2, solution 6.64:

Posterior distribution of θ is proportional to: $\pi(\theta)$ f(11) = 150/(**11** + θ)^4, $5 < \theta < \infty$.

$$\int_5^\infty \frac{150}{(11+\theta)^4}\, d\theta = -50/(11+\theta)^3 \Big]_{\theta=5}^{\theta=\infty} = 25/2048.$$

Posterior density of θ is: {150/(**11** + θ)^4} / (25/2048) = 12,288/(**11** + θ)^4, $5 < \theta < \infty$.

The posterior probability that θ exceeds 10 is:

$$\int_{10}^\infty \frac{12{,}228}{(11+\theta)^4}\, d\theta = -4096/(11+\theta)^3 \Big]_{\theta=10}^{\theta=\infty} = 4096/21^3 = 0.442.$$

4, solutions 4.29 and 4.30 are numbered incorrectly; switch them

5, solution 9.2:

such that $\frac{1}{n}\sum_{i=1}^{n} z_{i1}^2$ is **maximized**,

5, revise Q. 9.4 and solution:
You were given the following six observations with three variables:

Observation	X_1	X_2	X_3
1	21	17	10
2	3	4	14
3	11	20	23
4	33	8	12
5	20	15	6
6	31	6	24

The variables were standardized to have a mean of zero and standard deviation of one.
Then Principal Components Analysis was performed.
The principal component score vectors turned out to be:

PC1	PC2	PC3
-0.87157902	0.6604353	0.3923182
-0.06263353	-0.8361690	-1.8811591
-0.96048588	-1.4958783	1.0176317
0.97745723	1.0816114	0.1072682
-0.93929808	1.0966693	-0.1505856
1.85653927	-0.5066686	0.5145266

Determine the proportion of variance explained by each principal component.

9.4. The variance explained by the m^{th} principal component is: $\frac{1}{n}\sum_{i=1}^{n} z_{im}^2$.

For the first principal component:

$$\frac{(-0.87157902)^2 + (-0.6263353)^2 + (-0.96048588)^2 + 0.97745723^2 + (-0.93929808)^2 + 1.85653927^2}{6}$$

= 1.1618.

For the second principal component:

$$\frac{0.660453^2 + (-0.8361690)^2 + (-1.4958783)^2 + 1.0816114^2 + 1.0966693^2 + (-0.506686)^2}{6}$$

= 1.0004.

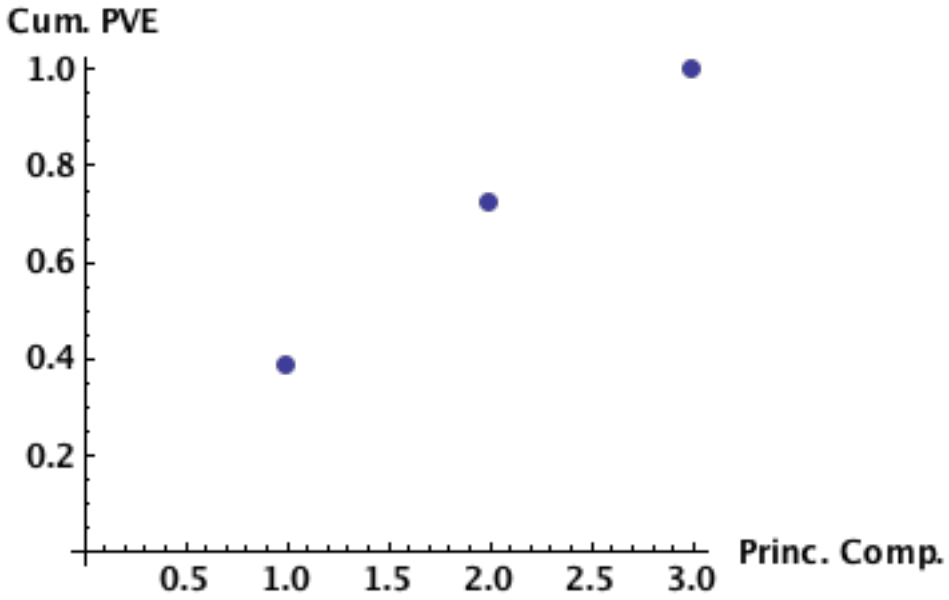
For the third principal component:

$$\frac{0.3923182^2 + (-1.8811591^2) + 1.0176371^2 + 0.1072682^2 + (-0.1505856^2) + 0.5145266^2}{6}$$

= 0.8379.

The proportions of variance explained are:
(1.1618, 1.0004, 0.8379) / (1.1618 + 1.0004 + 0.8379) = **38.73%, 33.35%, 27.93%**.

Comment: While Formula 12.9 in the textbook and the R function princomp each divide by n, the R function prcomp divides by n -1.
As long as one is consistent, the resulting proportions of variance explained are the same.
A graph of the cumulative proportions of variance explained:

