1/2) (1 - 1/2) = 2250. $\sqrt{2250} \cong 47$.
This is a standard deviation of 47 games lost, or 47 / 9000 = 0.5% in losing percentage.
Thus if all the teams' results were drawn from the same distribution, using the Normal Approximation, approximately 95% of the teams would have an average losing percentage between 49% and 51%.

Thus if all the teams' results were drawn from the same distribution, approximately 95% of the teams would have an average losing percentage between 49% and 51%.

Table 3 in the Paper:

Team	1	2	3	4	5	6	7	8
NL	53.4	49.9	47.3	51.8	44.7	56.5	47.8	48.8

Team	1	2	3	4	5	6	7	8
AL	49.5	49.4	47.0	48.5	42.6	52.9	56.4	53.5

Only 3 of 16 teams have losing percentages in that range. The largest deviation from the grand mean is 15 times the expected standard deviation if the teams all had the same underlying probability distribution.

"There can be no doubt that the teams actually differ. It is therefore a meaningful question to ask whether a given team is better or worse than average. A team that has been worse than average over one period of time is more likely to be worse than average over another period of time."

Mahler tests whether experience in one period has predictive power for other periods. Specifically, Mahler shows that **there is a significant correlation between the results of years close in time. Thus recent years can be usefully employed to predict the future.**[12] Thus this is a useful data set to use to investigate experience rating.

---

[10] See pages 229 and 234 in Mahler.
[11] See 9, 11/98, Q.25.
[12] See page 236 of Mahler.

Chi-Square Test of Whether the Risk Parameters Shift Over Time:

Mahler uses two methods to test whether risk parameters shift over time:
(i) He does chi-square tests.
(ii) He examines the correlations in pairs of years separated by a constant period.

The Chi-Square Test as used by Mahler can be summarized as follows:[13]

- Applied to the data of one team.
- H_0: The expected losing percentage is the same over time for this team.
- Group data into appropriate intervals.
  Mahler groups the 60 years into 5 year non-overlapping intervals.
- Calculate the mean losing percentage for the team over the 60 years.
- Then calculate for each interval: $(A - E)^2/E$,
  where A = actual observation = (5 year mean losing percentage)(5 years)(150 games),
  and E = expected observation = (60 year mean losing percentage)(5 years)(150 games).
- Sum up the contributions for all 12 intervals in order to get the chi-square statistic.
- If the statistic is greater than the critical value for number of intervals - 1 = 11 degrees of freedom, then reject the null hypothesis that parameters do not shift over time.

For each team, Mahler finds that there is less than a 0.2% chance that the different five-year segments were all drawn from the same distribution. Therefore, he rejects the hypothesis that the means are the same over time, in favor of the hypothesis that the parameters shift (at a noticeable amount) over time.

Chi-Square Statistics and p-values[14] [15]

NL1	NL2	NL3	NL4	NL5	NL6	NL7	NL8
107	45	98	35	39	73	114	119
7×10^{-18}	5×10^{-6}	4×10^{-16}	0.025%	5×10^{-5}	5×10^{-1}	3×10^{-19}	3×10^{-20}

[13] See Table 4 in Mahler. This is an application of material covered on preliminary exams.
[14] The values of the Chi-Square statistic are taken from Mahler's Table 4. I have added the probability values. Note that all of the p-values are less than 0.2%.
[15] The teams are identified in footnote 6 on page 229 of the paper by Mahler.
For example, AL5 is the New York Yankees.

Correlations Test of Whether the Risk Parameters Shift Over Time:

Here is a description of Mahler's correlation test, as applied to insurance data.

Suppose we have N similar risks and T years. We denote the manual loss ratio for risk n in year t as $LR_{n,t}$. (We use manual loss ratios, not standard loss ratios.)
For each year t, we have N loss ratios $\{LR_{1,t}, LR_{2,t}, ... , LR_{N,t}\}$.

For the one year differential, we examine the correlation of the T - 1 sets of pairs:
$\{LR_{1,1}, LR_{2,1} ... , LR_{N,1}\}$ with $\{LR_{1,2}, LR_{2,2}, ... , LR_{N,2}\}$
$\{LR_{1,2}, LR_{2,2} ... , LR_{N,2}\}$ with $\{LR_{1,3}, LR_{2,3}, ... , LR_{N,3}\}$
etc.
We take the average correlation for the one year differential.

We do the same for the two year differential, using the correlation of the T - 2 sets of pairs:
$\{LR_{1,1}, LR_{2,1} ... , LR_{N,1}\}$ with $\{LR_{1,3}, LR_{2,3}, ... , LR_{N,3}\}$
$\{LR_{1,2}, LR_{2,2} ... , LR_{N,2}\}$ with $\{LR_{1,4}, LR_{2,4}, ... , LR_{N,4}\}$
etc.
We take the average correlation for the two year differential.

We do the similar calculation for the other differentials in years.[16]

If the risk parameters do not shift over time, the average correlation should not differ significantly between the one year differential, two year differential, and so forth. If the risk parameters shift over time, the average correlation should be highest for the one year differential, second highest for the two year differential, and so forth. The rate at which the correlation drops as the differential widens measures how fast the risk parameters shift over time.[17]

Mahler's results in his Table 5 indicate that the risk parameters are shifting at a high rate for the baseball data examined.

---

[16] Results are shown in Table 5 in Mahler.
[17] This is discussed further at pages 640 to 642 of "A Markov Chain Model of Shifting Risk Parameters," by Howard C. Mahler, PCAS 1997, not on the syllabus.

Table 5 from the Paper

	Correlations	
Years Separating Data	NL	AL
1	0.651	0.633
2	0.498	0.513
3	0.448	0.438
4	0.386	0.360
5	0.312	0.265
6	0.269	0.228
7	0.221	0.157
8	0.190	0.124

The correlations decline as the separation increases.

Years further apart are less correlated than years closer together.

Data from last year is more valuable to predict the coming year, than data from 5 years ago.

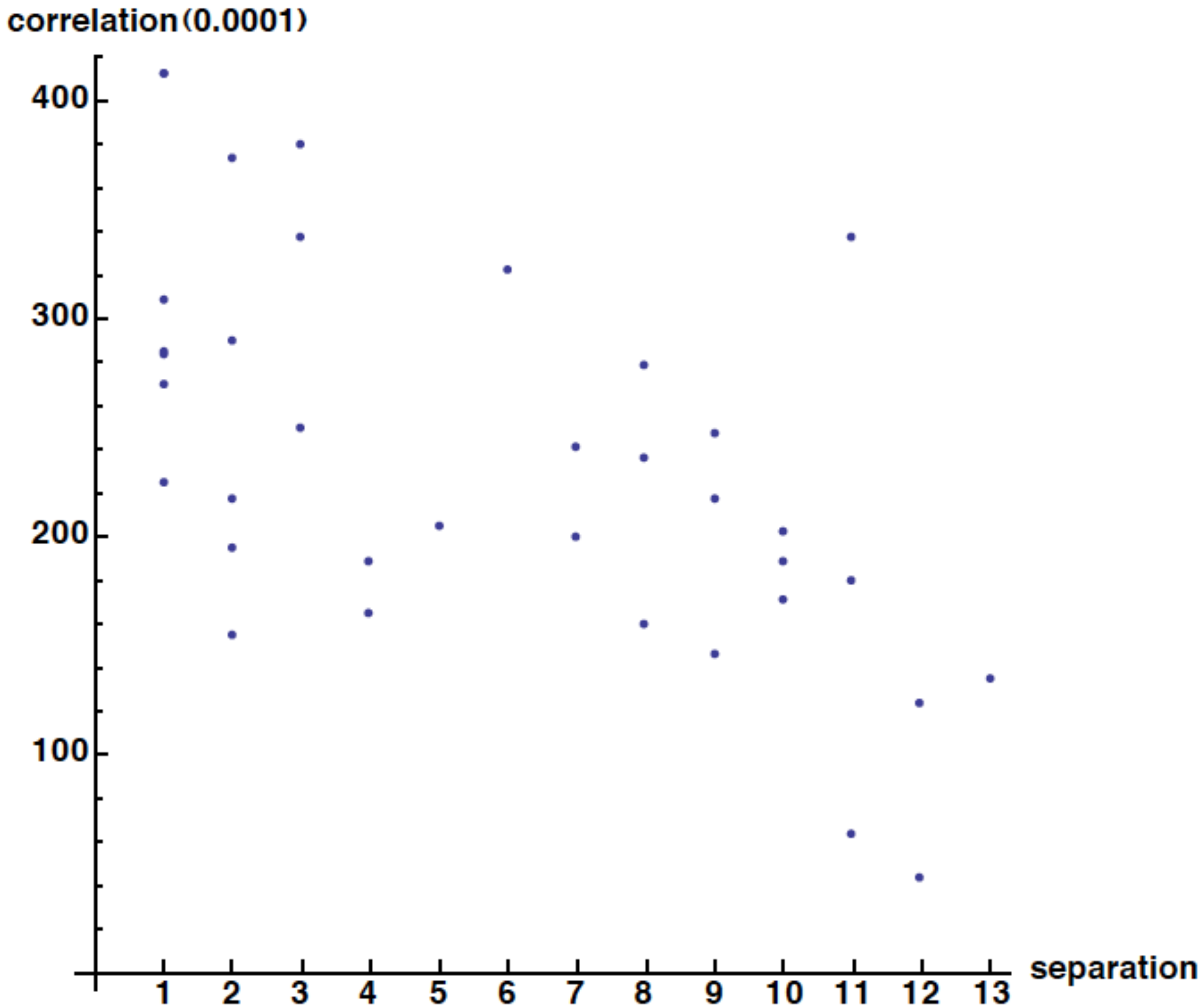
Thus the NCCI Experience Rating Plan, which assuming equal volume of data for each year gives equal weight to each year of data, is an approximation to the theoretically most accurate plan.[18]

[18] There are other complications such as the maturity of the data.
See "Credibility With Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty," by Howard C. Mahler. PCAS 1998, not on the syllabus.

California Driver Data:[19]

A similar correlations test has been performed on data for drivers in California.
The data show the number of accidents annually in 1961-1963 and 1969-1974, for a sample of drivers licensed from 1961 to 1974. There were 54,165 drivers divided between male and female.

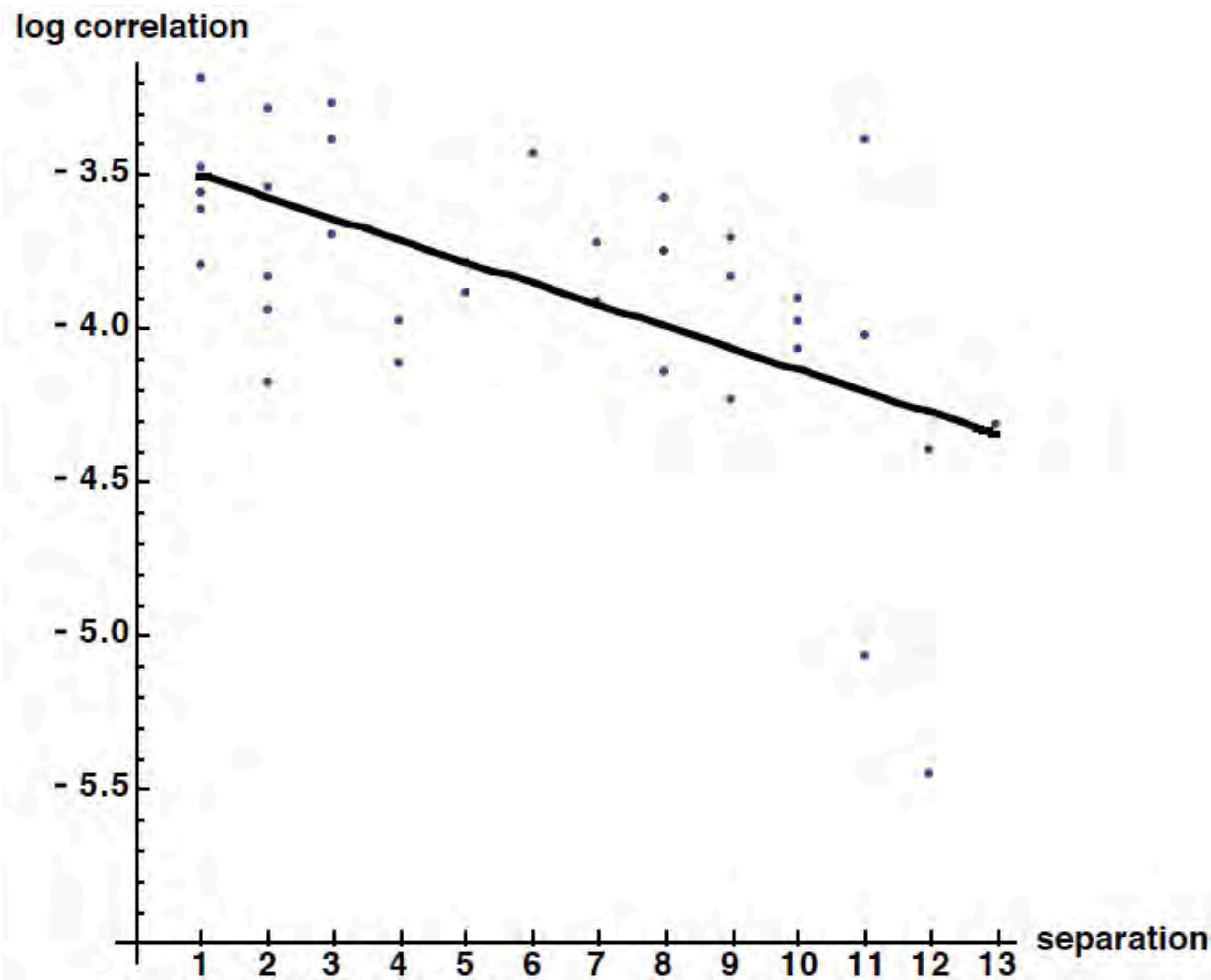
Correlations were computed for pairs of years of data separated by different numbers of years. For example, 1961 and 1962 are separated by one year, while 1961 and 1970 are separated by 9 years. Here is a graph of the results for female drivers:[20] [21]



Real insurance data can be messy; the data and thus the correlations between years are subject to significant random fluctuation. However, the correlations do appear to be declining as the separation between years increases.

[19] From Exhibit 1 in “The Credibility of a Single Private Passenger Driver,” by Howard C. Mahler, PCAS 1991.
[20] Due to the gap in the years of data, some separations have fewer values than one would otherwise expect.
[21] There were two cases where for a separation of one year the correlation was 0.0412.

Here is the same data on a log scale. Also shown is the least squares line fit to the logs of the correlations, -3.435 - 0.06999 x.[22]

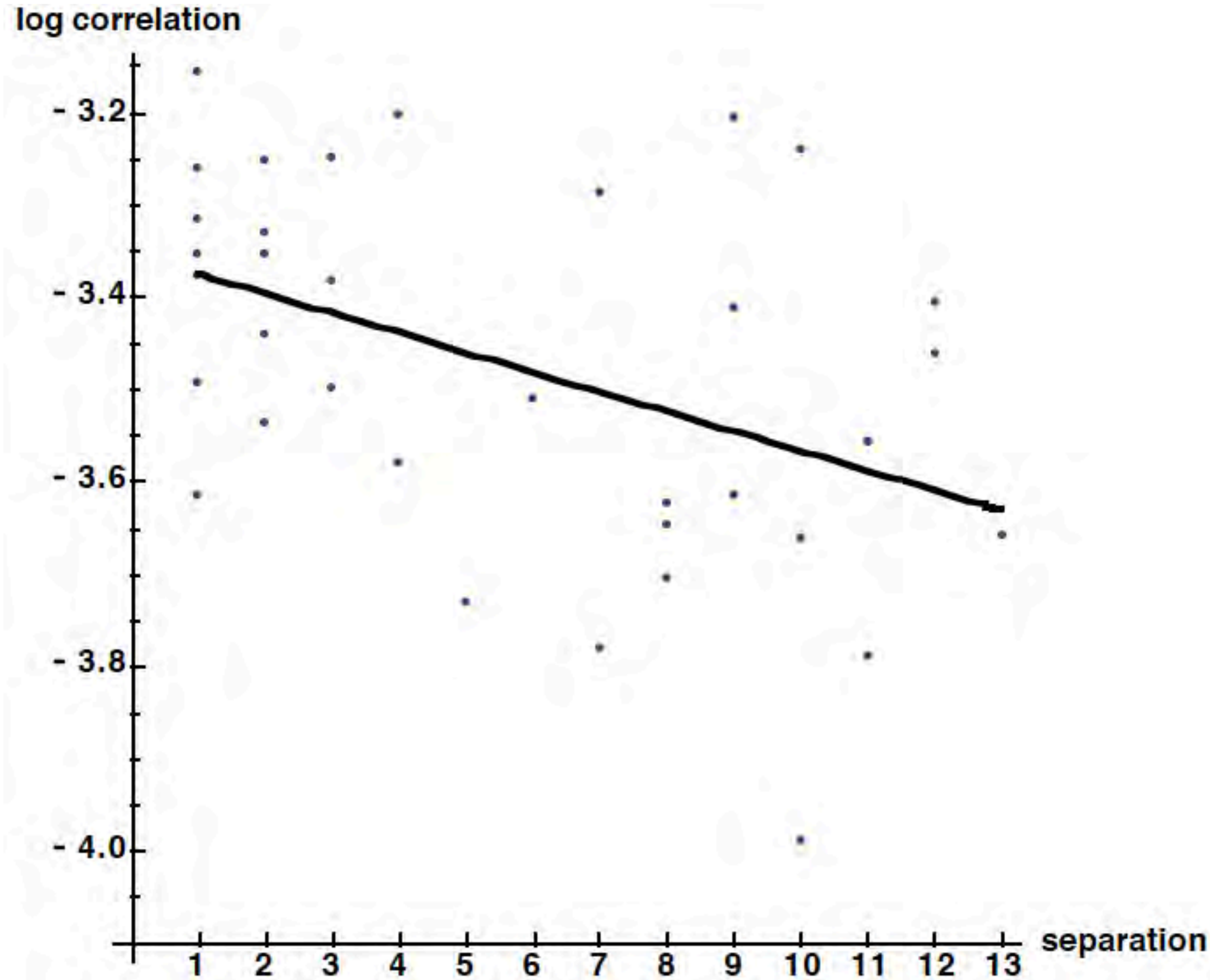


Thus there is evidence that the correlations are declining with separation and thus that parameters are shifting over time. The least squares line is: ln(corr) = -3.435 - 0.06999 x. ⇔ Correlation = (0.0322) (0.932^x), where x is the separation in years.[23]
The 0.932 measures the rate at which parameters are shifting; the further this base is from one, the more quickly parameters are shifting.

---

[22] R^2 = 0.36. The p-value for testing whether the slope is zero is 0.0001; so there is very good evidence that the slope is not zero. A negative slope corresponds to correlations declining with separation.
[23] There is a theoretical reason to expect correlations to follow this type of curve.
See "A Markov Chain Model of Shifting Risk Parameters," by Howard C. Mahler, PCAS 1997.

Here is a similar graph, but for the male drivers. The least squares line fit to the logs of these correlations is: -3.354 - 0.02140 x.[24]



Again there is evidence that the correlations are declining with separation and thus that parameters are shifting over time. The least squares line is: ln(corr) = -3.354 - 0.02140 x.⇔

Correlation = (0.0349) (0.979^x), where x is the separation in years.

The 0.979 measures the rate at which parameters are shifting. For females the similar base was 0.932, indicating parameters are shifting much more quickly for female drivers than for male drivers.[25]

For male drivers, the number of years of separation required for the correlation to decline to half of its original value is: ln(0.5) / ln(0.979) = 33; for females it is: ln(0.5) / ln(0.932) = 10.

---

[24] R^2 = 0.19. The p-value for testing whether the slope is zero is 0.007; so there is good evidence that the slope is not zero. A negative slope corresponds to correlations declining with separation.

[25] This conclusion based on this one data set should be taken with a grain of salt.

<u>The Effect of this Pattern of Correlations</u>:[26]

The correlation between years that are close together is higher than the correlation between years that are further apart. Therefore, the credibility assigned to more recent years of data should be higher for predicting the future.[27]

Delays in receiving data make estimates of the future less accurate.
Therefore, the optimal credibility decreases with increased delays in receiving the data.
When predicting year 5, it is better to have data for years 2, 3, and 4 than for years 1, 2, and 3.[28]

Up to a given point, using more year of data, with an optimal set of credibilities applied to each year, increases the accuracy of the estimate of the future. However, at a certain point adding more older years of data, no longer increases (measurably) the accuracy of the estimate.[29]

With equal weight to each year, at a certain point adding more older years of data, no longer increases the accuracy of the estimate; instead in this case at some point adding older years of data <u>decreases</u> the accuracy of the estimate.[30]

<u>Estimators</u>:

A credibility weighting formulas (credibility estimator) might be 60% of last year's loss ratio plus 40% of the overall average loss ratio. This estimator has two terms; a simple estimator has a single term, such as the overall mean, last year's experience, or the experience from two years ago.

Mahler's credibility estimators are:
1. A linear combinations of a few simple estimators.
2. Unbiased for the set of teams as a whole.
3. Analogous to experience rating.

---

[26] See for example, 9, 11/01, Q.1.
[27] In this paper, Mahler assumes the different years contain the same volume of data.
[28] Ignoring possible complications such as loss development.
See for example, Sections 7.10, 7.11, 7.12 and 10.10 of “Credibility With Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty,” by Howard C. Mahler. PCAS 1998, <u>not</u> on the syllabus.
See also the 9th and 10th pages of “Workers' Compensation Classification Credibilities”, by Howard C. Mahler, Fall 1999 CAS Forum.
[29] See Table 19 in Mahler. The slower the rate of shifting parameters, the longer it takes to reach such a point of diminishing returns.
[30] See Table 19 in Mahler. The slower the rate of shifting parameters, the longer it takes to reach a point where one should stop adding older years.

<u>The Three Criteria</u>:[31]

Mahler discusses the use of three criteria to determine optimal credibilities:
1. Least Squares Error.[32]
2. Small chance of a large error.[33]
3. Meyers/Dorweiler

If the predicted value is E (expected) and the observed value is O, the squared error is $(E - O)^2$.
To find the optimal credibility formula, we write SE = squared error = $\Sigma (E - O)^2$ as a function of the credibility Z and we set to zero the partial derivative of the squared error with respect to Z. The estimator (credibility formula) that gives the smallest squared error, on average, is the best. We minimize the <u>expected</u> squared error, not the squared error for a particular estimate.

The chance of a large error is the probability that the absolute value of (E - O)/E is more than a given number k. The absolute error is: $\left| \frac{\text{observed}}{\text{predicted}} - 1 \right|$.[34]

Small chance of large error chooses the credibility formula that minimizes
Prob[|(Expected - Observed)/Expected | > k].
The estimator (credibility formula) that gives the smallest number of large errors is the best.

Meyers/Dorweiler is different.[35] Perhaps the optimal experience rating plan uses 3 years of data and 40% credibility, but we use a plan with 6 years of data and 50% credibility. We fear that there may be patterns in the errors, meaning that the underwriter prefers to write either risks with credit modifications or risks with debit modifications. No matter the magnitude of the errors in the experience rating plan, the plan passes the Meyers/Dorweiler test if underwriters are indifferent between credit risks and debit risks.[36]

Meyers/Dorweiler criterion is concerned with the pattern of the errors. Unlike the other two criteria, large errors are not an issue for the Meyers/Dorweiler criterion, as long as there is no pattern relating the errors to the experience modification.

---

[31] See Section 7 of Mahler. Know the three methods, how they work, and any unusual characteristics.
These methods - and particularly the Meyers/Dorweiler method - form the basis of likely exam questions.
[32] The basis of Buhlmann Credibility or greatest accuracy credibility.
[33] The idea behind Classical Credibility.
[34] “The second criterion deals with the probability that the observed result will be more than a certain percent different than the predicted result. The less this probability. the better the solution.”
This verbal description matches the formula with predicted (expected) rather than observed in the denominator.
[35] Taken from Glenn G. Meyers in “An Analysis of Experience Rating”, PCAS 1985, based upon the ideas of Paul Dorweiler.
[36] Dorweiler's view is quoted in “Workers Compensation Experience Rating, What Every Actuary Should Know,” by Gillam at page 218: “A necessary condition for proper credibility is that the credit risks and debit risks equally reproduce the permissible loss ratio.”
See also “Experience Rating - Equity and Predictive Accuracy,” by Venter, at page 7: “On a standard premium basis . . . the loss ratios should be less dispersed, and, ideally, all equal for a better working plan.” and at page 2: “From the viewpoint of the insurer, after experience rating, all insureds have the same expected profit potential, regardless of their past loss history.”

The Meyers/Dorweiler criterion uses Kendall's τ (tau), a measure of correlation.[37]
The optimal credibility using the Meyers/Dorweiler criterion has a Kendall's tau of 0.
We measure the correlation of:
1. (actual losing percentage)/(predicted losing percentage), and
2. (predicted losing percentage)/(overall average losing percentage).

Item #2 is analogous to the experience modification.[38]
Item #1 is analogous to the modified loss ratio, the ratio of losses to modified premium.[39] [40]
Thus the Meyers/Dorweiler criterion desires that the correlation between the experience modification and the modified loss ratio be zero.

If this correlation were positive, then debit risks, those with modifications greater than 1, would tend to have larger modified loss ratios. In other words, after applying the experience rating plan, underwriters would on average not want to write debit risks. Credit risks would tend to have smaller modified loss ratios. In other words, after applying the experience rating plan, underwriters would on average want to write credit risks.

If this correlation were negative, then debit risks, would tend to have smaller modified loss ratios. In other words, after applying the experience rating plan, underwriters would on average want to write debit risks. Credit risks would tend to have larger modified loss ratios. In other words, after applying the experience rating plan, underwriters would on average not want to write credit risks.

Unlike the other two criteria, the Meyers/Dorweiler criterion can not be used to distinguish between using different number of years of data. For each value of N, there is a value of Z such that the correlation is zero.[41]

---

[37] The details of computing Kendall's tau are in Appendix B of Mahler, <u>not</u> on the syllabus.
It involves comparing the ranked order of the two vectors.
[38] If for example, the predicted losing percentage is 60%, then the ratio to the average losing percentage is 60%/50% = 1.2. This is similar to an experience modification factor of 1.2; this team (insured) is predicted to be worse than average. Similarly, a predicted losing percentage of 45% corresponds to an experience modification factor of 45%/50% = 0.9; this team (insured) is predicted to be better than average.
[39] Modified premium = (manual premium)(experience modification).
The modified premium is what would be called standard premium in Workers Compensation.
[40] Modified loss ratio = losses/{(manual premium)(experience modification)}.
Losses ⇔ actual losing percentage. manual premium ⇔ overall mean = 50%.
experience modification ⇔ (predicted losing percentage)/(overall average losing percentage).
Therefore Item #1 = (actual losing percentage)/(predicted losing percentage) ⇔
Losses/{(manual premium)(experience modification)} = Modified loss ratio.
[41] See page 249 of Mahler.

Testing an Experience Rating Plan:

There are several ways to test an experience rating plan:

- We examine whether credit risks or debit risks are more profitable. If credit risks are more profitable than debit risks, then the experience rating credibility is too low; we should give credit risks bigger credits. If credit risks are less profitable than debit risks, then the experience rating credibility is too high; we should give credit risks smaller credits.
- The efficiency test is conceptually the same as the credit vs debit above, but it uses five categories of risks, ranked in order of the experience modifications, instead of two.[42]
- *The ratio of variances generalizes the efficiency test: we rank the risks by their modifications into N groups, from lowest mods to highest mods. We determine the average manual and standard loss ratios in each group, and we compute the variance of the average standard loss ratios divided by the variance of the average manual loss ratios.*
  The lower the ratio of the variances, the better the experience rating plan.
- The Meyers-Dorweiler test uses the Kendall τ statistic for the correlation between the actual loss ratio relativities and the indicated loss ratio relativities.
- The minimum squared error test sums the squared errors between the actual loss ratio relativity and the indicated loss ratio relativity; the lower the sum of the squared errors, the better the experience rating plan. *Alternatives to the minimum squared error test are the minimum χ^2 test and the minimum absolute error test.*
- *Let μ be the expected loss ratio for an insured prior to experience rating. Let $M = E[\mu]$ over a group of insureds. Let F be the estimator of* m*, in this context the result of using an experience rating plan. Then the efficiency is of F is:* $1 - \frac{E[(m - F)^2]}{E[(m - M)^2]}$.

  The higher the efficiency, the better the experience rating plan.[43]

---

[42] See "Experience Rating - Equity and Predictive Accuracy," by Gary G. Venter.
According to William R. Gillam at pages 219-220 of "Workers Compensation Experience Rating: What Every Actuary Should Know", "The test statistic for each size group is the variance of the modified ratios divided by the variance of the unmodified ratios. A low test statistic indicates a plan that has eliminated much of the between variance (in risk theoretic terms) or made risks of differing experience more equally desirable."

[43] This is the efficiency test of Glenn G. Meyers in "An Analysis of Experience Rating", PCAS 1985, mentioned at page 220 of "Workers Compensation Experience Rating: What Every Actuary Should Know," by William R. Gillam. Meyers applies efficiency to models where the risk parameters vary between insureds within a group. In such models we are assumed to know the expected loss ratio for each insured, and we see how well the experience rating plan works for a set of data generated from this group.

Rating plans which do well on one test often do well on other tests. But the tests examine different characteristics of the rating plan. Some tests check for bias, often referred to as patterns of errors (credit-debit; quintiles; Meyers-Dorweiler) and some tests check for accuracy (minimum squared error, minimum χ^2, minimum absolute error, ratios test).

A plan is biased if the experience modification helps us select among risks. For example, suppose we gave all risks 10% credibility, but the proper credibility is higher. A risk with a credit modification is overpriced, since the true experience modification would be lower with greater credibility, and a risk with debit modification is underpriced, since the true experience modification would be higher with greater credibility.

In a perfect plan, the loss ratio relativity predicted by the plan would be the expected relativity.[44]

- The plan is <u>unbiased</u> if no value of the experience modification is a predictor of rate redundancy or inadequacy versus other risks after application of the modification, in other words with respect to standard premium.
- The plan is <u>accurate</u> if the difference between the predicted loss ratio and the expected loss ratio is zero; any differences between the predicted loss ratio and the actual loss ratio stem from random loss fluctuations.

We desire an experience rating plan that is as close to unbiased and as accurate as practical.[45]

Whether a risk is a debit risk or a credit risks depends on the plan. It is tempting to presume that the credit risks are the risks with expected loss ratio relativities less than one and the debit risks are the risks with expected loss ratio relativities more than one. This is not correct, since we do not know the expected losses for any risk.

Rather, the credit risks are the risks with experience modifications below one. Whether a risk is a credit risk or a debit risk depends on the plan parameters, such as the expected losses, the state accident limit, the primary-excess split, and the credibilities. For a particular plan, we desire that most of the credit risks actually are better than average, and that most of the debit risks are actually worse than average.

---

[44] No actual experience rating plan is ever perfect.

[45] As mentioned by Venter in "Experience Rating - Equity and Predictive Accuracy," we also want the experience rating plan to provide incentives for the insured to reduce losses.

<u>An Example of Comparing Experience Rating Plans</u>:

Consider five prototypical insureds of similar size. We show the experience modifications predicted by two rating plans, P and Q. We show the subsequently observed loss ratio to manual premium relativities, for the period of time predicted by the experience rating plans.[46] [47]

Risk	Experience Modification		Subsequently Observed Manual Loss Ratio Relativity
	P	Q	
1	0.75	0.86	0.71
2	0.80	0.90	0.79
3	0.91	0.94	0.94
4	1.05	1.02	1.14
5	1.44	1.26	1.42

Both plans P and Q seem to do a reasonable of predicting which risks will be better than average and which risks will be worse than average. Either plan would be better than no experience rating.

Let us look at the loss ratios to standard premium relativities for each plan.
For example, for Risk 1 for Plan P, 0.710/0.750 = 0.947.

Risk	Manual L.R.	Mod for P	Standard L.R. for P	Mod for Q	Standard L.R. for Q
1	0.710	0.750	0.947	0.860	0.826
2	0.790	0.800	0.988	0.900	0.878
3	0.940	0.910	1.033	0.940	1.000
4	1.140	1.050	1.086	1.020	1.118
5	1.420	1.440	0.986	1.260	1.127

Ideally we would like the loss ratios to standard premium to be similar for debit and credit risks; in other words after the application of experience rating all risks should ideally have the same expected loss ratio.[48]

---

[46] Any useful comparison would involve thousands of insureds. We show 5 solely for illustrative purposes.
[47] Analogous to Exhibit 3, Part 2 of "Parameterizing the Workers Compensation Experience Rating Plan," by William R. Gillam, <u>not</u> on the syllabus.
The values shown in Gillam are for risks grouped into quintiles.
The lowest quintile for Plan P, would be the insureds with the 20% lowest modifications using Plan P.
The lowest quintile for Plan Q, would be the insureds with the 20% lowest modifications using Plan Q.
The lowest quintile for Plan P, would be similar to the lowest quintile for Plan Q, but would consist of a somewhat different set of insureds.
[48] For thousands of risks, the observed loss ratio for a quintile would be close to the expected loss ratio.
For a single insured, this need not be the case.

We see that for the best and worst risks, Plan P does a better job of this than Plan Q.
Plan P appears to be more responsive than Plan Q; in other words Plan P assigns a higher credibility to the insureds own experience.[49] [50]

The Meyers/Dorweiler criteria would compute the correlation between the loss ratios to standard premium and the experience modification. We prefer this correlation to be close to zero.

Meyers and Mahler use the Kendall τ statistic; however, how to compute that is in Appendix B of Mahler, not on the syllabus.[51] For illustrative purposes we can use the usual sample correlation:
$r = \hat{Cov}[X, Y] / (s_X s_Y) = \Sigma(X_i - \bar{X})(Y_i - \bar{Y}) / \sqrt{\Sigma(X_i - \bar{X})^2 \ \Sigma(Y_i - \bar{Y})^2}$.

We take X_i to be the standard loss ratio relativities, and Y_i to be the experience modification for a given plan.

Exercise: For Plan P, calculate the sample correlation between the loss ratios to standard and the experience modifications.
[Solution: X = (0.947, 0.988, 1.033, 1.086, 0.986).
Y = (0.75, 0.8, 0.91, 1.05, 1.44).
$\bar{X}$ = 1.008. $\bar{Y}$ = 0.99.
$s_X^2 = \{(0.947 - 1.008)^2 + ... + (0.986 - 1.008)^2\} / (5 - 1) = 0.0028285$. $s_Y^2 = 0.07655$.
$\hat{Cov}[X, Y] = \{(0.947 - 1.008)(0.75 - 0.99) + ... + (0.986 - 1.008)(1.44 - 0.99)\} / (5 - 1) = 0.002805$.
$r = 0.002805 / \sqrt{(0.0028285)\ (0.07655)} = 0.1906$.]

For Plan P, r = 0.19, while for Plan Q, r = 0.83. Thus by this criterion, very similar to the Meyers/Dorweiler criteria, Plan P is better than Plan Q.

The efficiency test would compare the squared deviations from the mean before and after experience rating.[52] We would like the squared deviations after experience rating to be small; the experience rating plan should adjust for risk differences.

We group the insureds into five equally sized groups. The first group would contain those insureds with the smallest modifications under the given experience rating plan. The final group would contain those insureds with the largest modifications under the given experience rating plan.[53]

We would like the squared deviations before experience rating to be large;
the experience rating plan should identify risk differences.

---

[49] There may be other differences between the plans, such as whether and how they split the losses into primary and excess.
[50] Plan R that assigned even more credibility to this size of risk than Plan P, might do worse than Plan P. Higher credibility is not always better!
[51] See pages 300-302 of Meyers, "An Analysis of Experience Rating," PCAS 1985, not on the syllabus.
[52] See Section 2.10 of "Individual Risk Rating."
[53] In this illustrative example, we only have five insureds, so each quintile has just one insured.

For the manual premium loss ratio relativities, the average is 1.00.
The sum of squared differences from the mean is:
$(0.71 - 1)^2 + (0.79 - 1)^2 + (0.94 - 1)^2 + (1.14 - 1)^2 + (1.42 - 1)^2 = 0.3278$.

For the standard premium loss ratio relativities for Plan P, the average is 1.008.
The sum of squared differences from the mean is:
$(0.947 - 1.008)^2 + (0.988 - 1.008)^2 + (1.033 - 1.008)^2 + (1.086 - 1.008)^2 + (0.986 - 1.008)^2 = 0.0113$.

For the standard premium loss ratio relativities for Plan Q, the average is 0.990.
The sum of squared differences from the mean is:
$(0.826 - 0.990)^2 + (0.878 - 0.990)^2 + (1.000 - 0.990)^2 + (1.118 - 0.990)^2 + (1.127 - 0.990)^2 = 0.0747$.

The efficiency test statistic is: [54]

$$\frac{\text{sum of squared differences from the mean of the standard ratios}}{\text{sum of squared differences from the mean of the manual ratios}} = \frac{\text{the variance of the standard ratios}}{\text{the variance of the manual ratios}}.$$

Smaller efficiency test statistic is better.

For Plan P, the efficiency test statistic is: 0.0113/0.3278 = 0.034.
For Plan Q, the efficiency test statistic is: 0.0747/0.3278 = 0.228.

Therefore, based on the efficiency test, Plan P works better than Plan Q.[55]
After the application of the experience modifications from Plan P, the loss ratios to standard vary less among the quintiles of insureds, than they do for Plan Q.

[54] See Section 2.10 of "Individual Risk Rating."
[55] For this size of risk, and for the limited number of risks looked at for illustrative purposes.

Kendall's Tau:[56]

Kendall's tau is a measure of correlation that depends on ranks.
Kendall's tau is not as sensitive to strong outliers as Pearson's correlation coefficient

In order to compute Kendall's tau, the first step is to order the first elements of the pairs from smallest to largest. Then list the resulting ranks of the second elements.[57]

For example let us take eight pairs of heights of fathers and their adult son:

Father	Son	Rank	Concordant	Discordant
53	56	1	7	0
54	58	2	6	0
57	61	4	4	1
58	60	3	4	0
61	63	6	2	1
62	62	5	2	0
63	65	8	0	1
66	64	7		
Sum			25	3

The number concordant listed in a row is the number of ranks below it in the column that are greater than the given rank. The number discordant listed in a row is the number of ranks below it in the column that are less than the given rank.

Then $\tau = \frac{C - D}{C + D} = \frac{25 - 3}{25 + 3} = 0.7857$.
Note that the denominator, 25 + 3 = 28 = (8)(8 - 1) / 2 = n (n-1)/2.

[56] See Appendix B of Mahler, not on the syllabus.
[57] One would get the same Kendall's correlation by instead ordering the second elements of the pairs from smallest to largest, and then listing the resulting ranks of the first elements.

When there are no ties, in order to calculate Kendall's tau:[58]

1. Order the first elements of the pairs from smallest to largest.
2. List the resulting ranks of the second elements of the pairs.
3. The number concordant listed in a row is the number of ranks below it in the column that are greater than the given rank. C = sum of concordants.
4. The number discordant listed in a row is the number of ranks below it in the column that are less than the given rank. D = sum of discordants.
5. $\tau = \frac{C - D}{C + D} = \frac{C - D}{n(n-1)/2}$.

Exercise: You are given six risks of similar size.

Risk	Experience Modification	Subsequent Loss Ratio to Standard Premium
A	1.00	74%
B	0.70	66%
C	1.20	107%
D	1.40	88%
E	0.80	71%
F	0.90	63%

Calculate Kendall's tau between the experience modifications and the subsequent loss ratios to standard premium.
[Solution: Order the risks by the rank of their experience modification.

Risk	Experience Mod.	Loss Ratio to Stand. Prem.	Concordant	Discordant
B	0.70	66%	4	1
E	0.80	71%	3	1
F	0.90	63%	3	0
A	1.00	74%	2	0
C	1.20	107%	0	1
D	1.40	88%		
Sum			12	3

$\tau = \frac{C - D}{C + D} = \frac{12 - 3}{12 + 3} = 0.6.$

<u>Comment</u>: If Kendall's tau were close to zero, that would indicate that an Experience Rating Plan is working well according to the Meyers/Dorweiler criterion. For a practical application, we would look at many more than 6 insureds.]

[58] Things get a little more complicated when there are ties.

Assuming independence, and thus that the actual correlation is zero, Kendall's tau has a mean of zero and variance of: $\frac{2\,(2n+5)}{9n\,(n-1)}$.

Thus one can use Kendall's Rank Correlation Coefficient and the Normal approximation to test the hypothesis that there is no relationship between the two samples.[59]

$$Z = \tau \sqrt{\frac{9n\,(n-1)}{2\,(2n+5)}}.$$

One can perform the usual two-sided and one-sided tests.

For the heights example, with a sample size of 8, τ = 0.7857.

H_0: The correlation of the joint distribution from which the paired samples were drawn is zero.

H_1: The correlation of the joint distribution from which the paired samples were drawn is positive.

$$Z = 0.7857 \sqrt{\frac{(9)(8)(7)}{(2)(21)}} = 2.722.$$

Thus for this one-sided test, the probability-value is: 1 - Φ[2.722] = 0.32%.
We reject H_0 at a 0.5% level.
In other words, at a 0.5% significance level we conclude that there is a positive correlation between the heights of fathers and sons; taller fathers tend to have taller sons.

Exercise: For 200 experience rated risks of a similar size, an actuary calculates Kendall's tau between the experience modifications and the subsequent loss ratios to standard premium.
τ = -0.03.

H_0: The correlation between the experience modifications and the subsequent loss ratios to standard premium is zero.

H_1: The correlation between the experience modifications and the subsequent loss ratios to standard premium is not zero.

What is the probability-value of this test?

[Solution: Var[τ] = $\frac{2\,(2n+5)}{9n\,(n-1)} = \frac{2\,(405)}{(9)\,(200)\,(199)}$ = 0.002261.

Using the Normal approximation, Z = -0.03 / $\sqrt{0.002261}$ = -0.631.

For this two-sided test, the probability-value is: 2 Φ[-0.63] = 53%.

<u>Comment</u>: Since τ is close to zero, by the Meyers-Dorweiler criterion, the experience rating plan is doing a good job of correcting for risk differences.]

---

59 In practical applications, one would use an exact statistical table for sample sizes of 10 or less.
For example, for a one-sided test with n = 7, one would reject at 5% for τ > 11/21, and reject at 1% for τ > 15/21.

Geometrically Declining Weights:[60]

Parallel to one of the examples in Mahler, pure premium for a class are projected based on the formula:[61]

$E = Z X + (1 - Z) P$, where
X = the most recent accident year's pure premium
P = the prior estimate of the most recent accident year
Z = the credibility assigned to the most recent accident year

Exercise: Assume no delay in obtaining data and Z = 20%.
What is the weight given to accident year 2007 data in the estimate of accident year 2009?
[Solution: $P_{2009} = Z X_{2008} + (1 - Z) P_{2008} = Z X_{2008} + (1 - Z) \{Z X_{2007} + (1 - Z) P_{2007}\} =$
$Z X_{2008} + (1 - Z) Z X_{2007} + (1 - Z)^2 P_{2007} = 0.2 X_{2008} + 0.16 X_{2007} + 0.64 P_{2007}$.
The weight given to AY 2007 data is 16%.
Comment: The weight given to AY 2008 data is 20%.
The remaining weight of 64% is given to the prior estimate of the 2007 pure premium; this estimate was based in turn on data from years prior to 2007.]

In general, let the credibility be Z for the latest experience.
If we forecast for year t, and there is no delay, then the weight given to each past year of data is:[62]

Year	Weight
t-1	Z
t-2	$Z (1 - Z)$
t-3	$Z (1 - Z)^2$
t-4	$Z (1 - Z)^3$
t-n	$Z (1 - Z)^{n-1}$

As $Z \to 100\%$, we give full weight to the most recent experience and no weight to older experience. As $Z \to 0\%$, the weights for each year become similar and very small.

[60] See page 255 of Mahler. See 9, 11/05, Q.2.
[61] The most recent experience has been developed to ultimate, and has been adjusted for trend and any other changes. The prior estimate has been been adjusted for trend and any other changes. Such complications do not occur with the baseball data.
[62] These weights are from a Geometric Distribution. The weight given to year t-n is f(n-1).
Using the notation in Loss Models, $\beta/(1 + \beta) = 1 - Z$, or $\beta = (1 - Z)/Z = 1/Z - 1$.
This is an example of (single) exponential smoothing.

If risk parameters shift at a faster rate, then all of the past years become a worse predictor of the future. However, more recent experience becomes relatively more useful than older experience to predict the future. For example, as a predictor of 2009, 2007 data is more affected by an increase in the rate of shifting than is 2008 data. Since all of the weight is being applied to some past year of data, the weight to the most recent year of data increases.

Therefore, if the risk parameters shift at a faster rate, then Z increases.
If instead the risk parameters shift at a slower rate, then Z decreases.

Geometrically Declining Weights with Delay:

This form of estimator is similar to pure premium ratemaking, where the credibility weighted pure premium is: Z (the indicated pure premium) + (1 - Z) (the underlying pure premium).
In insurance applications we usually have a delay in getting information.

Exercise: Assume a delay in obtaining data. For example, we have year 2007 data available to predict year 2009, but do not have 2008 data available at that time. Z = 20%. What is the weight given to accident year 2005 losses in the estimate of accident year 2009 losses?
[Solution: $P_{2009} = Z\,X_{2007} + (1 - Z)\,P_{2008} = Z\,X_{2007} + (1 - Z)\,\{Z\,X_{2006} + (1 - Z)\,P_{2007}\} =$
$Z\,X_{2007} + (1 - Z)\,Z\,X_{2006} + (1 - Z)^2\,P_{2007} =$
$Z\,X_{2007} + (1 - Z)\,Z\,X_{2006} + (1 - Z)^2\,\{Z\,X_{2005} + (1 - Z)\,P_{2006}\} =$
$Z\,X_{2007} + (1 - Z)\,Z\,X_{2006} + (1 - Z)^2\,Z\,X_{2005} + (1 - Z)^3\,P_{2006}\}$.
The weight given to 2005 losses is: $(1 - Z)^2\,Z = (0.8^2)(0.2) = 12.8\%$.]

Least Squares Credibilities:[63]

The least squares credibilities minimize the expected squared error between the estimate and the observation.[64] [65] The least squares credibilities depend on the years used in the estimator as well as the assumed covariance structure.[66] Table 16 shows the resulting credibilities for the covariance structure underlying the years of baseball data, with no delay.[67]

Portion of Table 16 in the Paper					
Number of Years of Data Used	Years Between Data and Estimate				
	1	2	3	4	5
1	66.0%	-	-	-	-
2	57.7%	12.6%	-	-	-
3	56.1%	4.8%	13.5%	-	-
4	55.6%	4.6%	11.5%	3.5%	-
5	55.7%	5.1%	11.7%	6.0	-4.4%

Let us interpret this table. Let us assume we are trying to predict 1960.

If we use 1959, Z = 66.0%.
We give the remaining weight of 34.0% to the overall mean relativity of 1.
If instead we use 1958 and 1959, then we weight 1958 12.6% and weight 1959 57.7%.
We give the remaining weight of 29.7% to the overall mean relativity of 1.

If we use 1957, 1958 and 1959, then we weight 1957 13.5%, 1958 4.8% and 1959 56.1%.
We give the remaining weight of 25.6% to the overall mean relativity of 1.
If we use 1956, 1957, 1958 and 1959, then we weight 1956 3.5%, 1957 11.5%, 1958 4.6%, and 1959 55.6%. We give the remaining weight of 24.8% to the overall mean relativity of 1.

We notice that using two years of data, due to shifting risk parameters over time, the more recent year 1959 is given more weight than the more distant year 1958. This follows from the fact that 1959 is more closely correlated with 1960 than is 1958.

The pattern for more years of data gets more complicated. Some of that is due to the specific values for the covariances used here in the paper.[68]

---

[63] See Section 11 of the paper.
[64] I subsequently show how to solve the linear equations in order to solve for the least squares credibilities.
[65] For the given form of linear estimator. So for example, we would specify in advance that we using a linear combination of years 1, 2 and 3 and the overall mean, in order to estimate year 4.
[66] Unlike in Table 9 in the paper, we allow different years of data to be given different weight.
Unlike in Table 9 in the paper, here we work with an assumed covariance structure based on the baseball data, rather than working directly with the baseball data.
[67] I will subsequently discuss the linear equations that are solved for the least squares credibilities.
[68] Subsequently, I show a similar example with a more regular pattern of credibilities.

For three years of data, 1959 is given by far the most weight of 56.1%, but 1957 is given weight of 13.5%, which is more than the 4.8% given to year 1958. This is due to an "edge effect". 1957 is more closely correlated with 1956 and earlier years than is 1958. By giving somewhat more weight to 1957, we in some sense capture some information about years 1956 and prior. Thus we end up getting a better estimate of 1960.

For five years of data, one of the weights is negative. This can happen; there is nothing in the mathematics to prevent it. In some cases, giving negative weight to one year allows one to give more weight to another year and reduce the expected squared error. If one desired, one could constrain each of the weights to be at least zero and no more than one, as one would want in the case of items labeled credibilities.

Table 19 in the paper compares the mean squared errors of different situations.[69]

Portion of Table 19 in the Paper, Using the Credibilities from Table 16	
Number of Years of Data Used	Mean Squared Error (0.0001)
1	52
2	51
3	49
4	48
5	48

We note that for example, using 2 years of data is a special case of the using three years of data with one of the credibilities constrained to be zero. Thus as we use more years of data, with varying credibilities by year, the minimum expected squared error declines.[70]

With varying credibilities by year, using more years of data leads to a smaller mean squared error.[71]
Given the number of years of data to be used, we solve for the least squares credibilities, with separate credibilities assigned to each year. Using the most recent two years of data is the same as using three years and setting Z = 0 for the most distant year. We can do at least as well and usually better if we solve for the best credibilities when we use three years of data, rather than setting one of them equal to zero.[72]

When using <u>varying</u> weights by year, including more years of data usually decreases the minimum expected mean squared error, although eventually it stays the same.

---

[69] Subsequently, I will discuss how to calculate the expected mean squared error.
[70] The minimum mean squared error for using 5 years of data is slightly less than that for using 4 years of data, even though in the table they round to the same value. For this particular example, after about 6 years of data one reaches a point of extremely small improvement from using more years of data.
[71] See also page 260 of the paper by Mahler
[72] This is similar to the idea that the loglikelihood for the maximum likelihood Gamma Distribution must be at least as good as the loglikelihood for the maximum likelihood Exponential Distribution, since the Exponential is a special case of the Gamma with $\alpha = 1$.

Rather than separate credibilities by year, instead one could give each year the same weight.

Portion of Table 17 in the Paper		
Number of Years of Data Used	Credibility	Z/N
1	66.0%	66.0%
2	70.3%	35.2%
3	72.9%	24.3%
4	73.6%	18.4%
5	72.2%	14.4%

Thus for example, if estimating 1960 using three years data with equal weights, we would give each of 1957, 1958, and 1959 weight 24.3%.

Table 19 in the paper compares the mean squared errors of these different situations.[73]

Portion of Table 19 in the Paper		
Number of Years of Data Used	Mean Squared Error (0.0001)	
	Differing Credibilities	Equal Weights
1	52	52
2	51	54
3	49	55
4	48	57
5	48	60

Using one year of data, the two cases are identical. Using two or more years of data, having the weights constrained to be equal is a special case of varying weights, and thus can not do as well.

Thus when we use equal weights, the minimum expected squared error is greater than or equal to that using weights that are not necessarily equal. For example, for two years of data 54 > 51.

With equal weights, using more years of data is not a special case of using fewer years of data. Thus the mean squared errors do not necessarily decrease as we increase the number of years used. In fact, **due to shifting risk parameters, when using equal weights, eventually including more years of data increases the minimum expected mean squared error**.[74]
In fact, in this case, two years with equal weights does worse than using one year of data.

---

[73] Subsequently, I will discuss how to calculate the mean squared error.
[74] At page 245 of the syllabus reading, discussing the mean squared error criterion when applying equal weights to each year of data: "The results of applying the first criterion are shown in Table 6. Based on most actuarial uses of credibility, an actuary would expect the optimal credibilities to increase as more years of data are used. In this example they do not. In fact, using more than one or two years of data does an inferior job according to this criterion. This result is to be expected, since the parameters shift substantially over time. Thus the use of older data (with equal weight) eventually leads to a worse estimate."

Table 18 in the paper shows credibilities with no weight given to the overall mean, so that the credibilities are constrained to add to one.[75]

Portion of Table 18 in the Paper					
Number of Years of Data Used	Years Between Data and Estimate				
	1	2	3	4	5
1	100.0%	-	-	-	-
2	72.6%	27.4%	-	-	-
3	66.1%	10.3%	23.6%	-	-
4	63.5%	9.1%	16.0%	11.4%	-
5	63.1%	8.7%	15.8%	9.5%	2.9%

For example, if using 1957, 1958, and 1959 to estimate 1960, we would give 1957 weight 23.6%, 1958 weight 10.3%, and 1959 weight 66.1%.

Table 19 in the paper also compares the mean squared errors depending on whether there is weight given to the overall mean or not.

Portion of Table 19 in the Paper		
Number of Years of Data Used	Mean Squared Error (0.0001)	
	Weight to Overall Mean	No Weight to Overall Mean
1	52	63
2	51	58
3	49	54
4	48	52
5	48	52

Again using fewer years of data is a special case of using more years of data. Thus the minimum mean squared errors decline as more years of data are used, with very limited improvement eventually. Having the weight to the overall mean constrained to be zero is a special case of using the optimal weight on the overall mean, so the mean squared error is at least as big. For example, for two years, 58 > 51.

[75] These are calculated using equations 11.6 and 11.7, which you are extremely unlikely to be asked about.

Here is a somewhat different set of least squares credibilities, based on modeling the same baseball data via Markov Chains.[76] We allow each year to have a different credibility, give the remaining weight to the overall mean, and have no delay in getting data; thus these credibilities are similar to those shown in Table 16 in the syllabus reading.

Number of Years of Data Used	Years Between Data and Estimate				
	1	2	3	4	5
1	67.0%	-	-	-	-
2	55.1%	17.7%	-	-	-
3	54.3%	15.0%	4.9%	-	-
4	54.2%	14.8%	4.2%	1.4%	-
5	54.2%	14.8%	4.1%	1.2%	0.4%

This model has smoothed out the peculiarities of the covariances of the baseball data that are due to random fluctuation. Thus we see a much more regular pattern of credibilities. More distant years get less credibility than more recent years, declining in a nice pattern. Due to the high rate at which parameters shift in the baseball data, the credibilities for distant years get small quickly. The sum of the credibilities approaches 74.7%.[77] Unlike Table 16, there are no negative credibilities.

[76] Taken from Table 7, in "A Markov Chain Model of Shifting Risk Parameters," by Howard C. Mahler, PCAS 1997, not on the syllabus. In terms of number of games lost, the covariances between years can be approximated by: $\text{Cov}[X_i, X_j] = (170)(0.818^{|i-j|}) + 37\ \delta_{ij}$, where δ_{ij} is zero if $i \neq j$ and one if $i = j$.

[77] With shifting risk parameters, the limit of the sum of the credibilities as N approaches infinity is less than 1. The faster the rate of shifting, the smaller is this limit.
For 10 years of data, the least squares credibilities are: 54.2%, 14.8%, 4.1%, 1.2%, 0.3%, 0.1%, 0, 0, 0, 0.

<u>An Example of Solving for the Least Squares Credibility</u>:[78]

The paper uses the following notation:
τ^2 = between variance.
C(k) = covariance for data of the same risk, k years apart = "within covariance"
C(0) = "within variance".

For a data set, you are given: $\tau^2 = 8$, C(0) = 50, C(1) = 20, C(2) = 15, and C(3) = 10.[79]

For two different years: $\text{Cov}[X_i, X_j] = \tau^2 + C(|i - j|)$.
For example, $\text{Cov}[X_1, X_4] = \tau^2 + C(3) = 8 + 10 = 18$.
For a single year of data: $\text{Cov}[X_i, X_i] = \text{Var}[X_i] = \tau^2 + C(0) = 8 + 50 = 58$.

Thus the covariance matrix is:

$$\begin{matrix}\text{Year 1}\\ \text{Year 2}\\ \text{Year 3}\\ \text{Year 4}\end{matrix}\begin{pmatrix} 58 & 28 & 23 & 18 \\ 28 & 58 & 28 & 23 \\ 23 & 28 & 58 & 28 \\ 18 & 23 & 28 & 58 \end{pmatrix}.$$

Use data from year 3 to predict year 4.
Give weight Z to the relativity for year 3, and weight 1 - Z to the overall mean relativity of 1.

Equations 11.3 are the linear equations for the least squares credibility:[80]

$$\sum_{j=1}^{N} Z_j \ \text{Cov}[X_i, X_j] = \text{Cov}[X_i, X_{N+\Delta}]$$, where we are predicting year N + Δ, using years 1 to N.

With just one year of data from year 3 (N = 1) used to predict year 4:[81]
$Z \ \text{Cov}[X_3, X_3] = \text{Cov}[X_3, X_4]$.

Then $Z = \dfrac{\text{Cov}[X_3, X_4]}{\text{Var}[X_3]}$ = 28/58 = 48.3%.

Instead let us use data from years 2 and 3 to predict year 4.

In other words, we will give weight Z_2 to the relativity for year 2, weight Z_3 to the relativity for year 3, and weight $1 - (Z_2 + Z_3)$ to the overall mean relativity of 1.

---

[78] Based on recent exams, this is unlikely to be asked.
[79] These are illustrative values. Note that the variance of a single year is more than the covariance between two different years. Also, the covariance between years further apart is less than between years that are closer together.
This is the pattern we get with shifting risk parameters over time.
[80] These are the Normal Equations for credibility; see equations 20.25 & 20.26 in <u>Loss Models</u>, <u>not</u> on the syllabus.
Note that if all of the covariances are multiplied by the same constant, the credibilities remain the same.
[81] Since we are using year 3 to predict year 4, Δ = 1.

Here N = 2 and Δ = 1, and we get two linear equations in two unknowns:[82]
$58Z_2 + 28Z_3 = 23$.
$28Z_2 + 58Z_3 = 28$.

Solving: Z_2 = 55/258 = 21.3%, and Z_3 = 49/129 = 38.0%. Thus we give weight 21.3% to year 2, weight 38.0% to year 3, and the remaining weight of 40.7% to the overall mean relativity of 1. Due to shifting risk parameters, Year 2 is less correlated with year 4 than is year 3. Year 3 is more useful for predicting year 4 than is year 2; year 3 is given more weight.

Exercise: Assume that a team has a losing percentage of 0.453 in year 2, and a losing percentage of 0.411 in year 3. Predict the losing percentage for this team in year 4.
[Solution: (21.3%)(0.453) + (38.0%)(0.411) + (40.7%)(0.500) = 0.456.
Comment: One could divide everything by the overall mean losing percentage of 0.5 in order to put everything in terms of relativities with respect to average.]

Let us instead assume we give weight Z to the average of the relativities for years 2 and 3, and weight 1 - Z to the overall mean relativity of 1.[83]

$\text{Cov}[(X_2 + X_3)/2, X_4] = \{\text{Cov}[X_2 , X_4] + \text{Cov}[X_3 , X_4]\} / 2 = (23 + 28)/2 = 25.5$.
$\text{Var}[(X_2 + X_3)/2] = \{\text{Var}[X_2] + \text{Var}[X_3] + 2\ \text{Cov}[X_2, X_3]\} / 2^2 = \{58 + 58 + (2)(28)\} / 4 = 43$.

Thus the linear equation for Z, analogous to equation 11.3 is:
$43\ Z = 25.5. \Rightarrow Z = 25.5 / 43 = 59.3\%$.

Exercise: Assume that a team has a losing percentage of 0.453 in year 2, and a losing percentage of 0.411 in year 3. Predict the losing percentage for this team in year 4.
[Solution: (59.3%) (0.453 + 0.411)/2 + (1 - 59.3%) (0.500) = 0.460.
Comment: Differs slightly from the previous estimate using separate credibilities by year.]

The previous separate credibilities were: Z_2 = 21.3%, and Z_3 = 38.0%.
They sum to 59.3%, the same as the single Z applied to the average.[84]

Note that using a single Z, we constrained the weights given to years two and three to be equal. Thus this is a special case of solving for the least squares Z_2 and Z_3. Thus the expected squared error using the best single Z must be greater than or equal to that from using the best Z_2 and Z_3.

---

[82] The coefficients on the lefthand side are the second and third rows and columns of the covariance matrix. The values on the righthand side are the second and third rows of column four, since we are predicting Year 4.
[83] In this paper, every year of data has the same volume of data, so we are giving weight Z/2 to each year.
[84] One can show algebraically, that this will be true in general when using only two years of data.

Repeating the covariance matrix:

Year 1	58	28	23	18
Year 2	28	58	28	23
Year 3	23	28	58	28
Year 4	18	23	28	58

Let us instead use data from years 1 and 2 to predict year 4. There is now a one year delay. We give weight Z_1 to the relativity for year 1, weight Z_2 to the relativity for year 2, and weight 1 - $(Z_1 + Z_2)$ to the overall mean relativity of 1.

Equations 11.3 are the linear equations for the least squares credibility:

$$\sum_{j=1}^{N} Z_j \text{ Cov}[X_i , X_j] = \text{Cov}[X_i, X_{N+\Delta}],$$ where we are predicting year N + Δ, using years 1 to N.

Exercise: Write down the linear equations for the least squares credibilities.
[Solution: Here N = 2 and Δ = 2, and we get two linear equations in two unknowns:
$58Z_1 + 28Z_2 = 18$.
$28Z_1 + 58Z_2 = 23$.
Comment: The coefficients on the lefthand side are the first two rows and columns of the covariance matrix. The righthand side is the first two rows of column four, since we are predicting Year 4.]

Solving these linear equations: Z_1 = 20/129 = 15.5%, and Z_2 = 83/258 = 32.2%.
Thus we give weight 15.5% to the relativity for year 1, 32.2% weight to the relativity for year 2, and the remaining weight of 52.3% to the overall mean relativity of 1.

This compares to the previous case with no delay when Z_2 = 21.3% and Z_3 = 38.0%.
With no delay, we give more weight to the data: 21.3% > 15.5%, and 38.0% > 32.2%.

Due to shifting risk parameters, more distant years are less useful for predicting the future. Thus with a delay the credibilities are smaller. The bigger the delay, the smaller the credibilities.

Interestingly, with the delay the weight assigned to year 2 is larger than it was without the delay. That is because least squares credibility is a relative concept. The weight assigned to a year of data depends on how good an estimator is each of the other years being used.

Year 3 is a better estimator of year 4 for than is year 2; thus when using year 3 and year 2 this tends to decrease the weight given to year 2. In contrast, year 1 is a worse estimator of year 4 than is year 2; thus when using year 1 and year 2 this tends to increase the weight given to year 2.

Let us instead assume we give weight Z to the average of the relativities for years 1 and 2 and weight 1 - Z to the overall mean relativity of 1.[85]

$\text{Cov}[(X_1 + X_2)/2, X_4] = \{\text{Cov}[X_1 , X_4] + \text{Cov}[X_2 , X_4]\} / 2 = (18 + 23)/2 = 20.5.$

$\text{Var}[(X_1 + X_2)/2] = \{\text{Var}[X_1] + \text{Var}[X_2] + 2\ \text{Cov}[X_1, X_2]\} / 2^2 = \{58 + 58 + (2)(28)\} / 4 = 43.$

Thus the linear equation for Z, analogous to equation 11.3 is:

$43\ Z = 20.5. \Rightarrow Z = 20.5 / 43 = 47.7\%.$

The previous separate credibilities were: $Z_1 = 15.5\%$, and $Z_2 = 32.2\%$.

They sum to 47.7%, the same as the single Z applied to the average.

Due to shifting risk parameters, Z = 47.7% when using years 1 and 2 to predict year 4 is smaller than Z = 59.3% when instead using years 2 and 3 to predict year 4.

Repeating the covariance matrix:

$$\begin{matrix} \text{Year 1} \\ \text{Year 2} \\ \text{Year 3} \\ \text{Year 4} \end{matrix} \begin{pmatrix} 58 & 28 & 23 & 18 \\ 28 & 58 & 28 & 23 \\ 23 & 28 & 58 & 28 \\ 18 & 23 & 28 & 58 \end{pmatrix}.$$

Exercise: We will use years 1, 2, and 3 to predict year 4.
Write down the linear equations for the least squares credibilities.
[Solution: We get three linear equations in three unknowns:

$58Z_1 + 28Z_2 + 23Z_3 = 18.$

$28Z_1 + 58Z_2 + 28Z_3 = 23.$

$23Z_1 + 28Z_2 + 58Z_3 = 28.$

Comment: The coefficients on the lefthand side are the first 3 rows and columns of the covariance matrix. The righthand side is the first 3 rows of column four, since we are predicting Year 4.]

Solving these linear equations:

$Z_1 = 170/2191 = 7.8\%$, $Z_2 = 115/626 = 18.4\%$, and $Z_3 = 796/2191 = 36.3\%$.[86]

Thus we give weight 7.8% to the relativity for year 1, 18.4% weight to the relativity for year 2, 36.3% weight to the relativity for year 3, and the remaining weight of 37.5% to the overall mean relativity of 1.

[85] In this paper, every year of data has the same volume of data, so we are giving weight Z/2 to each year.

[86] You will not be asked to solve three linear equations on your exam.

Let us instead assume we give weight Z to the average of the relativities for years 1, 2, and 3 and weight 1 - Z to the overall mean relativity of 1.[87]

$\text{Cov}[(X_1 + X_2 + X_3)/3, X_4] = \{\text{Cov}[X_1, X_4] + \text{Cov}[X_2, X_4) + \text{Cov}[X_3, X_4]\} / 3 =$
$(18 + 23 + 28)/3 = 23.$
$\text{Var}[(X_1 + X_2 + X_3)/3] =$
$\{\text{Var}[X_1] + \text{Var}[X_2] + \text{Var}[X_3] + 2\ \text{Cov}[X_1, X_2] + 2\ \text{Cov}[X_1, X_3] + 2\ \text{Cov}[X_2, X_3]\} / 3^2 =$
$\{58 + 58 + 58 + (2)(28) + (2)(23) + (2)(28)\} / 9 = 332/9.$

Thus the linear equation for Z, analogous to equation 11.3 is:
$(332/9)\ Z = 23. \Rightarrow Z = 207 / 332 = 62.3\%.$

Alternately, we could use equation 11.4:[88]

$$Z = N \frac{N\,\tau^2 + \sum_{I=1}^{N} C(N+\Delta-i)}{N^2\,\tau^2 + \sum_{j=1}^{N}\sum_{I=1}^{N} C(|i-j|)}.$$

With N = 3 and Δ = 1:

$$Z = (3)\ \frac{(3)(8) + C(3) + C(2) + C(1)}{(9)(8) + C(0) + C(1) + C(2) + C(1) + C(0) + C(1) + C(2) + C(1) + C(0)} =$$

$$(3)\ \frac{24 + 10 + 15 + 20}{72 + 50 + 20 + 15 + 20 + 50 + 20 + 15 + 20 + 50}$$

$= (3)(69) / 332 = 207 / 332 = 62.3\%.$

The separate credibilities were: Z_1 = 7.8%, Z_2 = 18.4%, and Z_3 = 36.3%.
These sum to 62.5%, close to but different than the single credibility of 62.3%.[89]
In general, we expect them to be similar but not identical.

---

[87] In this paper, every year of data has the same volume of data, so we are giving weight Z/3 to each year.
[88] I would not memorize this equation.
[89] The sum is: 391/626 = 0.62460, while the single Z = 207/332 = 0.62349.

Expected Squared Errors:[90]

The least squares credibilities minimize the expected squared error between the estimate and the observation.[91] The expected squared error is given by equation 11.2:[92]

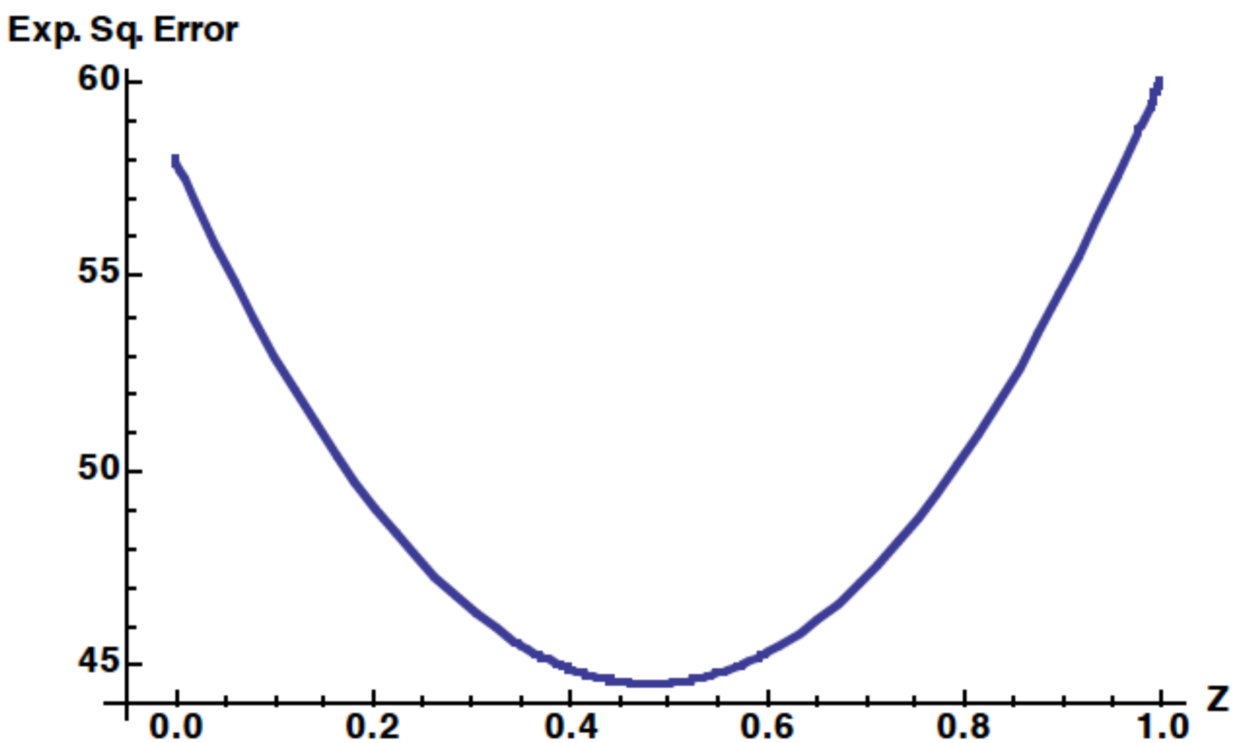
$$V(\vec{Z}) = \sum_{i=1}^{N}\sum_{j=1}^{N} Z_i Z_j \{\tau^2 + C(|i-j|)\} - 2\sum_{i=1}^{N} Z_i\{\tau^2 + C(N+\Delta-i)\} + \tau^2 + C(0).$$

For the previous example, we had: $\tau^2 = 8$, $C(0) = 50$, $C(1) = 20$, $C(2) = 15$, and $C(3) = 10$.

If we are using year 3 to estimate year 4, then $N = 1$ and $\Delta = 1$, and:
$V(Z) = Z^2\{\tau^2 + C(0)\} - 2\,Z\{\tau^2 + C(1)\} + \tau^2 + C(0) = 58Z^2 - 56\,Z + 58$.

Setting the derivative of V(Z) equal to zero in order to minimize the expected squared error:
$(2)(58)\,Z - 56 = 0. \Rightarrow Z = 28/58 = 48.3\%$, matching a previous result.

At Z = 48.3%, the expected squared error is: $(58)(0.483^2) - (56)(0.483) + 58 = 44.48$.
Here is a graph of the expected squared error as a function of Z, a parabola:



[90] Based on recent exams, this is unlikely to be asked about in any detail.
[91] For the given form of linear estimator. So for example, we would specify in advance that we using a linear combination of years 1, 2 and 3 and the overall mean, in order to estimate year 4. We would also need to specify whether we will give each year the same weight or instead apply separate credibilities to each year of data.
[92] I would not memorize this equation.

Exercise: We are instead using year 2 to estimate year 4.
Determine the expected squared error as a function of Z, and find its minimum.
[Solution: $V(Z) = Z^2\{\tau^2 + C(0)\} - 2 Z\{\tau^2 + C(2)\} + \tau^2 + C(0) = 58Z^2 - 46Z + 58$.
Setting the derivative of V(Z) equal to zero in order to minimize the expected squared error:
$(2)(58) Z - 46 = 0. \Rightarrow Z = 46/116 = 39.7\%$.
At Z = 39.7%, the expected squared error is: $(58)(0.397^2) - (46)(0.397) + 58 = 48.88$.]

Note that when using year 2 rather than year 3, the credibility is smaller while the expected mean squared error is larger. Specifically, the minimum squared error is now 48.88 compared to 44.48.
Due to shifting parameters over time, year 2 is a worse predictor of year 4 than is year 3, and thus the minimum expected squared error is greater if we use year 2 rather than year 3.

Now let us use data from years 2 and 3 to predict year 4. Then using equation 11.2:
V(Z) =
$Z_2^2\{\tau^2 + C(0)\} + 2Z_2Z_3\{\tau^2 + C(1)\} + Z_3^2\{\tau^2 + C(0)\} - 2Z_2\{\tau^2 + C(2)\} - 2Z_3\{\tau^2 + C(1)\} + \tau^2 + C(0) =$
$58Z_2^2 + 56Z_2Z_3 + 58Z_3^2 - 46Z_2 - 56Z_3 + 58$.

It turns out that Equation 11.2 can be rewritten in matrix form,
Mean Squared Error = $V(Z) = Z^T C Z$.
C is the matrix of covariances for the years of data.
Z is the (column) vector with credibilities in the years used to estimate, -1 in the year being estimated, and zeros in any other years. Z^T is the transpose of Z.

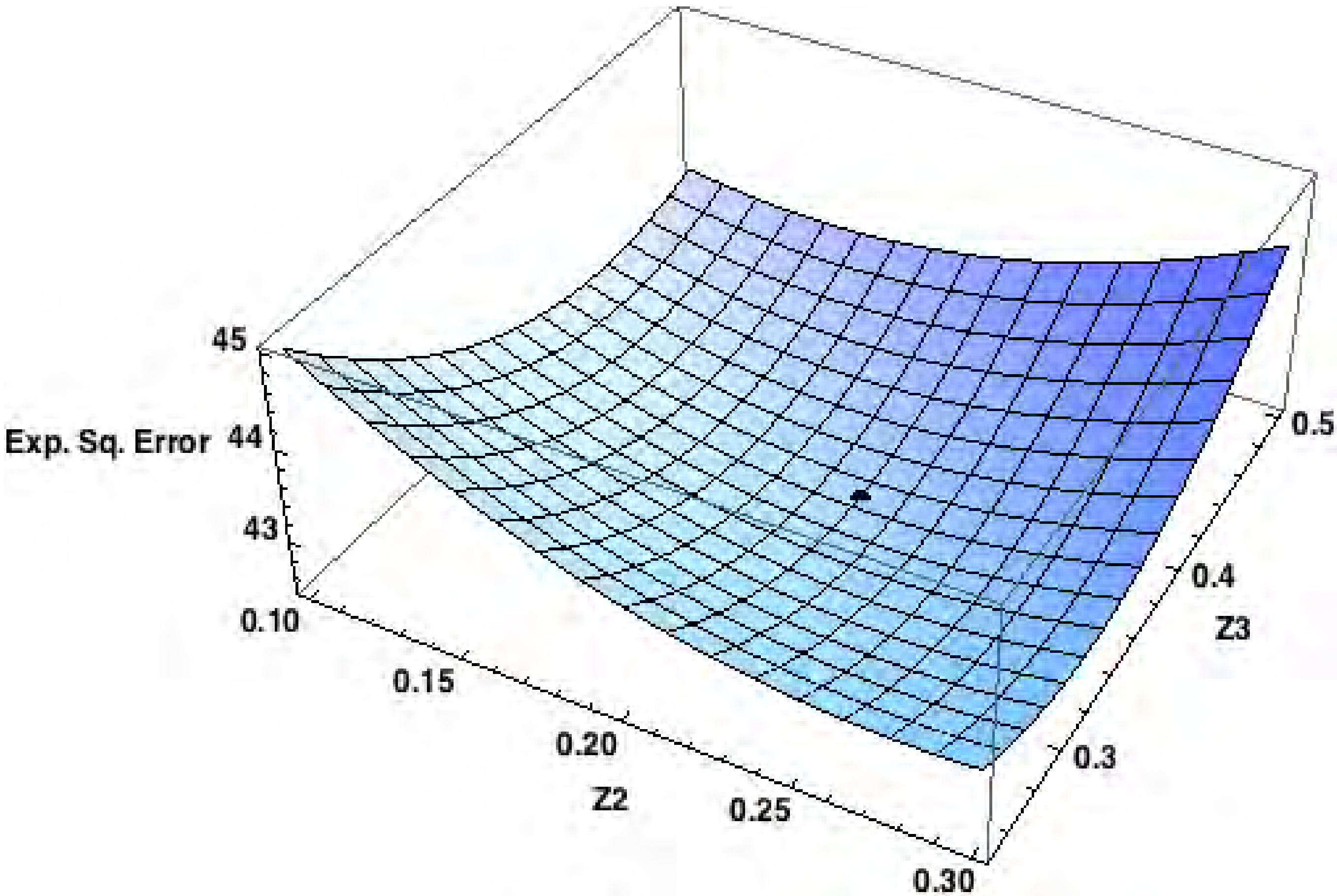
In this case: $V(Z) = (Z_2 , Z_3, - 1) \begin{pmatrix} 58 & 28 & 23 \\ 28 & 58 & 28 \\ 23 & 28 & 58 \end{pmatrix} \begin{pmatrix} Z_2 \\ Z_3 \\ -1 \end{pmatrix} = (Z_2 , Z_3, - 1) \begin{pmatrix} 58Z_2 + 28Z_3 - 23 \\ 28Z_2 + 58Z_3 - 28 \\ 23Z_2 + 28Z_3 - 58 \end{pmatrix}$

$= 58Z_2^2 + 56Z_2Z_3 + 58Z_3^2 - 46Z_2 - 56Z_3 + 58$.

Setting the partial derivative with respect to Z_2 equal to zero:
$116 Z_2 + 56 Z_3 = 46$.
Setting the partial derivative with respect to Z_3 equal to zero:
$56 Z_2 + 116 Z_3 = 56$.

These are equivalent to the two equations we got before for this situation, and the solution is: $Z_2 = 21.3\%$ and $Z_3 = 38.0\%$. For these least squares credibilities, the expected squared error is: $(58)(0.213^2) + (56)(0.213)(0.380) + (58)(0.380^2) - (46)(0.213) - (56)(0.380) + 58 = 42.46$.

Here is a graph of the expected squared error as a function of Z_1 and Z_2,
with the minimum shown as a dot, (0.213, 0.380, 42.46):



"The optimal credibilities are uniquely determined given the covariance structure. However, there are many other sets of credibilities which produce expected squared errors very close to minimal."[93]

Note that using just year 3 is a special case of using years 2 and 3, with Z_2 constrained to be zero. Thus the minimum expected squared error using just year 3 has to be greater than or equal that from using both years 2 and 3. In this example, 44.48 is greater than 42.46.

As shown previously, when giving the same weight to years 2 and 3, the least squares credibility is Z = 59.3%. In other words, when constrained to be equal, $Z_2 = Z_3 = 59.3\%/2$. Then, V(Z) = $(58)(0.2965^2) + (56)(0.2965^2) + (58)(0.2965^2) - (46)(0.2965) - (56)(0.2965) + 58 = 42.88$.
As it has to be, the minimum expected squared error when the weights are constrained to be equal is greater than that when the weights are allowed to be different; 42.88 > 42.46.

93 See page 265 of "An Example of Credibility and Shifting Risk Parameters," by Howard C. Mahler.

Half-Life:[94]

One can model shifting risks parameters via a covariance structure between years of data that is of the form: $Cov[X_i, X_j] = a\ \rho^{|i-j|} + b\ \delta_{ij}$, where δ_{ij} is zero if $i \neq j$ and one if $i=j$.[95]

$\rho < 1$ measures the speed at which risk parameters shift. The correlations between years decline by a factor of ρ when the separation between those years of data increases by a year.

Define the "half-life" as the length of time for the correlations between years to decline by a factor of one-half: $\rho^{\text{half-life}} = 0.5$. $\Rightarrow$ half-life = $\ln(0.5) / \ln(\rho)$.
The half-life is a somewhat more intuitive way to quantify the rate at which parameters shift.

Some examples, with the approximate values of ρ and the corresponding half-lives:[96] [97] [98]

Example	ρ	Half-life
Baseball Win/Loss Data by Team	0.82	3.5 years
California P.P. Auto Driving Data	0.95	13.5 years
Workers Compensation Classes	0.94	11.2 years
Workers Comp. Experience Rating	0.82	3.5 years

We see that the rate of shifting in the baseball example in the syllabus reading is much faster than that in the first two insurance examples. While this made the baseball data set a good one to use to develop these ideas and illustrate the results, the effect of shifting risk parameters will be significantly smaller in many applications to insurance data.

---

[94] See "A Markov Chain Model of Shifting Risk Parameters," by Howard C. Mahler, PCAS 1997, not on the syllabus.
[95] For $\rho < 1$, this models shifting risk parameters over time. This is an approximation to the form in "A Markov Chain Model of Shifting Risk Parameters," by Howard C. Mahler, PCAS 1997.
[96] For the California driving data, with ρ approximately 0.94 for Female Drives and 0.97 for Male drivers. It is not clear whether this difference between males and females is significant or just due to random fluctuations in the data set. See "The Credibility of a Single Private Passenger Driver", by Howard C. Mahler, PCAS 1991.
[97] For Workers Compensation classes, classification relativities for the Manufacturing Industry in Massachusetts, for classes with expected annual losses between \$300,000 and \$1 million.
See page 535 of "Credibility With Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty," by Howard C. Mahler, PCAS 1998.
The rate of shifting risk parameters may be more rapid for smaller classes than for larger classes.
In "Workers' Compensation Classification Credibilities", by Howard C. Mahler, Fall 1999 CAS Forum, the equivalent of $\rho = 0.85$ for the very smallest classes and $\rho = 0.99$ for the very largest classes were selected.
[98] For the Workers Compensation experience rating data, see page 589 of "Credibility With Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty," by Howard C. Mahler, PCAS 1998. Based on experience rating data for Massachusetts. The selected values differ somewhat between primary and excess and by size of insured.

Conclusions:

Know well the three paragraphs on page 280, the conclusions of the paper:

When shifting parameters over time is an important phenomenon, older years of data should be given substantially less credibility than more recent years of data. The more significant this phenomenon, the more important it is to minimize the delay in receiving the data that is to be used to make the prediction.

Three different criteria were examined that can be used to select the optimal credibility: least squares, limited fluctuation, and Meyers/Dorweiler. In applications, one or more of these three criteria should be useful. While the first two criteria are closely related, the third criterion can give substantially different results than the others.

Generally the mean squared error can be written as a second order polynomial in the credibilities. The coefficients of this polynomial can be written in terms of the covariance structure of the data. This in turn **allows one to obtain linear equation(s) which can be solved for the least squares credibilities in terms of the covariance structure**.

Further Reading and Resources:[99]

"A Markov Chain Model of Shifting Risk Parameters", by Howard C. Mahler, PCAS 1997.
www.casact.org/pubs/proceed/proceed97/97581.pdf
This 1997 paper expands on "An Example of Credibility and Shifting Risk Parameters."
Pages 629-639 revisit the baseball example.

"Credibility With Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty," by Howard C. Mahler, PCAS 1998. www.casact.org/pubs/proceed/proceed98/980455.pdf
This 1998 paper expands on the 1997 paper.
Pages 615-617 briefly discuss the baseball example.

"Workers' Compensation Classification Credibilities", by Howard C. Mahler,
Fall 1999 CAS Forum. www.casact.org/pubs/forum/99fforum/99ff433.pdf
This paper applies the ideas of the 1998 paper to a practical example of classification ratemaking.

Stuart A. Klugman, "Credibility with Shifting Risk Parameters," SOA Study Note, 2014.

[99] Not on the syllabus.

Problems:

1.1. (1 point) All but which of the following are reasons Mahler uses baseball data to study experience rating?
A. Each baseball team plays the same number of games.
B. The won-loss data are accurate, final, and readily available.
C. The set of teams does not change over the period of time studied.
D. Baseball teams win a similar percentage of games over a decade.
E. All of A, B, C, and D are true.

1.2. (1 point) Which of the following did Mahler conclude regarding differences between teams?
1. A team that had been worse than average over one period of time is more likely to be better than average over the subsequent period of time.
2. Observed differences between teams over six decades are greater than could be attributed to chance alone if teams were inherently equal.
3. The fact that one team's loss is another team's win has a material effect on the distribution of losing percentages in the baseball analogy.

1.3. (1 point) Which of the following among Mahler's conclusions regarding changes in the inherent winning potential of the teams over time?
1. For all the teams, a Chi-Square test showed differences over time that are significant at the 1 % level.
2. Significant correlations exist between a team's results in one year and its results in other years less than ten years before or after.
3. A team's experience in recent years is useful in predicting its experience in the upcoming year.

Use the following information for the next two questions:
We are estimating optimal credibility for experience rating of NBA basketball teams under three versions of the draft rule:
(1) Team with poorest record gets the first draft pick.
(2) Team with best record gets the first draft pick.
(3) The order of draft picks is chosen randomly.
Let Z(i) is the credibility under rule i (i = 1, 2, 3).

1.4. (1 point) Rank the experience rating credibility under these three rules.

1.5. (1 point) Which of the following are true?
1. $0 \le Z(1) \le 1$
2. $0 \le Z(2) \le 1$
3. $0 \le Z(3) \le 1$

1.6. (1 point) Which of the following statements are true of Mahler's credibility estimators?
1. They are linear combinations of a few simple estimators.
2. They are unbiased for the set of teams as a whole.
3. They are more analogous to schedule rating than to experience rating.

1.7. (3 points) Eleven insureds have the following relative loss ratios in two consecutive years:

1	2	3	4	5	6	7	8	9	10	11
0.90	0.93	0.96	0.98	0.99	1.00	1.01	1.02	1.04	1.07	1.10
0.98	0.93	1.00	0.90	0.98	1.02	0.99	0.99	1.07	1.10	1.04

Based on the least squares criterion, what is the proper credibility for these insureds?

1.8. (1 point) We are designing an experience rating system which weights the class mean with the unweighted mean of the risk's latest N years of data. Which of the following criteria can be used to select optimal values for the credibility Z and the number of years N?
1. Least squared error
2. Small chance of large error
3. Meyers/Dorweiler

1.9. (1 point) Mahler in "An Example of Credibility and Shifting Risk Parameters," concludes that to predict baseball losing percentages, a reasonable method is to use three years of data with Z_1 = 10%, Z_2 = 10%, Z_3 = 55%, and the remaining weight to the a priori mean.
A baseball team had the following record:
2005: won 67 games and lost 95 games.
2006: won 61 games and lost 101 games.
2007: won 66 games and lost 96 games.
Using the above method, in 2008, what is the predicted record for this team for its first 88 games?

1.10. (1 point) Which of the following are true about optimal credibility estimators?
1. When one combines the unweighted N-year average with the grand mean, the estimate continues to improve as N increases.
2. The exact values of the optimal credibility weights $\{Z_i\}$ are not critical as long as one is close to the optimal set.
3. The ideal credibility estimator would reduce the mean squared error between the estimated and observed values to zero.

1.11. (1 point) Let a set of Z_i's be the credibility factors that minimize the expected squared error as determined by the covariance structure. Which of the following are true?
1. The expected squared error is a linear function of the Z_i.
2. The optimal Z_i are all nonnegative.
3. It is necessary to distinguish among three sources of variance: variance between risks (τ^2), the process variance excluding the effect of shifting parameters over time (δ^2), and the portion of the process variance due to shifting parameters over time (ζ^2).

1.12. (1 point) The optimal credibility weights for an experience rating plan depend on the variance between risks (τ^2), the process variance excluding the effect of shifting parameters over time (δ^2), and the portion of the within variance due to shifting parameters over time (ζ^2).
A certain plan uses five years of experience and a two step credibility procedure: the risk's own experience gets credibility Z of 50%, divided between five years of experience are weighted 10%, 15%, 20%, 25%, and 30%.
We are updating the credibility weights, based on new estimates of τ^2, δ^2, and ζ^2.
Which of the following statements are correct?
1. As τ^2 increases, the value of Z decreases.
2. As δ^2 increases, the value of Z decreases.
3. As ζ^2 increases, the weight for year 1 (now 10%) decreases.

1.13. (2 points) The optimal credibility weights depend on the variance between risks (τ^2), the process variance excluding the effect of shifting parameters over time (δ^2), and the portion of the within variance due to shifting parameters over time (ζ^2). Briefly explain how each of these three elements differs between class ratemaking and experience rating.

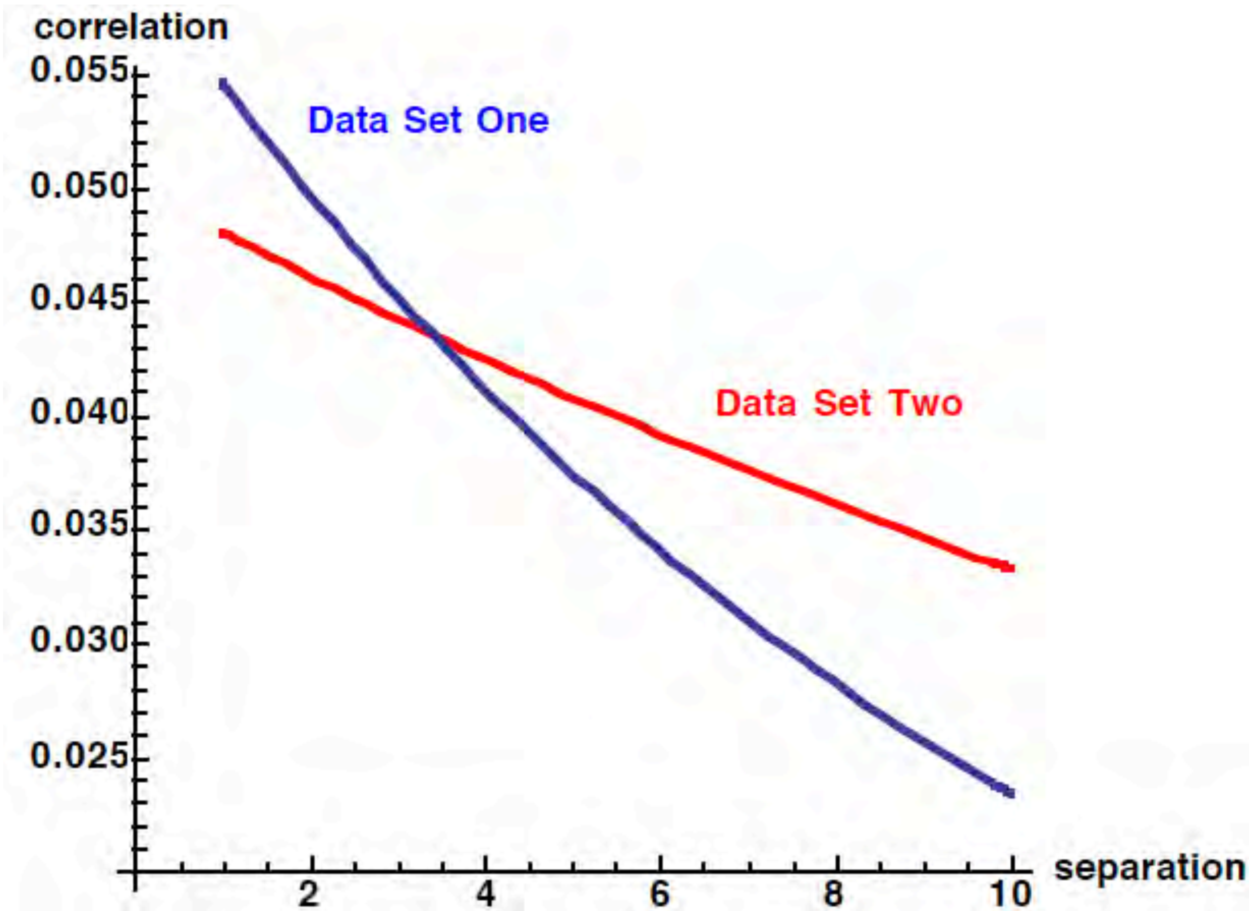
1.14. (3 points) You are analyzing an experience rating plan.
Briefly explain how each of the changes affect the following items:
Between Variance (τ^2), Within Variance ($\delta^2 + \zeta^2$), Effect of Shifting Risk Parameters, and Credibility (Z).
(a) Change from a no-split experience rating plan to one with a reasonable primary-excess split.
(b) Use 2 years of data instead of 5 in the experience rating plan.
(c) Refine the classification plan.

1.15. (1 point) Which of the following are true of the Meyers/Dorweiler criterion?
1. The criterion assures that debit and credit risks are equally attractive to insurers.
2. As credibility approaches zero, the Kendall τ statistic approaches one.
3. An experience rating plan that satisfies the criterion is an acceptable plan.

1.16. (1 point) Using Mahler's terminology,
1. Let $X(\theta, t)$ be the observation for risk θ at time t.
2. Let $\mu(\theta, t)$ be the expected value for risk θ at time t.
3. Let $\mu(\theta)$ be the long-term expected value for risk θ.
4. Let M be the long-term all-risk grand mean.
Which of the above may be directly observed in an insurance pricing situation?

1.17. (1 point) What problems are caused by delays in obtaining data?
1. Delays in obtaining data degrade the performance of experience rating plans.
2. The optimal credibility decreases as the delay increases.
3. If the delay exceeds three years, then for the data set examined in Mahler's "An Example of Credibility and Shifting Risk Parameters," the predictive value of the data is close to zero.

1.18. (1 points) You are looking at two different data sets, each consisting of many years of data. In each case, you calculate correlations between pairs of years of data. You then fit a curve to the correlations as a function of the separation of the years of data. Here is a graph of the results:



What conclusions do you draw and why?

1.19. (1 point) Mahler discusses an estimate of the form F = Z X + (1 - Z) P, where X is the most recent data point, P is the previous estimate, and Z is a selected weight.
Assume that there is no delay in obtaining the data.
If Z = 55%, what weight is given to the data for 2006 in estimating losses for 2009?

1.20. (1 point) A rate indication for 2009 uses weighted experience from 2002 through 2006. Based on the considerations outlined by Mahler, which of the following statements are true?
1. It is appropriate to assign equal weights to the years of data.
2. It is not appropriate to assign nonzero weight to data from 2002.
3. The traditional weights of 10%, 15%,20%, 25%, 30% perform significantly worse than an optimal set of weights derived using Mahler's techniques.

1.21. (1 point) According to Mahler in "Credibility and Shifting Risk Parameters," which of the following statements are correct?
1. The mean squared error can generally be written as a second order polynomial in the credibilities.
2. The coefficients of this polynomial can be written in terms of the covariance structure of the data.
3. This in turn allows one to obtain quadratic equations which can be solved for the least squares credibilities in terms of the covariance structure.

1.22. (2 points) (For baseball fans) You are updating the study in Mahler's paper using similar baseball data from 1961 to the present.
(a) Mention two complications that would occur that Mahler did not have to deal with.
(b) Would you expect shifting risk parameters to have a bigger effect or smaller effect than in Mahler's study? Why?

1.23. (15 points) In "An Example of Credibility and Shifting Risk Parameters," Mahler uses the following notation:
τ^2 = between variance.
C(k) = covariance for data of the same risk, k years apart = "within covariance"
C(0) = "within variance".
For a data set, you are given τ^2 = 5, C(0) = 50, C(1) = 10, C(2) = 8, C(3) = 6, and C(4) = 4.
One will be using least squares credibility, with the complement of credibility given to the grand mean and varying weights to each year of data.
In each case, determine the optimal credibilities to be assigned to each year of data.
(a) (1 point) Use data for Year 1 to Predict Year 2.
(b) (1 point) Use data for Year 1 to Predict Year 3.
(c) (1 point) Use data for Year 1 to Predict Year 4.
(d) (2 points) Use data for Years 1 and 2 to Predict Year 3.
(e) (2 points) Use data for Years 1 and 2 to Predict Year 4.
(f) (4 points) Use data for Years 1, 2, and 3 to Predict Year 4.
(g) (4 points) Use data for Years 1, 2, and 3 to Predict Year 5.

1.24. (10 points) In the previous question, in parts (d) through (g), instead require that the weight given to each year be the same. Calculate the resulting least squares credibility.

1.25. (2 points)
(a) Define the phenomena of shifting risk parameters.
(b) Pick a line of insurance and give one reason why risk parameters would shift over time for an insured.
Do not discuss something that would likely result in a change in classification or territory.

1.26. (2 points) (For football fans) You are updating the study in Mahler's paper using data from the National Football League from 1961 to the present.
(a) Mention two issues that Mahler's study did not have to deal with.
(b) Would you expect shifting risk parameters to have a bigger effect or smaller effect than in Mahler's study? Why?

1.27. (10 points) In “An Example of Credibility and Shifting Risk Parameters,” Mahler uses the following notation:
τ^2 = between variance.
C(k) = covariance for data of the same risk, k years apart = “within covariance”
C(0) = “within variance”.
For a data set, you are given τ^2 = 10, C(0) = 30, C(1) = 15, C(2) = 10, C(3) = 6, and C(4) = 3.
One will be using least squares credibility, with the complement of credibility given to the grand mean and varying weights to each year of data.
In each case, determine the optimal credibilities to be assigned to each year of data.
(a) (1 point) Use data for Year 1 to Predict Year 2.
(b) (1 point) Use data for Year 1 to Predict Year 3.
(c) (1 point) Use data for Year 1 to Predict Year 4.
(d) (1 point) Use data for Year 1 to Predict Year 5.
(e) (2 points) Use data for Years 1 and 2 to Predict Year 3.
(f) (2 points) Use data for Years 1 and 2 to Predict Year 4.
(g) (2 points) Use data for Years 1 and 2 to Predict Year 5.

1.28. (10 points) For each of the parts of the previous question, calculate the corresponding minimum expected squared error.

1.29. (1 point) In “An Example of Credibility and Shifting Risk Parameters,” one of the techniques used by Mahler is least squares credibility, with the complement of credibility given to the grand mean and varying weights to each year of data.
For an example, Mahler determines the optimal credibilities to be assigned to each year of data and displays them in Table 16.
Which of the following statements is true about these optimal credibilities?
A. They are not negative.
B. More distant years are given less weight than more current years.
C. As more years of data are used, the credibility assigned to the first year of data does not increase.
D. As more years of data are used, the expected squared error does not increase.
E. None of A, B, C, or D.

1.30. (6 points) Use the following information:
- You are using data from years 1 through 5 in order to predict year 6.
- The variance of each year of data is 6.
- The covariance between different years of data is:

$$\mathrm{Cov}[X_i, X_j] = 0.9^{|i-j|}.$$

In “An Example of Credibility and Shifting Risk Parameters,” one of the techniques used by Mahler is least squares credibility, with the complement of credibility given to the grand mean and varying weights to each year of data.
Determine the credibilities to assign to each of the five years of data.
(Use a computer to help you with the computations.)

1.31. (2 points) Three experience rating plans are being compared.
You are trying to evaluate which is optimal.
Each rating plan has been tested on the same five different policies of similar size.
You compare the modification factor for each plan calculated before the policy period to the subsequent experience during the policy period.
The following tables summarize the indicated modifications and policy period experience.

Policy Number	Rating Plan 1 Modification Factor	Rating Plan 2 Modification Factor	Rating Plan 3 Modification Factor	Policy Period Experience
1	0.80	0.87	0.81	0.85
2	0.90	0.87	0.83	0.85
3	1.00	1.00	1.00	1.00
4	1.10	1.03	1.09	1.05
5	1.20	1.23	1.27	1.25

Which is the preferred plan based on the Meyers/Dorweiler criterion? Why?
Which is the preferred plan based on the least squared error criterion? Why?

1.32. (3 points) You are using N years of data without any delay in order to estimate the next year. The remaining weight will be given to the grand mean.
Allowing the credibilities to differ by year, the following least squares credibilities were determined, with year 1 being the most recent year.
Also shown is the corresponding minimum mean squared error (0.00001):

Year	N=1	N=2	N=3	N=4	N=5	N=6	N=7	N=8	N=9	N=10
1	3.43%	3.32%	3.23%	3.16%	3.10%	3.05%	3.01%	2.98%	2.95%	2.93%
2		3.11%	3.01%	2.94%	2.87%	2.82%	2.78%	2.74%	2.71%	2.69%
3			2.82%	2.74%	2.67%	2.61%	2.57%	2.53%	2.50%	2.47%
4				2.56%	2.49%	2.43%	2.38%	2.34%	2.30%	2.27%
5					2.33%	2.26%	2.21%	2.16%	2.13%	2.10%
6						2.12%	2.06%	2.01%	1.97%	1.94%
7							1.93%	1.87%	1.83%	1.79%
8								1.75%	1.71%	1.67%
9									1.60%	1.55%
10										1.45%
Total	3.43%	6.43%	8.24%	11.40%	13.46%	15.29%	16.93%	18.38%	19.69%	20.86%
MSE	3835	3832	3829	3826	3824	3822	3821	3820	3819	3818

Note that the values shown in a column may not sum to the total shown due to rounding.

Fully discuss the results shown.

1.33. (2 points)
Compare and contrast the following 3 covariance structures between years of data.

$$\begin{matrix}\text{Year 1}\\ \text{Year 2}\\ \text{Year 3}\end{matrix}\begin{pmatrix}200 & 200 & 200\\ 200 & 200 & 200\\ 200 & 200 & 200\end{pmatrix}\quad \begin{matrix}\text{Year 1}\\ \text{Year 2}\\ \text{Year 3}\end{matrix}\begin{pmatrix}200 & 140 & 140\\ 140 & 200 & 140\\ 140 & 140 & 200\end{pmatrix}\quad \begin{matrix}\text{Year 1}\\ \text{Year 2}\\ \text{Year 3}\end{matrix}\begin{pmatrix}200 & 140 & 110\\ 140 & 200 & 140\\ 110 & 140 & 200\end{pmatrix}$$

Which one corresponds to a situation of shifting risk parameters over time? Explain why.

1.34. (20 points)
You are using N years of data without any delay in order to estimate the next year.
The remaining weight will be given to the grand mean.
Allow the credibilities to differ by year.
The covariance between different years of data is:
$Cov[X_i, X_j] = (127.5)\ 0.75^{|i-j|} + (42.5)\ 0.965^{|i-j|} + 37\ \delta_{ij}$, where δ_{ij} is zero if $i \neq j$ and one if $i=j$.
With the aid of a computer, for N = 1, 2, 3, ... , 10, in each case determine the least squares credibilities and the corresponding minimum mean squared errors.

1.35. (2 points) Mahler performs a chi-square test on his baseball data.
(a) (0.5 points) What is the purpose of this test?
(b) (1.5 points) Fully describe how this test is performed.

1.36. (3 points) Risk parameters are shifting over time.
In order to estimate the next year, you are using N years of data without any delay.
The remaining weight will be given to the grand mean.
In each case, the least squares credibilities have been determined as well as the corresponding minimum expected squared errors.
In one set of calculations, the credibilities were allowed to differ by year.
In a second set of calculations, the credibilities were the same for each year of data used.
The resulting minimum expected squared errors were as follows:

	N=1	N=2	N=3	N=4	N=5	N=6	N=7	N=8	N=9	N=10
Creds. Differ	58.71	58.00	57.60	57.36	57.23	57.14	57.10	57.07	57.05	57.04
Creds. the Same	58.71	58.01	57.63	57.45	57.37	57.36	57.40	57.47	57.55	57.64

Discuss fully these results.

1.37. (3 points) The won-loss record of the Boston Red Sox baseball team by decade:

Decade	Wins	Losses	Percent
1900s	691	634	0.52151
1910s	857	624	0.57866
1920s	595	938	0.38813
1930s	705	815	0.46382
1940s	854	683	0.55563
1950s	814	725	0.52891
1960s	764	845	0.47483
1970s	895	714	0.55625
1980s	821	742	0.52527
1990s	814	741	0.52347
2000s	920	699	0.56825
2010s	788	670	0.54047
Total	9518	8830	0.51875

Conduct a chi-square test with an α value of 0.10 on actual vs. expected wins to confirm whether or not risk parameters have shifted over time. Use the following table of critical values:

Degrees of Freedom	Critical Value α = 0.10
1	2.706
2	4.605
3	6.251
4	7.779
5	9.236
6	10.645
7	12.017
8	13.362
9	14.684
10	15.987
11	17.275
12	18.549

1.38. (3 points) The following are the losing relativities for the Boston Red Sox baseball team for 2000 to 2018: 0.951, 0.981, 0.852, 0.827, 0.79, 0.827, 0.938, 0.815, 0.827, 0.827, 0.901, 0.889, 1.148, 0.802, 1.124, 1.037, 0.852, 0.852, 0.667.
(A relativity of 1, would mean that they lost half of their games that year.
In 2000, they lost 47.53% of their games.)
One will estimate the relativity for the coming year as: Z (current year) + 1 - Z.
Determine the mean squared error for Z = 60%.

1.39. (3 points) For this question, use a spreadsheet.
For three risks of the same size from the same class,
the following are loss ratios relative to that for their class:

Year	Risk 1	Risk2	Risk 3
2013	0.93	0.89	0.75
2014	1.12	0.94	1.39
2015	0.83	1.10	0.98
2016	0.90	0.73	1.08
2017	0.81	0.89	1.56
2018	1.11	1.06	0.89
2019	1.27	0.71	1.00
2020	0.65	0.86	0.76
2021	0.59	0.94	0.99

To estimate the 2022 relative loss ratios for each risk, an actuary gives equal weight to the 3 most recent years (N = 3), and the complement (weight 1 - Z) is given to 1.
a. (1.5 points) Calculate the credibility, to the nearest tenth (Z = 0, 0.1, 0.2, etc ...), that minimizes the mean squared error of the actuary's prediction.
b. (0.5 points) Using the credibility from part (a), estimate the relative loss ratios for 2022 for each of the three risks.
c (1.0 point) Using the Small Chance of Large Errors Criterion with an error threshold of 10%, determine the most appropriate credibility, to the nearest tenth (Z = 0, 0.1, 0.2, etc ...)
SHOW ALL WORK.

1.40. (9, 11/95, Q.10) (1 point) Which of the following are conclusions of Mahler in "An Example of Credibility and Shifting Risk Parameters"?

1. When parameter shift is present, the optimal credibility (based on least squares criterion) for the most recent available year of data increases as the delay in receiving the data increases.
2. Older years of data receive greater credibility when parameter shift is present than when it is not.
3. When parameter shift is present, use as many years of data as possible to maximize the accuracy of the prediction.

1.41. (9, 11/95, Q.31) (3 points) List and describe the three (3) evaluation methods used by Mahler in his paper "An Example of Credibility and Shifting Risk Parameters" to arrive at estimates for optimal credibility. Which one does Mahler suggest might give results that disagree with the others, and why might this be?

1.42. (9, 11/96, Q.20) (1 point) According to Mahler's "An Example of Credibility and Shifting Risk Parameters," which of the following are true ?

1. The best that can be done using credibility to combine two estimates is to reduce the mean squared error between the estimated and observed values to 50% of the minimum of the squared errors from either relying solely on the data or ignoring the data.
2. It is desirable to have the correlation between the experience modification and the loss ratio modified by experience modification to be zero.
3. Mahler recommends using as many years of data as there are available.

Note: I have rewritten statement #1 in order to match the current syllabus.

1.*43. (9, 11/97, Q.44)* *(3 points): Using Mahler's "An Example of Credibility and Shifting Risk Parameters," calculate the proportion of the total variance due to parameter shifting for the following scenario:*
- *There are 10 baseball teams.*
- *Each team plays 200 games.*
- *5 teams have true mean losing potential of 0.4.*
- *The other 5 teams have true mean losing potential of 0.6.*
- *The number of losses is Poisson distributed around its true mean losing potential.*
- *The actual number of losses for each team are:*

Team	*Number of Losses*
1	*75*
2	*115*
3	*61*
4	*110*
5	*94*
6	*133*
7	*139*
8	*98*
9	*81*
10	*94*

Show all work.

1.44. (9, 11/97, Q.45) (3 points) A major retailer, the Unlimited, has contracted you to project their loss ratio for general liability. The previous actuary was fired by the Unlimited, because she would rely only on the industry loss ratio to make the projection. The Unlimited has asked you to consider giving half of the credibility weight to the industry loss ratio and half to its own loss ratio from the previous year. The industrywide loss ratio is 65%. Using the least squares criterion, do you agree or disagree with your client? The Unlimited's historical data is as follows:

Policy Year	Loss Ratio
1/1 - 12/31/92	75%
1/1 - 12/31/93	70%
1/1 - 12/31/94	65%
1/1 - 12/31/95	60%
1/1 - 12/31/96	55%

1.45. (9, 11/97, Q.46) (2 points) Mahler's "An Example of Credibility and Shifting Risk Parameters" describes the Meyers/Dorweiler criterion for evaluating methods of assigning credibility to past data in order to predict future performance.

a. (1 point) In utilizing this criterion, what are the two ratios Mahler calculates to evaluate the predictors of baseball losing percentages?

b. (1 point) Within the contest of an experience rating plan, what quantities would be the equivalents to each of the ratios given in part (a)?

1.46. (9, 11/98, Q.13) (1 point) Following the approach described by Mahler in "An Example of Credibility and Shifting Risk Parameters" and given the following data, use exponential smoothing to calculate the expected 1999 loss ratio for the Increasingly Risky Corporation.
Increasingly Risky has produced the following historical loss ratios:

1998	100%
1997	90%
1996	80%
1995	70%
1994 and Prior	60%

Credibility Z = 30%

1.47. (9, 11/98, Q.14) (1 point) In "An Example of Credibility and Shifting Risk Parameters," Mahler discusses the maximum reduction in the mean squared error of an estimate that can be accomplished by using credibility.
You are given the following estimates based upon one year of data:
Mean squared error relying solely on the data = 80.
Mean squared error ignoring the data = 100.
What is the best mean squared error that can be achieved using a linear weighted average of the two estimates?

1.48. (9, 11/98, Q.25) (4 points) For the past 25 years, the Bermuda Captives have battled in the highly competitive Island Sunshine League. Their losses in each individual 100 game season are shown below, in five year intervals. Also shown below are the 25 year average losing percentages for each team in the Island Sunshine League. Each team played 100 games in each of the 25 years.

Bermuda Captives Loss Record	5 Year Subtotal
Seasons 1 - 5	160
Seasons 6 -10	170
Seasons 11 - 15	294
Seasons 16 - 20	330
Seasons 21 - 25	296

Team	25 Year Average Loss %
Bermuda Captives	50.0%
Barbados Bombers	60.0%
Jamaica White Sox	55.0%
Trinidad Hurricanes	45.0%
Cayman Cubs	40.0%

Critical Chi-Square statistic at 95% confidence level: 9.488
In Mahler's paper "An Example of Credibility and Shifting Risk Parameters," the author discusses three tests to perform on the data sets being observed. Use Mahler and the data above to answer the following questions.

a. (0.5 point) Mahler performs a test using the binomial distribution on the data set. What is the purpose of this test?
b. (0.75 point) Perform the binomial test at the 95% confidence level using the standard normal approximation, and give your conclusion of that test with respect to the above data. Show all work.
c. (0.5 point) Mahler performs a chi-square test on the data set. What is the purpose of this test?
d. (0.75 point) Perform the chi-square test described by Mahler at the 95% confidence level, and give your conclusion of that test with respect to the above data. Show all work.
e. (0.5 point) Mahler performs a correlation test on the data set. What is the purpose of this test?
f. (1 point) Describe how one would perform the correlation test on the above data set. What would the likely conclusion be on the above data set?

1.49. (9, 11/99, Q.48) (4 points) In Mahler's "An Example of Credibility and Shifting Parameters," the author gives the following equation:

$$V(Z) = \sum_{i=1}^{N}\sum_{j=1}^{N} Z_i\, Z_j\, (\tau^2 + C(|i - j|)) - 2\sum_{i=1}^{N} Z_i\, (\tau^2 + C(N + \Delta - i)) + \tau^2 + C(0)$$

where Z_1 is the credibility for the earliest year used.

a. (1 point) Define the following terms:

i) V(Z)

ii) τ^2

iii) C(k)

iv) Δ

b. (3 points) The Cayman Island Captives play in the Actuarial Baseball League. Using the following information, predict the Captives' winning percentage in the year 2000, based on least squares credibility as described by Mahler with N = 2 years of data. Show all work.

Year	Winning Percentage
1997	55%
1998	40%
1999	45%
Grand Mean	50%

τ^2	0.1000
C(0)	0.8000
C(1)	0.5000
C(2)	0.3500

1.50. (9, 11/00, Q.34) (2 points) Answer the following based on Mahler's "An Example of Credibility and Shifting Risk Parameters."

a. (1.5 points) Briefly describe three criteria used to compare the performance of credibility methods.

b. (0.5 point) Mahler states that one criterion differs from the other two criteria on a conceptual level. Which criterion is that? Briefly state in what way it differs from the others.

1.51. (9, 11/01, Q.1) (1 point) In Mahler's "An Example of Credibility and Shifting Risk Parameters," the author evaluates various estimates for baseball teams' future losing percentages using historical losing percentages He discusses the impact of shifting parameters over time in this context. According to Mahler, which of the following statements regarding shifting risk parameters is <u>false</u>?

A. The correlation between years that are close together is significantly less than the correlation between years that are further apart.

B. With delays in receiving historical data, the resulting estimates of the future will be less accurate.

C. Based on the least squares criterion, the optimal credibility decreases with increased delays in receiving the data.

D. If the data available to predict the next year, Year_{x+1}, included only data from Year_{x-1}, there is a significant increase in the squared error as compared to what would result if the data available included Year_{x}.

E. Older years of data should be given substantially less credibility than more recent years of data.

1.52. (9, 11/03, Q.21) (1 point)
Briefly describe two methods to test whether risk parameters shift over time.

1.53. (9, 11/04, Q.3) (1 point) Three experience rating plans have been developed and you are trying to evaluate which is optimal. Each rating plan has been tested on four different risks. The following tables summarize the indicated modifications and the resulting errors.

Plan 1		
Risk Number	Predicted Modification Factor	Error
1	1.30	40%
2	1.30	40%
3	0.70	30%
4	0.70	30%

Plan 2		
Risk Number	Predicted Modification Factor	Error
5	1.30	10%
6	1.30	-10%
7	0.70	-20%
8	0.70	20%

Plan 3		
Risk Number	Predicted Modification Factor	Error
9	1.30	4%
10	1.20	2%
11	0.80	-2%
12	0.70	-4%

Which of the following summarizes the preferred plan based on the Meyers/Dorweiler criterion and the least squared error criterion?

	Meyers/Dorweiler Criterion	Least Squared Error Criterion
A.	Plan 1	Plan 2
B.	Plan 1	Plan 3
C.	Plan 2	Plan 1
D.	Plan 2	Plan 3
E.	Plan 3	Plan 3

1.54. (9, 11/05, Q.2) (3 points)
a. (1.5 points) Expected losses for a risk within a class are projected based on the formula:
E = Z X + (1 - Z) P, where
X = the most recent accident year's losses
P = the prior estimate of the most recent accident year
Z = the credibility assigned to the most recent accident year
Assume:
- No delay in obtaining data
- Z = 10%

What is the difference in the weight given to accident year 2001 losses in accident year 2002's estimate and the weight given to accident year 20011 losses in accident year 2005's estimate?
b. (1.5 points) If there are significant shifts in risk parameters that require Z to be reevaluated, will the answer to part a. above increase, decrease, or remain constant. Explain your answer. Assume that there are no changes other than the shifts in risk parameters.

***1.55*. (9, 11/07, Q.6)** (2 points)
The actuary for an insurance company has been asked by senior management to determine whether the company's expected frequency has been shifting over time.
The actuary knows that the company has maintained a constant number of exposures and a uniform mix of business since 1997.
Based on an assumption that expected frequency has remained constant during the period, the actuary has compiled the following data.

Year	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Actual Claims	475	420	460	500	490	525	515	510	540	575
Expected Claims	500	500	500	500	500	500	500	500	500	500

Discuss two methods that the actuary could use to test whether the expected frequency has been shifting over time. Describe any assumptions, calculations, or additional information that would be necessary to completely formulate and carry out each test.

1.56. (8, 11/12, Q.3) (1.75 points) The table below shows property claim frequency by year for the last five years. Assume that claim frequencies are Poisson distributed with a mean of 1.5.

Year	Exposures	Frequency
2011	118	1.5
2010	132	1.7
2009	121	1.3
2008	109	1.6
2007	97	1.3

The critical value for the relevant chi-squared distribution is 9.49.

a. (1.25 points) Calculate the chi-squared test statistic for whether the claim frequency is shifting over time. Interpret the result.

b. (0.5 point) Describe a second method for testing whether the claim frequency is shifting over time.

1.57. (8, 11/15, Q.4) (2.25 points)

An actuary is reviewing an account that has been with the company for over ten years.

Given the following:

- The claim frequency for this account follows a Poisson distribution, with $\lambda = 0.012$
- The recorded frequency for the last five years is as follows:

Year	Exposures	Frequency
2010	9,500	0.011
2011	11,000	0.010
2012	13,000	0.013
2013	10,500	0.012
2014	12,000	0.010

- The critical value for the relevant Chi-squared distribution is 9.49

a. (1.5 points) Use the Chi-squared test to evaluate whether the claim frequency is shifting over time. Include the hypotheses, test statistic, and provide an interpretation of the result.

b. (0.75 points)

Fully describe another method for determining whether claim frequency is shifting over time.

1.58. (CAS Sample Q.11, taken from the Fall 2021 Exam 8) (3.5 points)
The following are historical loss ratios for three risks in a class:

Year	Risk 1	Risk2	Risk 3
2012	63.1%	72.5%	52.3%
2013	59.0%	52.6%	48.9%
2014	63.5%	69.7%	53.9%
2015	74.3%	73.8%	50.1%
2016	45.9%	61.7%	50.9%
2017	42.3%	57.8%	46.6%
2018	58.9%	67.2%	48.6%
2019	60.2%	56.5%	50.7%
2020	52.8%	58.3%	46.1%

Grand Mean 57%
To estimate 2021 loss ratios for each risk, an actuary gives equal weight to the 3 most recent years (N = 3), and the grand mean of 57% is used as the complement (weight 1 - Z).
a. (1.5 points) Calculate the credibility, to the nearest tenth (Z = 0, 0.1, 0.2, etc ...), that minimizes the mean squared error of the actuary's prediction.
b. (1 point) Using the Small Chance of Large Errors Criterion with an error threshold of 5%, determine if the credibility calculated in Part a is the most appropriate.
c. (0.75 point) Justify which of the above two criteria the actuary should use to determine credibility, and use that credibility to estimate the loss ratio for 2021 for each of the three risks.
d. (0.25 point) Briefly describe a test that can be used to determine if the recommendation in Part c is reasonable.
SHOW ALL WORK.

Solutions:

1.1. D.

1.2. Statement 1 is false.
Only statement number 2 is correct.
The losing percentages of the various teams are not random; there are better teams and worse teams (statement 2).
One might still argue: "Maybe teams are not the same, but perhaps past performance is a poor predictor of future performance." So Mahler shows that experience in one period has predictive power for other periods.
Statement 3 is false. Insurance risks are independent; a loss for one insured does not imply anything about losses for other insureds. (Certain lines of business, such as wind losses for homeowners, are exceptions.) Baseball differs from insurance by the constraint on the overall losing percentage: it is always 50%. If there were only two teams, the won-loss percentage of one team tells us the won-loss percentage of the other team, but with enough teams and games played each year, this constraint is not material.

1.3. Statement 1 is true; Mahler shows this for all of the teams (see page 236 of the text).
Statements 2 and 3 are Mahler's conclusions from his data; he says (page 239):
On the other hand, there is a significant correlation between the results of years close in time. Thus recent years can be usefully employed to predict the future.

1.4. D. Under rule #1, a bad record one year is less likely to produce a bad record the following year than under rule #3. Therefore, $Z(1) < Z(3)$.
Under rule #2, a good record one year is more likely to produce a good record the following year than under rule #3. Therefore, $Z(2) > Z(3)$.
Thus, $\mathbf{Z(1) < Z(3) < Z(2)}$.
Comment: None of these is the rule used by the NBA, but rule #1 is the closest.

1.5. Statement 1: With the current draft rules, where the worst team gets the first draft pick, the optimal credibility may be less than zero. To see this, suppose there were just one player on each team and 100 teams. The team that does worst one year, gets the first draft pick, and it becomes one of the best teams.
For example, a team with a losing percentage of 80% one year, might be expected the next year to have a losing percentage of 40%. Z (80%) + (1 - Z)(50%) = 40%, would imply Z = -1/3, so the credibility is less than zero.
With 12 players on a team (though only five starters) and only two to three dozen teams, the worst team does not necessarily become one of the best from a single draft pick, but it may move up to above average, which also implies negative credibility.
Statement 2: If the better teams get higher draft picks, the teams which are good one year are expected to become even better the next year, and the teams which are poor one year are expected to become even worse the next year.
For example, a team with a losing percentage of 40% one year, might be expected the next year to have a losing percentage of 35%. Z (40%) + (1 - Z)(50%) = 35%, would imply Z = 1.5, so the credibility is greater than one.
Comment: The period of time studied in Mahler's paper was prior to the 1965 introduction of baseball's draft of players.
For a history of the rules for the NBA draft see:
http://www.nba.com/history/draft_evolution.html

1.6. Statement 1: Mahler uses linear combinations of the previous years' loss ratios and the overall average loss ratio.
Statement 2: Each estimator gives an overall 50% losing percentage, so it is unbiased. In the baseball analogy, the overall expected won-loss ratio is 50%. Similarly, the overall mean is 50%, the average of last year's won-loss ratio for all teams is 50%, and so forth.
Rating procedures can be biased for several reasons; we give illustrations:
Trended or developed losses are generally biased, though we may not know size or even the direction of the bias. A rate review may use an 8% loss cost trend; if the trend is higher or lower than 8%, the trended loss ratio is biased.
A loss ratio credibility weighted with loss ratios from other states or other insurers is biased, since other states or insurers may have higher or lower expected loss ratios.
Statement 3: These estimators are like experience rating; they use past experience to predict future experience. An analogy for schedule rating would be to look at the recent draft picks to predict the changes in next year's losing percentages.

1.7. With a credibility of Z and a relative loss ratio the previous year of L_1, the predicted relative loss ratio for the second year is: $Z L_1 + (1 - Z)$.
The squared error is: $\{Z L_1 + (1 - Z) - L_2\}^2 = \{Z (L_1 - 1) - (L_2 - 1)\}^2$.
Taking the sum of the squared errors, and setting the derivative with respect to Z equal to zero:
$\Sigma\ 2 \{Z (L_1 - 1) - (L_2 - 1)\} (L_1 - 1) = 0$.
$\Rightarrow Z = \Sigma (L_1 - 1)(L_2 - 1) / \Sigma (L_1 - 1)^2$.
$\Sigma (L_1 - 1)(L_2 - 1) = (-0.10)(-.02) + (-0.07)(-0.07) + ... + (0.10)(0.04) = 0.0226$.
$\Sigma (L_1 - 1)^2 = (-0.10)^2 + (-0.07)^2 + ... + (0.10)^2 = 0.0340$.
The credibility is 0.0226/ 0.0340 = **66.5%**.
Comment: The mean of each years relative loss ratios is 1, by definition.
The credibility is the slope of the regression line, which is the linear curve of best fit to the data points, using least squares.

1.8. 1 and 2 only. See page 249 of Mahler.

1.9. The a priori mean is 50%, which is given the remaining 25% weight.
The teams predicted losing percentage is:
(10%){95/(67 + 95)} + (10%){101/(61 + 101)} + (55%){96/(66 + 96)} + (25%)(50%) = 0.572.
Out of 88 games, this team is expected to lose: (0.572)(88) = 50.3 games.
Therefore, the predicted record is about: **38 wins and 50 losses**.
Comment: This data is for the Tampa Bay Rays of the American League.
Through 7/7/08 inclusive, their record in 2008 was 55 wins and 33 losses.
This is an example of a large prediction error.
It is impossible to avoid some large prediction errors, particularly when using a simple technique based solely on past losing percentages.
Hopefully, such large prediction errors are rare in experience rating.

1.10. Statement 1 uses an unweighted average. If risk parameters shift over time, last year's losses may be a good predictor of next year's losses, but losses from five years ago may be a poor predictor. Using more years of an unweighted average may not improve the estimate. See section 8.3 of Mahler, page 245, second to last paragraph from bottom:
“Based on most actuarial uses of credibility, an actuary would expect the optimal credibilities to increase as more years of data are used. In this example they do not. In fact, using more than one or two years of data does an inferior job according to this criterion.”
The use of older data with equal weight eventually leads to a worse outcome.
Table 17 on page 266 of Mahler shows empirical results. With a 1 year unweighted average, the optimal credibility is 66.0%. The optimal credibility increases to a maximum of 73.6% with 4 years of data, and then decreases as the years increase. At 10 years, the credibility has decreased to 66.9%.
Statement 2 says that if we know the general range in which the optimal credibility value lies, it doesn't make much difference what value we pick from that range.
In the past, some actuaries believed it was important to choose the proper full credibility standard, since different credibility values may produce more or less accurate rates. Not so, says Mahler. If the credibility values is near the optimal value, there is little difference in the accuracy of the rates.
Statement 3 is false. The best we can expect is to reduce the mean squared error to about 75% of the lower of the original estimates. See the third paragraph on page 252 of Mahler:
“In the current case, the best that can be done using credibility to combine two estimates is to reduce the mean squared error between the estimated and observed values to 75% of the minimum of the squared errors from either relying solely on the data or ignoring the data.”
<u>Comment</u>: If Statement 1 were changed to a weighted N-year average, it be would true.
A weighted average using N years is a special case of the weighted average using N+1 years, since it is an N+1 year average with a weight of 0 for the oldest year. Since the N year average is one instance of an N+1 year average, the optimal N+1 year average must be at least as good as the optimal N year average.
In statement 2, Mahler is not saying that the credibility values do not affect the rate indication. Different credibility values give different indications. Suppose the indicated pure premium is \$5.00 per \$100 of payroll and the underlying pure premium is \$2.00 per \$100 of payroll. A credibility of 60% gives a rate of (60%)(\$5.00) + (40%)(\$2.00) = \$3.80, and a credibility of 40% gives (40%)(\$5.00) + (60%)(\$2.00) = \$3.20. This is a difference of about 15 to 20% in the rates. Different credibility values give different indicated rates. But the two sets of rates may have about the same expected squared error.
Suppose the true pure premium is \$4.00 per \$1 00 of payroll. The \$3.80 rate has a squared error of $(4.00 - 3.80)^2 = 0.040$ and the \$3.20 rate has a squared error of $(4.00 - 3.20)^2 = 0.640$. This is an enormous difference. However, we are speaking about the <u>expected</u> squared error. For a given size, we are choosing between 40% and 60% credibility. Sometimes the 40% credibility gives a higher rate and sometimes the 60% credibility gives a higher rate. Mahler says that the mean squared error won't differ much, as long as both values are close to the optimal value. If the optimal value is between 40% and 60%, both credibilities give about the same expected squared error.
Let's change the scenario. Suppose we don't know the true pure premium. One credibility value gives an indicated pure premium of \$3.20, whereas another credibility value gives an indicated pure premium of \$3.80. We don't know which estimate is closer to the true pure premium. Sometimes the first estimate is better, and sometimes the second is better.

We ask, over the full distribution of true pure premiums, which estimate is better? Mahler says: As long as the credibility is near the optimal value, there is not much difference.
Suppose a credibility estimator gives a pure premium that just equals the future loss costs. The squared error is zero, which is less than 75% of the minimum. However, Mahler is talking about the expected squared error, not the actual squared error in any particular instance. The actual squared error may be 0% by happenstance.
If the mean squared error is zero, the estimator is right on the mark; predicting future experience perfectly. This never happens, since there is random fluctuation in the losses.
You might recall the 75% figure as follows. If the optimal credibility is close to 0% or to 100%, it doesn't reduce the mean squared error much from the lower of relying entirely on the data or not relying on the data at all. If the optimal credibility is 10%, the difference between 0% and 10% is not great. The largest effect occurs when the optimal credibility is 50%. In that case, we should be using 50% of each estimator instead of 100% of either the experience or the overall mean. The mean squared error is reduced by the square of 50%, or 25%; this is the complement of 75%. $1 - 0.5^2 = 0.75$.
The previous paragraph is obviously not a mathematical derivation of the 75% result; see Appendix E of Mahler's paper, not on the syllabus, for the derivation. It simply shows the intuition.

1.11. None of these statements are true!
Statement 1 should say second order function, not linear. Mahler says on the top of page 264: "Equation 11.2 shows that the squared error is a second order polynomial in the Z_i. This equation is the fundamental result for analyzing least squares credibility."
In equation 11.2 on page 263, V(Z) is the expected squared error. The second order polynomial comes from the $Z_i Z_j$ product in the first double summation.
Statement 2 is false. Table 16 on page 264 shows the solution of the matrix equations, and many of the credibilities are negative. In footnote 45 of page 265, Mahler says:
"Giving negative weight to some years allows a larger weight to be given to other years. The net effect is to reduce the expected squared error."
This result is counterintuitive. One might think that the negative credibilities stem from random loss fluctuations. But the negative credibilities all show up in the same columns (columns 5, 7, and 8 of Table 16), so there is some systematic effect, though finding an intuitive explanation is difficult.
Statement 3 is false. Mahler divides the variance into three parts, but he does not use this division for least squares credibilities. As he says at the top of page 263:
"It is possible to divide the within variance into two parts. The first part is the process variance excluding the effect of shifting parameters over time. The second part is that portion of the within variance due to shifting parameters over time. While this division may aid our understanding, it is not necessary for the calculation of the least squares credibilities."
Illustration: Suppose we examine a group of 100 large employers with $100 million of payroll apiece. The average employer has five workers' compensation claims a month. If we draw a sample of ten months, with each month taken from a random employer, what is the expected variance? That is, if we look at Employer #23 for January, Employer #41 for February, and so forth, what is the expected variance among the number of monthly claims?
Suppose first that there is no process variance and no between variance. If each employer has 5 claims each month, the variance of the number of claims in the sample is zero.
Suppose there is process variance but no between variance: that is, each employer has a Poisson distribution of claims with a mean of five. The process variance for each employer is 5, and the expected variance of the number of claims in the sample is 5.
Suppose the process variance is zero: that is, each employer has its expected number of claims each month, but the between variance is not zero: the expected claims differ by employer. If the between variance is 8, then the expected variance of the sample is 8.
Suppose the between variance is zero (all employers are identical) and the process variance at any moment in time is zero (the employer always realizes the expected number of claims). If the expected number of claims changes over time, then the variance of the claims in the ten samples is more than zero. This is what is captured by Mahler's ζ^2.
<u>Comment</u>: To prepare for the exam, know equations 11.1, 11.2, 11.3, and 11.4. The derivation of equation 11.2 is in Appendix C of Mahler, <u>not</u> on the syllabus. Question 48 of the Fall 1999 exam asked a question about equation 11.2, which was given to you.

1.12. Statement 1 is false. Individual risk rating is most important (credibility is highest) when classes are heterogeneous; as the between variance increases, the credibility Z increases. For individual risk rating, the between variance is the variance among the risks in the class, not the variance between the classes.

If $\tau^2 = 0$, then the class is perfectly homogeneous; the expected losses of risks within the class do not differ. In this case, the class rate is the proper rate for each risk. The loss history of any risk gives us extra noise, not extra information; the proper experience rating credibility is zero.

If the class is exceedingly heterogeneous, in other words if τ^2 is large, then the class rate tells us little about the proper rate for any insured.

The insured's experience is a combination of noise and information. We use credibility to separate the information from the noise.

Statement 2 is true. Statement 2 deals with the flip side of this issue. Individual risk rating is useful if the risk's experience gives us information about that risk's loss propensities. Suppose that the within variance is zero; i.e., the losses don't change from year to year. If last year the loss costs were $100,000, they are $100,000 this year as well (adjusted for exposure changes and loss cost trend). As the within variance goes to zero, the credibility goes to 100%.

If the within variance is high, the loss history doesn't tell us much about expected losses. As the within variance increases, the credibility decreases. Small insureds have high within variance, so their credibility is low. Large insureds have low within variance, so their credibility is high. The within variance is $\delta^2 + \zeta^2$.

Statement 3: As the variance due to shifting parameters increases, we give less weight to older accident years and more weight to more recent years: as the variance due to shifting parameters increases, the weight given to year 1 (now 10%) decreases.

If instead $\zeta^2 = 0$, then there are no shifting risk parameters, and every year would be equally good for predicting the future.

Comment: The weights sum to 100%, so the weight given to year 5 (now 30%) increases if ζ^2 increases. We can't say anything about the weights for years 2, 3, and 4. Presumably, the weight for year 2 decreases, and the weight for year 4 increases, but we can't say this with certainty.

The within variance would also be called the Expected Value of the Process Variance (EPV).

The between variance would be also called the Variance of the Hypothetical Means (VHM).

1.13. The variance between risks (τ^2) is the variance of the true class rates for class ratemaking, or the variance between the individual risk propensities within a class for experience rating. The process variance excluding the effect of shifting parameters over time (δ^2) is the random fluctuation of the class loss costs for class ratemaking, or the random fluctuation of an individual risk's loss costs for experience rating. The portion of the within variance due to shifting parameters over time (ζ^2) is the variance of the average class pure premium stemming from changes in the class risk parameters over time, or the variance in the individual risk's expected pure premium stemming from changes in the insured's attributes over time.

1.14. (a) A rating plan that uses a reasonable primary-excess split has less variance of the individual risk's credibility weighted experience, whereas a no-split rating plan has higher variance. Thus this change in the rating plan doesn't change the variance of the hypothetical means, but it effectively decreases the process variance of the individual risk's experience. This decreases the K value in the credibility formula and increases the credibility. There is no change in the effect of shifting risk parameters.
(b) Using 2 years instead of 5 in the rating plan does not change the variance of the hypothetical means. If we use a weighted average of the years, using only 2 years degrades the rating plan and increases the process variance, so the experience rating credibility decreases.
Using only the most recently available 2 years of data reduces the effects of shifting risk parameters on the experience rating plan. (Which years we use does not change the covariance structure of the entire data set.) If we use an unweighted average of the years of data, using only 2 years may improve the rating plan compared to using 5 years, if the effect of shifting risk parameters is large. (See Table 6 in Mahler.) In that case, the experience rating credibility could be larger for either 2 or 5 years, depending on the details. (See Table 9 in Mahler.)
(c) Refining the classification plan decreases the variance of the hypothetical means, since all risks within any class are more alike. There is no change on the process variance of any individual risk. The K value in the credibility formula increases and credibility decreases. We can rely more on the class rate and less on the individual's experience, since the class rate is now a better estimate of the individual's loss potential than it was before the classification plan was refined.
There is no change in the effect of shifting risk parameters.

1.15. Only statement 1 is correct. At the top of page 285, Mahler writes:
“If an experience rating plan works properly, then after the application of experience rating, an insurer should be equally willing to write debit and credit risks. In other words, the modified loss ratio of expected losses to modified premiums should be the same for debit and credit risks.”
Underwriters sometimes say: "We don't want to give this risk such a large debit, we don't want to punish him too much for one accident." This sentence is confused; the confusion stems from common parlance. The experience rating plan is not rewarding or punishing a risk for good or poor experience. Rather, the past experience helps predict future loss costs, and the modified rates are the best estimate of the future loss costs.
Since the standard premium, which includes the debit or credit mod, is the best estimate of future loss costs, insurers should be indifferent between debit and credit risks.
Statement 3 is false. Meyers/Dorweiler solely deals with if there is a pattern in the errors. For any experience rating plan, there is some credibility that satisfies Meyers/Dorweiler. If the plan successfully identifies good and poor risks, the credibility should be high; if the plan can not identify good and poor risks, the credibility should be low. In each case, the proper credibility gives a Kendall tau statistic of zero and satisfies the Meyers-Dorweiler criterion.
The plan may have enormous errors, but if there is no pattern, the ideal credibility satisfies Meyers-Dorweiler. See page 271 of Mahler.
Statement 2 is false. The Kendall τ statistic reflects the correlation in the order of two series. If two series are from uncorrelated distributions, the expected Kendall τ statistic is zero, and the actual correlation is symmetrically distributed on [-1, +1]. The same statements are true for the Kendall τ statistic as for the statistical correlation; see page 286 of Mahler's paper in Appendix B, <u>not</u> on the syllabus. If the credibility approaches zero, past experience is not used at all. The modification is one for all risks, and the correlation with the actual loss costs is zero.
The expected value of τ is zero; the actual value in any instance is symmetrically distributed over [-1, +1].

1.16. 1 only.
We do not directly observe expected values; that eliminates choices 2 and 3. Insurance, unlike baseball, has no constraint on the grand mean. We estimate the mean by observing all risks over a long period. However, that is still an estimate subject to random fluctuation.
For insurance situations where we are interested in relativities compared to average, then by definition M = 1, however it is not directly observed.

1.17. 1 and 2 only.
Statement 1 is true. Mahler says on pages 252-253:
“If there is a delay before the data are available for use in experience rating, the resulting estimate of the future will be less accurate.
As the delay increases, the squared error increases significantly.
Statement 2 is correct as well. As Mahler says on page 254:
"The optimal credibility (as determined using the least squares criterion) decreases as the delay increases. Less current information is less valuable for estimating the future.”
Statement 3 is false for the data set examined. The predictive value declines slowly as the delay increases, and it takes many years before it gets close to zero. Table 11 on page 254 shows the figures. The average credibility is about 70% with a 1 year lag between latest data point and future prediction and about 45% with a 4 year lag. Statement 3 might be true for some data set where risk parameters were shifting significantly faster than in the baseball data examined by Mahler.
However, this is extremely unlikely to be the case for insurance data; insurance data tend have parameters that are more stable than in Mahler’s baseball data.
Comment: What is the relation between delays and shifting risk parameters?
Suppose we predict the pure premium for year 6 using 3 years of experience data.
If there is no delay, we use years 3, 4, and 5.
If there is a short delay, we use years 2, 3, and 4.
If there is a long delay, we use years 1, 2, and 3.
If the risk parameters don't shift over time, all three methods should have similar expected squared errors.
If the risk parameters shift over time, then the first method is best, and the last method is worst.

1.18. In both cases, the correlations decline with increasing separation of the years.
This indicates that parameters are shifting over time.
The rate of decline in correlations is swifter for data set one, indicating that parameters are shifting more quickly for data set one than for data set two.

1.19. Let P_{2009} = estimate of 2009. Let X_{2008} = observation for 2008.
$P_{2009} = Z\ X_{2008} + (1 - Z)\ P_{2008}$.
Similarly, $P_{2008} = Z\ X_{2007} + (1 - Z)\ P_{2007}$.
$P_{2007} = Z\ X_{2006} + (1 - Z)\ P_{2006}$.
Therefore, $P_{2009} = Z\ X_{2008} + (1 - Z)\ P_{2008} = Z\ X_{2008} + (1 - Z)Z\ X_{2007} + (1 - Z)^2\ P_{2007}$
$= Z\ X_{2008} + (1 - Z)Z\ X_{2007} + (1 - Z)^2\ Z\ X_{2006} + (1 - Z)^3\ P_{2006}$.
The coefficient for X_{2006} is $(1 - Z)^2\ Z$.
When Z = 55%, $(1 - Z)^2\ Z = (1 - 0.55)^2\ (0.55)$ = **11.1 %.**
Comment: The weights applied to years of data decline geometrically.
This form of estimator is similar to what is done in pure premium ratemaking.
For pure premium ratemaking, the credibility weighted pure premium is:
Z (the indicated pure premium) + (1 - Z) (the underlying pure premium).
In loss ratio ratemaking, the credibility weighted loss ratio is:
(Z) the experience loss ratio + (1 - Z) (the permissible loss ratio).

1.20. None of 1, 2, or 3 is said by Mahler.
Comment: See Mahler at page 272.
Statement 3 is one of the most practical implications from Mahler's paper: if we know the approximate credibility, a refined figure doesn't make much of a difference. For example, any credibility figure between 40% and 70% might give about the same expected squared error. Actuaries sometimes argue whether the full credibility standard should be 2,500 claims or 3,000 claims. In many cases, it doesn't make much difference.

1.21. Statements 1 and 2 are correct; the mean squared error is the expected squared error. Statement 3 is false. To solve the second order equation, Mahler takes partial derivatives. This creates linear equations, which can be solved for the credibilities.
Comment: See Mahler at page 280.

1.22. (a) (1) Some new teams entered the leagues due to expansion. Mahler had the same 8 teams in each league throughout. We would have varying numbers of teams. For example, in 1969 the Kansas City Royals and Seattle Pilots (now the Milwaukee Brewers) joined the American League. These new teams were worse than average. Thus the existing teams seemed to improve on average between 1968 and 1969.
(2) Some seasons were shortened by strikes. Thus there are some years where a significantly smaller number of games were played.
(3) Leagues were split into divisions, and in recent seasons, teams play teams within their division more frequently. Thus unlike in Mahler's study, teams do not play approximately the same number of games against each other team in their league. If in a given season a certain division is significantly stronger than average, then the teams in that division play opponents who are stronger than average. Therefore, the expected winning percentages for teams in that division would be lower than it would otherwise be if there was a balanced schedule.
(4) Interleague play was introduced recently. While only about 10% of games involve play between the two leagues, this complication was not present in Mahler's Study.
The average winning percentage for a league is no longer 50% each year.
(For example, in 2006 the American League won 154 out of 252 interleague games;
154/252 = 61%. Thus that year, the average winning percentage for the American League was greater than 50%.) Also the expected winning percentage of a team is effected by which teams it is scheduled to play that season. Each season, a team only plays some of the teams in the other league and that varies from year to year.
(b) Since Mahler's study, free agency was introduced. Thus players switch teams more frequently now. Thus I would expect the effect of shifting risk parameters to be greater than in Mahler's study.
Alternately, the difference between the best and the worst teams is usually less than in Mahler's study; there is more parity among the teams. Therefore, there is a smaller region in which the winning percentages can vary from year to year. Thus I would expect the effect of shifting risk parameters to be less than in Mahler's study.
Alternately, since Mahler's study, baseball has instituted a draft. Teams with the worst record get to draft earlier. This will tend to allow bad teams to get better more quickly. Conversely good teams will have a harder time staying good for a long time. Therefore, parameters may shift more quickly than in the era in Mahler's study.
<u>Comment</u>: There are many possible reasonable answers. In part (a) only give two reasons.

1.23. For two different years, $\text{Cov}[X_i, X_j] = \tau^2 + C(|i - j|)$.

For example, $\text{Cov}[X_1, X_3] = \tau^2 + C(2) = 5 + 8 = 13$.

For a single year of data, $\text{Cov}[X_i, X_i] = \text{Var}[X_i] = \tau^2 + C(0) = 5 + 50 = 55$.

A covariance matrix is:

$$\begin{matrix} \text{Year 1} \\ \text{Year 2} \\ \text{Year 3} \\ \text{Year 4} \\ \text{Year 5} \end{matrix} \begin{pmatrix} 55 & 15 & 13 & 11 & 9 \\ 15 & 55 & 15 & 13 & 11 \\ 13 & 15 & 55 & 15 & 13 \\ 11 & 13 & 15 & 55 & 15 \\ 9 & 11 & 13 & 15 & 55 \end{pmatrix}.$$

$\sum_{j=1}^{N} Z_j \text{ Cov}[X_i, X_j] = \text{Cov}[X_i, X_{N+\Delta}]$, where we are predicting year $N + \Delta$, using years 1 to N.

(a) Using data for Year 1 to Predict Year 2, the equation is:

$55Z = 15. \Rightarrow Z = 15/55 =$ **27.3%**.

(b) Using data for Year 1 to Predict Year 3, the equation is:

$55Z = 13. \Rightarrow Z = 13/55 =$ **23.6%**.

(c) Using data for Year 1 to Predict Year 4, the equation is:

$55Z = 11. \Rightarrow Z = 11/55 =$ **20.0%**.

(d) Using data for Years 1 and 2 to Predict Year 3, the equations are:

$55Z_1 + 15Z_2 = 13.$

$15Z_1 + 55Z_2 = 15.$

The coefficients on the lefthand side are the first two rows and the first two columns of the covariance matrix, since we are using data from Years 1 and 2. The values on the righthand side are the first two rows of column three, since we are predicting year 3.

Solving, $\mathbf{Z_1 = 17.5\%}$ and $\mathbf{Z_2 = 22.5}\%$.

(e) Using data for Years 1 and 2 to Predict Year 4, the equations are:

$55Z_1 + 15Z_2 = 11.$

$15Z_1 + 55Z_2 = 13.$

The values on the righthand side are the first two rows of column four, since we are predicting Year 4.

Solving, $\mathbf{Z_1 = 14.6\%}$ and $\mathbf{Z_2 = 19.6\%}$.

(f) Using data for Years 1, 2, and 3 to Predict Year 4, the equations are:

$55Z_1 + 15Z_2 + 13Z_3 = 11.$

$15Z_1 + 55Z_2 + 15Z_3 = 13.$

$13Z_1 + 15Z_2 + 55Z_3 = 15.$

The coefficients on the lefthand side are the first three rows and the first three columns of the covariance matrix, since we are using data from Years 1, 2, and 3. The values on the righthand side are the first three rows of column four, since we are predicting Year 4.

Solving, $\mathbf{Z_1 = 11.0\%}$, $\mathbf{Z_2 = 15.0}\%$, and $\mathbf{Z_3 = 20.6}\%$.

(g) Using data for Years 1, 2, and 3 to Predict Year 5, the equations are:
$55Z_1 + 15Z_2 + 13Z_3 = 9.$
$15Z_1 + 55Z_2 + 15Z_3 = 11.$
$13Z_1 + 15Z_2 + 55Z_3 = 13.$
The values on the righthand side are the first three rows of column five, since we are predicting Year 5.
Solving, $\mathbf{Z_1 = 8.6\%}$, $\mathbf{Z_2 = 12.7}\%$, and $\mathbf{Z_3 = 18.1}\%$.
Comment: Parts f and g are beyond what you should be asked on your exam.
See Equation 11.3 in Mahler.
These linear equations are called the Normal Equations, as discussed in Loss Models.
The notation used in the syllabus paper by Mahler, written in 1988 and published in 1990, is unnecessarily complex. All one really needs is the covariance matrix. The least squares credibilities are determined by the relative sizes of the elements of the covariance matrix.
With no delay in getting data, $\Delta = 1$, similar to Mahler's Table 16:

	Years Between Data and Estimate		
Number of Years of Data Used (N)	1	2	3
1	27.3%		
2	22.5%	17.5%	
3	20.6%	15.0%	11.0%

With a delay in getting data, $\Delta = 2$:

	Years Between Data and Estimate		
Number of Years of Data Used (N)	2	3	4
1	23.6%		
2	19.6%	14.6%	
3	18.1%	12.7%	8.6%

1.24. Use data for Years 1 and 2 to Predict Year 3.
$\text{Cov}[(X_1 + X_2)/2, X_3] = \{\text{Cov}[X_1, X_3] + \text{Cov}[X_2, X_3]\} / 2 = (13 + 15)/2 = 14.$
$\text{Var}[(X_1 + X_2)/2] = \{\text{Var}[X_1] + \text{Var}[X_2] + 2\ \text{Cov}[X_1, X_2]\} / 2^2 = \{55 + 55 + (2)(15)\} / 4 = 35.$
Thus the weight given to the average of the years is: Z = 14/35 = **40%**.
(Thus 20% weight is given to each year. When giving different weights we got: Z_1 = 17.5% and Z_2 = 22.5%. Note that 17.5% + 22.5% = 40%.)
Use data for Years 1 and 2 to Predict Year 4.
$\text{Cov}[(X_1 + X_2)/2, X_4] = \{\text{Cov}[X_1, X_4] + \text{Cov}[X_2, X_4]\} / 2 = (11 + 13)/2 = 12.$
$\text{Var}[(X_1 + X_2)/2] = \{\text{Var}[X_1] + \text{Var}[X_2] + 2\ \text{Cov}[X_1, X_2]\} / 2^2 = \{55 + 55 + (2)(15)\} / 4 = 35.$
Thus the weight given to the average of the years is: Z = 12/35 = **34.3%**.
Use data for Years 1, 2 and 3 to Predict Year 4.
$\text{Cov}[(X_1 + X_2 + X_3)/3, X_4] = \{\text{Cov}[X_1, X_4] + \text{Cov}[X_2, X_4] + \text{Cov}[X_3, X_4]\} / 3 =$
$(11 + 13 + 15)/3 = 13.$
$\text{Var}[(X_1 + X_2 + X_3)/3] =$
$\{\text{Var}[X_1] + \text{Var}[X_2] + \text{Var}[X_3] + 2\ \text{Cov}[X_1, X_2] + 2\ \text{Cov}[X_1, X_3] + 2\ \text{Cov}[X_2, X_3]\} / 3^2 =$
$\{55 + 55 + 55 + (2)(15) + (2)(13) + (2)15)\} / 9 = 251/9.$
Thus the weight given to the average of the years is: Z = 13 / (251/9) = **46.6%**.
Use data for Years 1, 2 and 3 to Predict Year 5.
$\text{Cov}[(X_1 + X_2 + X_3)/3, X_5] = \{\text{Cov}[X_1, X_5] + \text{Cov}[X_2, X_5] + \text{Cov}[X_3, X_5]\} / 3 =$
$(9 + 11 + 13)/3 = 11.$
$\text{Var}[(X_1 + X_2 + X_3)/3] =$
$\{\text{Var}[X_1] + \text{Var}[X_2] + \text{Var}[X_3] + 2\ \text{Cov}[X_1, X_2] + 2\ \text{Cov}[X_1, X_3] + 2\ \text{Cov}[X_2, X_3]\} / 3^2 =$
$\{55 + 55 + 55 + (2)(15) + (2)(13) + (2)(15)\} / 9 = 251/9.$
Thus the weight given to the average of the years is: Z = 11 / (251/9) = **39.4%**.

<u>Comment</u>: One could use equation 11.4: $Z = N \dfrac{N\,\tau^2 + \sum\limits_{I=1}^{N} C(N+\Delta-i)}{N^2\,\tau^2 + \sum\limits_{j=1}^{N}\sum\limits_{I=1}^{N} C(|i-j|)}.$

Requiring that the weights applied to each year be equal results in a larger minimum mean squared error than allowing the weights to vary.

1.25. (a) Shifting Risk Parameters: The parameters defining the risk process for an individual insured are not constant over time. There are (a series of perhaps small) permanent changes to the individual insured's risk process as one looks over several years.
(b) 1. Private Passenger Automobile Insurance:
A driver's risk parameters might shift if he changed the location to which he commutes to work.
(He drivers the same distance, but it is over different type of roads.)
So for example, if he now drives to work over local streets while before he mostly drove on a highway, his expected pure premium changes.
2. Workers Compensation Insurance:
There might be a change in the attitude of management with regard to workplace safety.
If management of the company paid more attention to workplace safety, then the expected pure premium would go down.
3. Homeowners Insurance:
The number of children in the neighborhood changes over time.
(The insured remains in the same house and there are no children living there.)
As the number of neighborhood children increases, there is more chance of a liability claim; the expected pure premium for the liability coverage would increase.
<u>Comment</u>: There are many possible examples. one could give in part (b).
The risk parameters of a Workers Compensation class can shift over time, so that its relativity compared to average for its Industry Group changes over time. This could be due to changes in the manufacturing process, how the work is performed, or the nature of the job.
The automobile experience of a town relative to the rest of the state could shift as that town becomes more densely populated.
The insurance experience of a town relative to the rest of the state could shift as that town undertook an effective campaign against insurance fraud.

1.26. (a) (1) Some new teams entered the leagues due to expansion. Mahler had the same 8 teams in each league throughout. We would have varying numbers of teams. When new and weaker teams enter the league, the existing teams seem to improve compared to average.
(2) The number of games per season was increased over this period of time from 14 to 16. Thus the amount of data varies.
(3) Some seasons were shortened by strikes. Thus there is one year (1982) where a significantly smaller number of games were played.
(4) Unlike in Mahler's study, teams do not play approximately the same number of games against each other team in their league. If in a given season a certain division is significantly stronger than average, then the teams in that division play opponents who are stronger than average. Therefore, the expected winning percentages for teams in that division would be lower than it would otherwise be if there was a balanced schedule.
(5) Each season a team plays at most 16 games, compared to about 150 in Mahler's study. Thus there is much more random fluctuation in the data than in Mahler's Study.
(b) Since the average career of a star football player is shorter than the average career of a star baseball player, I would expect the quality of a team to change more quickly in football. Thus, I would expect shifting risk parameters to have more effect on football data.
Alternately, since there are more players on a football team than a baseball team, the effect on the quality of the team from replacing one player is less than for baseball. I would expect the quality of a team to change less quickly in football. Thus, I would expect shifting risk parameters to have less effect on football data.
<u>Comment</u>: There are many possible reasonable answers. In part (a) only give two reasons.
Feel free to make up a similar question to answer based on some other team sport you may prefer, such as basketball, hockey, soccer, etc.

1.27. For two different years, $Cov[X_i , X_j] = \tau^2 + C(|i - j|)$.

For example, $Cov[X_2, X_5] = \tau^2 + C(3) = 10 + 6 = 16$.

For a single year of data, $Cov[X_i, X_i] = Var[X_i] = \tau^2 + C(0) = 10 + 30 = 40$.

A covariance matrix is:

Year 1	40	25	20	16	13
Year 2	25	40	25	20	16
Year 3	20	25	40	25	20
Year 4	16	20	25	40	25
Year 5	13	16	20	25	40

$\sum_{j=1}^{N} Z_j \ Cov[X_i , X_j] = Cov[X_i, X_{N+\Delta}]$, where we are predicting year N + Δ, using years 1 to N.

(a) Using data for Year 1 to Predict Year 2, the equation is: 40Z = 25.

⇒ Z = 25/40 = **5/8 = 62.5%**.

(b) Using data for Year 1 to Predict Year 3, the equation is: 40Z = 20.

⇒ Z = 20/40 = **1/2 = 50.0%**.

(c) Using data for Year 1 to Predict Year 4, the equation is: 40Z = 16.

⇒ Z = 16/40 = **40.0%**.

(d) Using data for Year 1 to Predict Year 5, the equation is: 40Z = 13.

⇒ Z = 13/40 = **32.5%**.

(e) Using data for Years 1 and 2 to Predict Year 3, the equations are:

$40Z_1 + 25Z_2 = 20$.

$25Z_1 + 40Z_2 = 25$.

The coefficients on the lefthand side are the first two rows and the first two columns of the covariance matrix, since we are using data from Years 1 and 2. The values on the righthand side are the first two rows of column three, since we are predicting year 3.

Solving, $\mathbf{Z_1 = 7/39 = 17.9\%}$ and $\mathbf{Z_2 = 20/39 = 51.3}\%$.

(f) Using data for Years 1 and 2 to Predict Year 4, the equations are:

$40Z_1 + 25Z_2 = 16$.

$25Z_1 + 40Z_2 = 20$.

The values on the righthand side are the first two rows of column four, since we are predicting Year 4.

Solving, $\mathbf{Z_1 = 28/195 = 14.4\%}$ and $\mathbf{Z_2 = 16/39 = 41.0\%}$.

(g) Using data for Years 1 and 2 to Predict Year 5, the equations are:

$40Z_1 + 25Z_2 = 13$. $25Z_1 + 40Z_2 = 16$.

Solving, $\mathbf{Z_1 = 8/65 = 12.3\%}$ and $\mathbf{Z_2 = 21/65 = 32.3\%}$.

<u>Comment</u>: See Equation 11.3 in Mahler.

1.28. (a) $V(Z) = Z^2\{\tau^2 + C(0)\} - 2\,Z\{\tau^2 + C(1)\} + \tau^2 + C(0) =$
$40Z^2 - (2)(25)Z + 40 = (40)(5/8)^2 - (50)(5/8) + 40 =$ **24.375**.
(b) $V(Z) = Z^2\{\tau^2 + C(0)\} - 2\,Z\{\tau^2 + C(2)\} + t^2 + C(0) =$
$40Z^2 - (2)(20)Z + 40 = (40)(1/2)^2 - (40)(1/2) + 40 =$ **30**.
(c) $V(Z) = Z^2\{\tau^2 + C(0)\} - 2\,Z\{\tau^2 + C(3)\} + \tau^2 + C(0) =$
$40Z^2 - (2)(16)Z + 40 = (40)(0.4)^2 - (32)(0.4) + 40 =$ **33.6**.
(d) $V(Z) = Z^2\{\tau^2 + C(0)\} - 2\,Z\{t^2 + C(4)\} + \tau^2 + C(0) =$
$40Z^2 - (2)(13)Z + 40 = (40)(0.325)^2 - (26)(0.325) + 40 =$ **35.775**.
(e) Years 1 and 2 predicting year 3.

$$V(Z) = (Z_1\ , Z_2, -1, 0, 0) \begin{pmatrix} 40 & 25 & 20 & 16 & 13 \\ 25 & 40 & 25 & 20 & 16 \\ 20 & 25 & 40 & 25 & 20 \\ 16 & 20 & 25 & 40 & 25 \\ 13 & 16 & 20 & 25 & 40 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ -1 \\ 0 \\ 0 \end{pmatrix} =$$

$$(7/39, 20/39, -1, 0, 0) \begin{pmatrix} 40 & 25 & 20 & 16 & 13 \\ 25 & 40 & 25 & 20 & 16 \\ 20 & 25 & 40 & 25 & 20 \\ 16 & 20 & 25 & 40 & 25 \\ 13 & 16 & 20 & 25 & 40 \end{pmatrix} \begin{pmatrix} 7/39 \\ 20/39 \\ -1 \\ 0 \\ 0 \end{pmatrix} =$$

(7/39, 20/39, -1, 0, 0) . (0, 0, -920/39, -463/39, -369/39) = 920/39 = **23.59**.
(f) Years 1 and 2 predicting year 4.

$$V(Z) = (Z_1\ , Z_2, 0, -1, 0) \begin{pmatrix} 40 & 25 & 20 & 16 & 13 \\ 25 & 40 & 25 & 20 & 16 \\ 20 & 25 & 40 & 25 & 20 \\ 16 & 20 & 25 & 40 & 25 \\ 13 & 16 & 20 & 25 & 40 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ -1 \\ 0 \\ 0 \end{pmatrix} =$$

$$(28/195, 16/39, 0, -1, 0) \begin{pmatrix} 40 & 25 & 20 & 16 & 13 \\ 25 & 40 & 25 & 20 & 16 \\ 20 & 25 & 40 & 25 & 20 \\ 16 & 20 & 25 & 40 & 25 \\ 13 & 16 & 20 & 25 & 40 \end{pmatrix} \begin{pmatrix} 28/195 \\ 16/39 \\ 0 \\ -1 \\ 0 \end{pmatrix} =$$

(28/195, 16/39, 0, -1, 0) . (0, 0, -463/39, -5752/195, -1077/65) = 5752/195 = **29.50**.

(g) Years 1 and 2 predicting year 5.

$$V(Z) = (Z_1\ , Z_2, 0, 0, -1) \begin{pmatrix} 40 & 25 & 20 & 16 & 13 \\ 25 & 40 & 25 & 20 & 16 \\ 20 & 25 & 40 & 25 & 20 \\ 16 & 20 & 25 & 40 & 25 \\ 13 & 16 & 20 & 25 & 40 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ 0 \\ 0 \\ -1 \end{pmatrix} =$$

$$(8/65, 21/65, 0, 0, -1) \begin{pmatrix} 40 & 25 & 20 & 16 & 13 \\ 25 & 40 & 25 & 20 & 16 \\ 20 & 25 & 40 & 25 & 20 \\ 16 & 20 & 25 & 40 & 25 \\ 13 & 16 & 20 & 25 & 40 \end{pmatrix} \begin{pmatrix} 8/65 \\ 21/65 \\ 0 \\ 0 \\ -1 \end{pmatrix} =$$

(8/65, 21/65, 0, 0, -1) . (0, 0, -123/13, -1077/65, -432/13) = 432/13 = **33.23**.

<u>Comment</u>: The expected squared error is given by equation 11.2:

$$V(\vec{Z}) = \sum_{i=1}^{N}\sum_{j=1}^{N} Z_i Z_j\{\tau^2 + C(|i-j|)\} - 2\sum_{i=1}^{N} Z_i\{\tau^2 + C(N+\Delta-i)\} + \tau^2 + C(0).$$

It turns out that Equation 11.2 can be rewritten in matrix form,
Mean Squared Error = V(Z) = Z^T C Z. C is the matrix of covariances for the years of data.
Z is the (column) vector with credibilities in the years used to estimate, -1 in the year being estimated, and zeros in any other years. Z^T is the transpose of Z.
Note that year 1 is a worse predictor of year 3 than it is of year 2. Therefore, the mean square error is larger for predicting year 3 than it is for predicting year 2; 30 > 24.375.
The longer the delay in getting data, the larger the mean squared error.
Using years 1 and 2 to predict year 3 is a better estimator than using just year 2. Therefore, the mean square error is larger using just year 2 than it is using years 1 and 2; 30 > 23.59.

1.29. D. While we would like statement A to be true, with several years of data and a particular set of covariances, the optimal weights can be negative.
While statement B sounds like it should be true, with several years of data and a particular set of covariances, the optimal weight for 1953 data may be more than that for 1954 data.
While statement C sounds like it should be true, the optimal weight assigned to the most recent year of data may be slightly more when using for example 6 years of data rather than 5 years of data.
Statement D is true. See Table 19 in Mahler.
For example using 1952 and 1953 to predict 1954 is a special case of using 1951, 1952, and 1953 to predict 1954, where 1951 is given a weight of zero. Thus the minimum expected squared error from the latter can not be more than that from the former.
<u>Comment</u>: If the covariance matrix was more structured, one could usually avoid negative credibilities. See for example, Tables 4 and 7 in “A Markov Chain Model of Shifting Risk Parameters”, by Howard C. Mahler, PCAS 1997.

1.30. Var[X] = 6. $Cov[X_1, X_2] = 0.9$. $Cov[X_1, X_3] = 0.9^2 = 0.81$.

$Cov[X_1, X_4] = 0.9^3 = 0.729$. $Cov[X_1, X_5] = 0.9^4 = 0.6561$. $Cov[X_1, X_6] = 0.9^5 = 0.59049$.

The covariance matrix between the years of data is:

$$\begin{pmatrix} 6 & 0.9 & 0.81 & 0.729 & 0.6561 & 0.59049 \\ 0.9 & 6 & 0.9 & 0.81 & 0.729 & 0.6561 \\ 0.81 & 0.9 & 6 & 0.9 & 0.81 & 0.729 \\ 0.729 & 0.81 & 0.9 & 6 & 0.9 & 0.81 \\ 0.6561 & 0.729 & 0.81 & 0.9 & 6 & 0.9 \end{pmatrix}$$

Therefore, the equations for the least squares credibilities (the Normal Equations) are:

$6Z_1 + 0.9Z_2 + 0.81Z_3 + 0.729Z_4 + 0.6561Z_5 = 0.59049$.

$0.9Z_1 + 6Z_2 + 0.9Z_3 + 0.81Z_4 + 0.729Z_5 = 0.6561$.

$0.81Z_1 + 0.9Z_2 + 6Z_3 + 0.9Z_4 + 0.81Z_5 = 0.729$.

$0.729Z_1 + 0.81Z_2 + 0.9Z_3 + 6Z_4 + 0.9Z_5 = 0.81$.

$0.6561Z_1 + 0.729Z_2 + 0.81Z_3 + 0.9Z_4 + 6Z_5 = 0.9$.

Solving: $Z_1 = 5.525\%$, $Z_2 = 6.373\%$, $Z_3 = 7.560\%$, $Z_4 = 9.150\%$, $Z_5 = 11.228\%$.

5.525% + 6.373% + 7.560% + 9.150% + 11.228% = 39.836%.

The remaining weight of 60.164% is given to the a priori mean.

<u>Comment</u>: Beyond what you will be asked on your exam.

The older years are less correlated with year 6, the year we wish to estimate, and thus their data is given less weight.

See "A Markov Chain Model of Shifting Risk Parameters," by Howard C. Mahler, PCAS 1997, <u>not</u> on the syllabus.

1.31. For each set of predictions we calculate the errors: predicted - observed.

Policy Number	Rating Plan 1 Modification Factor	Error
1	0.80	-0.05
2	0.90	+0.05
3	1.00	0
4	1.10	+0.05
5	1.20	-0.05

Policy Number	Rating Plan 2 Modification Factor	Error
1	0.87	+0.02
2	0.87	+0.02
3	1.00	0
4	1.03	-0.02
5	1.23	-0.02

Policy Number	Rating Plan 3 Modification Factor	Error
1	0.81	-0.04
2	0.83	-0.02
3	1.00	0
4	1.09	+0.04
5	1.27	+0.02

Plan 2 has positive errors for debit risks and negative errors for credit risks.
Plan 3 has negative errors for debit risks and positive errors for credit risks.
In both cases, the errors are correlated with the experience modifications.
In the case of Plan 1, the errors have a correlation close to zero with the experience modifications.
Thus **by the Meyers/Dorweiler criterion, we prefer Plan 1**.
Plan 1 has a larger average squared error than plan 3, which has a larger average squared error than plan 2. Thus **by the least squared error criterion we prefer plan 2**.
Comment: Intended as an improvement on the less than completely logical past exam question: 9, 11/04, Q.3.
One would do such testing on thousands of policies rather than just 5.

1.32. This probably is based on a situation with shifting risk parameters.
More distant years are given less weight than more recent years.
For example, if we were using 2003, 2004, and 2005 to predict 2006, we would give 2003 weight 2.82%, 2004 weight 3.01% and 2005 weight 3.23%. (The remaining weight of 90.94% would be given to the overall mean.) This makes sense, since 2003 is less correlated with 2006 than is 2005, and thus is a worse predictor of 2006 than is 2005.
Using fewer years of data is a special case of using more years of data, where some of the credibilities have been constrained to be zero. (The credibilities by year are allowed to differ.)
Thus using more years of data does a better job than using fewer years of data. Thus the minimum mean squared errors should decline as we use the least square credibilities for more years of data. This is in fact what we observe. For example, for 4 years of data the minimum mean squared error is 0.03826, while for 5 years of data it is 0.03824. (As more years are added, the MSE continues to improve, but only very slowly. Eventually there will no longer be any significant improvement from adding years.)
The sum of the credibilities increases as the number of years of data increase; we give less weight to the overall mean. The sum of credibilities increases at a decreasing rate. (With shifting risk parameters, as the number of years of data approaches infinity, the sum of credibilities will approach a value less than one. This differs from the Buhlmann Credibility formula,
$Z = N / (N+K)$, where the limit is one.)
<u>Comment</u>: Based on an approximation to the model of California Female P. P. Auto Drivers in "A Markov Chain Model of Shifting Risk Parameters", by Howard C. Mahler, PCAS 1997.
The credibilities shown by year are similar to those in Table 4 of that paper.
The results shown in the question were based on: $\text{Cov}[X_i, X_j] = (0.0014)\ 0.94^{|i-j|} + 0.037\ \delta_{ij}$,
where δ_{ij} is zero if $i \neq j$ and one if $i=j$.
The sum of the credibilities approaches about 32% as the number of years approach infinity.
(Similar to Figure 12 in the Markov Chain paper.)
The minimum mean squared error approaches 0.03814 as the number of years approach infinity.
This model has smoothed out the peculiarities of the covariances of the data that are due to random fluctuation. Thus we see a regular pattern of credibilities. More distant years get less credibility than more recent years, declining in a nice pattern. This female driver data has a slower rate of shifting risk parameters than does the baseball data, thus the credibilities for distant years decline more slowly for the driver data than the baseball data.

1.33. The first covariance matrix has all of its elements equal. The variance of a year is the same as the covariance of two different years. Thus all of the years of data are perfectly correlated. Whatever the observed relativity (for a baseball team, or insured, or class) is in one year, it is the same in every other year. This is not a reasonable model for insurance.
In the second covariance matrix, the variance of year of data is 200, while the covariance between different years is 140.
Therefore, the correlation of any two different years of data is: 140/200 = 70%.
The correlation between different years does not depend on how far apart they are.
In the third covariance matrix, the variance of a year of data is 200, while the covariance between consecutive years is 140, and the covariance of year 1 and year 3 is 110.
Therefore, the correlation of consecutive years of data is: 140/200 = 70%, the correlation of years 1 and 3 is: 110/200 = 55%. The correlation between years further apart is less than the correlation of years closer together. This is what we expect with shifting risk parameters over time.
The third matrix corresponds to a situation of shifting risk parameters over time.
Comment: The second matrix is an example of the Buhlmann covariance structure.

1.34. For years 1, 2, and 3, the covariance matrix is: $\begin{pmatrix} 207 & 136.638 & 111.296 \\ 136.638 & 207 & 136.368 \\ 111.296 & 136.368 & 207 \end{pmatrix}$.

Thus if we are using years 1 and 2 to predict year 3, the least squares credibilities satisfy:
$207 Z_1 + 136.638 Z_2 = 111.296$.
$136.638 Z_1 + 207 Z_2 = 136.368$.
Solving $Z_1 = 0.1807$ and $Z_2 = 0.5408$.
Then using equation 11.2, the minimum expected squared error is:
$207 Z_1^2 + 207 Z_2^2 + (2)(136.638) Z_1 Z_2 - (2)(111.296) Z_1 - (2)(136.638) Z_2 + 207 = 112.994$.
With 1 being the most recent year, proceeding in a similar manner we get:

Year	N=1	N=2	N=3	N=4	N=5	N=6	N=7	N=8	N=9	N=10
1	66.0%	54.1%	52.9%	52.7%	52.7%	52.6%	52.6%	52.6%	52.5%	52.5%
2		18.1%	14.7%	14.2%	14.1%	14.1%	14.0%	14.0%	14.0%	14.0%
3			6.3%	4.6%	4.3%	4.2%	4.2%	4.1%	4.1%	4.1%
4				3.2%	2.0%	1.7%	1.6%	1.6%	1.6%	1.6%
5					2.3%	1.3%	1.0%	0.9%	0.9%	0.9%
6						2.0%	1.0%	0.8%	0.7%	0.7%
7							1.8%	0.9%	0.7%	0.6%
8								1.7%	0.8%	0.6%
9									1.5%	0.8%
10										1.4%
Total	66.0%	72.1%	73.9%	74.7%	75.3%	75.8%	76.2%	76.6%	77.0%	77.3%
MSE	116.81	112.99	112.55	112.43	112.38	112.33	112.30	112.27	112.24	112.22

Note that the values shown in a column may not sum to the total shown due to rounding.
Comment: The covariances are on a basis of number of games lost for the baseball data; they are based on a model of the baseball data shown at page 661 of “Credibility With Shifting Risk Parameters, Risk Heterogeneity, and Parameter Uncertainty,” by Howard C. Mahler, PCAS 1998. (This is a mixture of two Markov Chains with different rates of shifting of risk parameters.) This model has smoothed out the peculiarities of the covariances of the data that are due to random fluctuation. Thus we see a regular pattern of credibilities. (Contrast here to Table 16 in the syllabus reading.) More distant years get less credibility than more recent years, declining in a nice pattern.
However, there is also an “edge effect”; the most distant year used tends to get more weight since it is correlated with even more distant years. For example, when using five years of data, for example 1951 to 1955, then the most distant year, 1951, contains useful information about years 1950, 1949, 1948, etc. Thus 1951 is given more weight than 1952; 2.3% > 2.0%.

1.35. (a) The purpose is to test whether risk parameters shift over time.
In other words, determine whether inherent loss potential (L%) is shifting over time for each team.
(b) The test is applied separately to the data of one baseball team.
H_0: The expected losing percentage is the same over time for this team.
Compute this team's losing percentage over the whole experience period (of 60 years).
Then group data for that team into appropriate intervals; Mahler groups the 60 years into 5 year non-overlapping intervals.
Calculate for each interval: $(A - E)^2/E$,
where A = actual observation = (5 year mean losing percentage)(5 years)(150 games),
and E = expected observation = (60 year mean losing percentage)(5 years)(150 games).
Sum up the contributions for all intervals in order to get the chi-square statistic.
Compare to the Chi-Square Distribution with number of degrees of freedom equal to the number of intervals minus one; in the paper Mahler compares to the Chi-Square with 11 degrees of freedom.
If the statistic is greater than the critical value for the appropriate significance level, for example 5%, then for this team we reject the null hypothesis that parameters do not shift over time.
Comment: See Table 4 in the paper. For each of the 16 teams, the p-value was less than 0.2%.

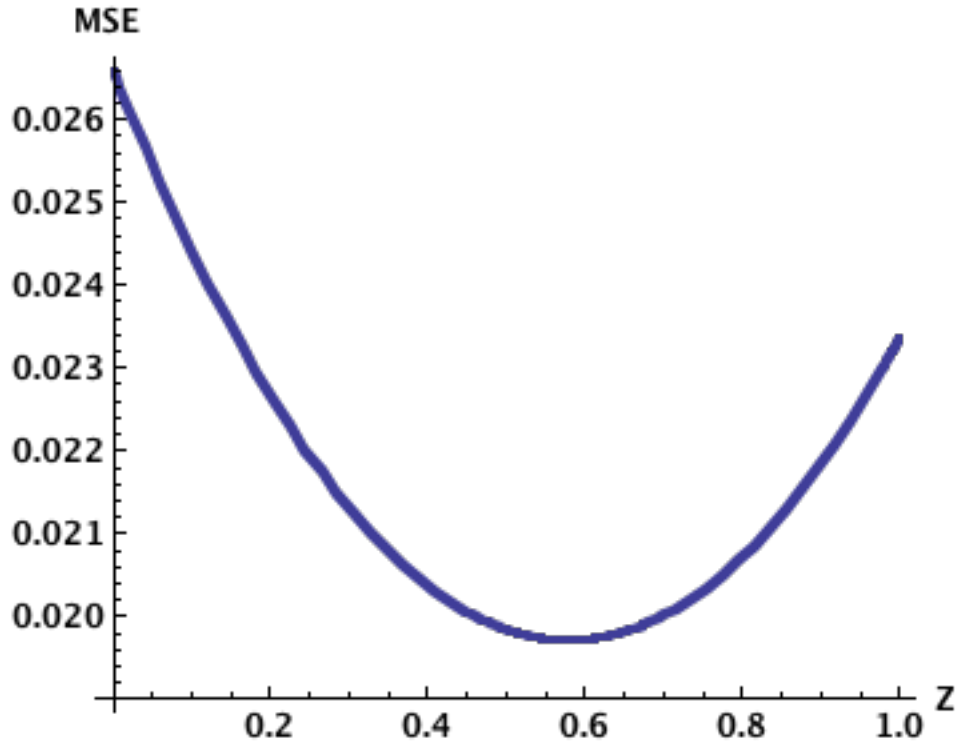
1.36. In the case of credibilities that can differ by year, using three years of data is a special case of using four years of data, with the credibility of the most distant year constrained to be zero. Therefore, the best credibilities for four years of data do at least as well and probably better than the best credibilities for three years of data. As expected, we see the minimum expected squared errors decline as we used more years of data. After a while we reach of point of diminishing improvement. For example, ten years with a MSE of 57.04 is only slightly better than nine years with a MSE of 57.05.
In each case, constraining the credibilities to be the same by year is a special case of allowing the credibilities to differ. Thus allowing the credibilities to vary does at least as well and probably better than requiring the credibilities to be the same by year. In fact, the mean squared errors are smaller for the case where the credibilities differ. (For N =1 the two methods are the same.) For example, for four years of data, 57.36 is better than 57.45.
In the case of credibilities that are the same by year, using fewer years of data is a not a special case of using more years of data. Distant years should be given very little weight, but we are requiring all years to be given the same weight. While adding year of data may be better, with shifting risk parameters, eventually adding years of data will be worse. In this case, the mean squared errors improve through 6 years of data. However, after that the mean squared errors increase. For example, using 7 years of data has a mean squared error of 57.40, worse than the 57.36 for 6 years of data.
Comment: The results shown in this question are based the covariance between different years of data being: $\text{Cov}[X_i, X_j] = (10)\ 0.88^{|i-j|} + 50\ \delta_{ij}$, where δ_{ij} is zero if $i \neq j$ and one if $i=j$.
Parameters shift more slowly than in the baseball example.
Given the covariances, one could solve for the least squares credibilities.
For example, using 3 three years of data: Z_1 = 8.3%, Z_2 = 9.9%, and Z_3 = 12.1%.
Using 3 three years of data instead with equal weights, each year is weighted 10.1%.

1.37. The overall average is 0.51875.
So the expected wins for each decade are: (0.51875)(games).
The contributions are: (observed - expected)^2 / expected.

Decade	Wins	Losses	Expected Wins	Contribution
1900s	691	634	687.34375	0.01945
1910s	857	624	768.26875	10.24802
1920s	595	938	795.24375	50.42172
1930s	705	815	788.50000	8.84242
1940s	854	683	797.31875	4.02946
1950s	814	725	798.35625	0.30654
1960s	764	845	834.66875	5.98330
1970s	895	714	834.66875	4.36084
1980s	821	742	810.80625	0.12816
1990s	814	741	806.65625	0.06686
2000s	920	699	839.85625	7.64776
2010s	788	670	756.33750	1.32548
Total	9518	8830	9518.02500	93.38002

We compare to the 10% critical value for 12 - 1 = 11 degrees of freedom: 17.275.
Since 93.380 > 17.275, there is evidence of shifting risk parameters.
<u>Comment</u>: Similar to 8, 11/18, Q. 1a.

1.38. For example the estimate for 2001 using 2000 is: (0.6)(0.951) + 0.4 = 0.9706.
The estimates are: 0.9706, 0.9886, 0.9112, 0.8962, 0.874, 0.8962, 0.9628, 0.889, 0.8962, 0.8962, 0.9406, 0.9334, 1.0888, 0.8812, 1.0744, 1.0222, 0.9112, 0.9112, 0.8002.
(The estimate of 0.8002 for 2019 can not be compared to an actual value.)
The squared error for estimating 2001 is: $(0.9706 - 0.981)^2$ = 0.00010816.
The average squared error is: **0.0197233**.
<u>Comment</u>: Similar to 8, 11/18, Q. 1b.
A graph of the mean squared error as a function of Z:



The minimum MSE is for Z = 58%.

1.39. a) Try different values of Z. The smallest squared error is for Z = **0.2**.
For example, for Z = 0.2, the estimate for year 2016 for Risk 1 is:
(0.2)(0.93 + 1.12 + 0.83)/3 + (1 - 0.2)(1) = 0.9920.
Squared error is: $(0.9920 - 0.90)^2$ = 0.00846.

0.2	Observed			Predictions			Squared Error		
Year	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3
2013	0.93	0.89	0.75						
2014	1.12	0.94	1.39						
2015	0.83	1.10	0.98						
2016	0.90	0.73	1.08	0.9920	0.9953	1.0080	0.00846	0.07040	0.00518
2017	0.81	0.89	1.56	0.9900	0.9847	1.0300	0.03240	0.00896	0.28090
2018	1.11	1.06	0.89	0.9693	0.9813	1.0413	0.01979	0.00619	0.02290
2019	1.27	0.71	1.00	0.9880	0.9787	1.0353	0.07952	0.07218	0.00125
2020	0.65	0.86	0.76	1.0127	0.9773	1.0300	0.13153	0.01377	0.07290
2021	0.59	0.94	0.99	1.0020	0.9753	0.9767	0.16974	0.00125	0.00018
							MSE	**0.05542**	

(To more decimal places, the smallest mean squared error is for Z = 21%.)

b) Predictions of the 2022 relative loss ratios using Z = 20%:
Risk 1: (0.2)(1.27 + 0.65 + 0.59)/3 + (1 - 0.2)(1) = **0.9673**.
Risk 2: (0.2)(0.71 + 0.86 + 0.94)/3 + (1 - 0.2)(1) = **0.9673**.
Risk 3: (0.2)(1.00 + 0.76 + 0.99)/3 + (1 - 0.2)(1) = **0.9833**.

c) The absolute error is: $\left|\frac{\text{observed}}{\text{predicted}} - 1\right|$.

Try different values of Z. The fewest large errors are for Z equal to either **0.3, 0.4, or 0.5**.
For example, for Risk 1 the prediction for 2015 is 0.9840. |0.90/0.9840 - 1| = 8.5%.

0.4	Observed			Predictions			Absolute Error		
Year	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3
2013	0.93	0.89	0.75						
2014	1.12	0.94	1.39						
2015	0.83	1.10	0.98						
2016	0.90	0.73	1.08	0.9840	0.9907	1.0160	8.5%	26.3%	6.3%
2017	0.81	0.89	1.56	0.9800	0.9693	1.0600	17.3%	8.2%	47.2%
2018	1.11	1.06	0.89	0.9387	0.9627	1.0827	18.3%	10.1%	17.8%
2019	1.27	0.71	1.00	0.9760	0.9573	1.0707	30.1%	25.8%	6.6%
2020	0.65	0.86	0.76	1.0253	0.9547	1.0600	36.6%	9.9%	28.3%
2021	0.59	0.94	0.99	1.0040	0.9507	0.9533	41.2%	1.1%	3.8%

For Z = 40%, out of 18 cases, there are 11 errors of size more than 10%.

Alternately, the absolute error is: $\left|\frac{\text{predicted}}{\text{observed}} - 1\right|$.

Try different values of Z. The fewest large errors are for Z equal to either **0.3, 0.4, or 0.5**.
For example, for Risk 1 the prediction for 2015 is 0.9840. |0.9840/0.90 - 1| = 9.3%.

0.4	Observed			Predictions			Absolute Error		
Year	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3
2013	0.93	0.89	0.75						
2014	1.12	0.94	1.39						
2015	0.83	1.10	0.98						
2016	0.90	0.73	1.08	0.9840	0.9907	1.0160	9.3%	35.7%	5.9%
2017	0.81	0.89	1.56	0.9800	0.9693	1.0600	21.0%	8.9%	32.1%
2018	1.11	1.06	0.89	0.9387	0.9627	1.0827	15.4%	9.2%	21.6%
2019	1.27	0.71	1.00	0.9760	0.9573	1.0707	23.1%	34.8%	7.1%
2020	0.65	0.86	0.76	1.0253	0.9547	1.0600	57.7%	11.0%	39.5%
2021	0.59	0.94	0.99	1.0040	0.9507	0.9533	70.2%	1.1%	3.7%

For Z = 40%, out of 18 cases, there are 11 errors of size more than 10%.
<u>Comment</u>: Similar to CAS Sample Q.11, taken from the Fall 2021 Exam 8.
I believe that the first solution matches the syllabus reading, while the second matches the CAS solution to Sample Q.11.

1.40. Statement 1 is backwards. As the delay in receiving data increases, its predictive value decreases and the credibility decreases.
Statement 2 is backwards.
Statement 3 is backwards. If one gives each year equal weight, as the number of years increases, eventually the accuracy will decrease. (If one determines separate optimal credibilities by year, as the number of years increases, eventually the accuracy will no longer increase significantly.)
Comment: The conclusions in this exam question are those of Bizarro-Mahler on a planet opposite of the real world.

1.41. 1. Least squared error.
Minimize the squared error or the mean squared error between the observed and predicted results. Analogous to Buhlmann credibility.
2. Small chance of large errors.
Minimize the probability that the observed results will be more than some chosen % different from the predicted. Analogous to classical credibility.
3. Meyers/Dorweiler.
Minimize, in other words make equal to zero, the correlation between:
$\frac{\text{observed}}{\text{predicted}}$ and $\frac{\text{predicted}}{\text{overall average}}$.
Use some correlation measure; Mahler uses Kendall's statistic, which counts inversions.
Meyers/Dorweiler results differ from the others because it's concerned with patterns rather than sizes of errors.
Comment: The results from the first two methods are usually very similar.

1.42. 1. False; should say 75% rather than 50%. See page 252 (and Appendix E.)
2. True. See page 271 (and Appendix B.)
3. False! See page 277.
Comment: The original statement #1 was false from Appendix E, no longer on the syllabus.
“Credibility methods reduce the squared error between the observed value and the estimated/ predicted value to a greater extent than they reduce the squared error between the true mean and the estimated predicted mean.”

1.43. For an individual team, the number of games lost is Poisson with mean nλ.
Therefore, for an individual team, the variance in number of games lost is also nλ.
The losing percentage is the number of games lost divided by n.
Therefore, the variance in losing percentage is: $n\lambda/n^2 = \lambda/n$.
Thus the expected value of the process variance in losing percentage is:
(0.5)(0.4/200) + (0.5)(0.6/200) = 0.0025.
The variance of the hypothetical mean losing percentages is:
$(0.5)(0.4 - 0.5)^2 + (0.5)(0.6 - 0.5)^2 = 0.01$.
The observed variance in losing percentages is:
$\{(75/200 - 0.5)^2 + ... + (94/200 - 0.5)^2\}/10 = 0.0138$.
Therefore, we can back out the amount of variance due to shifting risk parameters as:
0.0138 - 0.0025 - 0.01 = 0.0013.
The percentage of the total variance due to shifting risk parameters is: 0.0013/0.0138 = **9.4%**.
Comment: See page 297 of Mahler's Appendix D, no longer on the syllabus.
The paper observed 60 years and averaged the observed variances for the individual years.
The estimate from just one year of data is not reliable.
Also the paper assumed a Binomial Model.
There is no need to divide up the variance into three pieces in order to calculate credibilities.
This is something which may help your understanding, but is not necessary.
I would not have done this if I were rewriting the paper today.

1.44. For Z = 50%, the predicted loss ratios are:
For 1993: (65% + 75%)/2 = 70%. For 1994: (65% + 70%)/2 = 67.5%.
For 1993: (65% + 65%)/2 = 65%. For 1994: (65% + 60%)/2 = 62.5%
The total of the squared errors is: $(70 - 70)^2 + (65 - 67.5)^2 + (60 - 65)^2 + (55 - 62.5)^2 = 87.5$.
For Z = 0, the predicted loss ratios are all 65%, and the total of the squared errors is:
$(70 - 65)^2 + (65 - 65)^2 + (60 - 65)^2 + (55 - 65)^2 = 150$.
Since Z = 50% has a lower sum of squared errors than Z = 0, I agree with the client.
Comment: In practical applications one would not apply the least squares criterion to only 5 years of data from one insured. One could apply it to years of data from many similar insureds of similar size in order to determine which value of Z performs well.

1.45. a. Ratio 1 = (Team's actual losing percentage)/(Team's predicted losing percentage).
Ratio 2 = (Team's predicted losing percentage)/(grand mean of 50%).
b. Ratio 1 ⇔ The loss ratio to modified premium (loss ratio to standard premium).
Ratio 2 ⇔ The experience modification.
Comment: See Section 7.3 in the paper by Mahler.
Part b of this exam question is from Appendix B, which is no longer on the syllabus.
However, it would be a good idea to know this anyway.

1.46. New estimate = Z Latest year of data + (1 - Z)(prior estimate).
We start with an estimate of 60%.
Estimate of 1996 using data from 1995: (30%)(70%) + (1 - 30%) (60%) = 63%.
Estimate of 1997 using data from 1996: (30%)(80%) + (1 - 30%) (63%) = 68.1%.
Estimate of 1998 using data from 1997: (30%)(90%) + (1 - 30%) (68.1%) = 74.67%.
Estimate of 1999 using data from 1998: (30%)(100%) + (1 - 30%) (74.67%) = **82.269%**.
Comment: See Section 9.1 in the paper by Mahler.

1.47. The best that can be done using credibility to combine two estimates is to reduce the mean squared error between the estimated and observed values to 75% of the minimum of the squared errors from either relying solely on the data or ignoring the data.
(75%)(80) = **60**.
Comment: See Section 8.5 in the paper by Mahler.

1.48. a. To determine whether the data for each team was drawn from the same probability distribution. In other words, to determine whether an "inherent difference" in loss % exists between teams.
b. The variance in losing percentage in 2500 games would be: (0.5)(0.5)/2500 = 0.0001. standard deviation is: 1%.
If the data for each team was drawn from the same probability distribution, we would expect to see about 95% of the teams results between: 50% ± (2)(1%) = 48% to 52%.
In this case only 1 out of 5 teams is in that range.
(Two of the teams have losing percentages 5 standard deviations from average, while two team have losing percentages 10 standard deviations from average!)
Thus we conclude that the teams differ.
c. The purpose is to test whether risk parameters shift over time. In other words, determine whether inherent loss potential (L%) is shifting over time for each team.
d. The Bermuda Captives have an overall losing percentage of 50%.
The observed number of losses per 5 years for this team is: (5)(100)(50%) = 250.
(For this team this happens to also be the a priori mean.)
Chi-Square statistic is: $(160 - 250)^2/250 + (170 - 250)^2/250 + (294 - 250)^2/250 +$
$(330 - 250)^2/250 + (296 - 250)^2/250 = 99.808$.
(This statistic has: number of groups - 1 = 5 - 1 = 4 degrees of freedom.)
Since 99.808 > 9.488, we reject the null hypothesis at the 95% confidence level (5% significance level). We conclude that the risk parameters shift over time, at least for the Bermuda Captives.
e. The purpose is to test whether risk parameters shift over time.
f. For each year we have a vector of length 5 of losing percentages by team.
For the one year differential, we examine the correlation of the 24 sets of pairs of data separated by one year: year 1 versus year 2, year 2 versus year 3, etc.
Mahler uses Kendall's tau to measure the correlation.
We take the average of these 24 correlations for the one year differential.
We do the same for the two year differential, using the correlation of the 23 sets of pairs of data by two years. We take the average correlation for the two year differential.
We do the similar calculation for the other differentials in years.
If the risk parameters do not shift over time, the average correlation should not differ significantly between the one year differential, two year differential, and so forth. If the risk parameters shift over time, the average correlation should be highest for the one year differential, second highest for the two year differential, and so forth.
Given the results of the Chi-Square Test for the Bermuda Captives, the likely conclusion of this test is that the risk parameters shift over time.
Comment: See Section 4 in the paper by Mahler.

1.49. (a) V(Z) = the expected squared error between the observation and predication.
τ^2 = between variance.
C(k) = covariance for data for the same risk, k years apart = "within covariance."
Δ = the length of time between the latest year of data used and the year being estimated.
If Δ = 1, then there is no delay in receiving information.
(b) $V(Z) = Z_1^2(\tau^2 + C(0)) + Z_2^2(\tau^2 + C(0)) + 2\, Z_1Z_2(\tau^2 + C(1)) - 2\, Z_1(\tau^2 + C(2))$
$- 2\, Z_2(\tau^2 + C(1)) + \tau^2 + C(0)$
$V(Z) = 0.9\, Z_1^2 + 0.9\, Z_2^2 + 1.2\, Z_1Z_2 - 0.9\, Z_1 - 1.2\, Z_2 + 0.9.$
Setting the derivative of V with respect to Z_1 equal to zero:
$0 = 1.8Z_1 + 1.2Z_2 - 0.9.$
Setting the derivative of V with respect to Z_2 equal to zero:
$0 = 1.8Z_2 + 1.2Z_1 - 1.2.$ Solving, Z_1 = 10% and Z_2 = 60%.
Therefore, the weight given to the overall mean is: 1 - 10% - 60% = 30%.
Therefore, the estimate for the year 2000 is: (10%)(40%) + (60%)(45%) + (30%)(50%) = **46%**.
Alternately, for two different years, $Cov[X_i , X_j] = \tau^2 + C(|i - j|)$.
For example, $Cov[X_{1998}, X_{2000}] = \tau^2 + C(2) = 0.1000 + 0.3500 = 0.45$.
For a single year of data, $Cov[X_i, X_i] = Var[X_i] = \tau^2 + C(0) = 0.1000 + 0.8000 = 0.9000$.

A covariance matrix is:
$$\begin{array}{c} 1998 \\ 1999 \\ 2000 \end{array} \begin{pmatrix} 0.90 & 0.60 & 0.45 \\ 0.60 & 0.90 & 0.60 \\ 0.45 & 0.60 & 0.90 \end{pmatrix}.$$

$\Sigma\, Z_i\, Cov[X_i , X_j] = Cov[X_i, X_{N+\Delta}]$, where we are predicting year N + Δ, using years 1 to N.
Using data for Years 1998 and 1999 to Predict Year 2000, the equations are:
$0.9Z_1 + 0.6Z_2 = 0.45.$
$0.6Z_1 + 0.9Z_2 = 0.60.$
The coefficients on the lefthand side are the first two rows and the first two columns of the covariance matrix, since we are using data from Years 1998 and 1999. The values on the righthand side are the first two rows of column three, since we are predicting year 2000.
Proceed as before.
Comment: See page 263 and Equation 11.3 in Mahler. We give 1999 more weight than 1998.
Since N = 2, we do not use the information from 1997. In order to determine a least squares credibility to assign to 1997, we would need to be given C(3).
Mahler works with losing percentages. If one converted the data and the grand mean to losing percentages, the predicted losing percentage in 2000 would be:
(10%)(60%) + (60%)(55%) + (30%)(50%) = 54% = 1 - 46%.

1.50. a. 1. Least squares - minimize the total squared error between actual and predicted result.
2. Small chance of large error - minimize the likelihood that any one actual observation will be a certain % different from the predicted result.
3. Meyers/Dorweiler - minimize the correlation between the ratio of actual/predicted and the predicted/average actual.
b. Meyers/Dorweiler is different from the first two which focus on minimizing prediction error.
In contrast, Meyers/Dorweiler focuses on the pattern of the errors.

1.51. A. The principles for shifting risk parameters are:
Statements A and E. Years that are closer together have a higher correlation than years that are further apart, so credibility should be higher for more recent years. .
Statements B and C. Delays in receiving data make the experience less useful and it should receive less credibility.
The use of the current year of data to help predict next year increases the accuracy of the estimate, so statement D is true.

1.52. Correlation test:

- Group data by pairs based on time lag
- Calculate correlation for each pair
- Calculate the average correlation by time lag
- If the correlation decreases as time lag increases, then risk parameters shift over time.

Chi-Square Test:
Null Hypothesis - H_0: risk parameters do not shift over time

- Group data into appropriate intervals
- Calculate the overall expected value
- Then calculate for each interval, $(A - E)^2/E$,
  where A = actual observation and E = expected observation
- Sum up the contributions for all intervals in order to get the chi-square statistic.
- If the total statistic is greater than the critical value for number of intervals -1
  degrees of freedom, then reject the null hypothesis that parameters do not shift over time.

1.53. D. Under Plans 1 and 3, the risks with higher mods have larger errors.
Under Plan 2, there is no correlation between the mods and the errors; underwriters would be indifferent between writing credit or debit risks.
Therefore, Plan 2 does best under the Meyers/Dorweiler criterion.
Plan 3 has the smallest average squared error, so Plan 3 is preferred under the Least Squared Error Criterion.
<u>Comment</u>: This past exam questions was not really properly put together.
If plans 1 and 2 produce the same modification for the each risk, and they should have the same errors; they should perform the same.
Actual experience rating plans are tested on thousands of risks.
Based on this data, Plan 1 is a bad experience rating plan.

1.54. a. The weight given to accident year 2001 losses in accident year 2002's estimate is Z = 10%. We work out the weight given to accident year 2001 losses in accident year 2005's estimate as follows:

$P_{2005} = Z\ X_{2004} + (1 - Z)\ P_{2004} = Z\ X_{2004} + (1 - Z)\ \{Z\ X_{2003} + (1 - Z)\ P_{2003}\} =$

$Z\ X_{2004} + (1 - Z)\ Z\ X_{2003} + (1 - Z)^2\ P_{2003} =$

$Z\ X_{2004} + (1 - Z)\ Z\ X_{2003} + (1 - Z)^2\ \{Z\ X_{2002} + (1 - Z)\ P_{2002}\} =$

$Z\ X_{2004} + (1 - Z)\ Z\ X_{2003} + (1 - Z)^2\ Z\ X_{2002} + (1 - Z)^3\ P_{2002} =$

$Z\ X_{2004} + (1 - Z)\ Z\ X_{2003} + (1 - Z)^2\ Z\ X_{2002} + (1 - Z)^3\ \{Z\ X_{2001} + (1 - Z)\ P_{2001}\} =$

$Z\ X_{2004} + (1 - Z)\ Z\ X_{2003} + (1 - Z)^2\ Z\ X_{2002} + (1 - Z)^3\ Z\ X_{2001} + (1 - Z)^4\ P_{2001}$.

The weight given to X_{2001} is: $(1 - Z)^3\ Z$.

When Z = 10%, $(1 - Z)^3\ Z = (1 - 0.1)^3\ (0.1) = 7.29\%$.

The difference in the weight given to accident year 2001 losses in accident year 2002's estimate and the weight given to accident year 2001 losses in accident year 2005's estimate is:
10% - 7.29% = **2.71%**.

b. If there is a significant shift in risk parameters, then older years of data become much less predictive. Therefore, less weight is given to 2001 losses in the estimate of 2005 than when there was less shifting in risk parameters. This will make the difference in part (a) **increase**.

Comment: See page 255 of Mahler.

1.55. 1) χ^2 (Chi-Square Method).
The test statistic is: S(Actual – Expected)^2 / Expected.
Null Hypothesis: Expected number of claims is the same for each year.
Calculate the test statistic which sums the relative errors (squared)
Compare the test statistic to the critical value (from χ^2 distribution) with n-1 degrees of freedom.
If test statistic > critical value, then reject null and accept alternative, that risk parameters shift over time.
2) Correlation Test
Group data by pair for all possible combinations of time lag.
Calculate the correlation for each possible pair.
If the correlation decreases as the time lag increases, then there is a shifting of risk parameters over time.
<u>Comment</u>: Here is the result of the Chi-Square Test.
You would want the observed and assumed columns to add to the same amount,
thus the expected number of claims should be 501 rather than 500 as shown in the question.
(Using 500 would result in a statistic of 33.80.)

Year	Observed Number	Assumed Number	Chi Square
1997	475	501	1.35
1998	420	501	13.10
1999	460	501	3.36
2000	500	501	0.00
2001	490	501	0.24
2002	525	501	1.15
2003	515	501	0.39
2004	510	501	0.16
2005	540	501	3.04
2006	575	501	10.93
Sum	5,010	5,010	33.71

There are 10 years, and 10 - 1 = 9 degrees of freedom.
For 9 degrees of freedom, the critical value for 1/2% is 23.589.
(Value taken from the Chi-Square Table attached to a preliminary exam.)
Since 33.71 > 23.589, we reject the null hypothesis at 1/2%.
One could group data by interval of a few years (Mahler uses groups of 5 years over a period of 60 years.) He applies the test separately to each of the 16 teams.
The Chi-Square Test is shown by Mahler in his Table 4.
In item 2 of the solution, one would be calculating autocorrelations as per Time Series. See <u>Introductory Times Series with R</u>, by Cowpertwait & Metcalfe, not on the syllabus of this exam.
While that is a similar idea to what is done in the syllabus reading, it is <u>not</u> quite the same.
In the paper, one looks at the correlations of the vector of the losing percentages (each length 8) for 1901 and 1902. Then for 1902 and 1903. Then for 1903 and 1904. etc.
Then we average these results. This is the listed correlation for separation of 1 year.

Here the sample correlation is:

$$r = \frac{\text{estimated covariance of X and Y}}{\sqrt{(\text{estimated standard deviation of X})\ (\text{estimated standard deviation of Y})}}$$

$$= \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}.$$

For data separated by one year, the two vectors are:
X = (475, 420, 460, 500, 490, 525, 515, 510, 540). $\bar{X}$ = 492.778.
Y= (420, 460, 500, 490, 525, 515, 510, 540, 575). $\bar{Y}$ = 503.889.
(Which you call X and which you call Y is irrelevant.)
X - $\bar{X}$ = (-17.778, -72.778, -32.778, 7.222, -2.778, 32.222, 22.222, 17.222, 47.222).
Y - $\bar{Y}$ = (-83.889, -43.889, -3.889, -13.889, 21.111, 11.111, 6.111, 36.111, 71.111).

$\sum (X_i - \bar{X})(Y_i - \bar{Y})$ = 9127.8.

$\sum (X_i - \bar{X})^2$ = 10,805.6.

$\sum (Y_i - \bar{Y})^2$ = 16,138.9.

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} = \frac{9127.8}{\sqrt{(10{,}805.6)(16{,}138.9)}} = 0.691.$$

Alternately, one can fit a linear regression between X and Y using the stat functions on a calculator.
The output r is the desired correlation.
Similar to Mahler's Table 5, the autocorrelations for the data in this question are:

Separation	Correlation
1	0.691
2	0.528
3	0.717
4	0.470
5	0.654
6	0.770
7	-0.220

One would need more years of data, in order to draw a reliable conclusion from the correlation test. The paper has 60 years of data rather the 10 years here.

1.56. a) The expected number of claims in a year are: 1.5 times exposures.
The observed number of claims in a year are: (observed frequency)(exposures).

Year	Observed Number	Exposures	Expected Number	((Observed - Expected)^2)/Expected
2011	177	118	177.00	0.000
2010	224.4	132	198.00	3.520
2009	157.3	121	181.50	3.227
2008	174.4	109	163.50	0.727
2007	126.1	97	145.50	2.587
Sum	859.2	577	865.50	**10.060**

The Chi-square statistic is 10.060. 10.060 > 9.49, so we reject the null hypothesis.

⇒ The different years are not all drawn from the same Poisson Distribution.

⇒ The parameters are shifting over time.

b) Compute the correlations between different pairs of years of data for individuals.
Then average the correlations for years separated by a given number of years.
If the correlations decline as the separation increases, this indicates that parameters are shifting over time; the quicker the decline the more quickly parameters are shifting.
Comment: We have 5 - 1 = 4 degrees of freedom; the 5% critical value is 9.49.
See Tables 4 and 5 in Mahler.

1.57. (a) H_0: The expected frequency is 1.2% for each year.
H_1: Not H_0.
For 2011 the observed number is: (11,000)(0.010) = 110,
and the expected number is: (11,000)(0.012) = 132.
Contribution is: (Observed - Expected)^2/ Expected = $(110 - 132)^2$ / 132 = 3.6667

Year	Exposures	Frequency	Observed	Expected	Chi-Square Contribution
2010	9,500	0.011	104.5	114	0.79167
2011	11,000	0.010	110	132	3.6667
2012	13,000	0.013	169	156	1.0833
2013	10,500	0.012	126	126	0
2014	12,000	0.010	120	144	4
					9.54

Since the Chi-Square statistic is 9.54 > 9.49, at the corresponding significance level we reject the null hypothesis. This is evidence that (expected) claim frequency is shifting over time.
(b) For a given risk, compute the correlations between pairs of different years of data.
Average the correlations for all pairs with the same number of years between them.
If these average correlations decline quickly towards zero as the distance between pairs of years increases, then parameters are shifting at a significant rate.
Comment: 9.49 is the 5% critical value for a Chi-Square Distribution with 4 degrees of freedom.

1.58. a) Try different values of Z. The smallest squared error is for Z = **0.6**.
For example, for Z = 0.6, the estimate for year 2015 for Risk 1 is:
(0.2)(63.1%)+ (0.2)(59.0%) + (0.2)(63.5%) + (0.4)(57.0%) = 59.92%.
Squared error is: $(74.3\% - 59.92\%)^2 = 0.02068$.

0.6	Observed			Predictions			Squared Error		
Year	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3
2012	63.1%	72.5%	52.3%						
2013	59.0%	52.6%	48.9%						
2014	63.5%	69.7%	53.9%						
2015	74.3%	73.8%	50.1%	0.5992	0.6176	0.5382	0.02068	0.01450	0.00138
2016	45.9%	61.7%	50.9%	0.6216	0.6202	0.5338	0.02644	0.00001	0.00062
2017	42.3%	57.8%	46.6%	0.5954	0.6384	0.5378	0.02972	0.00365	0.00516
2018	58.9%	67.2%	48.6%	0.5530	0.6146	0.5232	0.00130	0.00329	0.00138
2019	60.2%	56.5%	50.7%	0.5222	0.6014	0.5202	0.00637	0.00132	0.00017
2020	52.8%	58.3%	46.1%	0.5508	0.5910	0.5198	0.00052	0.00006	0.00346
							MSE	**0.00667**	

b) The absolute error is: $\left| \frac{\text{predicted}}{\text{observed}} - 1 \right|$.

For example, for Risk 1 the prediction for 2015 is 59.92%. |59.92%/74.3% - 1| = 19.4%.

0.6	Observed			Predictions			Absolute Error		
Year	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3
2012	63.1%	72.5%	52.3%						
2013	59.0%	52.6%	48.9%						
2014	63.5%	69.7%	53.9%						
2015	74.3%	73.8%	50.1%	0.5992	0.6176	0.5382	19.4%	16.3%	7.4%
2016	45.9%	61.7%	50.9%	0.6216	0.6202	0.5338	35.4%	0.5%	4.9%
2017	42.3%	57.8%	46.6%	0.5954	0.6384	0.5378	40.8%	10.4%	15.4%
2018	58.9%	67.2%	48.6%	0.5530	0.6146	0.5232	6.1%	8.5%	7.7%
2019	60.2%	56.5%	50.7%	0.5222	0.6014	0.5202	13.3%	6.4%	2.6%
2020	52.8%	58.3%	46.1%	0.5508	0.5910	0.5198	4.3%	1.4%	12.8%

For Z = 60%, out of 18 cases, there are 13 errors of size more than 5%. 13/18 = 72.2%.

Trying other values of Z:

0.9	Observed			Predictions			Absolute Error		
Year	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3	Risk 1	Risk2	Risk 3
2012	63.1%	72.5%	52.3%						
2013	59.0%	52.6%	48.9%						
2014	63.5%	69.7%	53.9%						
2015	74.3%	73.8%	50.1%	0.6138	0.6414	0.5223	17.4%	13.1%	4.3%
2016	45.9%	61.7%	50.9%	0.6474	0.6453	0.5157	41.0%	4.6%	1.3%
2017	42.3%	57.8%	46.6%	0.6081	0.6726	0.5217	43.8%	16.4%	12.0%
2018	58.9%	67.2%	48.6%	0.5445	0.6369	0.4998	7.6%	5.2%	2.8%
2019	60.2%	56.5%	50.7%	0.4983	0.6171	0.4953	17.2%	9.2%	2.3%
2020	52.8%	58.3%	46.1%	0.5412	0.6015	0.4947	2.5%	3.2%	7.3%

For example, for Z = 90%, there are only 11 absolute errors greater than 0.05.
Thus by this criterion, Z = 90% is better than Z = 60%.

c) Many acceptable answers:

- I would recommend using the higher values of Z from the Small Chance of Large Errors Criterion, thereby putting more credibility on the 3 previous years, as opposed to the grand mean. This may indicate that years that are close together may be strongly correlated.
- I recommend the MSE criterion for selecting credibility, since changing the choice for P in the Small Chance of Large Error Criterion may change the best value of Z.
- I recommend the MSE criterion for selecting credibility, since it gives more stable estimates, as it puts more weight on the grand mean than the Small Chance of Large Errors Criterion.
- I recommend the MSE criterion for selecting credibility, since the Small Chance of Large Errors Criterion does not sharply distinguish between the different values of credibility. There is a broad range of credibilities all of which do reasonably well.

Predictions of the 2021 loss ratios using Z = 60%:
Risk 1: (60%)(58.9% + 60.2% + 52.8%)/3 + (40%)(57%) = **57.18%**.
Risk 2: (60%)(67.2% + 56.5% + 58.3%))/3 + (40%)(57%) = **59.20%**.
Risk 3: (60%)(48.6% + 50.7% + 46.1%)/3 + (40%)(57%) = **51.88%**.
Predictions of the 2021 loss ratios using instead Z = 90%:
Risk 1: (90%)(58.9% + 60.2% + 52.8%)/3 + (10%)(57%) = **57.27%**.
Risk 2: (90%)(67.2% + 56.5% + 58.3%))/3 + (10%)(57%) = **60.30%**.
Risk 3: (90%)(48.6% + 50.7% + 46.1%)/3 + (10%)(57%) = **49.32%**.

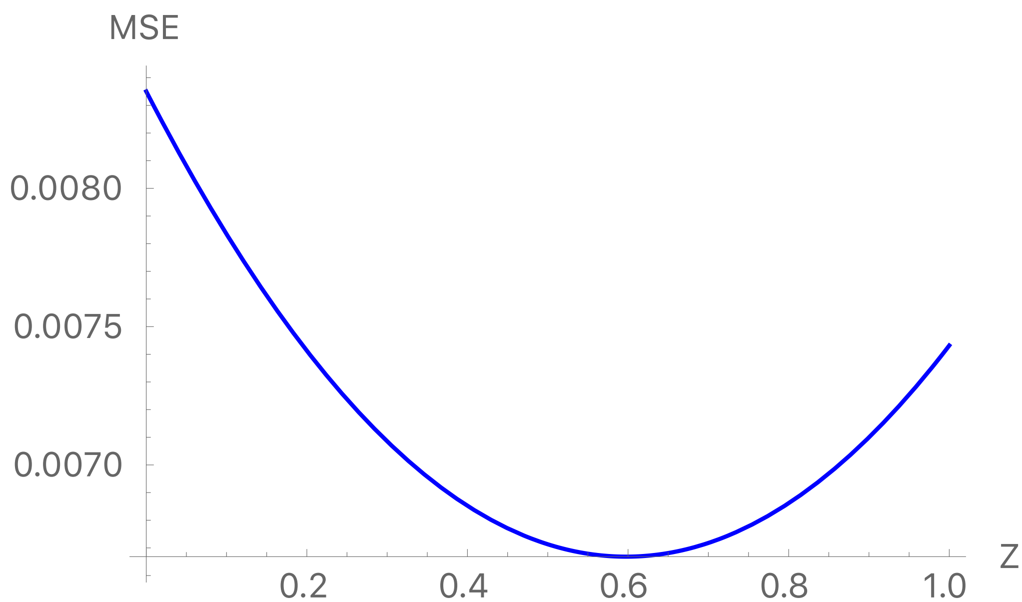
d) Different acceptable answers:

- We could test for shifting parameters over time by using either the chi-squared test or by checking the correlations for pairs of years with a given difference in time. If parameters are shifting quickly over time, then the smaller credibility indicated by the MSE criterion may be preferable; if not then the larger credibility indicated by the Small Chance of Large Errors may be preferable.
- Check the sensitivity of the best Z to different values of P in the Small Chance of Large Errors Criterion.

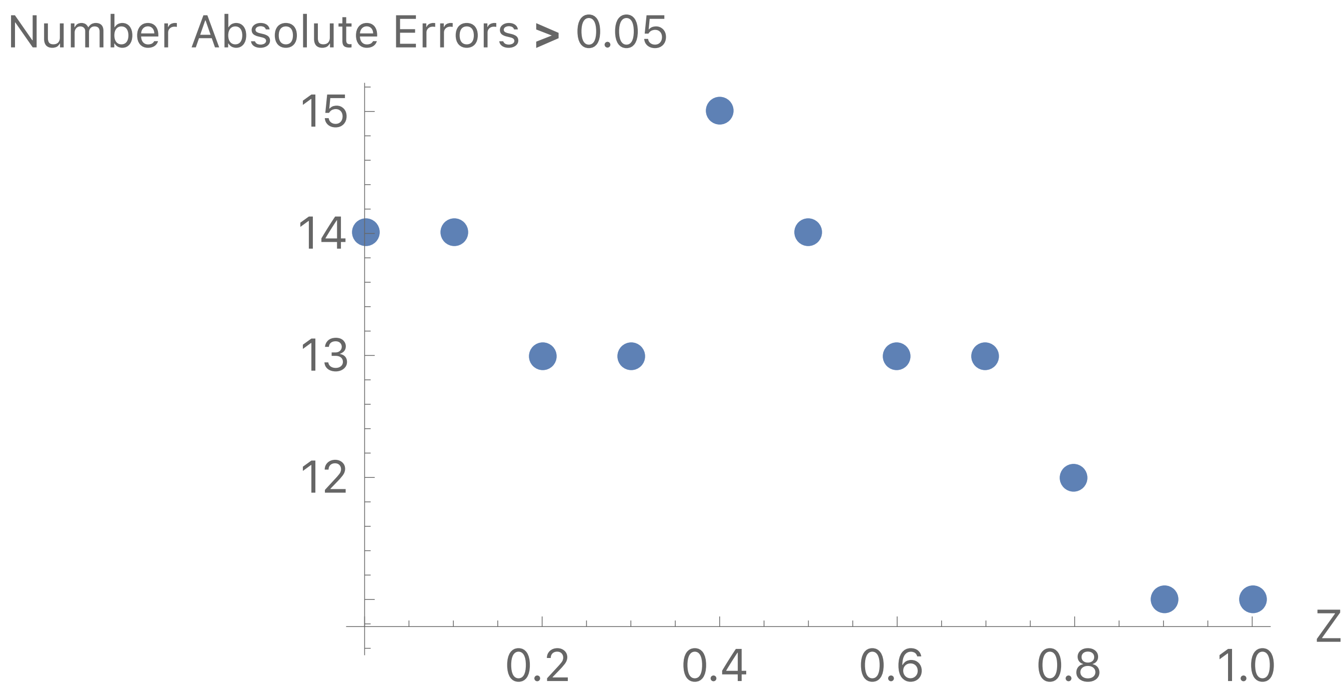
<u>Comment</u>: If the three risks are of different sizes, then the optimal credibility would differ between them.
Rather than working with historical loss ratios, one should at least adjust for any past rate changes.
MSE ⇔ Buhlmann Credibility

Small Chance of Large Errors ⇔ Classical Credibility

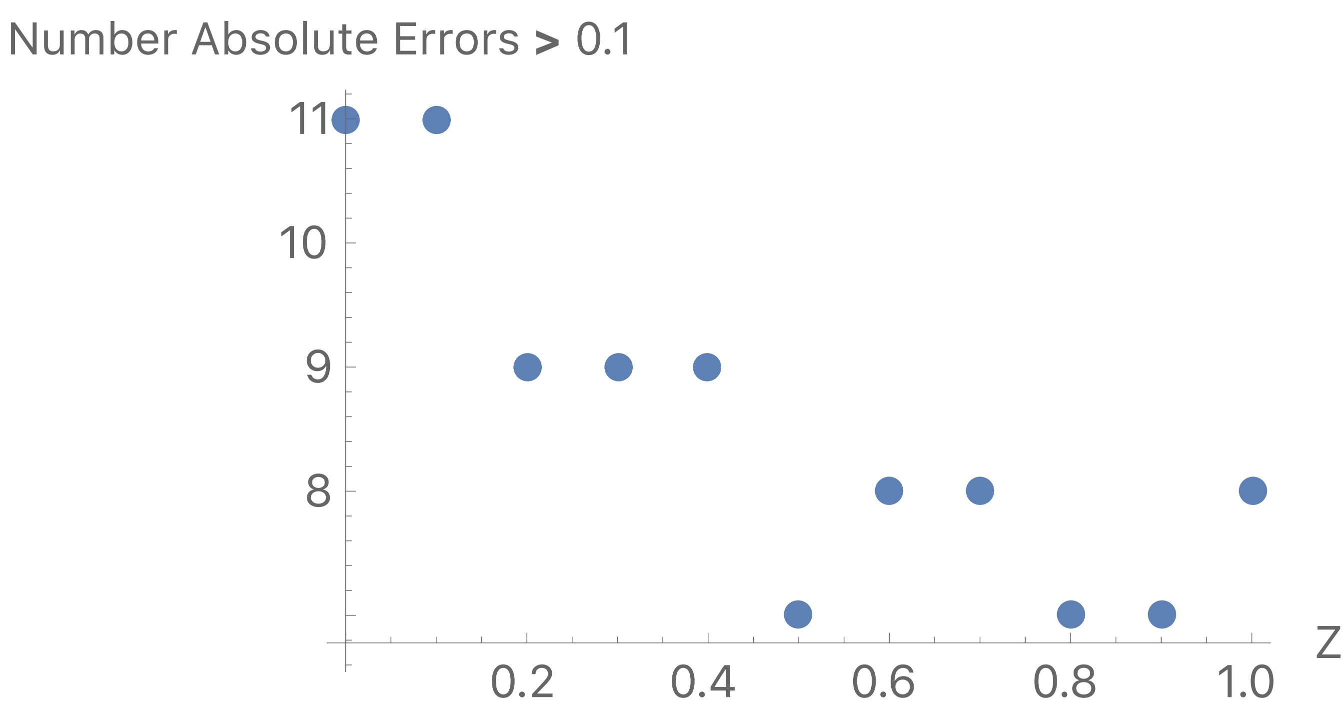
The paper does not discuss how to decide which criterion is preferable.
One was provided a spreadsheet for this Computer Based Test, which is essential.
The mean squared error as a function of the credibility Z:



The number of absolute errors greater than 0.05, as a function of the credibility Z:



Instead, the number of absolute errors greater than 0.10, as a function of the credibility Z:



Section 2, Bailey and Simon, Merit Rating[1 2]

In their classic paper, Bailey and Simon use Merit Rating data to determine the credibility to assign to the experience of a single private passenger car. The most important parts of this concise paper are Tables 2 and 3, and their conclusions.

A key concept is that when using credibility, Z is the discount compared to average given to an insured who is claims-free. This credibility varies by class and the number of years claims-free.

Merit Rating is a very simplified form of Experience Rating. As has been discussed previously, one way to analyze Experience Rating is to compare experience during a prior and subsequent period in order to determine how a plan would have worked in the past.[3]

Bailey-Simon compare a prior three year period to a subsequent one year period for Private Passenger Automobile Insurance in Canada.[4] They compare the subsequent frequency for groups with different numbers of years claims-free.[5] They found that Merit Rating has useful predictive ability beyond that of class and territory.[6]

---

[1] "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," by Robert A. Bailey and Leroy J. Simon, PCAS XLVI, 1959, pp. 159-164.
Including discussion of paper: Hazam, W. J., PCAS XLVII, 1960, pp. 150-152.
CAS Domains/Tasks A2 to A4.

[2] This excellent paper has been on the exam syllabus since shortly after it was written.
Thus it has been on the syllabus for over half a century!

[3] However, their method differs from those discussed by Gillam, Venter, and Mahler in their syllabus readings.

[4] For the prior period they only record how many years a car has been claims-free prior to the "present".

[5] In their much less important Table 4, they look at loss ratios, which involve dollars of loss.

[6] Since the average annual frequency is low, three years of private passenger auto data for a single car contains a lot of noise and relatively little signal; the credibilities are much smaller than those for a large commercial insured.

Merit Rating Plans:[7]

The Canadian Merit Rating Plan in place when Bailey and Simon wrote their paper is relatively simple.[8]

Those who are claim-free for only one year get a discount of 10%, Group Y.
Those who are claim-free for only 2 years get a discount of 20%, Group X.
Those who are claim-free for 3 or more years get a discount of 35%, Group A.[9] [10]
These discounts are off the base rate for those who are not claims-free, Group B.

Group A ⇔ no claims in the 3 year experience period has a claim.

Group X ⇔ the most recent 2 years claims free,
while the earliest year in the 3 year experience period has a claim.
For example, Merit Rating a 1958 policy: 1956 and 1957 claim free, but 1955 has a claim.

Group Y ⇔ the most recent 1 year claims free,
while the second year in the 3 year experience period has a claim.
For example, Merit Rating a 1958 policy: 1957 claim free, but 1956 has a claim.

As stated at the first page of Bailey-Simon:
Earned premiums are converted to a common rate basis by use of the relationship in the rate structure that A: X: Y: B = 65: 80: 90: 100.
Bailey-Simon put premiums on the level that would have been charged for Merit Rating Class B, those who are not claims free. For example, if the actual premiums for Merit Rating Group A were 6.5 million, then on a Group B basis they would be: 6.5 / (1 - 35%) = 10 million.

Currently in many states in the U.S., many insurers apply a simple form of Experience Rating to private passenger automobile insurance, often called Safe Driver Insurance Plans (SDIP).[11] They are usually somewhat more complex than the Merit Rating Plan discussed in Bailey-Simon. The number of moving traffic violations and/or at-fault claims will be determined for each driver over some recent period such as 3 or 5 years.[12] Those drivers with worse records will pay more than average, while those drivers with better records will pay less than average for their insurance.

---

[7] For background. You should not be tested on the details of Merit Rating Plans or Safe Driver Insurance Plans.
[8] See “The Canadian Merit Rating Plan for Individual Automobile Risks,” by Herbert E. Wittick, PCAS 1958.
[9] See page 159 of Bailey-Simon.
[10] All operators of a vehicle must be claim-free in order to get the discount; we are only looking at liability claims.
[11] See pages G-6 to G-9 of the ISO Personal Automobile Manual, not on the syllabus of this exam.
[12] For example, in Massachusetts as of 2004, the Safe Driver Insurance Plan (SDIP) uses 5 years of data on minor and major traffic law violations, and minor and major at fault accidents. Thus while this plan is largely based on frequency, there is a small component that depends on severity. Driving under the influence results in a larger surcharge than speeding. There are a number of additional complicated details in how this specific plan works.

The actuarial theory behind such plans is similar to that for the CGL and Workers Compensation Experience Rating Plans. However, the details differ. There is much smaller volume of data generated by a single private passenger automobile. The Canadian Merit Rating plan only uses frequency not severity.[13] Also most SDIPs use moving violations, so that someone who has no losses may still get a surcharge.

Also, the Canadian Merit Rating Plan differs from a plan that just added up the number of claims over the last three years. For example, let us assume the experience period is 1955, 1956, and 1957, and we are rating a 1958 policy.

Car	Number of Claims by Year		
	1955	1956	1957
1	0	0	0
2	1	0	0
3	2	0	0
4	1	1	0
5	0	1	0
6	0	0	1
7	1	1	1

Car 1 is put in Group A and gets the 35% discount.
Cars 2 and 3 are both put in Group X and get a 20% discount.
Cars 4 and 5 are both put in Group Y and get a 10% discount.
Cars 6 and 7 are both put in Group B and get no discount.

Note that insureds with different numbers of claims over the last three years may be charged the same amounts by the Canadian Merit Rating Plan. Also insureds with the same numbers of claims over the last three years can be charged different amounts by the Canadian Merit Rating Plan.[14]

In order to be put in Group A, the insured and/or principal operator must have been licensed for at least three years. In order to be put in Group X, the insured and/or principal operator must have been licensed for at least two years.

[13] As the volume of data declines, so does the optimal accident limit in an Experience Rating Plan; as the accident limit gets very low, the plan approaches a frequency only plan.
[14] While a plan that used the number of claims over three years might have more predictive power, Bailey-Simon is not comparing plans, but just trying to determine the predictive power of the plan then used in Canada.

<u>Claims-Free Discount</u>:

As mentioned, those who are claims-free get a discount; the longer the claims-free period the larger the discount. There are two ways to look at the size of the discounts. First there is the discount from the base rate. In the case of the Canadian Merit Rating Plan, the discount is off of the rate charged Group B, which is higher than average.

For example, from Table 1 in Bailey-Simon, for Class 1, the average discount is:

$$\frac{(35\%)(159{,}108) + (20\%)(7910) + (10\%)(9862)}{194{,}106} = 30.01\%.$$

The average Class 1 premium at Group B rates (in other words prior to any discounts) is: 194,106,000 / 3,325,714 = \$58.36.[15] However, after the effect of the Merit Rating discounts, the average rate paid is only: (\$58.36)(1 - 0.3001) = \$40.85

The advertised discount for three years claims-free is 35%; however, the discount off of the average rate is: 1 - (1 - 35%)/(1 - 30.01%) = 7.1%.

The advertised discount for two years claims-free is 20%; however, the surcharge above the average rate is: (1 - 20%)/(1 - 30.01%) - 1 = 14.3%.[16]

The Group X insureds are claims-free for only two-years; they had a claim three year ago. Thus they are worse than the insureds who have been claims-free for at least 3 years, and Group X pays more than average.

In the context of credibility theory, actuaries are interested in the experience and discounts with respect to <u>average</u>. Bailey-Simon will compute claims-free discounts compared to average for those who have been claims-free at least one year (A +X+Y), claims-free at least two years (A + X), and claims-free at least three years (A).

[15] Remember this is for 1957 and 1958. Also these are Canadian dollars.

[16] Public acceptance of the plan is much better if one advertises discounts off of the base rate, rather than making it obvious that some insureds are paying more than average.

Table 1, Bailey-Simon:

Private Passenger Auto Liability data from Canada (excluding Saskatchewan) from Policy Years 1957 and 1958.[17] The data is divided into five classes.[18] Their analysis will be performed separately on each class.

Within each class are four Groups, based on how long they have been claims-free:

A	3 or more years claims-free
X	2 years claims free
Y	1 year claims free
B	0 years claims-free
A + X	2 or more years claims-free
A + X + Y	1 or more years claims-free

We have exposures (earned car years), premiums (earned premiums at present Group B rates), and claims (number of claims incurred).

Then the number of claims is divided by premiums in $1000, rather than exposures.
For example, for Group A in Class 1: 217,151 / 159,108 = 1.365.

Bailey and Simon "have chosen to calculate Relative Claim Frequency on the basis of premium rather than car years. This avoids the maldistribution created by having higher claim frequency territories produce more X, Y, and B risks and also produce higher territorial premiums." [19]

[17] For an individual car, assume we have a policy written during 1958.
Then the Merit rating class (A, X, Y, B) would have been based on 3 past years of data.
(This may be 1955, 1956, and 1957 without a gap in obtaining data for Merit Rating.)
I believe the Class (1, 2, 3, 4, or 5) is the one for the 1958 policy.
For example, some insured who were in Class 1 (Pleasure - no male operator under 25) during 1958 would have been in different classes during 1955, 1956, or 1957.
The data was not scrubbed to remove insureds who switched classes over the relevant period.

[18] Class 1 is Pleasure - no male operator under 25.
Class 2 is Pleasure - Non-principal male operator under 25.
Class 3 is Business use.
Class 4 is Unmarried owner or principal operator under 25.
Class 5 is Married owner or principal operator under 25.

[19] Average premiums by territory within a class will vary due to differences in frequency, differences in severity, as well as to some extent incorrect territory relativities.

The use of premium based frequencies avoids double counting. If instead one used caryears as the denominator of frequency, the credibility calculation would account for both "within territory differences" and "between territory differences". However, territory relativities already account for the between territory differences.[20] [21] Compounding territory relativities with credibility would double count the between territory differences, and therefore the credibility would be overstated. Therefore, claims free drivers would be undercharged while other drivers would be overcharged.[22]

In order to remove this "double counting", we use premium as exposure. An assumption for this is that the premium differences should reflect the true pure premium differences between the territories.[23] [24]

The premiums are put on the basis of Group B, in other words prior to any discounts for Merit Rating.[25] [26] We are removing the effects of any current discounts due to Merit Rating in order to estimate the indicated discounts, rather than estimating a change in the current discounts. As discussed previously, Bailey-Simon will be estimating discounts compared to average.

Then Bailey-Simon divide the premium based frequency for a group by that for the whole class, in order to to get the relative claim frequency.
For example, for Group A in Class 1: 1.365/1.484 = 0.920

---

[20] Hazam points out in his discussion, we need to assume that the territory relativities are correct and reflect differences in frequency (per caryear) between the territories.

[21] If for some reason territorial rating is not used in spite of differences between territories in frequency, then there would be no double counting resulting from using car years as the denominator of frequency. See 8, 11/15, Q.1. In the absence of territorial rating, the appropriate Merit Rating credibilities are larger than they otherwise would be. In general, the less accurate the class/territory plan and relativities, the more predictive work there is for experience rating to do, and thus the larger the appropriate credibility for experience rating.

[22] Subsequently I have a detailed example illustrating this.

[23] See the discussion by Hazam.

[24] In the late 1950s, they used a very simple class system, as has been discussed. Since Bailey and Simon analyze data by class, this is accounted for. The only other rating variable is territory, which they approximately account for by using premium based frequency. It should be noted that within each class/territory cell there is still heterogeneity, which is why Merit Rating is useful.

[25] One could use another group as the base, and the relative claim frequencies would be the same.

[26] "Earned premiums are converted to a common rate basis by use of the relationship in the rate structure that A:X:Y:B = 65:80:90:100." In other words as mentioned previously, those in Merit Rating Class A currently get a 35% discount with respect to Merit Rating Class B, those in Merit Rating Class X currently get a 20% discount with respect to Merit Rating Class B, and those in Merit Rating Class Y currently get a 10% discount with respect to Merit Rating Class B. Thus for example, in order to be put on a Merit Rating Class B level, premium from an insured in Merit Rating Class A would be divided by 0.65.

Here is the calculation for the Class 2 data shown in their Table 1:

Class 2 - Pleasure - Non-principal male operator under 25					
Group	Years Claims-Free	Group B Premium	Number of Claims	Freq.	Rel. Freq.
A	3 or more	11,840,000	14,506	1.225	0.932
A+X	2 or more	12,552,000	15,507	1.235	0.940
A+X+Y	1 or more	13,496,000	16,937	1.255	0.955
Total		15,488,000	20,358	1.314	1.000

We need to combine Groups A and X in order to get those who are claims free for 2 years or more:
Claims-free at least 2 years = (3 or more years claims-free) + (2 years claims-free).
A + X + Y is those who are claims free for 1 year or more.[27]

<u>Table 2, Bailey-Simon</u>:

In their very important Table 2, for each class separately, the credibilities for one, two, and three years of data are calculated from the indicated claims-free discount compared to average.
For example, for Class 2, the overall frequency on a premium basis in Table 1 is:
20,358 / 15,488 = 1.314.
The frequency on a premium basis for Group A (3 years claims-free) is: 14,506 / 11,840 = 1.225.
Thus the indicated experience modification for Group A is: 1.225/1.314 = 0.932.
This is the relative claim frequency also shown in Table 1.

Then the claims free discount is: 1 - 0.932 = 6.8%.
This is the estimated credibility for three years of data shown in Table 2 for Class 2.
In general, for a given class and numbers of years or more claims-free:

$$1 - Z = M = \frac{\text{Premium Based Claim Frequency for Those Claims - Free N or More Years}}{\text{Overall Premium Based Claim Frequency for the Class}}.$$

Calculating in this manner the credibilities for one, two or three years is the most commonly asked exam question on this paper. For Class 2:[28]
The one-year credibility is: 1 - 1.255/1.314 = 1 - 0.955 = 4.5%.
The two-year credibility is: 1 - 1.235/1.314 = 1 - 0.940 = 6.0%.
The three-year credibility is: 1 - 1.225/1.314 = 1 - 0.932 = 6.8%.

---

[27] If they use the letters for the groups, we expect them to tell you their meaning in the question.
A = claims-free 3 or more years. X = claims-free 2 years. Y = claims-free one year. B = not claims-free.
[28] These match what is shown for Class 2 in Table 2 in Bailey-Simon.
These do <u>not</u> match the then current discounts in the Canadian plan, which as discussed were with respect to the base rate rather than with respect to average and were the same regardless of class.
The then current discounts in the Canadian plan preceded the study by Bailey and Simon.

Ratio of Credibility to Frequency:

In addition, in Table 2, for each class Bailey-Simon takes the ratio of the three-year credibility to the frequency.[29] For example for Class 2, the overall exposure based frequency is:
20.358 / 168,998 = 0.120. Then the ratio of the 3-year credibility to frequency is:
0.068 / 0.120 = 0.567.

The credibilities depend on the Expected Value of the Process Variance (EPV) and the Variance of the Hypothetical Means (VHM).[30] If each insured is Poisson, then the EPV is equal to the average frequency for the class. In any case, the EPV should be roughly proportional to the mean frequency.

If the Buhlmann Credibility formula holds, then the three-year credibility is
Z = 3 / (3 + K), with K = EPV / VHM.[31] [32]
For K big compared to 3, as it is in the situations in Bailey Simon: Z $\cong$ 3/K = (3) (VHM / EPV).

Let m be the overall mean frequency, which is also the mean of the hypothetical mean frequencies.
Assume the EPV is (approximately) proportional to the overall mean frequency: EPV = c μ.
Then the ratio of the credibility to the mean frequency is approximately:
(3)(VHM / EPV) / μ = (3/c) VHM / μ^2.

Thus the ratio of the credibility to the mean frequency is proportional to the square of the coefficient of variation of the hypothetical means: VHM / μ^2. Thus the smaller this ratio, the smaller the CV of the hypothetical means, and the less variation between the insureds within a class.

Thus the smaller this ratio of credibility to frequency, the more homogeneous the class.

The more homogeneous the class, the less the credibility assigned to the experience of an individual, as experience of an individual that differs from the average would more likely be random than a real difference. To take the extreme case, if all the risks in a class were known to be exactly alike, we would know that any variations in the experience of an individual from average for its class are random, and therefore should be given no credibility.

---

[29] I would prefer using the one-year credibility, since as will be discussed, the one-year credibility is less affected by shifting risk parameters over time than is the three-year credibility.
[30] Subsequently, I have a review of Buhlmann Credibility and some related material.
[31] As will be discussed subsequently, the Buhlmann Credibility formula does not hold for this data.
[32] A car that has been claims-free for at least three years may have many years of data. However if all we know is that it has been claim free for at least three years, then we are looking at the most recent three years of data.
N = 3.
Similarly, if we look at all the cars that has been claim free for at least the last two years (combining those that have been claims-free for exactly two years with those who have been claims-free for at leas 3 years), then N = 2.

All other things being equal, more claims means higher credibility. All other things being equal, one car for one year when the mean frequency for the class is 10% has more credibility than when the mean frequency is 5%; approximately, twice as much credibility in the first case than the second, all else being equal. Thus we divide by the mean frequency to adjust for its effect. This leaves the effect of homogeneity, which we are trying to compare between classes.

As shown in Table 2 of Bailey-Simon:

Class	Three-Year Credibility	Claim frequency per car-year	Ratio
1	8.0%	8.7%	0.920
2	6.8%	12.0%	0.567
3	8.0%	14.2%	0.563
4	9.9%	16.2%	0.611
5	5.9%	11.0%	0.536

With the highest ratio of credibility to mean frequency, Class 1 is the least homogeneous, in other words the most heterogeneous.[33] With the lowest ratio of credibility to mean frequency, Class 5 is the most homogeneous, although Classes 2 and 3 are nearly as homogeneous.[34]

"Classes 2, 3, 4 and 5 are more narrowly defined than Class 1, and the fact that the ratios in the last column of Table 2 for these classes are less than the ratio for Class 1 confirms the expectation that there is less variation of individual hazards in those classes. This also illustrates that **credibility for experience rating depends not only on the volume of data in the experience period but also on the amount of variation of individual hazards within the class**."[35]

[33] Class 1 is Pleasure - no male operator under 25.
[34] Class 2 is Pleasure - Non-principal male operator under 25, Class 3 is Business use,
Class 4 is Unmarried owner or principal operator under 25,
and Class 5 is Married owner or principal operator under 25.
[35] The homogeneity of classes is also discussed in the ASOP 12: Risk Classification.

<u>Table 3, Bailey-Simon</u>:

In their important Table 3, for each class separately, the two-year and three-year credibilities are compared to the one-year credibility.

As shown in Table 2 of Bailey-Simon:

Class	One-Year Credibility	Two-Year Credibility	Three-Year Credibility
1	4.6%	6.8%	8.0%
2	4.5%	6.0%	6.8%
3	5.1%	6.8%	8.0%
4	7.1%	8.5%	9.9%
5	3.8%	5.0%	5.9%

For Class 1, the ratio of the two-year to one-year credibility is: 6.8% / 4.6% = 1.48.

Then as shown in Table 3 of Bailey-Simon:

Class	Relative Credibility		
	One-Year	Two-Year	Three-Year
1	1.00	1.48	1.74
2	1.00	1.33	1.51
3	1.00	1.33	1.57
4	1.00	1.20	1.39
5	1.00	1.32	1.55

These credibilities go up much less than linearly as the number of years of data increase.

Bailey-Simon gives the following possible reasons:[36]
1. Risks entering and leaving the class.
2. An individual insured's chance for an accident changes from time to time within a year and from one year to the next.
3. The risk distribution of individual insureds has a marked skewness reflecting varying degrees of accident proneness.
4. The Buhlmann Credibility formula, Z = N / (N+K), increases somewhat less than linearly with N.

[36] To be discussed in more detail subsequently. The fourth reason is from the discussion by Hazam.

Indicated Merit Rating Factors:[37]

As mentioned, in the Canadian Merit Rating Plan, the Merit Rating Factors were with respect to Group B, those who are not claims-free. The base rate for each class and territory was charged to Group B, while those in the other groups were given a discount.

The Merit Rating Factors were 0.65 for Group A, 0.80 for Group X, and 0.90 for Group Y.

By comparing the relative premium based frequencies one can determine indicated Merit Rating Factors. For example, for the Class 1 data shown in their Table 1:

Class 1 - Pleasure - no male operator under 25					
Group	Years Claims-Free	Group B Premium	Number of Claims	Freq. per $1000 Prem.	Rel. Freq. to Group B
A	3 or more	159,108,000	217,151	1.365	**0.623**
X	2	7,910,000	13,792	1.744	**0.796**
Y	1	9,862,000	19,346	1.962	**0.896**
B	0	17,226,000	37,730	2.190	1.000

217,151 / 159,108 = 1.365. 37,730 / 17,226 = 2.190.
Thus the indicated Merit Rating Factor for Group A is: 1.365 / 2.190 = 0.623.

These indicated Merit Rating Factors are close to the then current factors.

It should be noted that for simplicity the same Merit Rating Factors are used regardless of class. Thus one would want to analyze the data for each class, as well as all of the data combined. Then one would select one set of Merit Rating Factors to apply to all classes.

Exercise: For Class 3, Business Use, the claim frequencies per $1000 of premium are:
Group A: 1.237. Group X: 1.511. Group Y: 1.555. Group B: 1.832.
Determine the indicated Merit Rating Factors.
[Solution: For Group A: 1.237/1.832 = 0.675.
For Group X: 1.511/1.832 = 0.825.
For Group Y: 1.555/1.832 = 0.848.]

[37] See 8, 11/18, Q. 3b.

Table 4, Bailey-Simon:

For the class with the most data, Class 1, Bailey-Simon also works with loss ratios rather than frequencies.[38] The denominator is the same premium at Group B rates. The numerator is incurred losses rather than number of claims.

The overall loss ratio is 43.6%.
The loss ratio for Group A (3 or more years claims-free) is 39.7%.
The relative loss ratio is: 39.7% / 43.6% = 0.911.
Thus the three-year credibility is: 1 - 0.911 = 5.5%.

The relative loss ratio for those who are claims free at least 2 years (A + X) is 0.924.
Thus the two-year credibility is: 1 - 0.924 = 7.6%.

The credibilities are:

1 Year	2 Year	3 Year
5.5%	7.6%	8.9%

These are similar to those for Class 1 based on frequency as shown in Table 2, but slightly bigger. This seems to indicate that those who are claims-free also have a lower expected future severity compared to those who are not claims-free.

The relative credibilities are:

1 Year	2 Year	3 Year
1.00	1.38	1.62

This is similar pattern as seen for the credibilities based on frequency. For Class 1, here the credibilities are slightly further from linear than were those based on frequency.

[38] The aggregate losses for an insured are affected by severity as well as frequency, and thus loss ratios are subject to more random fluctuation than are frequencies. Thus an analysis of loss ratios requires more data than a similar analysis of frequencies.

<u>An Alternate Way to Estimate the One-Year Credibility</u>:[39]

Bailey-Simon also backs out a one-year credibility by comparing the observed frequency in the prior year of those who were <u>not</u> claims-free (Merit Rating Group B) to their observed frequency in the subsequent year.

For example, as shown in Table 1, for Class 1 the observed overall frequency per <u>exposure</u> is: 288,019 / 3,325,714 = 0.0866. Assume that the overall frequency is Poisson with mean λ.
Then the mean number of claims for those who were not claim free (Group B) is:[40]
$\lambda / (1 - e^{-\lambda}) = 0.0866 / (1 - e^{-0.0866}) = 1.044$.

Thus Group B has a frequency relative to average within Class 1 of: $1 / (1 - e^{-\lambda}) = 1/ (1 - e^{-0.0866}) = 12.05$. However, based on its relative premium based frequency, in Table 1 we have an estimated modification for Group B in Class 1 of: 2.190 / 1.484 = 1.476.
Thus, 1.476 = (12.05) Z + (1)(1 - Z). $\Rightarrow$ Z = (1.476 - 1) / (12.05 - 1) = 4.3%.[41] This is similar to the 4.6% one-year credibility for Class 1 shown in Table 2 and based on the claims-free discount.

Let λ = the mean claim frequency (per exposure) for the class.
M = relative premium based frequency for risks with one or more claims in the past year.
Then, $M = Z / (1 - e^{-\lambda}) + (1 - Z)(1). \Rightarrow Z = \frac{M - 1}{1 / (1 - e^{-\lambda}) - 1} = (M - 1)(e^{\lambda} - 1)$.

Here are the similar results for all of the classes:

Class	Mean Freq. Overall	Mean Freq. For Group B	Prior Rel. For Group B	Subseq. Rel. For Group B	One Year Credibility	Table 2 1 year Z
1	8.66%	1.044	12.05	1.476	4.3%	4.6%
2	12.05%	1.061	8.81	1.307	3.9%	4.5%
3	14.24%	1.073	7.53	1.362	5.5%	5.1%
4	16.21%	1.083	6.68	1.247	4.3%	7.1%
5	10.96%	1.056	9.63	1.302	3.5%	3.8%

There is a reasonable match between the credibilities from looking at Group B and those from the claims-free discount, with the exception of Class 4. As will be discussed subsequently, there is an inherent problem with using the claim free discount to estimate credibilities for Class 4, which includes many drivers who have less than three years of driving experience. In any case, these two different techniques are expected to produce similar but somewhat different results, neither of which is equal to the least squares credibility.

---

[39] See page 160 and Appendix II in Bailey-Simon. See for example, 9, 11/05, Q.3, and 9, 11/09, Q.4a.
[40] See Appendix II in Bailey-Simon, to be discussed subsequently.
[41] Matching the result shown at the bottom of page 160 in Bailey-Simon.

Standard Method	Alternative Method
actual past claim frequency	theoretical past claim frequency
nonparametric	Poisson Distribution
claim frequency to premiums	claim frequency to exposures
claims-free risks	not claims-free risks
1, 2, and 3 year credibilities	one year credibility

Conclusions of Bailey-Simon:[42]

(1) The experience for one car for one year has significant and measurable credibility for experience rating.

(2) In a highly refined private passenger rating classification system which reflects inherent hazard, there would not be much accuracy in an individual risk merit rating plan, but where a wide range of hazard is encompassed within a classification, credibility is much larger.

(3) If we are given one year's experience and add a second year we increase the credibility roughly two-fifths. Given two years' experience, a third year will increase the credibility by one-sixth of its two-year value.

Conclusion number 1 has two parts. Bailey-Simon have demonstrated a practical and simple way to measure this credibility. Also they show in Table 2 that the credibility is big enough to make Merit Rating of Private Passenger Automobile Insurance practical and worthwhile from an actuarial point of view.[43]

Conclusion number 2 follows from general credibility theory applied to experience rating. The more homogeneous a class, the less credibility is given to the experience of an individual insured.

The key idea in conclusion number 3 is that based on their Table 3, the credibilities increase much less than linearly. The specific values are not anywhere near as important.

[42] They were writing more than half a century ago; things that may be obvious today were far from obvious then.
[43] As discussed by Hazam, including moving violations makes Merit Rating more worthwhile to use.

Buhlmann Credibility (Least Squares Credibility), Review:[44]

EPV = Expected Value of the Process Variance = $E_\theta[\,VAR[X \mid \theta]\,]$.

VHM = Variance of the Hypothetical Means = $VAR_\theta[\,E[X \mid \theta]\,]$.

Buhlmann Credibility Parameter = $\mathbf{K = \frac{EPV}{VHM}}$,

where the Expected Value of the Process Variance and the Variance of the Hypothetical Means are each calculated for a single observation of the risk process.
One calculates the EPV, VHM, and K prior to knowing the particular observation!

If one is estimating claim frequencies or pure premiums, then N is in exposures.
If one is estimating claim severities, then N is in number of claims.

For N observations, the Buhlmann Credibility Factor is: $\mathbf{Z = \frac{N}{N+K}}$.[45]

Estimate of the future = (Z) (Observation) + (1 - Z) (Prior Mean).

Assumptions:
- (1 - Z) is applied to the prior mean.
- The risk parameters and risk process do not shift over time.
- The expected value of the process variance (EPV) of the sum of N observations increases with N.
- The variance of the hypothetical means (VHM) of the sum of N observations increases with N^2.

For experience rating, we compare the individual relative to its class; the class has a relativity of one, and thus the estimated relativity = Z (observed relativity) + (1 - Z)(1).

Bayes Analysis, Review:[46]

The prior estimate is adjusted to reflect the new information.

Bayes' Theorem: $\mathbf{P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)}}$.

P(Risk Type | Observation) = $\mathbf{\frac{P(Observation \mid Risk\ Type)\, P(Risk\ Type)}{P(Observation)}}$.

[44] I do not expect you to be tested directly on any of this other than the use of the formula for Z.
[45] For the situations in Bailey-Simon, K is big compared to N, and thus Z should be approximately proportional to N.
[46] I do not expect you to be tested on any of this.

If $\pi(\theta)$ is the assumed prior distribution of the parameter θ,
then the posterior distribution of θ is proportional to: $\pi(\theta)$ P(Observation | θ).

The posterior distribution of θ is: $$\frac{\pi(\theta)\ \text{Prob[Observation} \mid \theta]}{\int \pi(\theta)\ \text{Prob[Observation} \mid \theta]\ d\theta}.$$

The Bayes estimate is: $$\frac{\int (\text{Mean given } \theta)\ \pi(\theta)\ \text{Prob[Obs.} \mid \theta]\ d\theta}{\int \pi(\theta)\ \text{Prob[Obs.} \mid \theta]\ d\theta}.$$

Bühlmann Credibility (Least Squares Credibility) is the weighted least squares line fit to the Bayesian estimate. In certain special mathematical situations, such as the Gamma-Poisson or the Beta-Binomial, the Bayesian analysis estimate is equal to that from Bühlmann Credibility (Least Squares Credibility). Due to the greater complexities of its probabilistic nature, Bayesian analysis is not used as commonly in practical applications in insurance as is Bühlmann credibility.[47]

Claims Free Discount Versus Least Squares (Buhlmann) Credibility:

Assume we are using credibility to estimate future frequency.
Then the estimated future frequency for an insured who had no claims is: $Z\ 0 + (1-Z)\mu = \mu - \mu Z$.
Thus as a percent, the estimated future frequency is Z less than average.
Thus Z is the claim free discount.

Bailey-Simon sets the credibility equal to the indicated claims free discount:
$$1 - Z = \frac{\text{observed frequency for those who were claims free}}{\text{overall frequency}}.$$ [48]

In general, the least squares credibility does not equal this indicated claims free discount. The least squares credibility is the linear estimator that best approximates the Bayes Estimates for all of the possible observations. In contrast, this indicated claims free discount only looks at the observed result for those with no claims. Usually, the claims free discount will be close to the least squares credibility.

One important special case is the Gamma-Poisson.[49] For the Gamma-Poisson the least squares credibility is equal to the Bayes Estimates; the Bayes Estimates are on a straight line. Thus, in this case, the claim free discount is equal to least squares credibility.[50]

On page 160, Bailey-Simon also backs out a one year credibility by comparing the observed frequency in the prior year of those who were not claim free to their observed frequency in the next year. Again, this will be very similar to but not identical to the least squares credibility.

---

[47] Bayes Analysis is harder to explain to nonactuaries.
[48] They do this separately for those who were claim free for at least a year, at least two years, and at least 3 years.
[49] The Gamma-Poisson is usually a pretty good model for private passenger auto frequencies.
[50] With finite data sets, the two ways to estimate the credibility will differ somewhat.

Review of the Mathematics Behind Experience Rating:

Assume that a insured has had no accidents over the last decade. This provides evidence that he is a safer than average insured; his expected claim frequency is lower than average for his class. Thus for automobile insurance one might give him a “safe driver discount” off of the otherwise applicable rate for his class.

This is an example of experience rating. Generally, experience rating consists of modifying the rate charged to an insured (driver, business, etc.) based on its past experience. While such plans can be somewhat complex in detail, in broad outline they all reward better than expected experience and penalize worse than expected experience. Depending on the particular circumstances more or less weight is put on the insured’s observed experience from the recent past.[51]

The new estimate of the insured’s frequency or pure premium is a weighted average of that for his classification and the observation. The amount of weight given to the observation is the credibility assigned to the individual insured’s data. In general, how much credibility to assign to an individual insured’s data should depend on:

1. What is being estimated. Pure Premiums are harder to estimate than frequencies.
Total Limits losses are harder to estimate than basic limits losses.
In a split plan, primary losses are easier to predict than excess losses.[52]

2. The volume of data. All other things being equal, the more data the more credibility is assigned to the observation.[53]

3. The Expected Value of the Process Variance. The more volatile the experience, the less credibility is assigned to it.

4. The variance of the hypothetical means within classes; the more homogeneous the classification the smaller this variance and the less credibility is assigned to the insured’s individual experience compared to that for the whole classification.

---

[51] The period of past experience used varies between the different Experience Rating Plans.
[52] For a thorough discussion of whether or not to use severity in addition to frequency, split versus non-split plans, and the choice of accident limits, see “An Analysis of Experience Rating,” by Glenn G. Meyers, PCAS 1985, and the discussion by Howard C. Mahler, PCAS 1987.
[53] For example, in Workers’ Compensation Insurance the data from a business with $10,000 in Expected Losses would be given much less credibility for Experience Rating than the data from a business with $1 million in Expected Losses.

The more homogeneous the classes, the less variation between the risks within the class, the less credibility assigned an individual's data and the more to the average for the class, when performing experience rating (individual risk rating.) The credibility is a relative measure of the value of the information contained in the observation of the individual versus the information in the class average. The more homogeneous the classes, the more value we place on the class average and the less we place in the individual's experience.

Thus low credibility is neither good nor bad. It merely reflects the relative values of two pieces of information. With a well designed class plan, the less we need to rely on the observations of the individual, compared to a poorly designed class plan. In auto insurance if we classified insureds based on their middle initials, we would expect to give the insureds individual experience a lot of credibility. A poor class plan leads one to rely more on individual experience.

Note that the role of the class in Experience Rating has changed from its role in Classification Ratemaking. In Experience Rating, the class experience receives the complement of credibility not given to the individual's experience. In the case of classification rating, the class experience gets the credibility while the complement of credibility is assigned to the experience of all classes combined. In Experience Rating, the insured is the smaller unit while the class is the larger unit. In Classification Ratemaking, the class is the smaller unit while the state is the larger unit. In both cases, the weight given to the classification's experience is larger the more homogeneous the class. Thus the more homogeneous the classes, the more credibility is given to the experience of each class for Classification Ratemaking. The more homogeneous the class, the less credibility is assigned to the individual's experience and therefore the more weight is given to the class experience for Experience Rating.

Simple models may help one to understand the mathematics behind experience rating.[54] The Gamma-Poisson frequency process is a good model for this purpose. Each insured's frequency is given by a Poisson Process. The mean frequencies of the insureds within a class are distributed via a Gamma Distribution. The variance of this Gamma Distribution quantifies the homogeneity of the class. The smaller the variance of this Gamma, the more homogeneous the class.

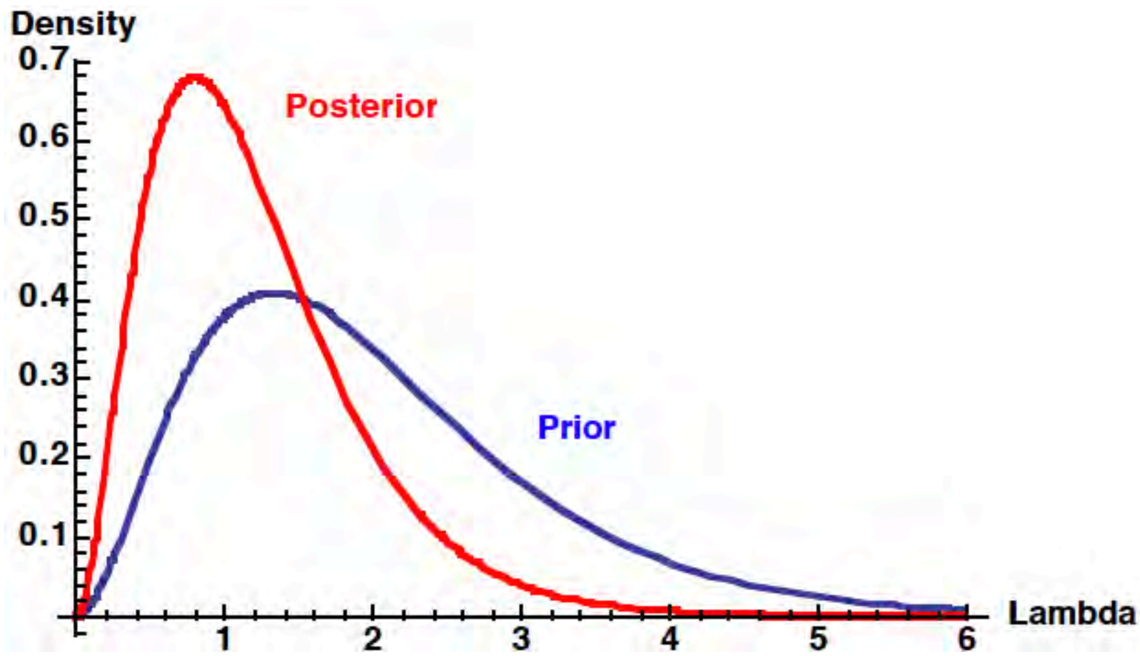
The observed experience of an insured can be used to improve the estimate of that insured's future claim frequency. We assume a priori that the average claim frequencies of the insureds in a class are distributed via a Gamma Distribution with $\alpha = 3$ and $\theta = 2/3$. The average frequency for the class is (3)(2/3) = 2.

If we observe no claims in a year, then the posterior distribution of that insured's (unknown) Poisson parameter is a Gamma distribution with $\alpha = 3$ and $\theta = 0.4$, with an average of: (3)(0.4) = 1.2.[55]
Thus the observation has lowered our estimate of this insured's future claim frequency.

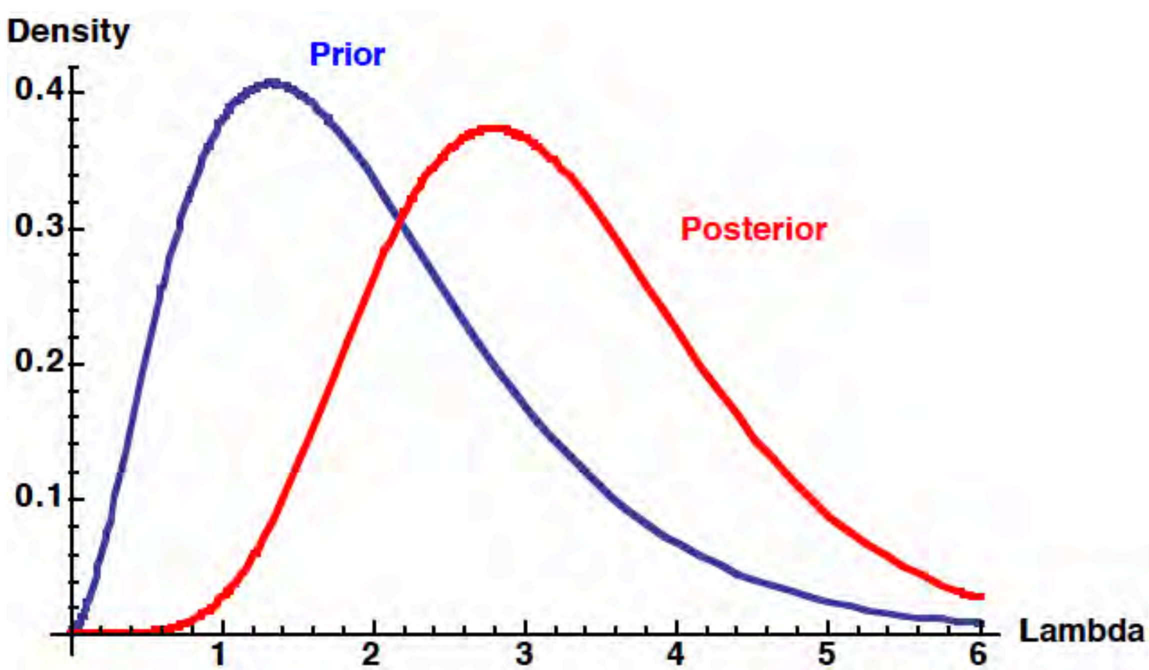
---

[54] See for example, "A Graphical Illustration of Experience Rating Credibilities," by Howard C. Mahler, PCAS 1998.
[55] The posterior alpha is 3 + 0 = 0. The posterior theta = 1/{1+1/(2/3)} = 1/ 2.5 = 0.4.

The prior Gamma with $\alpha = 3$ and $\theta = 2/3$, and the posterior Gamma with $\alpha = 3$ and $\theta = 0.4$, are shown:



If instead we observe 5 claims in a year, then the posterior distribution of that insured's (unknown) Poisson parameter is a Gamma distribution with $\alpha = 8$ and $\theta = 0.4$, with an average of: (8)(0.4) = 3.2.[56] Thus this observation has raised our estimate of this insured's future claim frequency.
The posterior Gamma in the case of this alternate observation is shown below:



56 The posterior alpha is 3 + 5 = 8. The posterior theta = 1/{1+1/(2/3)} = 1/ 2.5 = 0.4.

Shifting Risk Parameters:

One possible explanation for the credibilities increasing significantly less than linearly provided by Bailey-Simon is: "an individual insured's chance for an accident changes from time to time within a year and from one year to the next."[57] This is the concept of shifting risk parameters as discussed in the syllabus reading by Mahler, "An Example of Credibility and Shifting Risk Parameters."

When parameters shift over time, more distant years are worse predictors than they otherwise would have been. For example, let us assume 1956, 1957 and 1958 are available for predicting 1959. 1956 will be more affected by shifting risk parameters than would be 1958. Due to shifting risk parameters, all of the credibilities will be smaller than they otherwise would be, but the three year credibility (data from 1956 to 1958) is affected more than is the one year credibility (1958 data).

Thus we see a ratio of the three year to the one year credibility that is significantly less than 3. The more rapid the shifting, the larger the effect and thus the smaller this ratio.

A Model with No Shifting Risk Parameters:[58]

Assume there are no territories and we are looking at one class. Insureds do not move in and out of this class. Each insured is Poisson.
There are 100,000 insureds with λ = 10%, and 100,000 insureds with λ = 30%.

Of those with λ = 10%, the expected number claims free for one year is: 100,000 $e^{-0.1}$ = 90,484.
Of those with λ = 30%, the expected number claims free for one year is: 100,000 $e^{-0.3}$ = 74,082.
The expected future frequency for those who were claim free for one year is:

$$\frac{(90{,}484)(10\%) + (74{,}082)(30\%)}{90{,}484 + 74{,}082} = 19.00\%.$$

The overall frequency is 20%.
Thus, 1 - Z = 19.00% / 20%. $\Rightarrow$ Z = 5.00%.

[57] "The fact that the relative credibilities in Table 3 for two and three years are much less than 2.00 and 3.00 is partially caused by risks entering and leaving the class. But it can be fully accounted for only if an individual insured's chance for an accident changes from time to time within a year and from one year to the next, or if the risk distribution of individual insureds has a marked skewness reflecting varying degrees of accident proneness."
[58] Similar to a simple Bayes Analysis question on a preliminary exam.

Exercise: Using the technique in Bailey-Simon, determine the two-year credibility.
[Solution: Of those with λ = 10%, the number claims free for 2 years is: 100,000 $e^{-0.2}$ = 81,873.
Of those with λ = 30%, the number claims free for 2 years is: 100,000 $e^{-0.6}$ = 54,881.
The expected future frequency for those who were claims-free for two years is:

$$\frac{(81,873)(10\%) + (54,881)(30\%)}{81,873 + 54,881} = 18.026\%.$$

Thus, 1 - Z = 18.026% / 20%. ⇒ Z = 9.87%.
<u>Comment</u>: Bailey-Simon use data. We have applied their technique to the data we would expect to see if the given model were correct.]

Exercise: Using the technique in Bailey-Simon, determine the three-year credibility.
[Solution: Of those with λ = 10%, the number claims free for 3 years is: 100,000 $e^{-0.3}$ = 74,082.
Of those with λ = 30%, the number claims free for 3 years is: 100,000 $e^{-0.9}$ = 40,657.
The expected future frequency for those who were claims-free for three years is:

$$\frac{(74,082)(10\%) + (40,657)(30\%)}{74,082 + 40,657} = 17.087\%.$$

Thus, 1 - Z = 17.087% / 20%. ⇒ Z = 14.57%.]

The ratio of the two-year credibility to the one-year credibility is: 9.87% / 5% = 1.974.
The ratio of the three-year credibility to the one-year credibility is: 14.57% / 5% = 2.914.
Thus these credibilities increase slightly less than linearly, but much closer to linearly than those in Bailey-Simon.[59] This behavior can be explained by the Buhlmann Credibility Formula, Z = N / (N+K).

Exercise: Determine the Buhlmann Credibility Parameter, K, for this model.
[Solution: EPV = (10% + 30%)/2 = 0.2. VHM = $\{(0.1 - 0.2)^2 + (0.3 - 0.2)^2\}/2$ = 0.01.
K = EPV / VHM = 0.2 / 0.01 = 20.]

Comparing the Buhlmann (least squares) Credibilities with those from the claims-free discounts:

N	Credibility from Claims-Free	Buhlmann Credibility
1	5.00%	1/(1+20) = 4.76%
2	9.87%	2/(2+20) = 9.09%
3	14.57%	3/(3+20) = 13.04%

As expected the credibilities from the claims-free discounts are similar to those from Buhlmann Credibility, which increase somewhat less than linearly.

[59] In Table 3 of Bailey-Simon for Class 1, the ratios are 1.48 and 1.74.

Bayes Analysis versus Buhlmann Credibility:

For this simple example, let us assume we observe the total number of claims over three years for an individual insured of unknown type.

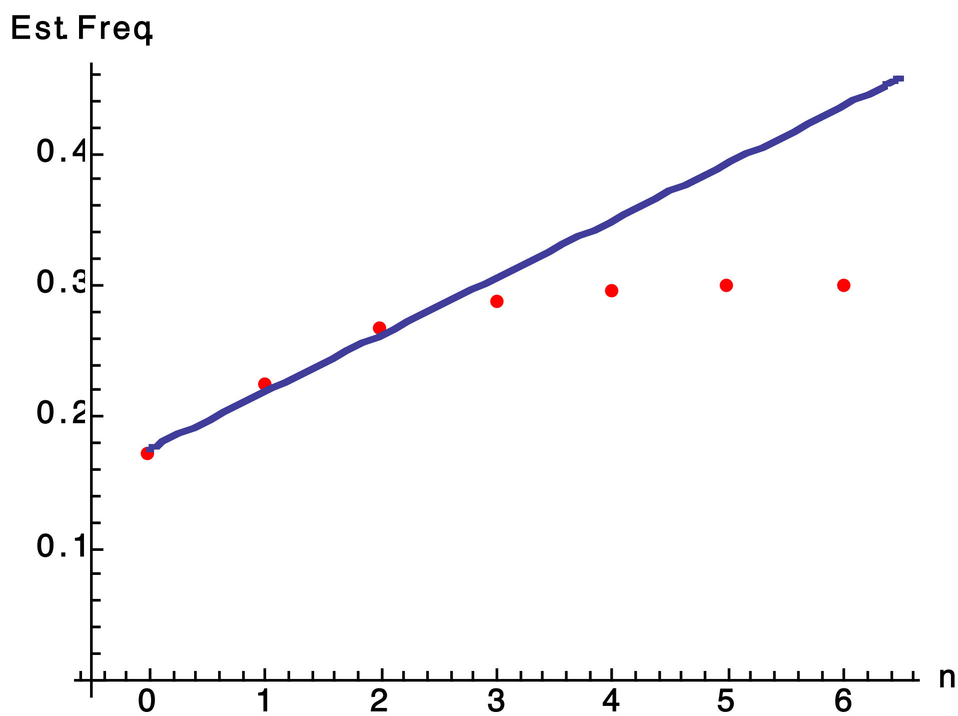
We had previously computed K = 20, Z = 3/23. Thus if we observe n claims in three years, the estimated future annual frequency is: (3/23) n + (20/23)(0.2).

Exercise: Assume we see one claim in three years.
Use Bayes Analysis to estimate the future annual frequency for that insured.
[Solution: Over three years we have a Poison with mean 3λ.
The chances of the observation are: $0.3\ e^{-0.3}$, and $0.9\ e^{-0.9}$.
Since the risk types are equally likely, the posterior probabilities are:

$$\frac{0.3\ e^{-0.3}}{0.3\ e^{-0.3} + 0.9\ e^{-0.9}} = 0.378, \text{ and } \frac{0.9\ e^{-0.9}}{0.3\ e^{-0.3} + 0.9\ e^{-0.9}} = 0.622.$$

Thus the estimated future estimated annual frequency for this insured is:
(0.378)(10%) + (0.622)(30%) = 22.44%.]

Proceeding in a similar manner, we can get the estimate from Bayes Analysis for other possible observations. Here is a graph with the Buhlmann Credibility Estimate as the straight line, and the estimates from Bayes Analysis as the dots, for n = 0, 1, ..., 6: [60]

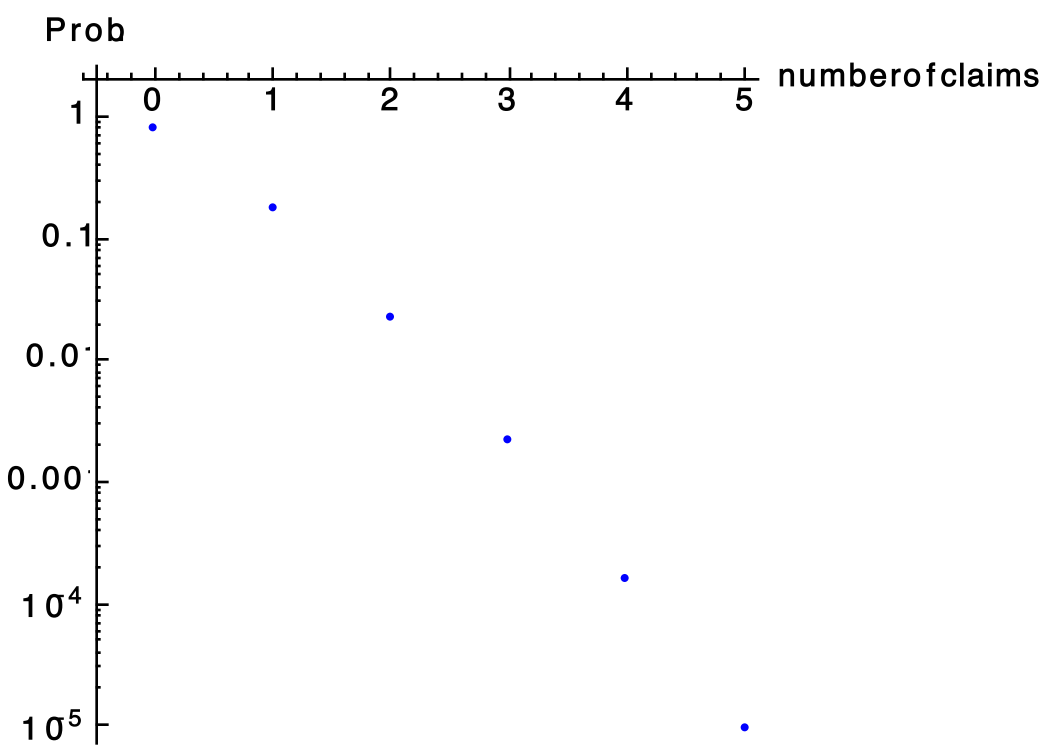


[60] We can observe more than 6 claims.

In general, the line formed by the Buhlmann Credibility estimates is the weighted least squares line to the Bayesian estimates, with the a priori probability of each outcome acting as the weights. The slope of this weighted least squares line to the Bayesian Estimates is the Buhlmann Credibility. Buhlmann Credibility is the Least Squares approximation to the Bayesian Estimates.

Exercise: Assume we see one claim in three years.
Use Bayes Analysis to estimate the probability of seeing two claims next year.
[Solution: From the previous exercise, the posterior probabilities are: 0.378, and 0.622.
Thus the probability that this insured will have 2 claims next year is:
$(0.378)(0.1^2\ e^{-0.1} / 2) + (0.622)(0.3^2\ e^{-0.3} / 2) = 2.24\%$.
Comment: A question for a preliminary exam, that you do not expect to be asked on this exam.]

For an insured who had one claim in three years, here its distribution of number of claims for the following year, with probability shown on a log scale:



Number of Insureds Claims-Free for Exact Numbers of Years.[61]

For the previous model, there are no territories and we are looking at one class.
Insureds do not move in and out of this class. Each insured is Poisson.
There are 100,000 insureds with λ = 10%, and 100,000 insureds with λ = 30%.

The expected number of insureds with no years claims-free, in other words who have at least one claim the first year is: 100,000 $(1 - e^{-0.1})$ + 100,000 $(1 - e^{-0.3})$ = 35,434.

Exercise: Determine the expected number of insureds claims-free for exactly one year.
[Solution: The expected number claims-free for at least one year is:
100,000 $e^{-0.1}$ + 100,000 $e^{-0.3}$ = 164,566.
The expected number claims-free for at least two years is:
100,000 $e^{-0.2}$ + 100,000 $e^{-0.6}$ = 136,754.
Thus the expected number claims-free for exactly one year is: 164,566 - 136,754 = 27,812.]

Exercise: Determine the expected number of insureds claims-free for exactly two years.
[Solution: 100,000 $(e^{-0.2} - e^{-0.3})$ + 100,000 $(e^{-0.6} - e^{-0.9})$ = 22,015.]

Here is a list of the expected number of insureds claims-free for exactly t years:[62]

t=0	t=1	t=2	t=3	t=4	t=5	t=6	
35,435	27,811	22,015	17,587	14,185	11,555	9507	
t=7	t=8	t=9	t=10	t=11	t=12	t=13	
7899	6627	5611	4791	4124	3574	3118	
t=14	t=15	t=16	t=17	t=18	t=19	t=20	More than 20
2735	2411	2135	1896	1690	1510	1352	12,433

[61] See 8, 11/16, Q.1.
[62] Assuming for simplicity that every insured has been driving for at least 21 years.

A Model with Shifting Risk Parameters:[63]

Alter the previous model so that each year an insured of a given type has a 20% chance of switching to the other type.[64] Thus an insured who has λ = 10% this year, has an expected frequency next year of: (80%)(10%) + (20%)(30%) = 14%. An insured who has λ = 30% this year, has an expected frequency next year of: (80%)(30%) + (20%)(10%) = 26%.

Of those with λ = 10%, the number claims-free for one year is: 100,000 $e^{-0.1}$ = 90,484.
Of those with λ = 30%, the number claims-free for one year is: 100,000 $e^{-0.3}$ = 74,082.
Thus the expected future frequency for those who were claims-free for one year is:

$$\frac{(90{,}484)(14\%) + (74{,}082)(26\%)}{90{,}484 + 74{,}082} = 19.402\%.$$

The overall frequency is 20%.
Thus, 1 - Z = 19.402% / 20%. $\Rightarrow$ Z = 2.99%.[65]

Of those with λ = 10% in the first year who were claims-free, the next year (0.8)(90,484) = 72,387 of them have λ = 10%, while (0.2)(90,484) = 18,097 of them have λ = 30%.
Of those with λ = 30% in the first year who were claims-free, the next year (0.2)(74,082) = 14,816 of them have λ = 10%, while (0.8)(74,082) = 59,266 of them have λ = 30%.
In summary, of those who were claims-free the first year, during the second year 72,387 + 14,816 = 87,203 will have λ = 10%, while 18,097 + 59,266 = 77,363 will have λ = 30%.

Of those who were claims-free in year one and with λ = 10% in year two, the number claims-free in year two is: 87,203 $e^{-0.1}$ = 78,905.
Of those who were claims-free in year one and with λ = 30% in year two, the number claims-free in year two is: 77,363 $e^{-0.3}$ = 57,312.
Thus the expected future frequency for those who were claims-free for two years is:

$$\frac{(78{,}905)(14\%) + (57{,}312)(26\%)}{78{,}905 + 57{,}312} = 19.049\%.$$

Thus, 1 - Z = 19.049% / 20%. $\Rightarrow$ Z = 4.76%.

---

[63] You are very unlikely to be asked a numerical question requiring you to work with such a model on your exam.
[64] This is relatively fast rate of shifting risk parameters over time. This is a extremely simplified and unrealistic version of the models in "A Markov Chain Model of Shifting Risk Parameters", by Howard Mahler, PCAS 1997.
[65] Due to shifting risk parameters over time, the one-year credibility has declined from 5.00% to 2.99%.

Exercise: Determine the credibility for three years claims-free.
[Solution: Of those with λ = 10% in the 2nd year who were claims-free, the next year (0.8)(78,905) = 63,124 of them have λ = 10%, and (0.2)(78,905) = 15,781 have λ = 30%.
Of those with λ = 30% in the 2nd year who were claims-free, the next year (0.2)(57,312) = 11,462 of them have λ = 10%, while (0.8)(57,312) = 45,850 of them have λ = 30%.
In summary, of those who were claims-free for 2 years, during the third year 63,124 + 11,462 = 74,586 will have λ = 10%, while 15,781 + 45,850 = 61,631 will have λ = 30%.
Of those who were claims-free in years one and two with λ = 10% in year three, the number claims-free in year three is: 74,586 $e^{-0.1}$ = 67,488.
Of those who were claims-free in year one and with λ = 30% in year two, the number claims-free in year two is: 61,631 $e^{-0.3}$ = 45,657.
Thus the expected future frequency for those who were claims-free for three years is:

$$\frac{(67,488)(14\%) + (45,657)(26\%)}{67,488 + 45,657} = 18.842\%.$$

Thus, 1 - Z = 18.842% / 20%. $\Rightarrow$ Z = 5.79%.]

Comparing credibilities from the claims-free discounts with and without shifting risk parameters:

N	No Shifting	Ratio to One-Year	With Shifting	Ratio to One-Year
1	5.00%		2.99%	
2	9.87%	1.97	4.76%	1.59
3	14.57%	2.91	5.79%	1.94

With shifting risk parameters, the credibilities increase much less than linearly. This is similar to the pattern in Bailey-Simon.[66]

[66] In Table 3 of Bailey-Simon for Class 1, the ratios are 1.48 and 1.74.

Number of Insureds Claims-Free for Exact Numbers of Years, Shifting Risk Parameters:[67]

The expected number of insureds with no years claims-free, in other words who have at least one claim the first year is: 100,000 $(1 - e^{-0.1})$ + 100,000 $(1 - e^{-0.3})$ = 35,434,
the same as without shifting risk parameters.

Of the original 200,000 insureds, over two years there are 4 groups:
80,000 with λ = 10% in both years, 20,000 with λ = 10% the first year and 30% the second year,
80,000 with λ = 30% in both years, 20,000 with λ = 30% the first year and 10% the second year.

Thus the expected number claims-free for exactly one year is:
80,000 $(e^{-0.1} - e^{-0.2})$ + 20,000 $(e^{-0.1} - e^{-0.4})$ + 80,000 $(e^{-0.3} - e^{-0.6})$ + 20,000 $(e^{-0.3} - e^{-0.4})$
= 28,349. This compares to 27,811 without shifting risk parameters.

This type of calculation quickly gets very tedious, so instead I simulated this situation. Here is comparison between no shifting and shifting risk parameters of the numbers claims-free:[68]

	No Shifting	With Shifting			No Shifting	With Shifting
t=0	35,435	35,623		t=11	4124	4354
t=1	27,811	28,543		t=12	3574	3711
t=2	22,015	22,698		t=13	3118	3009
t=3	17,587	19,179		t=14	2735	2578
t=4	14.185	15,662		t=15	2411	2133
t=5	11,555	12,991		t=16	2135	1772
t=6	9507	10,770		t=17	1896	1473
t=7	7899	9208		t=18	1690	1220
t=8	6627	7429		t=19	1510	1015
t=9	5611	6290		t≥20	13,785	5162
t=10	4791	5180				

The two patterns are similar. However, in the case of shifting risk parameters fewer insureds are claims-free for very long periods of time than when risk parameters are not shifting.[69]
It is not at all clear to me how one could use the information solely of the observed numbers of insureds claims-free for exactly t years in order to determine whether or not risk parameters are shifting and if so how quickly.

---

[67] See 8, 11/16, Q.1.
[68] The "No Shifting"column are the expected numbers calculated previously.
The "With Shifting" are from simulation.
Note that the first two simulated numbers differ somewhat from the expected numbers calculated previously.
[69] In this case, the two types of insureds have very different mean claim frequencies, and the rate at which parameters shift is large, in order to make the effects easier to spot.

Risks Entering and Leaving a Class:

One possible explanation for the credibilities increasing significantly less than linearly provided by Bailey-Simon is: “risks entering and leaving the class.”[70]

Let us assume that cars are frequently moving from one class to another.[71] For example, let us assume it is common for a car that is pleasure use one year to be business use the next year, or vice-versa. So in the Bailey-Simon data it is common to move from Class 1 to Class 3 or vice-versa.

For example, let us assume one car was pleasure use in 1956 to 1959, while another car was business use in 1956 and 1957, but pleasure use in 1958 and 1959. Then the three years of data 1956-1958 will be worse at predicting 1959 in the latter case than the former case. When we combine a whole bunch of data consisting of both situations, the credibility of three years of data for predicting the future will be lower than if all the cars had remained in the same class.

In this example, the data for 1958 is an equally good predictor of 1959 for both cars. However, there is another car, for which the class would be different in 1958 and 1959. So again, when we combine a whole bunch of data consisting of different situations, the credibility of one year of data for predicting the future will be lower than if all the cars had remained in the same class.

However, for a given car, its class in 1956 is more likely to be different than that in 1959, than is 1958 to be different than 1959. Thus the average effect on the credibility of more distant years is greater than that on more recent years. Thus the credibility of three years of data is more affected by shifting of classes than is the credibility of one year of data. Thus the credibilities go up less than linearly. The more frequently on average the classes of cars shift, the more the effect on the credibilities, and the lower is the ratio of the three year credibility to the one year credibility.

The effect of shifting classes is mathematically the same as shifting risk parameters. However, often in theoretical work on credibility, the term “shifting risk parameters” is restricted to those cases where there has been no change in the classifications and territories used for rating.

---

[70] “The fact that the relative credibilities in Table 3 for two and three years are much less than 2.00 and 3.00 is partially caused by risks entering and leaving the class. But it can be fully accounted for only if an individual insured’s chance for an accident changes from time to time within a year and from one year to the next, or if the risk distribution of individual insureds has a marked skewness reflecting varying degrees of accident proneness.”

[71] Frequently could be once every five years on average.
This can equally well be moving from one territory to another.

Marked Skewness Reflecting Varying Degrees of Accident Proneness:

One possible explanation for the credibilities increasing significantly less than linearly provided by Bailey-Simon is: "if the risk distribution of individual insureds has a marked skewness reflecting varying degrees of accident proneness."[72] While they provide no further explanation, I believe they are referring to the Gamma-Poisson, which was starting to be discussed about that time with respect modeling Merit Rating.[73] [74] Unfortunately, I do not believe that this is a possible cause of the observed behavior of the credibilities.[75]

The overall mean frequency (for a class) is observed, and thus constrained in any model of the Canadian data in Bailey-Simon. Similarly, we can determine the credibility applied to one year of data; in fact Bailey-Simon shows two complementary ways to do so. From this credibility one can back out the Buhlmann Credibility Parameter K. There is then a unique Gamma-Poisson model (for each class.)

In the absence of shifting risk parameters or insureds entering and leaving classes, we would have (for each class) a Gamma-Poisson model. The claim free discounts come from Bayes Analysis, which for the Gamma-Poisson is the same as the least squares (Buhlmann) credibility. $Z = N / (N + K)$. From the magnitude of the credibilities for one year, K must be relatively big. Therefore, the credibilities are approximately linear in N.

I do not see how having a Gamma-Poisson or some other model changes this, since K is backed out of the data, and does not depend on which particular model is used.

[72] "The fact that the relative credibilities in Table 3 for two and three years are much less than 2.00 and 3.00 is partially caused by risks entering and leaving the class. But it can be fully accounted for only if an individual insured's chance for an accident changes from time to time within a year and from one year to the next, or if the risk distribution of individual insureds has a marked skewness reflecting varying degrees of accident proneness."

[73] See for example, "Automobile Merit Rating and Inverse Probabilities," by Lester B. Dropkin, PCAS 1960. Bailey and Simon were each very involved in the literature on this and related subjects at this time.

[74] Each insured has a Poisson frequency with mean λ. However, across the class λ varies via a Gamma Distribution.

[75] To be fair Bailey and Simon were each pioneers in the development of credibility theory, and did not have the benefit we have of the many developments since they wrote their classical paper. By the way, Robert's father Arthur Bailey developed and published the mathematics of what would later be called "Buhlmann Credibility," about 15 years before Buhlmann published.

Appendix I:

In their Appendix I, Bailey-Simon demonstrate why one would expect the credibility to increase approximately linearly with the number of years of data, given certain assumptions.[76] They set up a discrete risk type model, the type of model which should be familiar from earlier exams.

Each insured has a Poisson frequency. For each insured their mean is the same every year.[77]

Percent of Insureds	Poisson Parameter (mean annual frequency λ)
40%	5%
40%	10%
20%	20%

Then the a priori mean frequency is: (40%)(5%) + (40%)(10%) + (20%)(20%) = 10%.

Assume an insured picked at random is claim free for one year, let us use Bayes Analysis to estimate that insured's future annual frequency.[78]

Percent of Insureds	λ	Chance of Observation	Posterior Chance of Risk Type
40%	5%	$e^{-0.05}$	$0.4e^{-0.05} / (0.4e^{-0.05} + 0.4\ e^{-0.1} + 0.2\ e^{-0.2}) = 41.989\%$
40%	10%	$e^{-0.1}$	$0.4e^{-0.10} / (0.4e^{-0.05} + 0.4\ e^{-0.1} + 0.2\ e^{-0.2}) = 39.941\%$
20%	20%	$e^{-0.2}$	$0.2e^{-0.20} / (0.4e^{-0.05} + 0.4\ e^{-0.1} + 0.2\ e^{-0.2}) = 18.070\%$

Thus the estimated future frequency for this insured is:
(41.989%)(5%) + (39.941%)(10%) + (18.070%)(20%) = 9.707%.[79]

This is lower than the 10% overall a priori frequency. Since the λ for each insured remains the same, and the proportion of risks of each type remains the same, the expected overall future annual frequency is also 10%.

Thus the modification for one year claims free is: 9.707/10 = 0.9707.
The credibility for one year claim free is: Z = 1 - 0.9707 = 2.93%.[80]

[76] A key conclusion of their paper is that the credibilities increase much less than linearly.
[77] No shifting risk parameters.
[78] I do not expect you to be asked to do Bayes Analysis on this exam.
[79] Matches the 0.09707 claim frequency after one year claim free shown in Bailey-Simon. What they have done is mathematically the same as Bayes Analysis, just assuming for convenience a total of 250,000 insureds.
[80] Matches the result shown in Bailey-Simon.

Exercise: An insured is picked at random and has two years claims free.
Use Bayes Analysis to estimate this insured's future annual claim frequency.
[Solution: The chance of the observation is Exp[-2λ].

Type	A Priori Probability	Poisson Parameter	Chance of Observation	Probability Weights	Posterior Probability	Mean
A	40%	5%	90.484%	0.36193	43.951%	5%
B	40%	10%	81.873%	0.32749	39.769%	10%
C	20%	20%	67.032%	0.13406	16.280%	20%
Sum	100%	10%		0.82349	100.000%	**9.430%**

Comment: Matches the result shown in Bailey-Simon for t = 2.]

Then for two years claim free: 1 - Z = 9.430%/10%. ⇒ Z = 5.70%.

Exercise: Using the technique in Bailey-Simon, determine the credibility for 3 years claims free.
[Solution: The chance of the observation is Exp[-3λ].

Type	A Priori Probability	Poisson Parameter	Chance of Observation	Probability Weights	Posterior Probability	Mean
A	40%	5%	86.071%	0.34428	45.882%	5%
B	40%	10%	74.082%	0.29633	39.491%	10%
C	20%	20%	54.881%	0.10976	14.628%	20%
Sum	100%	10%		0.75037	100.000%	**9.169%**

Then for three years claim free: 1 - Z = 9.169%/10%. ⇒ Z = 8.31%.
Comment: Matches the result shown in Bailey-Simon for t = 3.]

The three credibilities are: 2.93%, 5.70%, and 8.31%.
The ratio of the two year to the one year credibility is: 5.70%/2.93% = 1.945.
The ratio of the three year to the one year credibility is: 8.31%/2.93% = 2.836.

While these credibilities increase somewhat less than linearly, it is much closer to linear than the results Bailey-Simon get for the Canadian data, as shown in their Table 3. In Table 3, for example, for Class 1 the ratio of the two year to the one year credibility is 1.48, while the ratio of the three year to the one year credibility is only 1.74.

One could instead apply Buhlmann Credibility to their simple model in Appendix I.[81]

The process variance for each type is λ, so the Expected Value of the Process Variance is:[82]
(40%)(5%) + (40%)(10%) + (20%)(20%) = 10%.

[81] I do not expect you to be asked to do a Buhlmann Credibility problem on this exam.
[82] When mixing Poissons, the EPV is equal to the overall mean.

The first moment of the hypothetical means is the a priori overall mean:
(40%)(5%) + (40%)(10%) + (20%)(20%) = 0.1.
The second moment of the hypothetical means is:
$(40\%)(5\%^2) + (40\%)(10\%^2) + (20\%)(20\%^2) = 0.013$.
Therefore, the Variance of the Hypothetical Means is: $0.013 - 0.1^2 = 0.003$.

The Buhlmann Credibility Parameter is: K = EPV / VHM = 0.1 / 0.003 = 33.33.

Thus the Buhlmann (least squares) Credibility for one year of data is: $\frac{1}{1 + 33.33} = 2.91\%$.

The Buhlmann Credibility for two years of data is: $\frac{2}{2 + 33.33} = 5.66\%$.

The Buhlmann Credibility for three years of data is: $\frac{3}{3 + 33.33} = 8.26\%$.

Again these credibilities increase somewhat less than linearly.
As pointed out in the discussion by Hazam, in general $Z = \frac{N}{N + K}$ increases less than linearly; however, for K large compared to N this formula is not that far from linear.

We note that while the Buhlmann Credibilities are close to the claim free credits, they are not the same. For example, 8.26% ≠ 8.31%. Except in special mathematical cases where they are equal, the two types of credibilities will be close but not the same.

<u>Appendix II</u>:

Assume that the overall frequency is Poisson with mean λ.[83] The portion of insureds with no claims in a year is $e^{-\lambda}$. Then the portion of insureds with at least one claim in a year is: $1 - e^{-\lambda}$. Let x be the mean number of claims had by such insureds. Then since the overall mean is λ, we must have: $\lambda = (0)(e^{-\lambda}) + x(1- e^{-\lambda})$. $\Rightarrow x = \lambda/(1- e^{-\lambda})$.[84]

For example, as shown in Table 1, for Class 1 the observed overall frequency (per exposure) is: 288,019 / 3,325,714 = 0.0866. Thus we assume that those insureds who were not claim free during the most recent year, had an average number of claims of approximately:
$0.0866 / (1- e^{-0.0866}) = 1.044$.[85]

Note that in Appendix I, the model is instead a mixture of Poissons. For the example shown there, the overall frequency is 10%. Also the percentage claims free is:
226,544 / 250,000 = 0.9062.

Let y be the mean number of claims had by insureds who had at least one claim.
Then since the overall mean is 10%, we must have:
$0.10 = (0)(0.9062) + (y)(1 - 0.9062)$. $\Rightarrow y = 1.066$.

For the overall mean of 10%, and the technique Bailey-Simon uses, one would instead estimate the mean number of claims for those who have at least one claim to be instead:
$0.1 / (1 - e^{-0.1}) = 1.051$.

Thus the simple model in Appendix II would produce slightly different results than the more complicated model in Appendix I. For the limited purpose for which it is used by Bailey-Simon, the simpler method is okay.

More generally, let
y = mean frequency of those who have had at least one claim in the last year.
overall mean = 0 f(0) + y {1 - f(0)}. $\Rightarrow$ y = (overall mean) / {1 - f(0)}.

Let R = the ratio of the actual losses to the expected losses.
Then R = 1/{1 - f(0)}. Then the mod is: Z R + 1 - Z.

If the frequency is Poisson, then $f(0) = e^{-\lambda}$, and for those who have at least one accident $R = 1/(1 - e^{-\lambda})$. For example, if $\lambda = 0.0866$, then $R = 1/(1 - e^{-0.0866}) = 12.055$.

If instead the frequency is Negative Binomial parameterized as per Bahnemann, then $f(0) = (1-p)^r$, and for those who have at least one accident: $R = \frac{1}{1 - (1-p)^r}$.[86]

---

[83] A better model would be a mixture of Poissons as per Appendix I.
[84] For small λ, this is approximately: $\lambda / (\lambda - \lambda^2/2) = 2/(2-\lambda)$.
[85] Matching the result in Bailey-Simon.
[86] See 8, 11/19, Q. 3a.

<u>The Discussion by William J. Hazam</u>:[87]

The areas discussed are: use of premium based rather than exposure based frequencies, the Buhlmann Credibility formula, and the use of convictions for moving traffic violations.

As discussed, Bailey-Simon divide claims by premiums at the Group B rate, in order to get frequencies to compare.[88] This avoids double-counting. Hazam points out:
"that a premium base eliminates maldistribution only if (1) high frequency territories are also high premium territories and (2) if territorial differentials are proper."[89]

While most is due to frequency, some of the variation in premiums by territory is due to differences in severity.[90] Nevertheless, using premiums in the denominator is an improvement.[91]

When Bailey-Simon was written, all expenses were treated as variable. Currently, some expenses are treated as fixed. This would raise another issue with the use of premiums in the denominator .

The Buhlmann Credibility formula says $Z = N / (N+K)$.[92] For large K, the credibility increases only slightly less than linearly. While this does not explain the behavior observed by Bailey-Simon, it is one reason why the credibilities would go up less than linearly.

Given the one year credibilities in Table 2 of Bailey-Simon, we can back out a Buhlmann Credibility Parameter. For example for Class 1, $1/(1+K) = 4.6\%$. Thus $K = 20.7$. We can then use this K to calculate 2-year and 3-year credibilities.

Class	One-Year Credibility	K	Two-Year Credibility	Three-Year Credibility
1	4.6%	20.7	8.8%	12.7%
2	4.5%	21.2	8.6%	12.4%
3	5.1%	18.6	9.7%	13.9%
4	7.1%	13.1	13.2%	18.6%
5	3.8%	25.3	7.3%	10.6%

[87] This 3 page discussion of Bailey-Simon is also on the syllabus.
[88] "The authors have chosen to calculate Relative Claim Frequency on the basis of premium rather than car years. This avoids the maldistribution created by having higher claim frequency territories produce more X, Y, and B risks and also produce higher territorial premiums."
[89] In other words, if all expenses are treated as variable, then the expected loss ratios by territory should be equal.
[90] After adjusting for difference in the average class rating factor, most of the difference in average pure premiums between territories for Private Passenger Automobile is due to difference in average frequency. Some is due to difference in average severity. Based on my work on Massachusetts Private Passenger Automobile, I estimate that somewhere around 1/5 of the difference is due to severity while the remaining 4/5 is due to frequency.
[91] "However, premium, although not perfect, is an improvement over exposure as a base for this type of study. The fact that either or both of these inherent assumptions may not always exist does not detract from the qualitative nature of the conclusions but may alter somewhat the basic relative frequencies of Table 1 and the consequent values in Tables 2 and 3."
[92] Hazam's review was written before Buhlmann published his papers. This formula goes back to the 1918 PCAS.

We can see that these two-year and three-year credibilities are a poor match to those in Table 2 of Bailey-Simon.

Class	Two-Year Credibilities		Three-Year Credibilities	
	Buhlmann Formula	Table 2	Buhlmann Formula	Table 2
1	8.8%	6.8%	12.7%	8.0%
2	8.6%	6.0%	12.4%	6.8%
3	9.7%	6.8%	13.9%	8.0%
4	13.2%	8.5%	18.6%	9.9%
5	7.3%	5.0%	10.6%	5.9%

Due to shifting risk parameters and other possible causes mentioned by Bailey-Simon, the Buhlmann Credibility formula is <u>not</u> a good model for the credibilities for different numbers of years shown in Table 2 of Bailey-Simon.

Finally, Hazam mentions that many Merit Rating plans in the U.S. use moving traffic violations in addition to claims.[93] The addition of this useful information allows one to better distinguish between insureds within the same class, and therefore justifies larger credits and larger surcharges than when using just claims history.[94]

The amount of credibility depends as well on how refined the class plan is. The more homogeneous the classes, the less need there is for Merit Rating, and the smaller the credibility assigned to the data of an individual insured.

In any case, an actuary should use appropriate caution about extending the results on one set of data to other somewhat different situations.

---

[93] "It may be surmised from this approach to the Canadian results that, in a balanced merit rating plan, there is not enough credibility by class to warrant the magnitude of credits now being offered by many U. S. plans. We must remember, however, that these results are based strictly on claim frequencies, not claim frequencies plus convictions frequencies. Adding convictions no doubt helps substantiate larger credits but it is dubious that it will support current merit rating differentials, if the Canadian experience is at all indicative of what we might expect in this country."

[94] I looked extensively at such data for Massachusetts Private Passenger Automobile Insurance when I was involved in the redesign of the mandatory SDIP in the early 1980s. It was clear that for example someone who had recently been convicted of speeding had a higher expected future claim frequency than an otherwise similar driver who had not.

The Impact of Different Territories, and Why We Use Premiums in the Denominator:

Let us take an extremely simple model. There are two territories with equal exposures, and no classes. Each insured is Poisson, and λ does not vary over time. In Territory 1, half of the insureds have λ = 2% and the other half have λ = 8%.[95] In Territory 2, half of the insureds have λ = 6% and the other half have λ = 14%. The average severity for all insureds is $10,000.

The overall frequency is 7.5%. The overall pure premium is $750.
Territory 1 has a mean frequency of 5%, while Territory 2 has a mean frequency of 10%.
Thus Territory 1 has a pure premium of $500, while Territory 1 has a pure premium of $1000.
Assuming no fixed expenses, we charge Territory 2 twice as much on average as Territory 1.

Let us assume we give a percentage credit to those who are claim free for at least three years. Let us see what happens if we calculate the three-year credibility using exposures (rather than base class premiums) in the denominator. For convenience, assume 400,000 insureds in total.[96]

Type	Number who are 3 years claims-free	Number who are not 3 years claims-free
λ = 2%	$(100{,}000)(e^{-0.06})$ = 94,176	5,824
λ = 8%	$(100{,}000)(e^{-0.24})$ = 78,663	21,337
λ = 6%	$(100{,}000)(e^{-0.18})$ = 83,527	16,473
λ = 14%	$(100{,}000)(e^{-0.42})$ = 65,705	34,295

In total, there are 322,072 claims-free and 77,929 who are not.
The average future annual (exposure based) frequency for those who were claims free is:

$$\frac{(2\%)(94{,}176) + (8\%)(78{,}663) + (6\%)(83{,}527) + (14\%)(65{,}705)}{322{,}072} = 6.951\%.$$

The average future annual (exposure based) frequency overall is 7.5%.
Thus, 1 - Z = 6.951. ⇒ Z = 7.3%. ⇒ A 7.3% discount from average.

We wish to charge Territory 1 $500 on average.
Thus we wish to charge those who are claims-free: (0.927)($500) = $463.5.
Let the base rate be x.
There are claims-free 94,176 + 78,663 = 172,839, and not claims free: 5824 + 21,337 = 27,161.
27,161 x + (172,839)($463.5) = (200,000)($500). ⇒ x = $732.27.

---

[95] There is no way to distinguish the two types.
I have chosen the means to be very different for illustrative purposes.
[96] For an insured with lambda = 8%, the three years frequency is Poisson with lambda = 24%.
For simplicity assume each insured has been licensed for at least three years and no insured switches territories.

We wish to charge Territory 2 $1000 on average.
Thus we wish to charge those who are claims-free: (0.927)($1000) = $927.
Let the base rate be y.
There are claims-free 83,527 + 65,705 = 149,232, and not claims free:
16,473 + 34,295 = 50,768.
50,768 y + (149,232)($927) = (200,000)($1000). ⇒ y = $1214.58.

For those claims-free in Territory 1, the expected pure premium is:
($10,000) {(2%)(94,176) + (8%)(78,663)} / 172,839 = $473.07.
For those not claims-free in Territory 1, the expected pure premium is:
($10,000) {(2%)(5824) + (8%)(21,337)} / 27,161 = $671.34.

For those claims-free in Territory 2, the expected pure premium is:
($10,000) {(6%)(83,527) + (14%)(65,705)} / 149,232 = $952.23.
For those not claims-free in Territory 2, the expected pure premium is:
($10,000) {(6%)(16,473) + (14%)(34,295)} / 50,768 = $1140.42.

Let us compare the amount charged to the expected pure premiums:

Territory	Claims-free	Expected Pure Premium	Premium Charged
1	Yes	$473.07	$463.50
1	No	$671.34	$732.27
1	All	$500	$500
2	Yes	$952.23	$927.00
2	No	$1140.42	$1214.58
2	All	$1000	$1000

Using the exposure based frequencies to determine the claims-free credibility and discount, the pure premiums by cell do not match well to the premiums charged. Let us see what happens if instead we use premium based frequencies, as per Bailey-Simon.

There are 322,072 claims-free. The expected number of claims next year for these insureds is:
(2%)(94,176) + (8%)(78,663) + (6%)(83,527) + (14%)(65,705) = 22,387.

Assume that the current base rate for Territory 2 is twice that of Territory 1, 1000 versus 500.[97]
Then the annual premium at base rates next year for these insureds is:
(500)(94,176) + (500)(78,663) + (1000)(83,527) + (1000)(65,705) = 235,651,500.

The premium based frequency (per $1000) for those who were claims-free is:
22,387 / 235,651.5 = 0.09500.

The average premium based frequency overall is 7.5%/0.75 = 0.10000.[98]
Thus, 1 - Z = 0.09500/0.10000. ⇒ Z = 5%.

---

[97] All that is important is that the ratio is two to one, so that the current territory relativity is correct.
[98] The overall pure premium, and thus the assumed average premium is $750 or 0.75 thousand.

Taking into account the mix of the claims-free insureds by territory has resulted in this case in a smaller credibility of 5% rather than 7.3%.

We wish to charge Territory 1 \$500 on average.
Thus we wish to charge those who are claims-free: (0.95)(\$500) = \$475.
Let the base rate be x.
There are claims-free 94,176 + 78,663 = 172,839, and not claims free: 5824 + 21,337 = 27,161.
27,161 x + (172,839)(\$475) = (200,000)(\$500). ⇒ x = \$659.09.

We wish to charge Territory 2 \$1000 on average.
Thus we wish to charge those who are claims-free: (0.95)(\$1000) = \$950.
Let the base rate be y.
There are claims-free 83,527 + 65,705 = 149,232, and not claims free:
16,473 + 34,295 = 50,768.
50,768 y + (149,232)(\$950) = (200,000)(\$1000). ⇒ y = \$1146.97.

Now the comparison of the amount charged to the expected pure premiums is:

Territory	Claims-free	Expected Pure Premium	Premium Charged
1	Yes	\$473.07	\$475
1	No	\$671.34	\$659.09
1	All	\$500	\$500
2	Yes	\$952.23	\$950
2	No	\$1140.42	\$1146.97
2	All	\$1000	\$1000

As expected from Bailey-Simon, the premium based frequencies do a much better job of estimating appropriate claim-free discounts than do the exposure based frequencies.
The remaining discrepancy comes from having a single discount for both territories.

If we look at a single territory, then it will not matter whether we use premiums or exposures in the denominator. In Territory 1, the (exposure based) frequency for those who are claim free is:

$$\frac{(2\%)(94{,}176) + (8\%)(78{,}663)}{94{,}176 + 78{,}663} = 4.7307\%.$$

The overall frequency in Territory 1 is 5%.

Thus, 1 - Z = 4.7307%/5%. ⇒ Z = 5.385%.

In Territory 2, the (exposure based) frequency for those who are claim free is:

$$\frac{(6\%)(83{,}527) + (14\%)(65{,}705)}{83{,}527 + 65{,}705} = 9.522\%.$$

The overall frequency in Territory 2 is 10%.

Thus, 1 - Z = 9.522%/10%. ⇒ Z = 4.777%.

An Example Using Driver Data [99]

Here is example of the results of study similar to that in Bailey-Simon, which was done at the same time. However, the data was on drivers rather than cars.[100]

Some cars are driven by more than one driver, while some drivers drive more than one car. Also a much larger percent of licensed drivers do not drive during a year or drive only a minimal number of miles, compared to the percent of insured cars that are not driven or are only driven a minimal number of miles. So it makes some difference in the results whether one analyzes cars or drivers.

Drivers were grouped by the number of traffic violations they had over a three year period.[101] Then the number of accidents over a three year period by the drivers in the different groups was compared.[102] As expected, those drivers with more violations had a higher mean frequency.

Number of Violations	Number of Drivers	Mean Number of Accidents	Variance of Number of Accidents
0	55,757	0.087	0.096
1	20.613	0.194	0.207
2	8,753	0.274	0.299
3	4,320	0.354	0.395
4	2,297	0.426	0.501
5 or more	3,195	0.553	0.610
Total	94,935	0.163	0.193

Also the variance of the number of accidents within each group was computed.
If we assume that for each driver the number of accidents is Poisson distributed with mean λ, and the lambdas vary via Gamma Distribution, then the mixed distribution is Negative Binomial. As discussed on a preliminary exam, if the Gamma has parameters α and θ, then the Negative Binomial has parameters $r = \alpha$ and $\beta = \theta$.

It was found that in total, the number of accident data was fit well by a Negative Binomial. One can fit via the method of moments a Negative Binomial, to the total and to each group above. Set $r\beta$ = mean, and $r\beta(1+\beta)$ = variance. The results are shown below.

---

[99] Taken from "Some Considerations on Automobile Rating Systems Utilizing Individual Driving Records," by Lester B. Dropkin, PCAS 1959, not on the syllabus. See also the discussion by Robert A. Bailey in PCAS 1960. See also "Merit Rating in Private Passenger Automobile Liability Insurance and the California Driver Record Study," by Frank Harwayne, PCAS 1959.

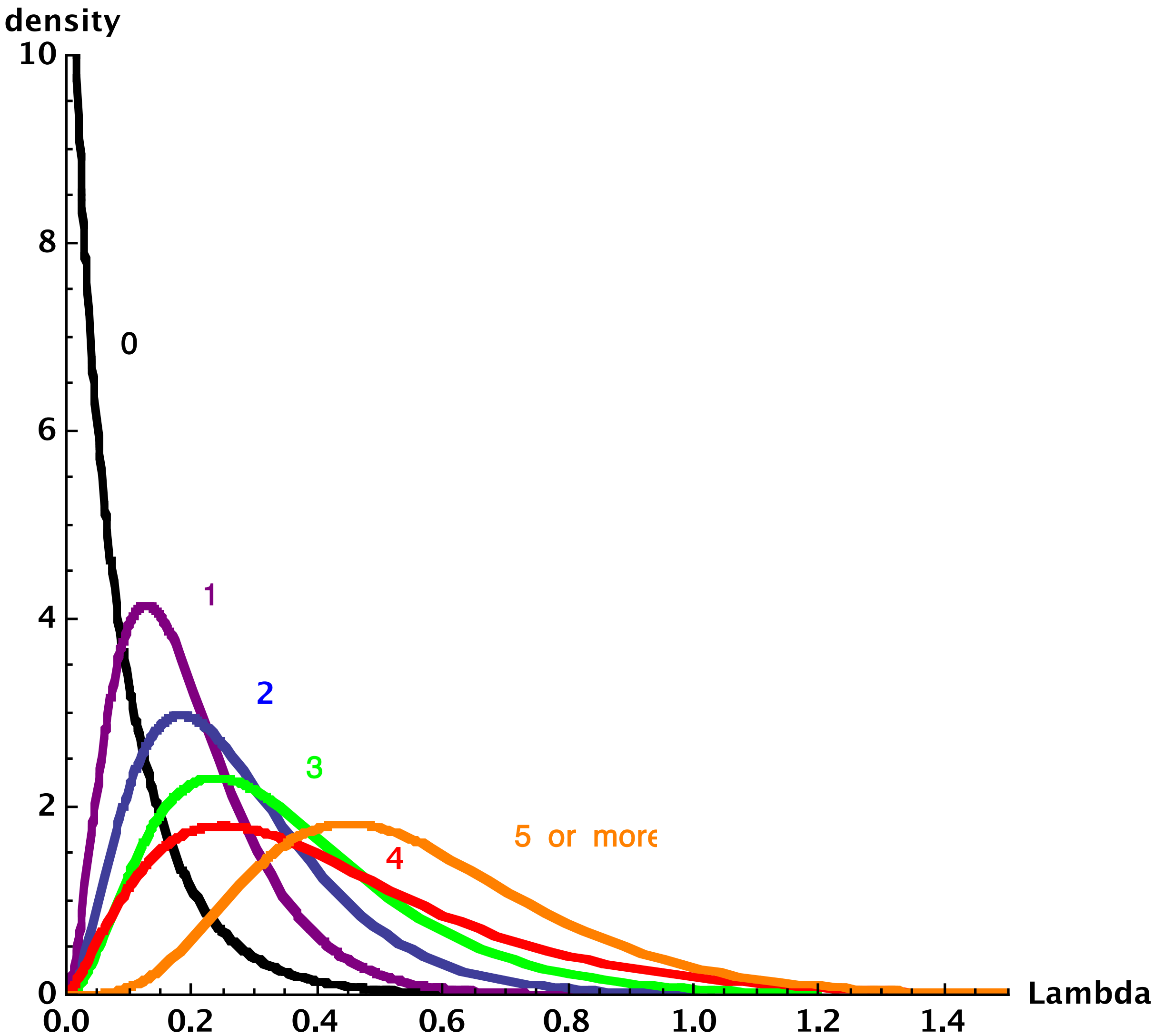
[100] Also the data was from California rather than Canada as in Bailey-Simon.

[101] Bailey-Simon instead looked at the number years a car had been claims-free.

[102] I believe it was over the same three year period as the violations.

Number of Violations	Fitted r	Fitted β
0	0.84	0.103
1	2.90	0.067
2	3.01	0.091
3	3.05	0.116
4	2.42	0.176
5 or more	5.37	0.103
Total	0.89	0.184

Then one can infer the parameters of the Gamma: $r = \alpha$ and $\beta = \theta$.
Here is a graph of the Gamma Distributions for the different groups of number of violations:[103]



[103] The group with no violations includes those licensed drivers who did not drive or drove only a minimal number of miles; this probably explains why its Gamma Distribution has a mode of zero. (Alpha is less than one.)

While each violations group is more homogeneous than the overall set of drivers, there is still lots of variation in expected mean frequency between drivers within a group.

One way to measure the homogeneity of each violations group is via the coefficient of variation (CV), the ratio of the standard deviation to the mean. For the Gamma Distribution the CV is $1/\sqrt{\alpha}$.

Number of Violations	Fitted alpha	CV
0	0.84	1.09
1	2.90	0.59
2	3.01	0.57
3	3.05	0.57
4	2.42	0.64
5 or more	5.37	0.43
Total	0.89	1.06

Based on the this measure, the group of those drivers with no violations is significantly more heterogeneous than the other groupings.[104]

Also as we would expect there is lots of overlap between the different groups. Here are the 10th and 90th percentile of the distributions of lambdas for the different groups.[105]

Number of Violations	Tenth Percentile	Ninetieth Percentile
0	0.006	0.209
1	0.070	0.347
2	0.101	0.486
3	0.131	0.625
4	0.134	0.793
5 or more	0.278	0.872
Total	0.014	0.388

Based on the Gamma Distribution inferred for all of the drivers in total, three years of accident data would be given 15.5% credibility for predicting future accident frequency.[106] This compares to three year credibilities in Bailey-Simon ranging from 5.9% to 9.9%.

---

[104] As mentioned before, the no violations group is mixture of those who did not drive a significant number of miles and those who did, making it more heterogeneous.

[105] Recall that these are three year accident frequencies.

[106] The overall Gamma Distribution of three year mean frequencies has $\theta = \beta = 0.184$.
The Buhlmann Credibility parameter is $K = 1/\theta = 5.435$.
However, here we have treated three years of data as one draw from the risk process.
$Z = 1 / (1 + 5.435) = 15.5\%$.

However, these credibilities are not comparable because of a number of reasons including:
1. Here we are looking at drivers rather than cars as in Bailey-Simon.
2. Different overall mean annual frequencies.
3. The 15.5% credibility would be in the absence of any classifications or territories, while in Bailey-Simon cars were divided between five classifications.[107]
4. Here we have the Buhlmann Credibility based on a Gamma-Poisson model, while in Bailey-Simon the credibility was based on the indicated claims-free discount.

Assuming, one divided the drivers into classes based on their number of violations, one could infer the credibility of three years of accident data from the Gamma fit to each violation group:[108]

Number of Violations	Fitted θ	Three Year Credibility
0	0.103	9.3%
1	0.067	6.3%
2	0.091	8.3%
3	0.116	10.4%
4	0.176	15.0%
5 or more	0.103	9.3%

Based on the above, we might give about 8% credibility to three years of accident data from drivers, if drivers were classified solely based on the number of violations they had over the last three years.

There is not enough information to infer what the credibility assigned to three years of either accident or violation data should be if there were a reasonable set of classifications and territories. The appropriate credibility given to the individual's experience would be less with a class and territory plan than in the absence of one. The better the class plan, the less credibility should be to assigned to the individual's experience in individual risk rating.

There is not enough information to infer what credibility should be assigned to three years of accident and violation data combined.[109] However, more credibility would be assigned than would be assigned to either the accident or violation data separately. Again the appropriate credibility given to the individual's experience would be less with a class and territory plan than in the absence of one.

---

[107] The better the class plan, the lower the credibility given to the experience of the individual.
[108] There are only 2300 to 4300 drivers in each of the last 3 categories, so there is considerable random fluctuation.
[109] One could just add the number of violations and accidents. Instead one could assign different numbers of "points" to different types of moving violations, and different numbers of points to different severities of at-fault and single vehicle accidents, as is done in some Safe Driver Insurance Plans.

Another Example, California Female Private Passenger Auto Drivers:[110]

Here is another example similar to that in Bailey-Simon. The data and analysis are different. Specifically, the data was on drivers rather than cars, tracked drivers over many more years than three, and the analysis was similar to that in Mahler's syllabus reading on shifting risk parameters.[111]

Number of Years of Data Used	Years Between Data and Estimate					
	1 (Most recent)	2	3	4	5	Total
1	3.2%	-	-	-	-	3.2%
2	3.1%	2.9%	-	-	-	6.0%
3	3.1%	2.8%	2.6%	-	-	8.5%
4	3.0%	2.7%	2.5%	2.3%	-	10.5%
5	3.0%	2.7%	2.4%	2.2%	2.1%	12.4%

Due to shifting risk parameters, the credibilities given to more distant years are less than those given to more recent years.

Also note that the total of the credibilities goes up more slowly than linearly with the number of years of data used. This is the same important pattern noted by Bailey and Simon for their data, although the increase is much closer to linear here than in Bailey-Simon.[112] Nevertheless, the increase in credibilities is less than would be expected from the Buhlmann Credibility formula: $Z = N / (N+K)$.

[110] Taken from Table 4, in "A Markov Chain Model of Shifting Risk Parameters," by Howard C. Mahler, PCAS 1997, not on the syllabus. These are least squares credibilities for no delay in receiving data. They were solved for in a manner parallel to "An Example of Credibility and Shifting Risk Parameters," by Howard C. Mahler.
[111] A fraction of drivers licensed in a state will not drive during a year, at least in that state. Some cars will be driven frequently by different drivers during a year. In any case, modeling licensed drivers is somewhat different than modeling the experience of insured cars.
[112] The credibilities depend on among other things the data used. Unlike the Bailey-Simon data from Canada, this data is not divided into classes and only includes female drivers. (There is another similar data set with male drivers.)

Fitting a Model of Shifting Risk Parameters to the Bailey-Simon Credibilities:

Assume that the covariance structure between years of data is of the form:[113]
$\mathrm{Cov}[X_i, X_j] = a\,\rho^{|i-j|} + b\,\delta_{ij}$, where δ_{ij} is zero if $i \neq j$ and one if $i=j$.
Since one can multiply the covariances by any constant and not change the least squares credibilities, for convenience let us take b = 1, so that $\mathrm{Cov}[X_i, X_j] = a\,\rho^{|i-j|} + \delta_{ij}$.[114]

Then the covariance matrix:
$$\begin{array}{c} \text{Year 1} \\ \text{Year 2} \\ \text{Year 3} \\ \text{Year 4} \end{array} \begin{pmatrix} 1+a & a\rho & a\rho^2 & a\rho^3 \\ a\rho & 1+a & a\rho & a\rho^2 \\ a\rho^2 & a\rho & 1+a & a\rho \\ a\rho^3 & a\rho^2 & a\rho & 1+a \end{pmatrix}.$$

Then applying credibility Z to the average of N years of data with no delay:[115]
$$Z = N \frac{\sum_{I=1}^{N} \mathrm{Cov}[X_i, X_{N+1}]}{\sum_{j=1}^{N}\sum_{I=1}^{N} \mathrm{Cov}[X_i, X_j]}.$$

For one year of data: $Z = \mathrm{Cov}[X_1, X_2] / \mathrm{Var}[X_1] = a\rho / (1+a)$.

For two years of data:
$$Z = 2 \frac{\mathrm{Cov}[X_1, X_3] + \mathrm{Cov}[X_2, X_3]}{\mathrm{Var}[X_1] + \mathrm{Cov}[X_1, X_2] + \mathrm{Cov}[X_2, X_1] + \mathrm{Var}[X_2]} = 2 \frac{a\rho + a\rho^2}{2 + 2a + 2a\rho}.$$

For three years of data:[116]
$$Z = 3 \frac{\mathrm{Cov}[X_1, X_4] + \mathrm{Cov}[X_2, X_4] + \mathrm{Cov}[X_3, X_4]}{3\mathrm{Var}[X] + 2\mathrm{Cov}[X_1, X_2] + 2\mathrm{Cov}[X_1, X_3] + 2\mathrm{Cov}[X_2, X_3]} = 3 \frac{a\rho + a\rho^2 + a\rho^3}{3 + 3a + 4a\rho + 2a\rho^2}.$$

---

[113] For $\rho < 1$, this models shifting risk parameters over time. This is an approximation to the form in "A Markov Chain Model of Shifting Risk Parameters," by Howard C. Mahler, PCAS 1997.
[114] Taking b = 1, then 1/a is similar to the Buhlmann Credibility Parameter K.
[115] Mathematically equivalent to equation 11.4 in "An Example of Credibility and Shifting Risk Parameters," by Howard C. Mahler.
[116] Note that if $\rho = 1$, in other words there are no shifting risk parameters, and replacing 1/a by K, then Z = (3)(3a) / (3 + 9a) = 3/(3+K), the usual Buhlmann Credibility formula.

For example, for Class 1 in the Bailey-Simon data, the credibilities are:[117]

One Year	Two Year	Three Year
4.6%	6.0%	8.0%

Setting these credibilities for one and two years equal to the previous formulas, we get two equations in two unknowns:

$a\rho / (1+a) = 0.046.$

$$2 \frac{a\rho + a\rho^2}{2 + 2a + 2a\rho} = 0.060.$$

Solving (with the aid of a computer): a = 0.09195, and ρ = 0.5463.

Plugging these values into the previous equation for the credibility for 3 years, we get:

$$3 \frac{a\rho + a\rho^2 + a\rho^3}{3 + 3a + 4a\rho + 2a\rho^2} = 7.9\%.$$

This is a reasonable match to the 8.0% in Bailey-Simon.

Proceeding in a similar manner for the other classes, we get:[118]

Class	Fitted a	Fitted ρ	Fitted 3 Year Credibility	Bailey-Simon 3 Year Cred.
1	0.09195	0.5463	7.9%	8.0%
2	0.12919	0.3933	6.5%	6.8%
3	0.09195	0.5463	7.4%	8.0%
4	0.33620	0.2822	8.7%	9.9%
5	0.11593	0.3658	5.4%	5.9%

There is a good match for Class 1, a fair match for Classes 2, 3, and 5, but a poor match for Class 4. This is due to an inherent problem in the use of Class 4 in the claims-free analysis of Bailey-Simon, which applies to a lesser extent, to Class 5.[119]

The definitions of the classes are given in the Bailey-Simon paper. Class 1 is Pleasure-No Male Operator under 25. Class 2 is Pleasure-Non-principal Male Operator under 25. Class 3 is Business Use. Class 4 is Unmarried Owner or Principal Operator under 25. Class 5 is Married Owner or Principal Operator under 25.

[117] See Table 2 in Bailey-Simon.
[118] Similar to Table 2.1 in Howard Mahler's Discussion of "An Analysis of Experience Rating" by Glenn Meyers, PCAS 1987, not on the syllabus.
[119] See Howard Mahler's Discussion of "An Analysis of Experience Rating," not on the syllabus.

There is an inherent problem in the use of Class 4 (Unmarried Owner or Principal Operator under 25) in the claims-free analysis of Bailey-Simon, which applies to a lesser extent, to Class 5 (Married Owner or Principal Operator under 25). The key point is that one cannot have three clean years of experience unless one has been licensed for at least three years. Class 4 includes many drivers who have less than three years of driving experience. Those risks with one year of experience go into Merit Rating Class Y (clean for one year) if they are clean, and Merit Rating Class B (clean for less than one year) if they are not.

Both Merit Rating Class A (clean for three years) and Merit Rating Class X (clean for two years) contain no risks with only one year of experience. We expect drivers with only one year of experience to be worse than the average for Class 4. Thus Merit Rating Class A (clean for three years) for driving Class 4, will have a lower frequency than the average for driving Class 4, merely because all of its drivers have at least three years of experience. Thus when we compare it to the remainder of driving Class 4, the resulting Bailey-Simon credibility for three years of data is overstated. The same is true to a lesser extent for the Bailey-Simon credibility for two years of data.

Note that in the fitted model, r is the rate at which the correlations decline as we increase the years of separation. For Class 1 in Bailey-Simon $\rho = 0.55$, which compares to an approximate value of $\rho = 0.95$ for the California Driver Data.[120]

Thus this would indicate that parameters are shifting much more quickly for the Canadian data in Bailey-Simon than the California Data. I find this unlikely, and suspect that something else explains the behavior in Bailey-Simon's data in addition to shifting risk parameters.[121]

Using the model fit to the Bailey-Simon credibilities for Class 1, the least squares credibilities with no delay are by year:[122] [123]

Number of Years of Data Used	Years Between Data and Estimate					
	1 (Most recent)	2	3	4	5	Total
1	4.6%	-	-	-	-	4.6%
2	4.5%	2.3%	-	-	-	6.8%
3	4.5%	2.3%	1.2%	-	-	7.9%
4	4.5%	2.2%	1.1%	0.6%	-	8.4%
5	4.5%	2.2%	1.1%	0.6%	0.3%	8.7%

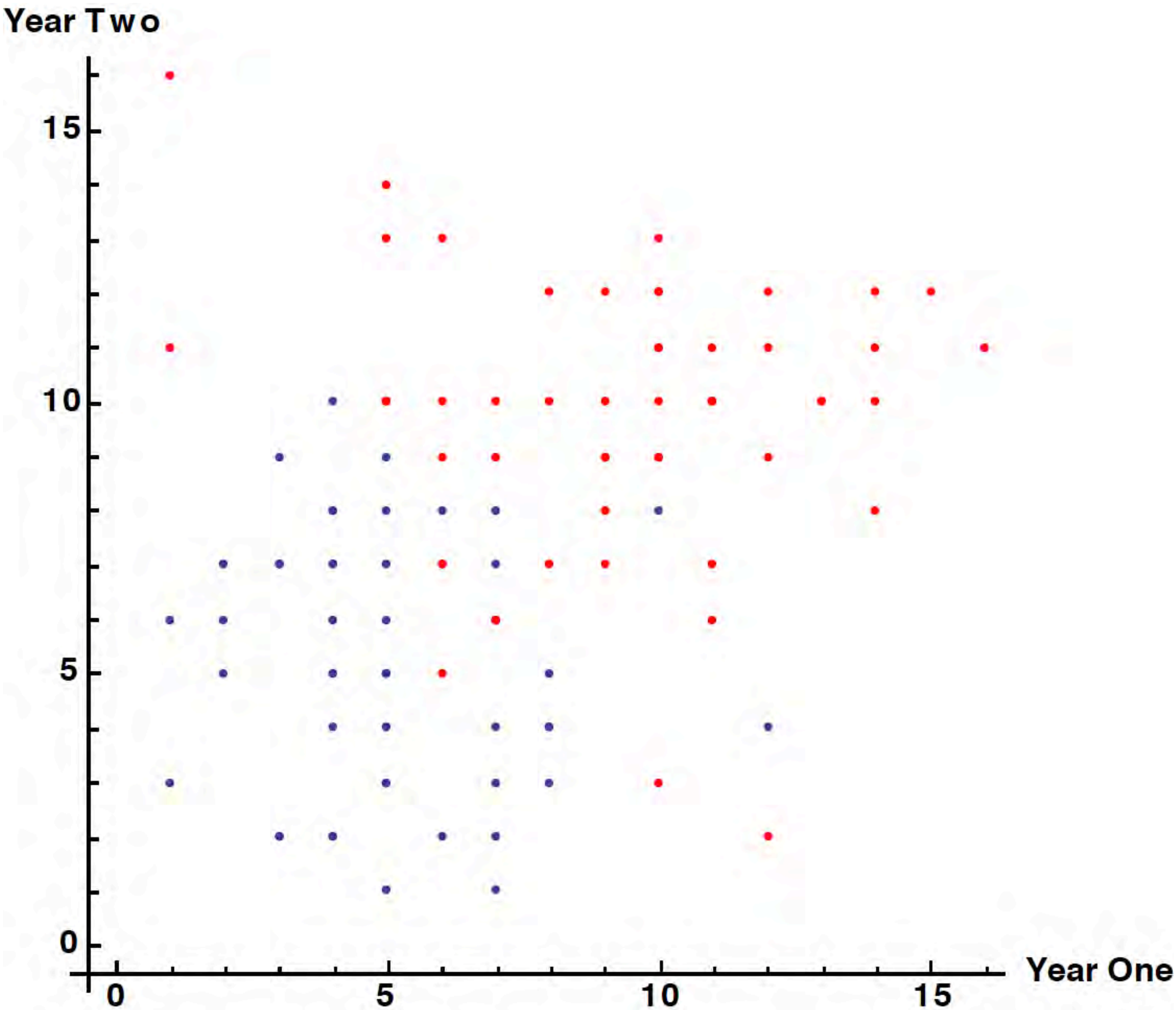
The total credibilities increase much less than linearly.

[120] With ρ approximately 0.94 for Female Drives and 0.97 for Male drivers. It is not clear that this difference between males and females is significant or just due to random fluctuations in the data set.
[121] In "An Analysis of Experience Rating," Glenn Meyers suggest parameter uncertainty is affecting the credibilities. Bailey-Simon mentions insureds switching classes and "the risk distribution of individual insureds has a marked skewness reflecting varying degrees of accident proneness" in addition to shifting risk parameters.
[122] Fit as per the method in "An Example of Credibility and Shifting Risk Parameters," by Howard C. Mahler.
[123] Total may differ from the sum of displayed values due to rounding of the displayed values.
For ten years the sum of the credibilities is 8.94%; the total approaches a limit of 8.95%.

<u>A Graphical Illustration</u>:[124]

Assume two type of insureds equally likely: Poisson with λ = 5 and Poisson with λ = 10.[125]

Let us simulate two years of frequency data from 50 insureds of each type.
Those with λ = 5 are shown as blue dots, while those with λ = 10 are shown as red pluses:



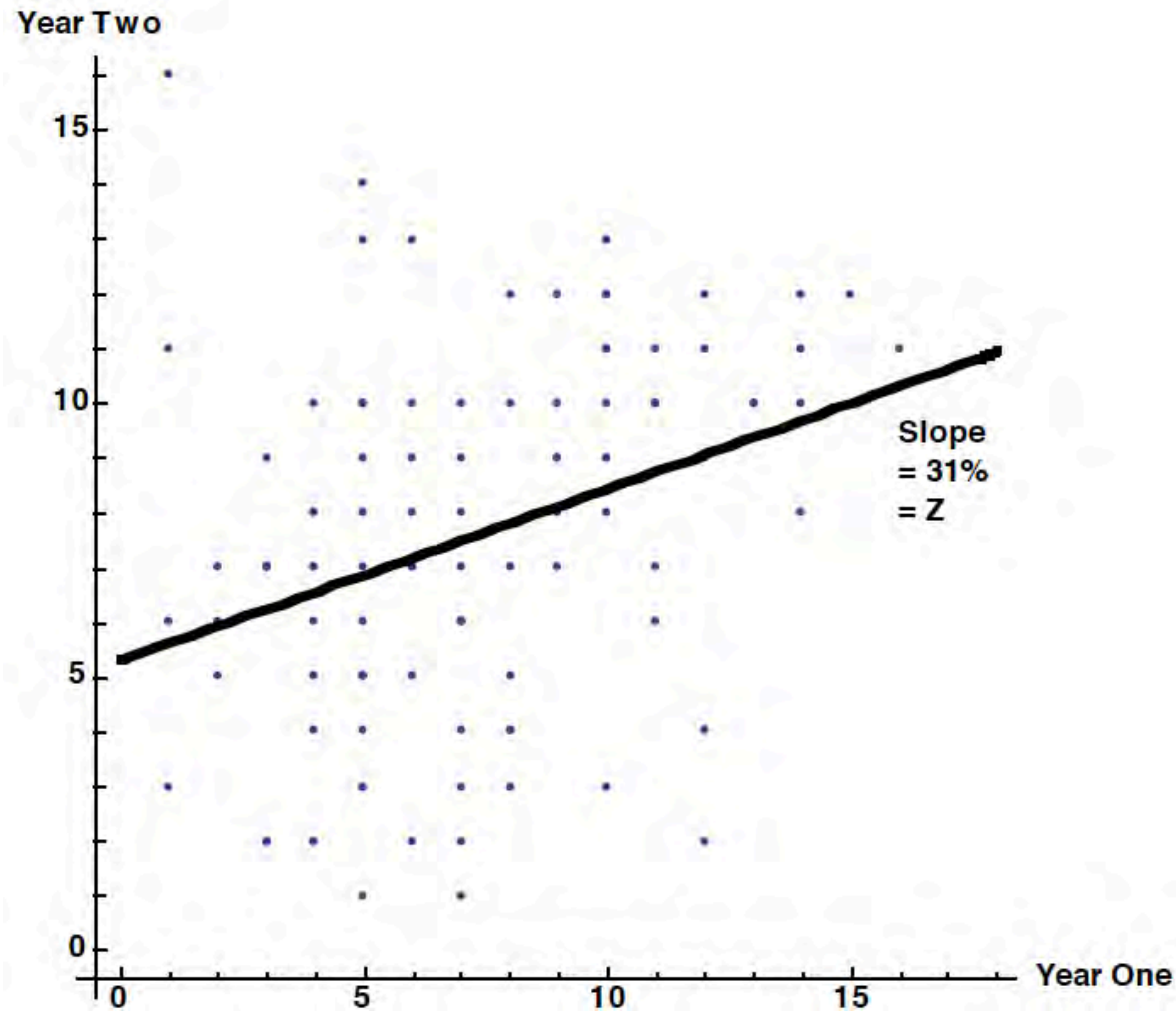
As expected, those with more claims than average in year 1 are more likely to have more claims than average in year 2. In other words, we can use past experience to predict future experience for an insured. This is the idea behind merit rating.

---

[124] See for example, "A Graphical Illustration of Experience Rating Credibilities," by Howard C. Mahler, PCAS 1998.
[125] The expected frequencies were chosen to be so large so that things would show up well in the graphs. Clearly this is not a model of Private Passenger Automobile Insurance.

In an insurance application of experience rating, we are assuming there is no way to distinguish the two types, other than through their past experience.

Here is a graph of the same data, without identifying the types of insureds:



A least squares line was fit to this data.[126] The slope of this fitted line is an estimate of the credibility of one year of data, in this case 31%.[127]

---

[126] You should not be asked to fit a regression on your exam.
[127] Due to limited data, this estimate of Z from data is subject to random fluctuation.

Important Ideas, Bailey-Simon:

Using claim frequency relative to premium instead of relative to exposures avoids distortions from maldistribution of merit rating classes between territory.
They use frequencies relative to average for those claims-free for various periods of time in order to estimate credibilities.

The alternative method of estimating a one-year credibility, compares frequencies relative to premiums vs. exposures for the group that is not claims-free.
Assuming Poisson frequency, the mean number of claims for those who were not claim free is: $\lambda / (1 - e^{-\lambda})$. Let λ = the mean claim frequency (per exposure) for the class.
M = relative premium based frequency for risks with one or more claims in the past year.

Then, $M = Z / (1 - e^{-\lambda}) + (1 - Z)(1)$. $\Rightarrow Z = \dfrac{M - 1}{1 / (1 - e^{-\lambda}) - 1} = (M - 1)(e^{\lambda} - 1)$.

The ratio of three year to one year credibility is much lower than three due to:
1. Marked skewness of the distribution of accident proneness.
2. Shifting risk parameters, which Mahler discusses in more detail.
3. Movement of insureds in and out of classes.
4. The nonlinearity of the credibility formula.

Merit rating credibility varies with claim frequency. A higher frequency is like a longer experience period for a Poisson distribution. Drivers with higher expected claim frequency have higher merit rating credibility, all else being equal.
The ratio of merit rating credibility to claim frequency varies by class. Homogeneous classes have higher class credibility and lower merit rating credibility. Merit rating extract the information after class rating has done its work. **A higher ratio of merit rating credibility to claim frequency in a class indicates greater heterogeneity of the drivers in that class**. As the class plan is more refined, classes are more homogeneous and the credibility of each risk declines.

The Three Conclusions of Bailey-Simon:
(1) The experience for one car for one year has significant and measurable credibility for experience rating.
(2) In a highly refined private passenger rating classification system which reflects inherent hazard, there would not be much accuracy in an individual risk merit rating plan, but where a wide range of hazard is encompassed within a classification, credibility is much larger.
(3) If we are given one year's experience and add a second year we increase the credibility roughly two-fifths. Given two years' experience, a third year will increase the credibility by one-sixth of its two-year value.

Problems:

2.1. (1 point)
Which of the following are conclusions reached by Bailey and Simon in their paper?
1. The experience for one car year has significant and measurable credibility for experience rating.
2. Merit rating adds a significant degree of accuracy to a private passenger rating system in which the classification system is highly refined, but it is of dubious value where a wide range of hazard is encompassed within a class.
3. If we are given one year's experience and add a second year, we increase the credibility roughly two-thirds.

2.2. (5 points) You are given the following data on the Adult Drivers Class for P.P. Auto Liability. Shown is the number of years they were without accident prior to 2010, the number of claims they had during 2010, and their loss cost premium during 2010 prior to the effects of Merit Rating:

Years since last accident	Premium ($ million)	Claims
5+	1520	134,200
4	70	8,900
3	80	10,400
2	90	12,500
1	100	14,400
0	140	19,600
Total	2000	200,000

a. (1 point) What is the credibility of 5 or more accident-free years of experience?
b. (1 point) What is the credibility of 4 or more accident-free years of experience?
c. (1 point) What is the credibility of 3 or more accident-free years of experience?
d. (1 point) What is the credibility of 2 or more accident-free years of experience?
e. (1 point) What is the credibility of 1 or more accident-free years of experience?

2.3. (2 points) Compare and contrast the Canadian Merit Rating Plan and the NCCI Experience Rating Plan, with respect to frequency and severity.

2.4. (1 point) Within a certain class and territory, you are given the following information for private passenger automobile insurance:

- Drivers with no claims in one year are expected to have 0.05 claims the next year.
- Drivers with 1 claim in one year are expected to have 0.12 claims the next year.

Determine the credibility of a single year of experience of a single private passenger car.

2.5. (2 points) You are an actuary at an insurer which writes private passenger automobile insurance. Alf Nadler is a critic of the insurance industry. Alf asks why for private passenger automobile insurance you use driver characteristics such as sex, age, marital status, principal place of garaging, and credit score, which are not socially acceptable, not controllable by the driver, and have no clear relation to future accidents. Alf proposes to the state legislature that insurers instead be required to use past accident history, which is socially acceptable, controllable by the insured, and has a clear relation to the expected future accidents. You are helping your company respond to Alf's proposal. What are some actuarial points you think your company's representative should make?

2.6. (1 point) Why do Bailey and Simon calculate claim frequency based on premiums rather than on car years when determining the credibility of claim-free experience?
A. Reliable data in terms of car years was not available at the level of detail required.
B. Premium as an exposure base adjusts for inflation from one year to another.
C. Because the same manual rates apply to each merit rating class, there was no material difference between the two exposure bases.
D. Premium as the denominator avoids distortion caused by variation in claim severity by territory.
E. Premium as the exposure base avoids distortion caused by variation in claim frequency by territory.

2.7. (1 point) The average pure premium in a territory for a class of private passenger automobile cars is $500 during 2012. You look at those cars within that class and territory that had no claims during 2011. The average pure premium for these cars during 2012 is $470. How much credibility would you give to a single private passenger car?

2.8. (2 points) Bailey and Simon present two methods of estimating a credibility for one year of data from a single private passenger car both based on data for the number of claims. Compare and contrast these two methods.

2.9. (1 point) If Bailey and Simon used claim frequency relative to car years instead of premium, their estimates of merit rating credibility would be:

	Cars with at least One Year Claim Free	Cars with No Claim Free Years
A.	understated	overstated
B.	overstated	understated
C.	understated	understated
D.	overstated	overstated
E.	unbiased	unbiased

2.10. (1 point) You are examining experience for private passenger automobile liability for 2 classes. Class 1 and 2 are similar, except class 1 has a mean frequency of 4%, while class 2 has a mean frequency of 12%.
Compare and discuss the Merit Rating credibilities of a single car from Class 1 for three years and a single car from Class 2 for one year.

2.11. (4.5 points) Based on Bailey and Simon's paper "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car" and the information given below, calculate the credibilities that can be assigned to the experience of a single private passenger car from each of the following two groups:
a. (1.5 points) The group of risks that have been claim free for one (1) or more years.
b. (1.5 points) The group of risks that have been claim free for no (0) years.
c. (1.5 points) Discuss why the techniques in parts (a) and (b) usually give different estimates of the credibility of one year of data.

Group	Number of Years Claim Free	Earned Car Years	Earned Premium at Present B Rates	Number of Claims Incurred
A	3 or more	185,000	225,000,000	18,200
X	2	12,000	15,000,000	1,400
Y	1	15,000	20,000,000	2,200
B	0	28,000	40,000,000	5,200
Total		240,000	300,000,000	27,000

2.12. (1 point) You are examining experience for private passenger automobile liability for 4 classes: retired drivers, young unmarried males, business use, and all others.
For which class would you expect to find the highest credibility for one year from a single car, relative to claim frequency? Briefly explain why.

2.13. (3 points) Les N. DeRisk is an actuary who is studying personal auto liability insurance for drivers aged 30 to 55.
Les assumes that personal auto claims are independent events; a claim in one week does not affect the likelihood of claims in other weeks.
Les correlates claims in years X and X+1 and finds that:
● Drivers with more claims in Year X are more likely to have claims in Year X+1.
● Across a large number of drivers, the correlation of the number of claims in Year X+1 and Year X for individual drivers is 10%.
a. (1 point) Does this correlation negate the assumption that claims are independent?
b. (1 point) How should Les test the assumption that claims are independent?
c. (0.5 point) Does the 10% correlation imply a 10% merit rating credibility for one year of data?
d. (0.5 point) How should Les infer the merit rating credibility?

2.14. (3 points) An insurance company has a private passenger auto book of business.
There is the following claims experience for Class 1 in State X:

Territory	Earned Premium at Present Rates Prior to Merit Rating	Earned Car Years	Number of Claims
A	$15,000,000	20,000	800
B	$25,000,000	28,000	1250
C	$30,000,000	30,000	1300
D	$25,000,000	23,000	1100
E	$20,000,000	17,000	800
Total	$115,000,000	118,000	5250

You will be trying to determine the credibility of a single private passenger car for Class 1 in State X, by comparing the experience of those who are claims-free for various periods of time to the experience of all cars in Class 1 in State X.
Which ratio would be more appropriate to use in this analysis:

$$\frac{\text{Number of Claims}}{\text{Number of Earned Car Years}} \text{ or } \frac{\text{Number of Claims}}{\text{Dollars of Earned Premiums}} ?$$

Justify your selection.
Is there some other ratio that you would use instead of these two?

2.15. (2 points) Using the procedures and formulas from Bailey and Simon's paper "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," determine which of the current classes exhibits less variation of individual hazards than the others.
Use the data shown below:

	Claim Frequency per $1,000 Earned Premium	Earned Premium per Earned Car Year	Credibility of 3 years of Data from a Single Car
Class 1	0.263	$300	5.8%
Class 2	0.369	$400	9.3%
Class 3	0.311	$350	8.1%

Assume that the earned premiums are adjusted to a common current rate level.
Show all work.

2.16. (4.5 points) Use the following information for private passenger automobile insurance in the province of Manaberta:

- There are two territories with the same number of car years in each.

Territory	Average Premium	Average Frequency Per Car Year	Average Severity
1	400	10%	2400
2	500	8%	3750

- For those cars that are claims free for at least the last 3 years:

Territory	Car Years	Premium	Subsequent Year Number of Claims
1	100,000	38 million	9000
2	110,000	53 million	8100

In each case, determine the credibility for three years of data.
(a) (0.5 point) Combining the data for the two territories, and using premiums as the denominator of "claim frequency".
(b) (0.5 point) Combining the data for the two territories, and using car years as the denominator of claim frequency.
(c) (1 point) For each territory separately, and using premiums as the denominator of "claim frequency".
(d) (1 point) For each territory separately, and using car years as the denominator of claim frequency.
(e) (1.5 points) Discuss the differences in the results in the previous parts.

2.17. (2 points)
You are given the following private passenger automobile results for the state of Fremont. Using the techniques from Bailey and Simon's "An Actuarial Note on the Credibility of a Single Private Passenger Car," answer the questions below:

Class	Claim Frequency per Car Year	One-year Credibility	Three-year Credibility
1	0.07	0.05	0.10
2	0.08	0.09	0.17
3	0.09	0.08	0.17

a. (1 point) For which class do its insured have more stable expected claim frequencies over the three year period?
Assume that there is no change in the exposures in each class during the three years and that the risk distribution in each class is not markedly skewed. Explain your answer.
b. (1 point) Which class has less variation in expected claim frequency between individual risks within its class? Explain your answer.

2.18. (2 points)
For a specific class, the following data shows the experience of a merit rating plan.

Merit Rating	Number of Accident-Free Years	Earned Premium at Present B Rates	Number of Incurred Claims
A	3 or More	$2400 million	12,000
X	2	$200 million	1200
Y	1	$220 million	1400
B	0	$380 million	2600
	Total	$3200 million	17,200

The base rate (for Merit Rating B) is $800 per exposure for this class.
Calculate the appropriate premium for an exposure that is accident free for one or more years.

2.19. (2.5 points) An insurance company has a private passenger auto book of business with the following claims experience:

Group	Number of Accident-Free Years	Earned Premiums ($ million)	Current Merit Rating Factor	Number of Claims Incurred
A	3 or more	1040	0.65	130,000
X	2	64	0.80	8,000
Y	1	90	0.90	12,000
B	0	170	1.00	C
Total		1358		150,000 + C

- Claim counts per exposure follow a Poisson distribution with parameter $\lambda = 0.08$.
- The credibility for one year of data estimated by examining the experience of those insureds who were accident free for zero years is equal to 4.4%.

a. (1.5 points) Calculate C, the number of claims incurred for Group B.
b. (1 point) Calculate the indicated merit rating factors.

2.20. (1 point) We are experience rating two insureds that are in different classes. The first insured has a higher volume of claims and more exposures than the second insured. Describe why the experience of the first insured may be given less credibility than that of the second insured.

***2.21*.** (1.5 points) An insurance company has a private passenger auto book of business with an experience modification factor in its rating plan.
Given the following:
● The expected claim frequency for the entire book of business is 0.06.
● The credibility for the group of risks that have had at least one claim in the last year is 4%.
In each case, calculate the experience modification factor for a policy that has had at least one claim in the last year.
a. (0.75 points) Annual claims for an individual driver follow a Poisson Distribution.
b, (0.75 points) Annual claims for an individual driver follow a Negative Binomial Distribution with r = 3.

For the Negative Binomial Distribution:

◦ $f(x) = \binom{x+r-1}{x} (1-p)^r p^x$

◦ $E[X] = \frac{pr}{1 - p}$

2.22. (2 points) You are an actuary who works for the Mycorona insurance company.
Mycorona writes private passenger automobile insurance.
Due to the COVID-19 outbreak and the associated lockdown measures, expected claim frequency in AY 2020 will decrease very significantly from AY 2019.
Mycorona has a merit rating plan (safe driver insurance plan).
Fully discuss whether its merit rating plan will adequately compensate its insureds for the impact of the pandemic on their expected losses under private passenger automobile insurance.

2.23. (1963, CAS Fellowship Exam IV, part b, Q.9)
In “An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car”, relative claim frequency was calculated on the basis of premium rather than car years.
(a) Why was this approach taken?
(b) What are the assumptions underlying this approach?

2.24. (9, 11/88, Q.11a) (1 point) The 1986 policy year collision experience of a sample of 100,000 cars, each of which had been insured for at least the preceding three years, was tabulated as follows:

Merit Rating Class Number of Years Claims-Free Prior to 1986 Policy Year	Policy Year 1986 Exposure (Car-Years)	Policy Year 1986 Number of Claims
3 or more	71,000	7,800
2	9,000	1,400
1	10,000	1,600
0	10,000	1,700
Total	100,000	12,500

Use the method of Bailey and Simon in their paper "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car" to estimate the credibility of the experience of one car for one year.

2.25. (9, 11/88, Q.12) (1 point) You are the actuary for the XYZ Insurance Company. Currently, you are considering implementing an experience rating program for your private passenger automobile insureds based on each insured's experience. Your analysis shows that, while an insured's past claim frequency is very credible in predicting future claim frequency, an insured's past loss ratio is not very credible in predicting the future loss ratio. Based on Bailey and Simon's paper "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car", list two potential nonrandom causes of this phenomenon.

2.26. (9, 11/94, Q.31) (2 points) Based on the methodology and notation used by Bailey and Simon in "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," and the table below calculate the credibility for category B risks (i.e., risks whose number of claims-free years equals zero) for a one-year experience period. (You can assume that the Poisson distribution reasonably approximates the distribution of observed claim counts among the risks from all merit rating groups combined.) Show all of your work.

Merit Rating (Number of Accident-Free Years)	Earned Car Years	Earned Premium at Present Category B Rates	Number of Claims Incurred
A(3+)	3,005,000	195,400,000	260,000
X(2)	148,000	10,700,000	18,000
Y(1)	184,000	13,200,000	25,000
B(0)	330,000	23,000,000	46,000
Total	3,667,000	242,300,000	349,000

2.27. (2 points) In the previous question, 9, 11/94, Q.31, assume instead that the Geometric distribution reasonably approximates the distribution of observed claim counts among the risks from all merit rating groups combined. Calculate the credibility for category B risks.

2.28. (9, 11/95, Q.6) (1 point) According to Hazam's discussion of Bailey and Simon's paper "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," which of the following are true?
1. For a study like that presented by Bailey and Simon, the use of premium as a base is an improvement over the use of exposure as a base.
2. Using a premium base eliminates the maldistribution only if high frequency territories are also high premium territories and if territorial differentials are proper.
3. Bailey and Simon's statement "the credibilities for experience periods of one, two, and three years would be expected to vary approximately in proportion to the number of years" holds largely true only for low credibilities.

Comment: I have rewritten this past exam question.

2.29. (9, 11/95, Q.30) (3 points) Based on Bailey and Simon's paper "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car" and the information given below, calculate the credibilities that can be assigned to the experience of a single private passenger car from each of the following two groups:
a. (1.5 points) The group of risks that have been claim free for two (2) or more years.
b. (1.5 points) The group of risks that have been claim free for no (0) years.
Show all work.

Group	Number of Years Claims Free	Earned Car Years	Earned Premium at Present D Rates	Number of Claims Incurred
A	3 or More	650,000	390,000,000	54,250
B	2	200,000	120,000,000	21,000
C	1	75,000	45,000,000	10,125
D	0	75,000	45,000,000	14,625
Total		1,000,000	600,000,000	100,000

***2.30.* (9, 11/95, Q.32)** (3 points) You have been retained as a consulting actuary for Hirisk Auto Insurance Company. The company has asked for you to determine if any of the three classifications in use is possibly in need of further refinement.
The only data available are shown below:

	Claim Frequency Per $1,000 Earned Premium
Class A Total	1.625
Class B Total	1.750
Class C Total	2.212

Only Risks with 3 or More Years Loss Free	Earned Premium Per Earned Car Year	Credibility of a Single Risk
Class A	$150	0.082
Class B	$148	0.046
Class C	$190	0.079

Using the procedures and formulas from Bailey and Simon's paper "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," determine whether one or more of the current classes exhibit(s) more variation of individual hazards than do(es) the other(s). Assume that the earned premiums are adjusted to a common current rate level. Show all work.

2.31. (9, 11/96, Q.50) (2 points)
You are given the following private passenger automobile results for a hypothetical state.
Using the techniques from Bailey and Simon's "An Actuarial Note on the Credibility of a Single Private Passenger Car," answer the questions below:

Class	Description
A	Pleasure Class - Unmarried Male Operator under age 25
B	Pleasure Class - Unmarried Female Operator under age 25
C	Pleasure Class - Operator over age 55

Class	1995 Claim Frequency	1995 One-year Credibility	1993-1995 Three-year Credibility
A	0.12	0.18	0.36
B	0.10	0.08	0.22
C	0.08	0.16	0.48

a. (1 point) Which class has a more stable claim frequency over the three year period?
Assume that there is no change in the exposures in each class during the three years and that the risk distribution in each class is not markedly skewed. Explain your answer.
b. (1 point) Which class has less variability in claim frequency within its class?
Explain your answer.

2.32. (9, 11/97, Q.19) (1 point) According to Bailey and Simon's "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," which of the following are true?
1. Relative claim frequency is calculated on a premium basis to avoid biases due to the fact that exposure based frequency varies by territory.
2. Credibility for experience rating depends only on the volume of data in the experience period.
3. The experience for one car for one year has significant and measurable credibility for experience rating.

2.33. (9, 11/98, Q.26) (3 points) Based on Bailey and Simon's "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," answer the following questions.
a. (2 points) Using the information below, calculate the number of claim incurred for Group C. Show all work.

Group	Number of Years Claim Free	Earned Car Years	Earned Premium at Present Group D Rates	Number of Claims Incurred
A	3 or more	700,000	$420,000	62,376
B	2	175,000	$105,000	15,955
C	1	100,000	$60,000	?????
D	0	25,000	$15,000	?????
Totals		1,000,000	$600,000	98,000

Credibility for the group of risks with one or more claim-free years (Z) = 0.086
b. (0.5 point) What conclusion do the authors reach with respect to merit rating using one year's worth of experience?
c. (0.5 point) In a highly refined private passenger rating classification system, what relative credibilities would the authors conclude should be assigned to the experience of an individual risk compared to the experience of a class?

2.34. (9, 11/99, Q.1) (1 point) In Bailey and Simon's "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," the authors state that under certain conditions, the credibilities associated with experience periods of one, two, and three accident-free years for insureds within a given class would be expected to vary approximately in proportion to the number of years. Which of the following are reasons why this would not be true?
1. Changes in an individual insured's chance for an accident within a year.
2. Skewness in the risk distribution of individual insureds.
3. The impact of risks entering and leaving the class.

2.35. (9, 11/00, Q.32) (3 points)
Based on Bailey and Simon's "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car" and the table below, answer the following.

	Private Passenger Automobile Liability - Non-Farmers				
	Class 3 - Business Use				
Merit Rating	Earned Car Years	Earned Premium at Present B Rates	Number of Claims Incurred	Claim Frequency per $1,000 of Premium	Relative Claim Frequency
A	247,424	$25,846,000	31,964	1.237	0.920
X	15,868	$1,783,000	2,695	1.511	1.123
Y	20,369	$2,281,000	3,546	1.555	1.156
B	37,666	$4,129,000	7,565	1.832	1.362
Total	321,327	$34,039,000	45,770	1.345	1.000

where: Class A - Three or more years claim free
Class X - Two years claim free
Class Y - One year claim free
Class B - Zero years claim free

a. (1.5 points) Calculate the credibilities for a single private passenger car for one year, two years, and three years. Show all work.
b. (0.5 point) Briefly describe the relationship that Bailey and Simon expect between the three credibilities from part (a).
c. (1 point) Do the credibilities calculated in part (a) follow the relationship described in part (b)? Briefly explain why or why not.

2.36. (9, 11/01, Q.2) (1 point) According to Bailey and Simon's "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," which of the following is false?
A. The experience for one car for one year has significant and measurable credibility for experience rating.
B. Credibility for experience rating depends on the variation of individual hazards within the class.
C. In a highly refined private passenger rating classification system that reflects inherent hazard, there would not be much accuracy in an individual risk merit rating plan.
D. In experience rating, an increase in the volume of data in the experience period increases the reliability of the indication in proportion to the square root of the volume.
E. None of A, B, C, or D are false.

2.37. (9, 11/01, Q.22) (2.5 points) Use Bailey and Simon's "An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car," and Hazam's discussion to answer the following questions.
a. (1.5 points) Using the information below, calculate the credibility for 1-year and 2-year claim free periods for Class 1. Show all work.

Number of Years Claim Free	Earned Premium at Present Rates	Number of Claims Incurred	Earned Car Years
2 or more	$5,000,000	7,000	15,000
1	$7,000,000	10,000	12,250
0	$1,000,000	2,000	400
Total	$13,000,000	19,000	27,650

b. (0.5 point) What exposure base do the authors use? Explain why.
c. (0.5 point) According to Hazam, what two conditions must be met to use the exposure base described in part (b)?

2.38. (9, 11/02, Q.47) (2 points)
a. (1.5 points)
Given the following data, calculate the credibilities for 1-year and 2-year claim free periods.

A represents 3 or more years since the most recent accident.
X represents 2 years since the most recent accident.
Y represents 1 year since the most recent accident.
B represents 0 years since the most recent accident.

	Earned Car Years	Earned Premium at Present Class B Rates	Number of Claims
A	50,000	$5,500,000	5,000
X	6,500	$682,500	1,000
Y	5,000	$535,000	850
B	4,500	$490,500	900
TOTAL	66,000	$7,208,000	7,750

b. (0.5 point)
Give two possible reasons that the 2-year credibility is less than 2 times the 1-year credibility.

***2.39.* (9, 11/03, Q.22)** (3 points) You are given the following data:

Class	Years since last accident	Actual Earned Premium at Present B Rates	Earned Car Years	Number of Claims
A	3+	375,000	2,500	200
X	2	15,000	100	12
Y	1	22,500	150	20
B	0	37,500	250	38

Assume that the same rate is charged to all insureds within a class and there have been no rate changes in or since the experience period.
a. (1 point) What is the credibility of 3 or more accident-free years of experience?
b. (1 point) What is the credibility of 1 or more accident-free years of experience?
c. (1 point) Give two possible reasons why the answer in part (a) is not 3 times the answer in part b.

2.40. (9, 11/04, Q.2) (1 point) Given the following information:

Class	Number of Years Since Most Recent Accident	Earned Car Years	Earned Premium at Present B Rates	Number of Claims
A	3 or more	10,000	$1,000,000	1000
X	2	7,000	$770,000	1155
Y	1	5,000	$625,000	1250
B	0	2,000	$400,000	1000
Total		24,500	$2,795,000	4405

Calculate the credibility of one or more accident-free years of experience.

2.41. (9, 11/05, Q.3) (3 points)
a. (2 points) Given the following information:
N = the number of drivers in the population
m = the mean claim frequency for all drivers
Mod = the credibility weighted modification factor for risks with one or more claims in the past year
Derive the formula for the credibility assigned to the experience of drivers with one or more claims in the past year. Assume that claim frequency follows a Poisson distribution.
b. (1 point) If there is a switch from a less refined class plan to a highly refined class plan, describe the likely change in the credibility assigned to an individual risk.

2.42. (9, 11/06, Q.2) (4 points)
a. (3 points) Given the following information about an automobile insurance portfolio:

Group	Number of Accident Free Years	Earned Premium at Present Group D Rates	Number of Claims Incurred
A	3 or more	$25,000,000	40,000
B	2	$8,000,000	15,000
C	1	$13,000,000	25,000
D	0	$8,000,000	30,000

Calculate the credibility of a single car for each of the following: one-year, two-year, and three-year accident-free periods.
b. (1 point) In performing the analysis in part (a) above, would using car years instead of earned premium as an exposure base be more preferable? Explain why or why not.

***2.43.* (9, 11/07, Q.2**) (3.5 points)
a. (2 points) The following data were compiled from the ABC automobile insurance portfolio:

Group	Number of Accident Free Years	Earned Premium at Present Group D Rates	Number of Claims Incurred
A	3 or more	$100,000,000	120,000
B	2	$10,000,000	25,000
C	1	$17,000,000	44,000
D	0	$10,000,000	36,000

Calculate the credibility of a single car for each of the following ranges of accident-free years:
I, 1 or more
ii. 2 or more
iii. 3 or more

b. (1 point)
The following table provides the single car credibility for the XYZ automobile insurance portfolio:

Accident-Free Years	Single Car Credibility
1 or More	0.06
2 or More	0.10
3 or More	0.14

Discuss two conclusions than can be drawn from the different credibility results of the ABC and XYZ portfolios.
c. (0.5 point) Explain why analysis of two portfolios with different classification plans could assign different values to the credibility of the experience of a single car.
Note: I have rewritten part (b) of this past exam question.

2.44. (9, 11/08, Q.5) (2 points)
A liability insurer collects the following data for a particular class of private passenger auto risks:

Accident-Free Years	Earned Exposures	Incurred Losses ($)
2 or more	2,500	1,000,000
1	500	500,000
0	1,000	2,500,000
Total	4,000	4,000,000

Assume the following:

- The base rate is $1,250 per exposure.
- An experience rating factor is the only factor applied to the base rate.

a. (1 point) Calculate the credibility of an exposure that is accident-free for 1 or more years.
b. (1 point) Calculate the premium for an exposure that is accident-free for 2 or more years.

2.45. (9, 11/09, Q.4) (3.5 points) The following information can be used to calculate the credibility assigned to the experience of a single private passenger car:

Group	Years Since Last Accident	Earned Car Years	Earned Premium at Present B Rates	Number of Claims
A	3 or more	650,000	400,000,000	50,000
X	2	230,000	150,000,000	20,000
Y	1	100,000	75,000,000	12,000
B	0	M	45,000,000	18,000
Total		980,000 + M	670,000,000	100,000

Assume claim counts follow a Poisson distribution.
a. (2.5 points) Calculate M, the earned car years for Group B, given that the credibility for an insured that has had no claim-free years is equal to 0.167.
b. (1 point)
Calculate the credibility for the group of risks that have been claim-free for two or more years.

2.46. (9, 11/10, Q.5) (1 point) An insurance company has a private passenger auto book of business with the following claims experience:

Group	Number of Accident Free Years	Earned Premium at Present Group D Rates	Number of Claims Incurred
A	3 or more	60,000,000	45,000
B	2	15,000,000	15,000
C	1	20,000,000	29,300
D	0	15,000,000	29,300
		100,000,000	108,000

Calculate the credibility of a single car for a driver with one or more accident-free years.

2.47. (8, 11/11, Q.1) (3 points) An insurance company is using a merit rating plan for drivers in two states. State X has the following claims experience:

Group	Number of Accident Free Years	Earned Premium at Present Group D Rates	Number of Claims Incurred
A	3 or more	$500,000	240
B	2	$150,000	125
C	1	$200,000	190
D	None	$300,000	300

State Y has the following relative claim frequencies for accident-free experience:

Number of Accident-Free Years	Relative Claim Frequencies to Total
3 or more	0.70
2 or more	0.77
1 or more	0.84

Assuming that no new risks enter or leave either state, use relative credibility to explain which state has more variation in an individual insured's probability of an accident.

2.48. (8, 11/12, Q.6) (2.5 points) An insurance company has a private passenger auto book of business with the following claims experience:

Territory	Years Since Last Accident	Earned Premium at Present Rates for Two Years Since Last Accident	Earned Car Years	Number of Claims	Incurred Loss
1	0	$15,000,000	15,000	5,000	$9,000,000
1	1	$125,000,000	125,000	41,000	$75,000,000
1	2+	$230,000,000	230,000	76,000	$138,000,000
2	0	$25,000,000	25,000	7,000	$16,000,000
2	1	$310,000,000	300,000	84,000	$187,000,000
2	2+	$550,000,000	535,000	147,000	$328,000,000
3	0	$10,000,000	10,000	4,000	$7,000,000
3	1	$80,000,000	100,000	35,000	$43,000,000
3	2+	$160,000,000	170,000	60,000	$100,000,000

Choose an appropriate exposure base for calculating credibility. Justify the selection.

2.49. (8, 11/14, Q.5) (2.5 points)
The following data shows the experience of a merit rating plan for a specific state.

Number of Accident Free Years	Earned Car Years	Earned Premium ($000)	Number of Incurred Claims
3 or More	250,000	250,000	1,200
2	300,000	100,000	625
1	25,000	100,000	750
0	12,000	150,000	1,500
Total	587,000	600,000	4,075

The base rate is $1,000 per exposure. No other rating variables are applicable.
a. (0.5 point) The typical exposure base used to develop the merit rating plan is earned premium. Briefly discuss two assumptions in selecting this exposure base.
b. (1.5 points) Calculate the ratio of credibility for an exposure with two or more years accident-free experience to one or more years accident-free experience.
c. (0.5 point) Calculate the premium for an exposure that is accident free for two or more years.

2.50. (8, 11/15, Q.1) (2.5 points)
An actuary is evaluating a merit rating plan for private passenger cars.
Given the following:

Number of Accident-Free Years	Earned Car Years	Number of Claims Incurred
2 or More	500,000	20,000
1	200,000	15,000
0	100,000	9,000
Total	800,000	44,000

- Frequency varies by territory.
- State law prohibits reflecting territory differences in rating.
- Annual claims for an individual driver follow a Poisson distribution.
- Claim cost distributions are similar across all drivers.

a. (0.5 point) Identify one potential issue with the exposure base used. Briefly explain whether or not earned premium would be a better choice for the exposure base.
b. (1.0 point) Calculate the credibility of one driver with one or more year's accident-free experience.
c. (1.0 point) Calculate the credibility of one driver with 0 Accident-Free years.

2.51. (8, 11/16, Q.1) (2.75 points) A group of insureds have different expected claim frequencies. The number of insureds claim-free for the past t years is as follows:

Expected Claim Frequency	t=0	t=1	t=2	t=3
0.05	50,000	47,500	45,000	44,000
0.10	50,000	45,000	43,000	36,000
0.20	25,000	20,500	16,500	14,000
Total	125,000	113,000	104,500	94,000

Determine whether the variation of an individual insured's chance for an accident changes over time.

2.52. (8, 11/17, Q.3) (1.5 points)
The following data shows the experience of a merit rating plan for private passenger vehicles. The merit rating plan uses multiple rating variables, including territory.

Number of Accident Free Years	Earned Car Years (000s)	Earned Premium ($000s)	Number of Incurred Claims
5 or More	250	500,000	15,000
3 and 4	100	90,000	13,500
1 and 2	80	90,000	8,000
0	70	90,000	10,500
Total	500	700,000	47,000

Territory	Frequency	Average Premium
A	0.05	1,500
B	0.10	2,000
C	0.15	1,250

a. (0.75 point) Recommend and justify an exposure base for this merit rating plan.
b. (0.75 point) Calculate the relative credibility of an exposure that has been three or more years accident-free using the exposure base from part (a) above.

2.53. (8, 11/18, Q.3) (2.75 points) An insurance company has a private passenger auto book of business with the following claims experience:

Group	Number of Accident-Free Years	Earned Premiums	Current Merit Rating Factor	Number of Claims Incurred
A	3 or more	216,000,000	0.60	25,000
X	2	135,000,000	0.75	18,000
Y	1	63,750,000	0.85	20,000
B	0	200,000,000	1.00	C
Total		614,750,000		63,000 + C

- Claim counts follow a Poisson distribution with parameter $\lambda = 0.05$.
- The credibility for the new policy period for an insured that has had no claim-free years is equal to 0.038.

a. (1.5 points) Calculate C, the number of claims incurred for Group B.
b. (0.75 point) Calculate the merit rating factor for an exposure that is accident-free for two or more years for the new policy period.
c. (0.5 point) Briefly explain two circumstances under which using earned premium as the exposure base would not correct for maldistribution.

2.54. (8, 11/19, Q.3) (1.75 points) An insurance company has a private passenger auto book of business with an experience modification factor in its rating plan.
Given the following:

- Annual claims for an individual driver follow a negative binomial distribution with r = 10.
- For the negative binomial distribution:
  - $f(x) = \binom{x+r-1}{x} (1-p)^r p^x$
  - $E[X] = \frac{pr}{1 - p}$
- The expected claim frequency for the entire book of business is 0.101.
- The credibility for the group of risks that have had at least one accident in the last year is 0.02.

a. (1.25 points) Calculate the experience modification factor for a policy that has had at least one accident in the last year.
b. (0.5 point) Describe why a class with a higher volume of claims and more exposures may have less credibility than a class with fewer claims and exposures.

Solutions:

2.1. Statement 1 is conclusion #1 of Bailey-Simon.
Statement 2 is backwards from conclusion #2 of Bailey-Simon.
Conclusion #3 of Bailey-Simon states that the credibility increases roughly by (only) two-fifths.
Only statement #1 is true.

2.2. The overall claim frequency on a premium basis is: 200,000/2000 = 100.
(a) Claim frequency on a premium basis for 5 or more years claim free: 134,200/1520 = 88.289.
1 - Z = 88.289 / 100. ⇒ Z = **11.7%**.
(b) Claim frequency on a premium basis for 4 or more years claim free:
(134,200 + 8900) / (1520 + 70) = 90.
1 - Z = 90 / 100. ⇒ Z = **10.0%**.
(c) Claim frequency on a premium basis for 3 or more years claim free:
(134,200 + 8900 + 10,400) / (1520 + 70 + 80) = 91.916.
1 - Z = 91.916 / 100. ⇒ Z = **8.1%**.
(d) Claim frequency on a premium basis for 2 or more years claim free:
(134,200 + 8900 + 10,400 + 12,500) / (1520 + 70 + 80 + 90) = 94.318.
1 - Z = 94.318 / 100. ⇒ Z = **5.7%**.
(e) Claim frequency on a premium basis for 1 or more years claim free:
(134,200 + 8900 + 10,400 + 12,500 + 14,400) / (1520 + 70 + 80 + 90 + 100) = 96.989.
1 - Z = 96.989 / 100. ⇒ Z = **3.0%**.
Comment: In part (b) those who have no claims in a 4 year window are:
those 4 years claims free plus those claims free for 5 or more years.
Different merit rating plans will have a different experience period.
Presumably this data was collected from a situation where the experience period was 5 years.
While all of these insureds are in the Adult Driver class, they may have different vehicle usage, different territories, etc. The premiums are prior to the impacts of any discounts for Merit Rating, and thus are analogous to the premium at current Group B rates in Bailey-Simon. Also they are loss costs premiums so as to get at the expected effects on frequency of the other rating factors without being distorted by fixed expenses. We could instead take premiums less expense fees.

2.3. The Merit Rating Plan results are not affected at all by severity. Claim frequency is only used in so far as the number of years claims-free. Using the total number of claims in the last three years for an individual driver would produce somewhat different results.
In the NCCI Experience Rating Plan, each claim is divided into primary and excess losses, with a split point of $5000. (Initially each loss is limited by the State Accident Limit.) The primary losses are more affected by frequency than severity, while the excess losses are more affected by severity than frequency. Primary credibilities are larger than excess credibilities. Therefore, the experience modification is more affected by frequency than severity. The ratio of excess credibility to primary credibility increases as the size of insured increases. Thus the modifications of larger insureds are more sensitive to severity than are those of smaller insureds.
Comment: Currently, some Merit Rating Plans (SDIPs) will reflect severity to a limited extent. For example, an at-fault claim of size more than $2000 might be assigned “4 points,” while an at-fault claim of size less than $2000 but at least $500 is only assigned “3 points.” (This is what is done in the Massachusetts SDIP.)
As one reduced the size of the split point in an experience rating plan, the primary losses became more and more like frequency; for a split point of $1, the primary losses would be the number claims. The optimal credibilities depend on the split point. Theoretically, the optimal split point depends on the size of the insured; smaller risks have a smaller optimal split point. Put another way, for smaller risks there is less predictive value to severity than for larger risks, all other things being equal. See for example, Howard C. Mahler’s discussion of “An Analysis of Experience Rating” by Glenn G. Meyers, PCAS 1987, not on the syllabus.
A single private passenger car generates the same amount of data as the very smallest Workers Compensation insureds. For example, a car might have expected annual liability losses of $1000.
A Workers Compensation insured with only $1000 in expected annual losses is too small to qualify for Experience Rating. (See the Table on page 16 of the plan. For example, a state might require $4000 in average annual premium, which correspond to about $3000 in expected annual losses.)
Such very small Workers Compensation insureds have too little data to fit into the Experience Rating Plan. However, predicting the future losses of such very small risks would benefit from a simple Merit Plan which just used frequency in a limited manner similar to the Canadian Merit Plan.

2.4. Let Y be the expected claim frequency for the average risk and Z be the one-year credibility for a single car. We have two equations:
Z x (0 claims) + (1 - Z) x (Y claims) = 0.05 claims.
Z x (1 claim) + (1 - Z) x (Y claims) = 0.12 claims.
Solving, Z = (0.12 - 0.05) / 1 = **7%**.
Comment: A similar idea to what Bailey and Simon do, but somewhat different.

2.5. Predictive accuracy is an important part of allowing any private insurance system to operate; it allows rates to be not unfairly discriminatory.
The driver characteristics we use all have been shown to have significant value in predicting the future losses of insureds. In addition, we use the past experience of an insured in our Safe Driver Insurance Plan in order to improve that prediction.
However, the past experience of a single private passenger car has a lot of randomness. Thus, there is not enough volume of data from a single private passenger car to by itself get an accurate prediction of future experience. (According to Bailey-Simon, the number of claims for one car over three years has about 6% to 10% credibility. The credibility would be higher in the absence of any class plan. Nevertheless, in the context of one car having 5 times or more the expected losses of another car in the same state, this is relatively small.) Past experience of a single private passenger car, including moving violations, is a useful supplement to a well designed, refined class plan. However, past experience can not replace the class plan.
Comment: Whether certain risk characteristics are socially acceptable is a matter of opinion, and not something actuarial. While controllability is desirable it is not necessary for a classification variable.

2.6. E. Low rated territories will have lower expected frequencies and thus more insureds who are claims-free for 3 or more years. Using premium based relative frequencies adjusts for this approximately.

2.7. \$470 = (0)Z + (1-Z)(\$500). Z = 1 - 470/500 = **6%**.
Comment: A similar idea to what Bailey and Simon do, but somewhat different.

2.8. In both methods one divides the data for a single class into Merit Rating Groups based on how many years the car has been claims-free.
In the first method, one compares the subsequent premium based frequency for those who are claims free for at least one year to the overall. This ratio is 1 - Z.
The second method instead compares those who are not claims free (Group B) to average.
One gets a modification M for Group B as in the first method.
Using a Poisson assumption and the average exposures based frequency for the class, one determines the average exposure based frequency for those in Group B for the experience period.
Using this together with M, one can back out a credibility for Group B:

$$Z = \frac{M - 1}{\text{(Group B experience period frequency relative to average)} - 1}.$$

Note that the first method uses neither exposures nor an assumption of the distributional form of the frequency.
Comment: The first method is also applied to estimate two and three year credibilities.
M = relative premium based frequency for risks with one or more claims in the past year.
Let λ = the mean claim frequency (per exposure) for the class.
Group B has a frequency relative to average within its class of: $1 / (1 - e^{-\lambda})$.

2.9. D. The claims-free group contains more insureds from low rated-territories, which makes their future exposure based frequency better than it otherwise would be; the estimated credibilities from comparing to average are too big. The not claims-free group contains more insureds from high
rated-territories, which makes their future exposure based frequency worse than it otherwise would be; the estimated credibilities are again too big.

2.10. The volume of data is the same in each case; (3)(4%) = 12%. However, shifting risk parameters make the more distant years of data less valuable for predicting the future. Therefore, I would expect the one year for a driver from Class 2 to have more credibility than three years of data for a driver from Class 1.

2.11. a. The overall premium based frequency is: 27,000 / 300 = 90.
The premium based frequency for those claims-free for 1 or more years (A + X + Y) is:
(18,200 + 1400 + 2200) / (225 + 15 + 20) = 83.85.
1 - Z = 83.85 / 90. ⇒ Z = **6.8%**.
b. The premium based frequency for those claims-free for 0 years (B) is: 5200 / 40 = 130.
Thus the modification for Group B is:
Future Relative Claim Frequency = 130 / 90 = 1.444.
Overall frequency per exposure is: 27,000 / 240,000 = 0.1125.
Given the Poisson assumption, the relative observed frequency for those who had at least one claim is: $1 / (1 - e^{-\lambda}) = 1 / (1 - e^{-0.1125}) = 9.398$.
Thus we must have: 1.444 = Z 9.398 + (1 - Z) 1.
⇒ Z = (1.444 - 1) / (9.398 - 1) = **5.3%**.
c. As always with finite data sets we have random fluctuation.
In addition, each technique makes assumptions and approximations. The premium based frequencies only approximately adjust for the maldistribution of the Groups by territory. In part (b) we had to make use of a Poisson assumption.
However, more fundamentally, we are measuring two somewhat different things. In part (a) we are attempting to back out the weight that would have done best in predicting the future experience of those insureds who had no claims this year (A + X + Y). In part (b), we are attempting to back out the weight that would have done best in predicting the future experience of those insureds who had at least one claims this year (B).
The Bayes Analysis estimates for different groups, those with 0 claims, those with 1 claim, those with 2 claims, etc. usually do not lie upon a straight line. (Only in special cases such as the Gamma-Poisson, are the Bayes estimates along a straight line, and thus Buhlmann Credibility equals Bayes Analysis.) Thus the optimal weight to use in each of these situations would be different.
Comment: The Buhlmann credibility is the slope of the weighted least squares line fit to the Bayes Estimates as function of the observations. Thus we would expect the estimates in parts (a) and (b) to differ from each other as well as the Buhlmann credibility.

2.12. The most heterogeneous class would have the highest credibility for a car. I would expect this to be the "all other" class, since it contains many different types of drivers with different potentials for loss.
Comment: The less homogeneous a class is, the more rely on the experience of an individual car within that class.

2.13. a. For any driver, claims may be independent.
However, the correlation compares claim rates of different drivers.
Some drivers have high claim frequency, with high expected rates in both years.
Some drivers have low claim frequency, with low expected rates in both years.
A high correlation implies that:
• Drivers are heterogeneous with stable risk parameters.
• Good drivers usually stay good for a second year; bad drivers usually stay bad for a second year.
A high correlation from year to year does not mean claims are not independent.
b. The actuary should examine the serial correlation in each driver's claim history.
If a given driver has a claim in Year X but not in Year Y, does claim frequency tend to be greater in Year X+1 than in Year Y+1?
Under independence, the answer should be no.
c. Merit rating is applied after class rating.
The 10% correlation implies that class rating plus merit rating has a 10% credibility.
In personal auto, class and territory rating separates drivers into relatively homogeneous classes.
Class and territory rating gets much of the credibility, leaving much less than 10% for merit rating.
d. The actuary should examine the correlation within classes.
Bailey-Simon examines a single class at a time; compare the future (premium based) frequency of those who were claims-free to that of the overall class.

2.14. Bailey-Simon uses $\frac{\text{Number of Car Years}}{\text{Dollars of Earned Premiums}}$, in order to adjust for the maldistribution that would result from low frequency territories having a larger portion of insureds who are claims-free.
It would be better to use premiums, provided the high rated territories have higher frequency and provided the territory relativities are correct.

Territory	Average Rate	Relative to Average	Frequency per Car-Year	Relative to Average
A	$750	0.769	4.00%	0.899
B	$893	0.916	4.46%	1.002
C	$1000	1.026	4.33%	0.973
D	$1086	1.114	4.78%	1.074
E	$1176	1.206	4.70%	1.056
Total	$975	1.000	4.45%	1.000

There is a tendency for the higher rated territories to have higher frequencies.
However, the relative average rates have a much wider spread than the relative average frequencies. Thus the average premiums largely reflect differences in severity and/or reflect incorrect territory relativities in the current rates.
Using for each subgroup (0 years claims-free, 1 year claims-free, 2 years claims free, etc.)
$\frac{\text{Number of Claims}}{\text{Dollars of Earned Premiums}}$ would adjust for the differences in frequency by territory, but would significantly over-adjust due to whatever is causing the wider differences in average premium.
Using $\frac{\text{Number of Claims}}{\text{Number of Earned Car Years}}$ would <u>not</u> adjust for the differences in frequency by territory.
In this case, the other reasons for differences in average premiums seem to have a bigger effect than differences in frequency. Thus on balance I would prefer to use
$\frac{\text{Number of Claims}}{\text{Number of Earned Car Years}}$ rather than $\frac{\text{Number of Claims}}{\text{Dollars of Earned Premiums}}$. We want to adjust for the different mixes of territory for the subgroups, due to the different frequencies by territory. If possible, it would probably be better to use for each subgroup (0 years claims-free, 1 year claims-free, 2 years claims free, etc.):

$$\frac{\text{Number of Claims}}{\sum_{\text{territories}} \text{(car years for subgroup in territory) (frequency within territory relative to whole class)}}.$$

The relative frequencies for the territories within Class 1 are: 0.899, 1.002, 0.973, 1.074, 1.056.
Assume that the subgroup that is claim free for at least 3 years has exposures within Class 1 by territory: 17,700, 24,500, 26,300, 19,900, 14,800. Then the above denominator would be:
(0.889)(17,700) + (1.002)(24,500) + (0.973)(26,300) + (1.074)(19,900) + (1.056)(14,800) = 102,876. This is less than the sum of exposures for this subgroup of 103,200, reflecting the somewhat higher proportion of low frequency territories in this subgroup than in all of Class 1.
<u>Comment</u>: Similar to 8, 11/12, Q.6.

2.15. For each class, we get the frequency per exposure by multiplying the frequency per $ premium times the premium per exposure.
For example, for Class 1: (0.000263)(300) = 7.89%.
Then take the ratio of the 3-year credibility to this frequency, as per Table 2 in Bailey-Simon.
For example, for Class 1: 5.8% / 7.890% = 0.7351.

Class	Cred.	Class Freq. per Prem.	Prem. per Expos.	Freq. per Expos.	Cred. / Freq.
1	5.8%	0.000263	300	7.890%	**0.7351**
2	9.3%	0.000369	400	14.760%	**0.6301**
3	8.1%	0.000311	350	10.885%	**0.7441**

A more homogeneous class will have a ratio of credibility for experience rating to frequency that is lower.
Thus Class 2 is more homogeneous than Classes 1 and 3;
Class 2 exhibits less variation of individual hazards than do the others.
Comment: Similar to 9, 11/95, Q.32.

2.16. (a) Assume each territory has x exposures.
Then total number of claims is: x 10% + x 8% = x 18%.
Total premium is: 400 x + 500 x = 900 x.
Overall premium based frequency is: 18% / 900 = 0.0002.
Claims-free premium based frequency is: 17,100 / 91 million = 0.000188.
Z = 1 - 0.000188 / 0.0002 = **6.0%**.
(b) Overall frequency is: (1/2)(10%) + (1/2)(8%) = 9%.
Claim free frequency is: (9000 + 8100) / (100,000 + 110,000) = 8.14%.
Z = 1 - 8.14% / 9% = **9.6%**.
(c) Overall frequency Territory 1 is: 10% / 400 = 0.00025.
Claims-free frequency Territory 1 is: 9000 / 38 million = 0.000237.
Territory 1 credibility is: 1 - 0.000237 / 0.00025 = **5.2%**.
Overall frequency Territory 2 is: 8% / 500 = 0.00016.
Claims-free frequency Territory 2 is: 8100 / 53 million = 0.000153.
Territory 2 credibility is: 1 - 0.000153 / 0.00016 = **4.4%**.
(d) Overall frequency Territory 1 is 10%.
Claim free frequency Territory 1 is: 9000 / 100,000 = 9%.
Territory 1 credibility is: 1 - 9%/10% = **10%**.
Overall frequency Territory 2 is 8%.
Claim free frequency Territory 2 is: 8100 / 110,000 = 7.36%.
Territory 2 credibility is: 1 - 7.36%/8% = **8.0%**.
(e) The pure premiums are: (10%)(2400) = 240, and (8%)(3750) = 300.
The ratio of pure premiums to average premium are: 240/400 = 60%, and 300/500 = 60%.
Thus the territory relativities appear to be correct.
However, the <u>higher</u> rated territory has the <u>lower</u> frequency.
Thus in part (a), using premiums in the denominator is not a good idea; it would not be adjusting for the differences in frequency between the territories.
Therefore, the result in part (b) is preferable to that in part (a).
There are difference in the classes within a territory in average premiums and frequencies.
If the higher rated classes are also higher frequency, then the results in part (c) using premiums in the denominator would be preferable to those in part (d) using car years in the denominator.
It makes sense that three years of data from the higher frequency territory 1 would have a larger credibility than three years of data from the lower frequency territory 2.
(However, in practical applications of a Safe Driver Insurance Plan, one would probably give the same credibility to a car year of data from any class and territory.)
<u>Comment</u>: The data in this question is not arranged in exactly the same way as in Bailey-Simon.
I do not have an opinion as to whether the results in part (b) or part (c) are preferable;
I would need to investigate further as to why they differ.
There are probably other reasonable answers to part (e).

2.17. (a) Bailey & Simon give 3 reasons why the credibilities increase less than linearly with number of year of data. The question has eliminated two of these reasons; the one that is left is shifting risk parameters. The faster parameters shift over time, the greater the effect of lowering the ratio of 3-year to 1-year credibility.
The ratios of three year to one year credibilities are for the given classes: 2, 1.9, and 2.1.
Thus Class 2 has been most affected by shifting risk parameters over time and Class 3 the least. Thus, the insureds in Class **3** have more stable expected claim frequencies from year to year.
(b) Less variation in individual hazard within its class is a smaller Variance of the Hypothetical Means. Such a class would have a smaller credibility all else being equal. However, a higher mean frequency would also produce a higher credibility, all else being equal.
Compare the one-year credibility to the mean frequency, the ratios are: 0.7, 1.1, and 0.9.
Thus Class **1** has less variability in expected claim frequency within its class.
As per Table 2 in Bailey-Simon, comparing the three-year credibility to the mean frequency, the ratios are: 1.4, 2.1, and 1.9. (I would prefer to use the one-year credibilities, which are less affected by shifting risk parameters.) A lower ratio indicates that lower relative credibility is assigned, meaning a more homogeneous class.
Thus Class **1** has less variability in expected claim frequency within its class.
<u>Comment</u>: Similar to 9, 11/96, Q.50. Part b is similar to 8, 11/11, Q.1.
See Table 2 in Bailey-Simon.

2.18. The indicated rate compared to average for those who are one or more years claims free is: $\frac{(12000 + 1200 + 1400) / (2400 + 200 + 220)}{17{,}200/3200} = 5.1773 / 5.375 = 0.9632$.

The indicated rate compared to average for those who are not claims free is:
$\frac{2600/380}{17{,}200/3200} = 6.8421 / 5.375 = 1.2729$.

Thus the appropriate premium for an exposure that is accident free for one or more years is:
(0.9632/1.2729) ($800) = **$605.36**.
Alternately, (5.1773/6.8421) ($800) = **$605.35**.
<u>Comment</u>: Similar to 8, 11/14, Q.5.

2.19. (a) Looking at the experience of those insureds who were accident free for 0 years:
$M = Z / (1 - e^{-\lambda}) + (1 - Z)(1) = (0.044)/(1 - e^{-0.08}) + (1 - 0.044) = 1.528$.
We need to adjust premiums to the Group B rate:
1040/0.65 = 1600. 64/0.8 = 80. 90/0.9 = 100.
1600 + 80 + 100 + 170 = 1950 million.
M is ratio of the premium based frequency for those in Group B to that overall.

$$1.528 = M = \frac{C / 170}{(150{,}000 + C) / 1950}.$$ ⇒ 229,200 + 1.528 C = 11.471 C. ⇒ C = **23,051**.

Alternately, $Z = (M - 1) / \{1/(1 - e^{-\lambda}) - 1\} = (M - 1)(e^{\lambda} - 1)$.
⇒ $0.044 = (M - 1) (e^{0.08} - 1)$. ⇒ M = 1.528. Proceed as before.

(b) The premium based frequency for Group B is: 23,051/170 = 135.594.
The premium based frequency for Group Y is: 12,000 / 100 = 120.
The indicated merit rating credit (with respect to Group B) for Group Y is: 120/135.594 = **0.884**.
The premium based frequency for Group X is: 8000 / 80 = 100.
The indicated merit rating credit (with respect to Group B) for Group X is: 100/135.594 = **0.737**.
The premium based frequency for Group A is: 130,000 / 1600 = 81.25.
The indicated merit rating credit for Group A is: 81.25/135.594 = **0.599**.
Comment: Similar to 8, 11/18, Q.3.

2.20. In experience rating, the credibility of an individual insured depends on: the volume of data for the insured, the expected value of the process variance, and the variance of the hypothetical means. More credibility is associated with: a higher volume of data, a smaller expected value of the process variance, and a larger variance of the hypothetical means between the individual insureds within the class.
If there is a larger variation between the individual insureds within a class, in other words if the class is heterogeneous, then the variance of the hypothetical means is large, and the experience rating credibility will be higher. If the first insured is in a relatively homogeneous class while the second insured is in a relatively heterogeneous class, then the experience of the first insured may be given less credibility even though it has a larger volume of data than the second insured.

2.21. Let y = mean frequency of those who have had at least one claim in the last year.
overall mean = 0 f(0) + y {1 - f(0)}. ⇒ y = (overall mean) / {1 - f(0)}.

Let R = the ratio of the actual losses to the expected losses.
Then R = 1/{1 - f(0)}. Then the mod is: Z R + 1 - Z = (4%)(R) + (1 - 4%).
(a) λ = 0.06. ⇒ R = $1/(1 - e^{-0.06})$ = 17.172. ⇒ M = (0.04)(17.172) + 1 - 0.04 = **1.647**.

(b) 0.06 = 3p/(1-p). ⇒ p = 0.02/1.02 = 0.01961. ⇒ f(0) = $(1 - p)^r = (1 - 0.01961)^3$ = 0.9423.

R = 1/(1 - 0.9423) = 17.338. ⇒ M = (0.04)(17.338) + 1 - 0.04 = **1.654**.

<u>Comment</u>: Similar to 8, 11/19, Q.3a.
Instead in the notation in <u>Loss Models</u>, β = 0.06/3 = 0.02.
f(0) = $1/(1+\beta)^r = 1/1.02^3$ = 0.9423.
In this case, there is not much difference between the results of part (a) and (b).
This is due to the fact that β is small. (As β approaches zero the Negative Binomial approaches a Poisson.)
Given the two different models used in parts (a) and (b), the appropriate credibility should differ somewhat between parts (a) and (b).

2.22. The decrease in frequency in 2020 will result in future years in more credits and fewer surcharges under the merit rating plan. (If the plan uses three years of data, insureds will benefit on average during 2021, 2022, and 2023.) However, the credibility assigned to one year of private passenger automobile data is small, for example 5%. Thus, only a small fraction of the reduction in frequency will find its way into the reduction in premiums for the insureds.
<u>Comment</u>: Merit Rating Plans are not designed to deal with unusual situations such as the COVID-19 pandemic. Mycorona could give its insureds a special dividend or premium reductions/rebates. The COVID-19 pandemic is an example of a contingency: an unexpected type of event that significantly affects insurance losses. In this case, the contingency reduces expected losses for this line of insurance. More commonly, contingencies increase the expected losses.
The question does not mention severity, but the pandemic could have for example raised the average severity somewhat.
See “Considerations for Handling Auto Insurance Data in the Era of COVID-19,”
Issue Brief from the American Academy of Actuaries, March 2021.

2.23. (a) "Earned premiums are converted to a common rate basis by use of the relationship in the rate structure that A: X : Y :B = 65 : 80 : 90 : 100.
The authors have chosen to calculate Relative Claim Frequency on the basis of premium rather than car years. This avoids the maldistribution created by having higher claim frequency territories produce more X, Y, and B risks and also produce higher territorial premiums."
In other words, Bailey and Simon were concerned about the inherent correlation of exposures between Merit Rating Groups and territories.
We would expect that Group B (not claims free) would have a larger percentage of exposures in territories with higher than average frequencies than would Group A (claims-free for at least three years). However, we are already charging insureds in those territories more than average. If we did not adjust for that here by dividing by premiums rather than exposures, we would be double counting. This adjustment removes the impact of things that are already included in the rate structure via territory relativities.
(b) We are assuming that the territory relativities underlying the current rates are a reasonably accurate reflection of differences in frequency between territories. We are assuming that little if any of the difference in territory rates are due to differences in average severity. Similarly, we are assuming that the effect of any other classification factors other than Merit Rating that underlay the current rates accurately reflect differences in frequency and do not reflect differences in severity.

2.24. The overall frequency is: 12,500 / 100,000 = 0.125.
The frequency for those who are claims-free for at least a year is: 10,800 / 90,000 = 0.120.
Their relative frequency is: 0.120 / 0.125 = 0.96.
$1 - Z = 0.96. \Rightarrow Z =$ **4.0%**.
Alternately, the subsequent frequency for those who are not claims-free is: 1700/10,000 = 0.17.
Assuming a Poisson frequency, the average number of claims for those who were not claims-free is: $\lambda / (1 - e^{-\lambda}) = 0.125 / (1 - e^{-0.125}) = 1.0638$.
$Z\ 1.0638 + (1 - Z)(0.125) = 0.170. \Rightarrow Z =$ **4.8%.**
Comment: Bailey-Simon uses premium based frequency. The first method is the intended solution.
Let, M = relative premium based frequency for risks with one or more claims in the past year.
Then, $Z = (M - 1)\ (e^{\lambda} - 1) = (0.17/0.125 - 1)\ (e^{0.125} - 1) = 4.8\%$.

2.25. 1. Loss ratios have premiums rather than exposures in the denominator. Premiums reflect class and territory differentials, which could account for most of the variance between the loss potential of individual insureds.
2. Severity is systematically opposite to frequency.
Comment: in the second response, we could for example have a model with two types:

Type	Mean Frequency	Mean Severity	Mean Pure Premium
1	5%	\$10,000	\$500
2	10%	\$5,000	\$500

2.26. Overall claim frequency: 349,000 / 3,667,000 = 0.0952.
Assuming Poisson, the average number of claims for Group B is:
$\lambda / (1 - e^{-\lambda}) = 0.0952 / (1 - e^{-0.0952}) = 1.0484$.
Relative frequency for Group B is: 1.0484 / 0.0952 = 11.01.
The overall premium based frequency is: 349,000 / 242,300 = 1.440.
The premium based frequency for Group B is: 46,000 / 23,000 = 2.
⇒ Modification for Group B is: 2/1.440 = 1.389.
Thus, 1.389 = Z 11.01 + (1-Z) 1. ⇒ Z = **3.9%**.

2.27. Take β equal to the overall mean of 0.0952.
The probability of no claims is: $1/(1+\beta) = 1/1.0952 = 0.9131$.
Let the average number of claims for Group B be x.
(0)(0.9131) + x(1 - 0.9131) = 0.0952. ⇒ x = 1.0955.
Relative frequency for Group B is: 1.0955 / 0.0952 = 11.51.
⇒ Modification for Group B is 1.389.
Thus, 1.389 = Z 11.51 + (1-Z) 1. ⇒ Z = **3.7%**.
Comment: The credibility depends only slightly on the Poisson versus Geometric assumption.

2.28. All three statements are true.

2.29. a) The overall premium based frequency is: 100,000/ 600,000 = 1/6.
The premium based frequency for those claims-free for 2 or more years (A+B) is:
(54,250 + 21,000) / (390,000 + 120,000) = 0.1475.
1 - Z = 0.1475 / (1/6). ⇒ Z = **11.5%**.
b) The premium based frequency for those claims-free for 0 years (D) is:
14,625 / 45,000 = 0.325.
Thus the modification for Group D is:
Future Relative Claim Frequency = (0.325) / (1/6) = 1.95.
Overall frequency per exposure is: 100,000 / 1,000,000 = 0.1.
Given the Poisson assumption, the relative observed frequency for those who had at least one claim is: $1 / (1 - e^{-\lambda}) = 1 / (1 - e^{-0.1}) = 10.51$.
Thus we must have: 1.95 = Z 10.51 + (1 - Z) 1.
⇒ Z = (1.95 - 1) / (10.51 - 1) = **10.0%**.
Comment: In part B, $Z = \frac{\text{(Future Relative Frequency) - 1}}{\text{(Past Relative Frequency) - 1}}$.

2.30. We are not given the average premium for each class.
I will estimate that the average premium for each class is approximately such that:
(average premium for class) (1 - Z) = (average premium for 3-years claims free and in class).
Thus the average premium for Class A is: 150 / (1 - 0.082) = 163.40.
For each class, we get the frequency per exposure by multiplying the frequency per $ premium times the premium per exposure.
For example, for Class A: (0.001625)(163.4) = 26.55%.
Then take the ratio of the 3-year credibility to this frequency, as per Table 2 in Bailey-Simon.
For example, for Class A: 8.2% / 26.55% = 0.3088.

Class	Z	Class Freq. per Prem.	Claims-Free Prem. per Expo.	Class Prem. per Expo.	Freq. per Expos.	Z / Freq.
A	8.2%	0.001625	$150	$163.40	26.55%	**0.3088**
B	4.6%	0.001750	$148	$155.14	27.15%	**0.1694**
C	7.9%	0.002212	$190	$206.30	45.63%	**0.1731**

A more homogeneous class will have a ratio of credibility for experience rating to frequency that is lower.
Thus Class A is more heterogeneous than Classes B and C;
Class A exhibits more variation of individual hazards than do the others.

2.31. (a) Bailey & Simon give 3 reasons why the credibilities increase less than linearly with number of year of data. The question has eliminated two of these reasons; the one that is left is shifting risk parameters. The faster parameters shift over time, the greater the effect of lowering the ratio of 3-year to 1-year credibility.
The ratios of three year to one year credibilities are for the given classes: 2, 2.75, and 3.
Thus Class A has been most affected by shifting risk parameters over time and Class C the least. Thus, assuming that the exam question meant in which class do the insureds have more stable <u>expected</u> claim frequencies from year to year, that is Class **C**.
(b) Less variation in individual hazard within its class is a smaller Variance of the Hypothetical Means. Such a class would have a smaller credibility all else being equal. However, a higher mean frequency would also produce a higher credibility, all else being equal.
Compare the one-year credibility to the mean frequency, the ratios are: 1.5, 0.8, and 2.
Thus Class **B** has less variability in claim frequency within its class.
As per Table 2 in Bailey-Simon, comparing the three-year credibility to the mean frequency, the ratios are: 3, 2.2, and 6. (I would prefer to use the one-year credibilities, which are less affected by shifting risk parameters.) A lower ratio indicates that lower relative credibility is assigned, meaning a more homogeneous class. Thus Class **B** has less variability in claim frequency within its class.
Alternately, assume the one-year credibility is 1/(1+K). ⇒ K = 1/Z - 1.
Also assume that the Expected Value of the Process Variance is equal to the mean.
(EPV = mean if each insured has a Poisson frequency.
For comparison purposes we need only assume it is proportional.)

Class	1995 Claim Frequency	1995 One-year Credibility	K = EPV/VHM	VHM
A	0.12	0.18	1/0.18 - 1 = 4.56	0.12/4.56 = 0.0263
B	0.10	0.08	1/0.08 - 1 = 11.5	0.10/11.5 = 0.0087
C	0.08	0.16	1/0.16 - 1 = 5.25	0.08/5.25 = 0.0152

Class B has smallest ratio of VHM / (mean freq.)^2. Thus Class **B** has less variability in claim frequency within its class, as measured by the square of the coefficient of variation.
<u>Comment</u>: Part b would have been better if it had been worded: “Which class has less variation in expected claim frequency between individual risks within its class?”

2.32. Statement #1 is true. We would expect that Class B (not claims free) would have a larger percentage of exposures in territories with higher than average frequencies than would Class A (claims-free for at least three years). To avoid double counting effects that are already reflected in the territory relativities, we divide by premiums at base class rates; a class with a higher than the average percentage of exposures in a high frequency territory will also have a higher than average base class premium.
2. Statement #2 is false. We would also be interested in the homogeneity of a class. To the extent that the insureds in a class are more similar, the credibility for experience rating (individual risk rating) is smaller.
3. Statement #3 is true. See conclusion #1 of the paper.
Comment: As discussed on a preliminary exam, for Buhlmann credibility we would be interested in the EPV, VHM, and volume of data. In this context, the variance of hypothetical means measures how different the insureds are within a class, the expected value of the process variance measures how much random fluctuation there is in the data, and the volume of data is the number of years from an individual car.

2.33. a) Let x be the number of claims for Group C.
The frequency on a premium basis for one or more claim free years is:

$$\frac{62{,}376 + 15{,}955 + x}{420{,}000 + 105{,}000 + 60{,}000} = \frac{78{,}333 + x}{585{,}000}.$$

The overall frequency on a premium basis is: 98,000 / 600,000.

We have: $1 - Z = 1 - 0.086 = M = \frac{\text{frequency for at least one year claims free}}{\text{overall frequency}}$. ⇒

$0.914 = \frac{78{,}333 + x}{585{,}000}$ / (98,000 / 600,000). ⇒ x = **9000**.

b) "The experience for one car for one year has significant and measurable credibility for experience rating."
c) "In a highly refined private passenger rating classification system which reflects inherent hazard, there would not be much accuracy in an individual risk merit rating plan, but where a wide range of hazard is encompassed within a classification, credibility is much larger."
If the class system is highly refined and each class is homogeneous (not much variation in hazard), then the majority of the credibility (weight) should be assigned to the class experience rather than the individual risk experience.
Comment: See the first and second conclusions of the paper.

2.34. All three statements are true.
"The fact that the relative credibilities in Table 3 for two and three years are much less than 2.00 and 3.00 is partially caused by risks entering and leaving the class. But it can be fully accounted for only if an individual insured's chance for an accident changes from time to time within a year and from one year to the next, or if the risk distribution of individual insureds has a marked skewness reflecting varying degrees of accident proneness."

2.35. a. For three years: 1 - Z = 0.920. ⇒ Z = **8.0%**.
For two or more years claim free, claim frequency is: (31,964 + 2695) / (25,846 + 1783) = 1.254.
1 - Z = 1.254 / 1.345. ⇒ Z = **6.8%**.
For one or more years claim free (A + X + Y), claim frequency is:
(31,964 + 2695 + 3546) / (25,846 + 1783 + 2281) = 1.277.
1 - Z = 1.277 / 1.345. ⇒ Z = **5.1%**.
b. If the chance of accident for an individual risk remains constant and no risks enter or leave, then the credibility should vary approximately in proportion to the number of experience years.
c. Comparing the credibilities for one year and two years: 6.8/5.1 = 1.33 ≠ 2.
Comparing the credibilities for two years and three years: 8.0/6.8 = 1.18 ≠ 1.5.
The credibilities do <u>not</u> follow the expected pattern. An individual insured's chance for an accident changes over time and/or risks may be entering or leaving.
<u>Comment</u>: Conclusion #3 of the paper: "If we are given one year's experience and add a second year we increase the credibility roughly two-fifths. Given two years' experience, a third year will increase the credibility by one-sixth of its two-year value."
In part (a) we could use the alternate method to get a one year credibility.
The premium based frequency for those claims-free for 0 years is given as 1.362.
Overall frequency per exposure is: 45,770 / 321,327 = 0.1424.
Given the Poisson assumption, the relative observed frequency for those who had at least one claim is: $1 / (1 - e^{-\lambda}) = 1 / (1 - e^{-0.1424}) = 7.534$.
Thus we must have: 1.362 = Z 7.534 + (1 - Z) 1.
⇒ Z = (1.362 - 1) / (7.534 - 1) = **5.5%**.
A somewhat different answer than using the other method.

2.36. D. Statement #1 is conclusion #1 from the paper and thus true.
Statement #2 is true. See page 160 of the paper: "This also illustrates that credibility for experience rating depends not only on the volume of data in the experience period but also on the amount of variation of individual hazards within the class."
Statement #3 is conclusion #2 from the paper and thus true.
While Statement #4 could be true, as per the square root rule from Classical Credibility, this is not what Bailey & Simon find for their particular data.
<u>Comment</u>: Conclusion #3 of the paper: "If we are given one year's experience and add a second year we increase the credibility roughly two-fifths. Given two years' experience, a third year will increase the credibility by one-sixth of its two-year value."
If it followed the square root rule, then the ratio of the credibilities for 3 years and 2 years would be: $\sqrt{3/2}$ = 1.225 rather than 7/6 = 1.167.

2.37. a) total frequency is: 19,000/13,000 = 1.462.
frequency for those 1 or more claim free is: (10,000 + 7,000) / (7,000 + 5,000) = 1.417.
⇒ One year credibility is: 1 - 1.417/1.462 = **3.1%**.
frequency for those 2 or more claim free is: 7,000/5,000 = 1.4.
⇒ Two year credibility is: 1 - 1.4/1.462 = **4.2%**.
b) The authors use earned premium as their exposure base to avoid the maldistribution caused when lower frequency territories produce a larger percentage of risks that are claims-free than higher frequency territories.
c) 1. Higher-frequency territories must also be higher-premium territories.
2. The territorial differentials in the current rates must be proper.
Comment: The given premiums should be prior to the effects of Merit Rating.

2.38. (a) Overall the claim frequency on a premium basis is: 7750 / 7208 = 1.0752.
For two or more years claim free (A + X), claim frequency is:
(5000 + 1000) / (5500 + 682.5) = 0.9705.
1 - Z = 0.9705 / 1.0752. ⇒ Z = **9.7%**.
For one or more years claim free (A + X + Y), claim frequency is:
(5000 + 1000 + 850) / (5500 + 682.5 + 535) = 1.0197.
1 - Z = 1.0197 / 1.0752. ⇒ Z = **5.2%**.
(b) 1. Individual insured’s chance for an accident changes from time to time within a year or from one year to the next.
2. Insureds are entering or leaving the class.
3. Individuals’ accident propensities in a class vary and are markedly skewed.
4. The Buhlmann Credibility formula is less than linear.

2.39. (a) Overall the claim frequency is:
(200 + 12 + 20 + 38) / (2500 + 100 + 150 + 250) = 0.09.
For three or more years claim free (A), claim frequency is: 200/2500 = 0.08.
1 - Z = 0.08 / 0.09. ⇒ Z = **11.1%**.
For one or more years claim free (A + X + Y), claim frequency is:
(200 + 12 + 20) / (2500 + 100 + 150) = 0.08436.
1 - Z = 0.08436 / 0.09. ⇒ Z = **6.3%**.
(c) 1. Individual insured’s chance for an accident changes from time to time within a year or from one year to the next.
2. Insureds are entering or leaving the class.
3. The risk distribution of individual insureds has a marked skewness reflecting varying degrees of accident proneness.
Comment: The volume of data in this question is way less than used by Bailey-Simon.
If every insured within a class is charged the same rate, then we can use the usual exposure based frequencies rather than the premium based frequencies used by Bailey-Simon. It makes no difference in the result, since consistent with the statement that every insured within a class is charged the same rate, each of the premiums at current class B rates are 150 times the exposures.

2.40. Overall the claim frequency on a premium basis is: 4405 / 2795 = 1.5760.
For one or more years claim free (A + X + Y), claim frequency is:
(1000 + 1155 + 1250) / (1000 + 770 + 625) = 1.4217.
1 - Z = 1.4217 / 1.5760. ⇒ Z = **9.8%.**

2.41. (a) Let Group B be those drivers with at least one claim last year.
Let x be the average number of claims per insured for Group B.
For a Poisson with mean μ, $f(0) = e^{-\mu}$.
Therefore, $\mu = (0)(e^{-\mu}) + (x)(1 - e^{-\mu})$. ⇒ $x = \mu / (1 - e^{-\mu})$.
Thus the relativity of Group B compared to average is: $x/m = 1/ (1 - e^{-\mu})$.
Then we have that the credibility weighted modification factor for risks in Group B is:
$M = Z / (1 - e^{-\mu}) + (1 - Z)(1)$.

$$\Rightarrow Z = \frac{M - 1}{1/(1 - e^{-\mu}) - 1} = \mathbf{(M - 1)\ (e^{\mu} - 1)}.$$

(b) Credibility for an individual risk is lowered when the class plan is highly refined, because it is more difficult to identify differences in the loss potential for a particular risk from the average risk in the class. In other words, the Variance of Hypothetical Means within a class is less, so that the Buhlmann Credibility Parameter K is larger, and Z is less.
Put another way, if the class plan is more refined, it does a better job of estimating the expected pure premium, and there is less need to rely upon the experience of an individual insured. The relative value of the information in the data from the individual has declined, and Z is less.
If a class plan were to get so refined that each class was homogeneous, in other words if every insured in the class had the same expected pure premium, there would be no need for merit rating (experience rating) and Z for the experience of the individual would be zero.
Comment: Part (a) tests the alternate technique at page 160 of Bailey-Simon based on looking at the relativity for those with at least one claim. For Class 1 in Bailey-Simon, from their Table 1, μ = 288,019 / 3,325,714 = 0.0866, and M = 1.476. Thus, $Z = (1.476 - 1)\ (e^{0.0866} - 1) = 0.043$.

2.42. (a) The overall claim frequency on a premium basis is: 110,000/54 = 2037.
Claim frequency on a premium basis for 3 or more claim free: 40,000 / 25 = 1600.
1 - Z = 1600 / 2037. ⇒ Z = **21.5%**.
Claim frequency on a premium basis for 2 or more claim free:
(40,000 + 15,000) / (25 + 8) = 1667.
1 - Z = 1667 / 2037. ⇒ Z = **18.2%**.
Claim frequency on a premium basis for 1 or more claim free:
(40,000 + 15,000 + 25,000) / (25 + 8 + 13) = 1739.
1 - Z = 1739 / 2037. ⇒ Z = **14.6%**.
(b) Using car years is <u>not</u> preferable to using earned premiums. Using earned premiums adjusts for the mix of business by territory; it adjusts for the effect of territories with higher than average expected frequencies by dividing by their higher than average premiums. Using cars years would not adjust for this, and thus we would be double counting the effect of territories via territorial rating factors and the experience of the insured via Merit Rating.
Specifically, a higher expected frequency territory would have a lower than average proportion of Group A and a higher than average proportion of Group D. Thus Group A would have a higher than average proportion of risks from lower rated territories. Thus the future experience of Group A would look better compared to the overall average than it otherwise would. We want to use Z in Merit Rating to adjust the estimated future frequency for an insured compared to its territory and class, not compared to the overall average. Thus, here what we want to do is compare the experience of group A to the average frequency in its mix of territories rather than the overall average. This is approximated by using earned premium in the denominator which adjusts for the expected frequency for the mix of territories in each Group.

2.43. (a) The overall claim frequency on a premium basis is: 225,000/137 = 1642.
Claim frequency on a premium basis for 3 or more claim free: 120,000/100 = 1200.
1 - Z = 1200 / 1642. ⇒ Z = **26.9%**.
Claim frequency on a premium basis for 2 or more claim free:
(120,000 + 25,000) / (100 + 10) = 1318.
1 - Z = 1318 / 1642. ⇒ Z = **19.7%**.
Claim frequency on a premium basis for 1 or more claim free:
(120,000 + 25,000 + 44,000) / (100 + 10 + 17) = 1488.
1 - Z = 1488 / 1642. ⇒ Z = **9.4%**.
(b) 1. The credibilities are smaller for XYZ than ABC. This is probably due to a more refined classification system for XYZ than ABC. This could also be due to a much lower mean frequency for ABC, so that one year from ABC contains less useful information than from XYZ.
For XYZ the ratio of the three year credibility to the one year credibility is: 14/6 = 2.33, while for ABC it is 26.9/9.4 = 2.86. Since for XYZ the credibilities are further from increasing linearly, there are probably more rapidly shifting risk parameters over time for XYZ than for ABC. This could instead or also be due to XYZ having more risks entering and leaving classes than for ABC.
(c) If one portfolio has a more refined class plan then the credibility assigned to the experience of a single car would be lower relative to the other portfolio which has a less refined plan.
<u>Comment</u>: The two-year credibility of 19.7% is more than twice the one-year credibility of 9.4%. Rather, we would expect the two-year credibility to be less than twice the one-year credibility.

2.44. (a) The overall pure premiums is: 4,000,000/4000 = 1000.
Pure premium for 1 or more years claims-free:
(1,000,000 + 500,000) / (2500 + 500) = 500.
1 - Z = 500 / 1000. ⇒ Z = **50.0%**.
(b) Pure premium for 2 or more years claims-free: 1,000,000/2500 = 400.
The overall pure premium is $1000.
Thus the premium for an exposure that is accident-free for 2 or more years is:
($1250)(400/1000) = **$500**.
Alternately, for a risk that is accident-free for 2 or more years:
1 - Z = 400/1000. ⇒ Z = 60.0%.
There are no losses during the two years, so that the mod is: (0)(0.6) + (1)(1 - 0.6) = 0.4.
Premium is: (0.4)($1250) = **$500**.
Comment: Bailey-Simon work with frequencies rather than pure premiums. All other things being equal, the credibility of one year for estimating future pure premiums is usually less than that for estimating frequencies. (It is easier to estimate future frequencies than pure premiums.)
The given earned exposures and incurred losses are for a subsequent year.
Given the assumptions, we are fine using exposures as the denominator of frequency.
We could instead use in the denominator 1250 times the exposures, making no difference in the estimated credibilities.
In part (b) we have assumed either that there are no fixed expenses, or that there is a separate expense fee which is not adjusted for Merit Rating and which we ignore.
In part (b) we might have an insured who is claim free in 2006 and 2007 and we are using these two years of experience to predict 2008; we give a claim-free discount of 60%.
The CAS sample solutions to part (b) make no sense to me.

2.45. (a) The overall (exposure based) frequency is m = 100,000 / (980,000 + M).
Assuming Poisson frequency, the mean number of claims for those in Group B is: $m / (1 - e^{-m})$.
The relative frequency for Group B is: $1 / (1 - e^{-m})$.
The premium based frequency for Group B is: 18,000/45,000,000.
The overall premium based frequency is: 100,000/670,000,000.
Therefore, the modification for Group B is: (18/45) / (100/670) = 2.68.
Thus we must have: $2.68 = (0.167)\{1 / (1 - e^{-m})\} + (1 - 0.167)(1)$. ⇒
$1 - e^{-m} = 0.0904168$. ⇒ 0.947688 = m = 100,000 / (980,000 + M). ⇒ M = **75,198**.
(b) Premium based frequency for those who are claim-free for two or more years:
(50,000 + 20,000) / (400 + 150) = 127.27.
Premium based frequency overall: 100,000/670 = 149.25.
1 - Z = 127.27/149.25. ⇒ Z = **14.7%**.
Comment: Part (a) tests the alternate technique at page 160 of Bailey-Simon based on looking at the relativity for those with at least one claim.
In part (a), I found it confusing that they used the letter M for the missing number of exposures.

2.46. Premium based frequency for those who are claims-free for one or more years:
(45,000 + 15,000 + 29,300) / (60 + 15 + 20) = 940.
Premium based frequency overall: 108,000/100 = 1080.
1 - Z = 940/1080. ⇒ Z = **13.0%**.

2.47. For State X we have the total claim frequency is: 855 / 1150 = 0.7435.

Number of Accident-Free Years	Relative Claim Frequencies to Total
3 or more	(240/500) / 0.7435 = 0.6456
2 or more	(365/650) / 0.7435 = 0.7553
1 or more	(555/850) / 0.7435 = 0.8782

In state X the ratio of three year to one year credibility is: (1 - 0.6456) / (1 - 0.8782) = 2.91.
In state Y the ratio of three year to one year credibility is: (1 - 0.70) / (1 - 0.84) = 1.875.
State Y credibilities go up much less than linearly, and thus state Y is more affected by shifting risk parameters.
State Y is more variation (over time) in an individual insured's probability of an accident.

2.48. It would be better to use premiums, provided the high rated territories have higher frequency and provided the territory relativities are correct.
The average rates are:
Territory 1: (15 + 125 + 230) / (15 + 125 + 230) = \$1000.
Territory 2: (25 + 310 + 550) / (25 + 300 + 535) = \$1029.
Territory 3: (10 + 80 + 160) / (10 + 100 + 170) = \$893.
(There is not a large spread of rates, but Territory 3 is the lowest rated.)
The average frequencies are:
Territory 1: (5 + 41 + 76) / (15 + 125 + 230) = 0.330.
Territory 2: (7 + 84 + 147) / (25 + 300 + 535) = 0.277.
Territory 3: (4 + 35 + 60) / (10 + 100 + 170) = 0.354.
While Territory 3 is the lowest rated, it has the highest frequency.
So using premiums as the denominator of frequency would not adjust for a maldistribution.
Thus, I would use **car-years** as the denominator of frequency in determining the credibility of a single private car using the general type of technique in Bailey and Simon.
Comment: When I read this question, it was very unclear to me what they were trying to get at.
If this happens to you on your exam, skip the question and come back later if you have time.
It would have helped me if they had said “choose an appropriate denominator to divide into the number of claims to use in determining the credibility of a single private car using the general type of technique in Bailey and Simon.” In my opinion, this was far from one of their better questions.
In Bailey-Simon, I would consider years as the exposure base for credibility; the more years of data for a car, the more credibility.
Bailey and Simon use premium as the denominator to eliminate maldistribution due to high frequency territories having a high territorial relativity and a lower number of accident free risks.
The purpose is to adjust for the mix of territories by subgroup (0 years claims-free, 1 year claims-free, 2 years claims free, etc.); we are concerned about the different relative claim frequencies by territory.
Hazam says that using premium as the denominator works only when:
High frequency territories are also high premium territories, and territorial relativities are proper.
(However, he does not say that when this is not the case we should use car-years as the denominator.)
We can check whether the territory current relativities are correct. The current loss ratios are:
Territory 1: (9 + 75 + 138) / (15 + 125 + 230) = 60%.
Territory 2: (16 + 187 + 328) / (25 + 310 + 550) = 60%.
Territory 3: (7 + 43 + 100) / (10 + 80 + 160) = 60%.
Thus the current territory relativities appear to be correct.
The average rates by years since last accident are for Territory 1 all \$1000.
In Territory 2, the average rate for those with zero years claims-free is \$1000, while for 2 years claims-free it is \$1028.
This is not the pattern we expect. We would assume that those who are claims-free for two years are on average in lower rated classes than those who have zero years claims free.

2.49. (a) I will assume we are analyzing data separately for each class, as per Bailey and Simon. (Here there do not seem to be any classes; "No other rating variables are applicable.")
Also I will assume as per Bailey and Simon that the premiums have been put on the current "Class B" rate level, in other words on the Merit Rating level of those with no years claims free; we need to remove the current impact of the Merit Rating Plan.
We assume that the current territory relativities are correct, and that differences in territory relativities are due to differences in expected frequency (per caryear) rather than expected severity. According to the review by Hazam: "a premium base eliminates maldistribution only if
(1) high frequency territories are also high premium territories and
(2) if territorial differentials are proper."
(b) Overall, frequency with respect to premium ($ million) is: 4075/600 = 6.792.
For two or more years claims free, frequency with respect to premium ($ million) is:
(1200 + 625) / (250 + 100) = 5.214.
Thus for two or more years claims free, Z = 1 - 5.214/6.792 = 23.2%.
For one or more years claims free, frequency with respect to premium ($ million) is:
(1200 + 625 + 750) / (250 + 100 + 100) = 5.722.
Thus for one or more years claims free, Z = 1 - 5.722/6.792 = 15.8%.
The ratio of these two credibilities is: 23.2% / 15.8% = **1.47**.
(c) Assume that the base rate is to be applied to an exposure which has zero years claim free.
For exposures who are zero years claims free, frequency with respect to premium ($ million) is:
1500/150 = 10. Thus we should charge an exposure that is accident free for two or more years:
(1000)(5.214/10) = **$521**.
Alternately, compared to average, we should give an exposure that is accident free for two or more years a discount of 23.2%.
Compared to average those with zero years claims free they should get a surcharge of:
10/6.792 - 1 = 47.2%.
Thus we should charge an exposure that is accident free for two or more years:
(1000/1.472)(1 - 23.2%) = $522.
Alternately, assuming that the base rate is the average rate, then we should charge an exposure that is accident free for two or more years: (1000) (1 - 23.2%) = **$768**.
<u>Comment</u>: In part (c), the examiners seem unaware that the base rate is for Merit Rating Class B, those who are zero years claims free. Rather they seem to assume that the base rate is the average rate, which is not how it is done in the real world. Bailey and Simon put all of their premiums on a Class B level; in other words they treat Merit Rating Class B as the base class.
In any case, the calculated mods are with respect to <u>average</u>.
The credibilities determined are unrealistically big.
The given data is very unusual and unrealistic, including the average premiums:

Number of Accident Free Years	Earned Car Years	Earned Premium ($000)	Average Premiums
3 or More	250,000	250,000	$1,000
2	300,000	100,000	$333
1	25,000	100,000	$4,000
0	12,000	150,000	$12,500
Total	587,000	600,000	$1,022

2.50. (a) Assume as in Bailey-Simon that this is data for one class.
Using car years may create maldistribution because some territories have higher frequency.
Using car years as the denominator of frequency, the credibility calculation would account for both "within territory differences" and "between territory differences". However, usually territory relativities already account for the between territory differences. We want Merit Rating to account for differences between cars not already accounted for by the class/territory relativities. Therefore using car years as the exposure base would double count territory differences, which usually would result in the credibility estimated for Merit Rating being too large.
However, since in this case state law prohibits reflecting territory differences in rating, using earned premium as the exposure base (dividing number of claims by earned premium) should work just as well as using earned exposures. Here using car years is appropriate due to the lack of territory differences in rating. Due to the rates not reflecting frequency differences between territory, the appropriate credibilities for Merit Rating are larger than they otherwise would be.
Alternately, premium may still be a stronger exposure base if nonterritorial factors are captured correctly, thereby reducing the maldistribution that exists using car years.
(b) Overall frequency is: 44/800 = 0.055.
Frequency of those with one or more years accident-free is:
(20 + 15) / (500 + 200) = 0.050.
Z = 1 - 0.05/0.055 = **9.09%**.
(c) Frequency of those with no years accident-free is: 9/100 = 9%.
$9\%/5.5\% = M = Z / (1 - e^{-0.055}) + (1 - Z)(1)$. ⇒ 17.69Z = 0.6364. ⇒ Z = **3.60%**.
Comment: For part (c) we are using the alternate method discussed at page 160 in Bailey-Simon.
It uses the Poisson assumption. Let λ = the mean claim frequency (per exposure) for the class.
M = relative premium based frequency for risks with one or more claims in the past year.
Then, $M = Z / (1 - e^{-\lambda}) + (1 - Z)(1)$. $\Rightarrow Z = \dfrac{M - 1}{1 / (1 - e^{-\lambda}) - 1} = (M - 1)(e^{\lambda} - 1)$.
The estimated credibilities in parts (b) and (c) are both for one year of data, and we would expect them to be more similar than they are here.
Bailey and Simon "have chosen to calculate Relative Claim Frequency on the basis of premium rather than car years. This avoids the maldistribution created by having higher claim frequency territories produce more X, Y, and B risks and also produce higher territorial premiums."

2.51. Here is the best solution I could come up with, expanding on the ideas in Appendix II of Bailey-Simon.
Let us assume that each insured is Poisson with mean λ, with the lambdas varying across the portfolio. Assume that over several years each insured has a constant expected frequency λ.
Then the probability of being claim free for zero years is: $1 - e^{-\lambda}$.
The probability of being claim free for at least one year is: $e^{-\lambda}$.
The probability of being claim free for at least two years is: $e^{-2\lambda}$.
Thus the probability of being claim free for exactly one year is: $e^{-\lambda} - e^{-2\lambda}$.
The probability of being claim free for at least two years is: $e^{-3\lambda}$.
Thus the probability of being claim free for exactly two years is: $e^{-2\lambda} - e^{-3\lambda}$.
Similarly, the probability of being claim free for exactly three years is: $e^{-3\lambda} - e^{-4\lambda}$.
Then for a subset of insureds with the same lambda:

$$\frac{\text{expected number claim free for exactly one year}}{\text{expected number not claim free}} = (e^{-\lambda} - e^{-2\lambda}) / (1 - e^{-\lambda}) = e^{-\lambda}.$$

$$\frac{\text{expected number claim free for exactly two years}}{\text{expected number claim free for exactly one year}} = (e^{-2\lambda} - e^{-3\lambda}) / (e^{-\lambda} - e^{-2\lambda}) = e^{-\lambda}.$$

$$\frac{\text{expected number claim free for exactly three years}}{\text{expected number claim free for exactly two years}} = (e^{-3\lambda} - e^{-4\lambda}) / (e^{-2\lambda} - e^{-3\lambda}) = e^{-\lambda}.$$

Thus within each of the given rows, if the assumptions are correct, we would expect these observed ratios to be close to equal. (Ignore the issue of how would one know the expected claim frequencies for the different rows of insureds.)
For the first row, the observed ratios are: 47,500/50,000 = 0.95, 45,000/47,500 = 0.947, and 44,000/45,000 = 0.978. The last ratio is dissimilar from the other two.
For the second row, the observed ratios are: 45,000/50,000 = 0.90, 43,000/45,000 = 0.956, and 36,000/43,000 = 0.837. These ratios are not similar to each other!
For the third row, the observed ratios are: 20,500/25,000 = 0.82, 16,500/20,500 = 0.805, and 14,000/16,500 = 0.849. These ratios are dissimilar from each other.
We do not see what we would expect; therefore something is wrong with the assumptions.
One or more of the following are true: individuals risk parameters are shifting over time, the frequency process is not Poisson, or insureds are entering and leaving the data base over the period of time studied.

Here is a sample solution from the CAS Examiner's Report that attempts to apply the ideas from Bailey-Simon (but fails to do so correctly, see my comments below):
Total insureds: 125,000 + 113,000 + 104,5000 + 94,000 = 436,500.
Insureds claims free for at least one year (in fact for those claims free for exactly 1, 2 or 3 years): 113,000 + 104,5000 + 94,000 = 311,500.
Insureds claims free for at least two years (in fact for those claims free for exactly 2 or 3 years): 104,5000 + 94,000 = 198,500.
Insureds claims free for at least three years (in fact for those claims free for exactly 3 years): 94,000.
Total expected claims: (186,500)(0.05) + (174,000)(0.10) + (76,000)(0.20) = 41,925.
Expected claims for those claims free exactly one year:
(47,500)(0.05) + (45,000)(0.10) + (20,500)(0.20) = 10,975.
Expected claims for those claims free exactly two years:
(45,000)(0.05) + (43,000)(0.10) + (16,500)(0.20) = 9,850.
Expected claims for those claims free exactly three years:
(44,000)(0.05) + (36,000)(0.10) + (14,000)(0.20) = 8,600.
Expected claims for insureds claims free for at least one year (in fact for those claims free for exactly 1, 2 or 3 years): 10,975 + 9,850 + 8,600 = 29,425.
Expected claims for insureds claims free for at least two years (in fact for those claims free for exactly 2 or 3 years): 9,850 + 8,600 = 18,450.
Expected claims for insureds claims free for at least three years (in fact for those claims free for exactly 3 years): 8,600.
Then for example, the expected frequency for those claims free for at least three years (in fact for those claims free for exactly 3 years): 8600/94,000 = 0.0915. Then, 0.0915/0.0960 = 0.9525.

n	# Claim free n or more years	Expected Claims	Expected Frequency	Relative Expected Frequency	"Credibility"
3	94,000	8,600	0.0915	0.9525	4.75%
2	198,500	18,450	0.0929	0.9677	3.23%
1	311,500	29,425	0.0945	0.9835	1.65%
Total	436,500	41,925	0.0960	1.0000	

For example, the "credibility" for three years is: 1 - 0.9525 = 0.0475.
If the variation of an insured's chance for an accident is <u>not</u> changing over time, then $\frac{\text{3 year credibility}}{\text{1 year credibility}}$ will be approximately equal to 3, and $\frac{\text{2 year credibility}}{\text{1 year credibility}}$ will be approximately equal to 2.

$\frac{\text{3 year credibility}}{\text{1 year credibility}}$ = 0.0475 / 0.0165 = 2.88. $\frac{\text{2 year credibility}}{\text{1 year credibility}}$ = 0.0323 / 0.0165 = 1.96.

The ratios are approximately 3 and 2, and therefore the chance for an accident is stable.

Here is a second sample solution from the CAS Examiner's Report that is a parody of the calculations in Bailey-Simon, demonstrating a lack of understanding of the ideas in Bailey-Simon.
Expected claims at t = 0 (actually for those not claims free):
(50,500)(0.05) + (50,000)(0.10) + (25,000)(0.20) = 12,500.
Expected claims at t = 1 (actually for those claims free exactly one year):
(47,500)(0.05) + (45,000)(0.10) + (20,500)(0.20) = 10,975.
Expected claims at t = 2 (actually for those claims free exactly two years):
(45,000)(0.05) + (43,000)(0.10) + (16,500)(0.20) = 9,850.
Expected claims at t = 3 (actually for those claims free exactly three years):
(44,000)(0.05) + (36,000)(0.10) + (14,000)(0.20) = 8,600.
Then the "frequency at t = 0": 12,000/125,000 = 0.1000.
The "frequency at t = 1": 10,975/113,000 = 0.09712.
The "frequency at t = 2": 9850/104,500 = 0.09426.
The "frequency at t = 3": 8600/94,000 = 0.09149.
The "frequency at t = 1 relative to t = 0": 0.09712/0.1000 = 0.9712.
$\Rightarrow$ One year "credibility": 1 - 0.9712 = 2.88%.
The "frequency at t = 2 relative to t = 0": 0.09426/0.1000 = 0.9426.
$\Rightarrow$ Two year "credibility": 1 - 0.9426 = 5.74%.
The "frequency at t = 3 relative to t = 0": 0.09149/0.1000 = 0.9149.
$\Rightarrow$ Three year "credibility": 1 - 0.9149 = 8.51%.
(Note this is <u>not</u> how Bailey-Simon calculates credibilities. Within a rating class, they compare for example the observed subsequent (premium based) frequency for those who are claims free for 2 years or more, to the overall observed subsequent (premium based) frequency.
Then Bailey-Simon are backing out the credibility for 2 years of data based on an observed credit appropriate for 2 or more years claims free.)
If the variation of an insured's chance for an accident is <u>not</u> changing over time, then $\frac{\text{3 year credibility}}{\text{1 year credibility}}$ will be approximately equal to 3, and $\frac{\text{2 year credibility}}{\text{1 year credibility}}$ will be approximately equal to 2.

$\frac{\text{3 year credibility}}{\text{1 year credibility}}$ = 8.51% / 2.88% = 2.95. $\frac{\text{2 year credibility}}{\text{1 year credibility}}$ = 5.74% / 2.88% = 1.99.

The ratios are approximately 3 and 2, and therefore the chance for an accident is stable.

Here is a third sample solution from the CAS Examiner's Report that is a parody of the calculations in the paper by Mahler, demonstrating a lack of understanding of the ideas in that paper. Determine the percent of the insureds in each column that are in each of the three rows.

t=0	t=1	t=2	t=3
50/125 = 40%	47.5/113 = 42.03%.	45/104.5 = 43.06%	44/94 = 46.81%
50/125 = 40%	45/113 = 39.82%	43/104.5 = 41.15%	36/94 = 38.30%
25/125 = 20%	20.5/113 = 18.14%	16.5/104.5 = 15.79%	14/94 = 14.89%

Now calculate the correlations between the various columns:

"lag 1"	t=0 vs t=1: 0.9965	t=1 vs t=2: 0.9998	t=2 vs t=3: 0.9806	AVG: 0.9923.
"lag 2"	t=0 vs t=2: 0.9980	t=1 vs t=3: 0.9845		AVG: 0.9913
"lag 3"	t=0 vs t=3: 0.9663			AVG: 0.9663

Since the correlations are decreasing with lag, this indicates that parameters are shifting over time.

Here is a fourth sample solution from the CAS Examiner's Report that is another parody of the calculations in the paper by Mahler, demonstrating a lack of understanding of the ideas in that paper.

Determine the expected claims for each entry in the rows and columns.

t=0	t=1	t=2	t=3
(0.05)(50,000) = 2500	2375	2250	2200
(0.10)(50,000) = 5000	4500	4300	3600
(0.20)(25,000) = 5000	4100	3300	2800

Then compute the correlations between the different columns.

"lag 1"	r(0,1) = 0.9842	r(1,2) = 0.9456	r(2,3) = 0.9954	Average = 0.9750
"lag 2"	r(0,2) = 0.8730	r(1,3) = 0.9909		Average = 0.8914
"lag 3"	r(0,3) = 0.8220			Average = 0.8220

Downward trending average correlation as lag increases. ⇒ Risk parameters are shifting.

Here is a fifth sample solution from the CAS Examiner's Report that is another parody of the calculations in the paper by Mahler, demonstrating a lack of understanding of the ideas in that paper.

Determine the ratios of the number of insureds in adjacent columns in each of the three rows.

For the first row, the observed ratios are: 47,500/50,000 = 0.95, 45,000/47,500 = 0.9474, and 44,000/45,000 = 0.9778.

For the second row, the observed ratios are: 45,000/50,000 = 0.90, 43,000/45,000 = 0.9556, and 36,000/43,000 = 0.8372.

For the third row, the observed ratios are: 20,500/25,000 = 0.82, 16,500/20,500 = 0.8049, and 14,000/16,500 = 0.8485.

Then take the correlations between these sets of ratios:

corr[{0.95, 0.90, 0.82}, {0.9474, 0.9556, 0.8049}] = 0.9049,

corr[{0.9474, 0.9556, 0.8049}, {0.9778, 0.8372, 0.8485}] = 0.3920.

corr[[{0.95, 0.90, 0.82}, {0.9778, 0.8372, 0.8485}]] = 0.748.

Average of correlations for "lag 1": (0.9049 + 0.3920)/2 = 0.6485.

Average of correlations for "lag 2": 0.748.

These correlations are not declining with increase in lags.

Thus there is no evidence that parameters are shifting over time.

Comment: This question does not follow any of the syllabus readings. Although this question bears a similarity to ideas in Bailey-Simon and to the shifting risk parameters paper by Mahler, the information needed to properly apply the ideas in those syllabus readings is not provided.
In my opinion, this is a terrible exam question, which demonstrates the lack of understanding of this material by its writer. I suspect those with a better understanding of this material did worse in attempting to somehow answer this exam question. For study purposes, I think this question has negative educational value. Of course, you might want to know how to mechanically reproduce one of the sample solutions in case this exact same form of question is repeated. Therefore, I have given the sample answers from the CAS Examiner's Report.
My commentary on the question and sample solutions follows.
How would one know the expected claim frequencies for the different subsets of insureds?
If an insured's individual chance of an accident changes over time, what could it mean to be in one of the given rows? If an insured's individual chance of an accident changes over time, the insureds in a given row can not have the same expected claim frequency over several years.
Although we are not shown the information, aren't there insureds who are claims-free for exactly four years, exactly five years, etc.? Thus, we do not know for example how many insureds were claims-free for 3 or more years.
In the first sample solution I showed, expected claims are calculated "at time t", by multiplying the number of insureds by the expected claim frequency. What does this mean? Yes if we have 50,000 insureds with an expected claim frequency of 0.05 then we would expect 2500 claims. However, these 50,000 insureds in the first column were not claim free, so they each had at least one claim. Perhaps this means we would expect 2500 claims the following year from these insureds; however; this would ignore the fact that those in a given (heterogeneous) group who are not claim free have higher than average expected future claim frequency compared to the group (the idea behind using credibility) and also that insureds claim propensity may change over time.
Rather as per Bailey-Simon, what we want to know is for a class of insureds the subsequent actual experience of those who were not claim-free, those who were claim free for at least one year, those who were claim free for at least two years, etc. Here we not given this vital information. The solution compares "expected" frequencies rather than as it should observed actual subsequent frequencies.
There is a comparison of the data for those claims free for exactly 1 to 3 years, those claims free for exactly 2 or 3 years, and those who are claims free for exactly 3. The correct comparisons would be between those claims free for at least one year, those who are claims free for at least 2 years, and those who were claims free for at least 3 years; we do not have that information.
Having performed a bunch of arithmetic, "credibilities" supposedly for one, two and three years of data are determined, which are not in fact credibilities in any meaningful sense. However, the conclusion drawn from these "credibilities" is correct. If risk parameters were shifting significantly over time, then the credibilities for one, two, and three years should increase significantly less than linearly.
In the paper by Mahler, the correlations are between different years of actual experience for a set of individual risks. After doing some arithmetic, the sample solutions compute correlations. However, these are not the type of correlations one would use to answer the question of whether we have shifting risk parameters.

The third sample solution works with correlations of the percent of insureds who are claims-free for exactly t years. There is no reason to assume that if risk parameters are constant, that this type of correlation will be independent of the differences in t. If these types of correlations decline as the difference in t (which is not the lag between different years of data) increases, this does not demonstrate that parameters are shifting.
The fourth sample solution and fifth sample solutions are also invalid.
Partial credit was also given for a Chi-Square approach, which is not shown. The Examiner's Report does not explain how one would know the expected number of insureds claims-free for exactly t years, to compare to the actual number. Nor does the Examiner's Report explain exactly how this has any relation to whether or not risk parameters shift.

2.52. (a) The use of premiums as the exposure base (as Bailey-Simon did) would make sense if the high rated territories are the high frequency territories. However, this is not the case here; territory C with the highest frequency has the lowest average premium.
(Different average severities seem to be responsible for a significant amount of the variation in premiums between territories.)
Thus I will use **earned car years as the exposure base**.
Note that in order to use premium as the exposure base to correct for maldistribution, one would also require that the territory differentials are properly priced; there is no way to determine whether or not that is the case here.
(b)

Number of Accident- Free Years	Car Years (000s)	Number of Claims	Frequency	Relative Freq.
3 or More	350	28,500	0.0814	0.866 = 814/940
1 or more	430	36,500	0.0849	0.903 = 849/940
Total	500	47,000	0.0940	1.000

Three year credibility is: 1 - 0.866 = 13.4%. One year credibility is: 1 - 0.903 = 9.7%.
Three year credibility relative to the one year credibility: 13.4% / 9.7% = **1.38**.
Alternately, one can estimate the credibility for one year of data from the experience of those who were not claim free. The frequency per car year for those who are not claim free is: 10,500 / 70,000 = 0.1500. Relative frequency is: 0.1500/0.0940 = 1.596.
Assume a Poisson frequency with mean equal to the overall mean: $\lambda = 0.0940$.
Then the average frequency for those who are not claim free is: $\lambda / (1 - e^{-\lambda})$.
Thus the relative frequency of those who are not claim free is: $1 / (1 - e^{-\lambda}) = 1 / (1 - e^{-0.094})$.
$\Rightarrow 1.596 = M = Z/ (1 - e^{-0.094}) + (1-Z)(1)$. $\Rightarrow$ credibility for one year of data = Z = 5.9%.

Three year credibility relative to the one year credibility: 13.4% / 5.9% = **2.27**.
Comment: For part (b), see Tables 1 and 3 in Bailey-Simon. In Bailey-Simon, the premiums have been adjusted to remove the effect of any discounts from the (current) Merit Rating Plan.
In part (a), the CAS allowed arguing that "while the frequencies do not appear to be in-line with premiums by territory, that premium may still be a better choice as it addresses some maldistribution and should be still used as the exposure base."
In that case, in part (b), one should have gotten:

Number of Accident- Free Years	Premium ($million)	Number of Claims	Frequency	Relative Freq.
3 or More	590	28,500	48.31	0.720 = 48.31/67.14
1 or more	650	36,500	56.15	0.836 = 56.15/67.14
Total	700	47,000	6,714	1.000

Three year credibility is: 1 - 0.720 = 28.0%.
One year credibility is: 1 - 0.839 = 16.4%.
Three year credibility relative to the one year credibility: 28.0% / 16.4% = 1.71.
The merit rating plan uses the number of years an insured is claims free.
The merit rating plan does not "use multiple rating variables, including territory." Rather the rating plan upon which merit rating is superimposed, uses multiple rating variables, including territory. These other rating variables should be controlled for. This is why Bailey and Simon apply this technique to data from each class separately.
Part (b) is unclear; it should have said "three year credibility relative to the one year credibility."

2.53. (a) Looking at the experience of those insureds who were accident free for 0 years:
$M = Z / (1 - e^{-\lambda}) + (1 - Z)(1) = (0.038)/(1 - e^{-0.05}) + (1 - 0.038) = 1.741$.
We need to adjust premiums to the Group B rate:
216/0.6 = 360, 135/0.75 = 180, 63.750/0.85 = 75. 360 + 180 + 75 + 200 = 815 million.
M is ratio of the premium based frequency for those in Group B to that overall.

$$1.741 = M = \frac{C / 200}{(63{,}000 + C) / 815}. \Rightarrow 109{,}683 + 1.741C = 4.075C. \Rightarrow C = \mathbf{46{,}994}.$$

Alternately, $Z = (M - 1) / \{1/(1 - e^{-\lambda}) - 1\} = (M - 1)(e^{\lambda} - 1)$.
$\Rightarrow 0.038 = (M - 1)(e^{0.05} - 1). \Rightarrow M = 1.741$. Proceed as before.

(b) The premium based frequency for two or more years claims free is:
(25,000 + 18,000) / (360 + 180) = 79.629.
The premium based frequency for Group B is: 46,994 / 200 = 233.32.
Indicated Merit Rating Factor = 79.629/233.32 = **0.34**.
(c) A premium base eliminates maldistribution only:
(1) If high frequency territories are also high premium territories.
(2) If territorial differentials are proper.
Thus using earned premium as the exposure base would not correct for maldistribution if:
(1) High premium territories are not also high frequency territories.
(2) or if the current territory differentials are not (approximately) correct.
Comment: The second bullet should have read instead "The credibility for one year of data estimated by examining the experience of those insureds who were accident free for zero years is equal to 0.038."
$\lambda = 0.05$ should be the mean frequency per exposure.
In Table 1 of Bailey-Simon, earned premiums have been adjusted to the Group B rate.
There is no educational value to making part (a) a backwards question.
In part (b), "candidates who calculated the mod relative frequency to total or the merit rating factor relative frequency to group B received full credit." Nevertheless, the merit rating factor, which is what was asked for, is gotten by measuring with respect to Group B, and differs from the mod which is relative to overall.
The current merit rating factors are for 2 years claims free or 3 or more years claims free. There is no current merit rating factor for 2 or more years claims free.

2.54. (a) $E[X] = \frac{pr}{1 - p}$. ⇒ 0.101 = 10 p / (1-p). ⇒ p = 0.101/10,101 = 0.01.

$f(x) = \binom{x+r-1}{x} (1-p)^r p^x \cdot$ $f(0) = (1-p)^r = (1 - 0.01)^{10} = 0.9044$.

Let y be the mean frequency for those insureds who had at least one accident.
(0.9044)(0) + (1 - 0.9044)y = 0.101. ⇒ y = 1.0565.

Experience mod for a policy that has had at least one accident in the last year:
(2%)(1.0565/0.101) + 1 - 2% = **1.189**.
(b) Assume that what they meant to ask was:
We are experience rating two individuals who are in different classes. The first individual has a higher volume of claims and more exposures than the second individual. Describe why the experience of the first individual may be given less credibility than that of the second individual.

In experience rating, the credibility of an individual depends on: the volume of data for the individual, the expected value of the process variance, and the variance of the hypothetical means. More credibility is associated with: a higher volume of data, a smaller expected value of the process variance, and a larger variance of the hypothetical means between the individuals within the class.
If there is a larger variation between the individuals within a class, in other words if the class is heterogeneous, then the variance of the hypothetical means is large, and the experience rating credibility will be higher. If the first individual is in a relatively homogeneous class while the second individual is in a relatively heterogeneous class, then the experience of the first individual may be given less credibility even though it has a larger volume of data than the second individual.
Comment: Part (a) is similar to what is done in Appendix II of Bailey-Simon; however, they assume a Poisson frequency rather than a Negative Binomial.
Let R = the ratio of the actual losses to the expected losses. Then, Bailey-Simon derive that for Poisson frequency with mean λ, for those who have at least one accident: $R = 1/(1 - e^{-\lambda})$.
More generally, one can derive that for those who have at least one accident: R = 1/{1 - f(0)}.
Thus for a Negative Binomial, for those who have at least one accident: $R = \frac{1}{1 - (1-p)^r}$.

In this exam question, $R = 1 / (1 - 0.99^{10}) = 10.485$.
Thus, the experience mod for a policy that has had at least one accident in the last year:
(2%)(10.485) + 1 - 2% = 1.189.
The information I used for the Negative Binomial Distribution was given in a formula sheet at the front of this exam; I have included it in the question.
Instead in the notation in Loss Models, $\beta = 0.101/10 = 0.0101$.
$f(0) = 1/(1+\beta)^r = 1/1.0101^{10} = 0.9044$.
Part (b) was poorly worded; I thought that they were asking about classification rating.
$$K = \text{Buhlmann Credibility Parameter} = \frac{\text{Expected Value of the Process Variance}}{\text{Variance of the Hypothetical Means}}.$$
Z = N / (N + K) or E / (E + K).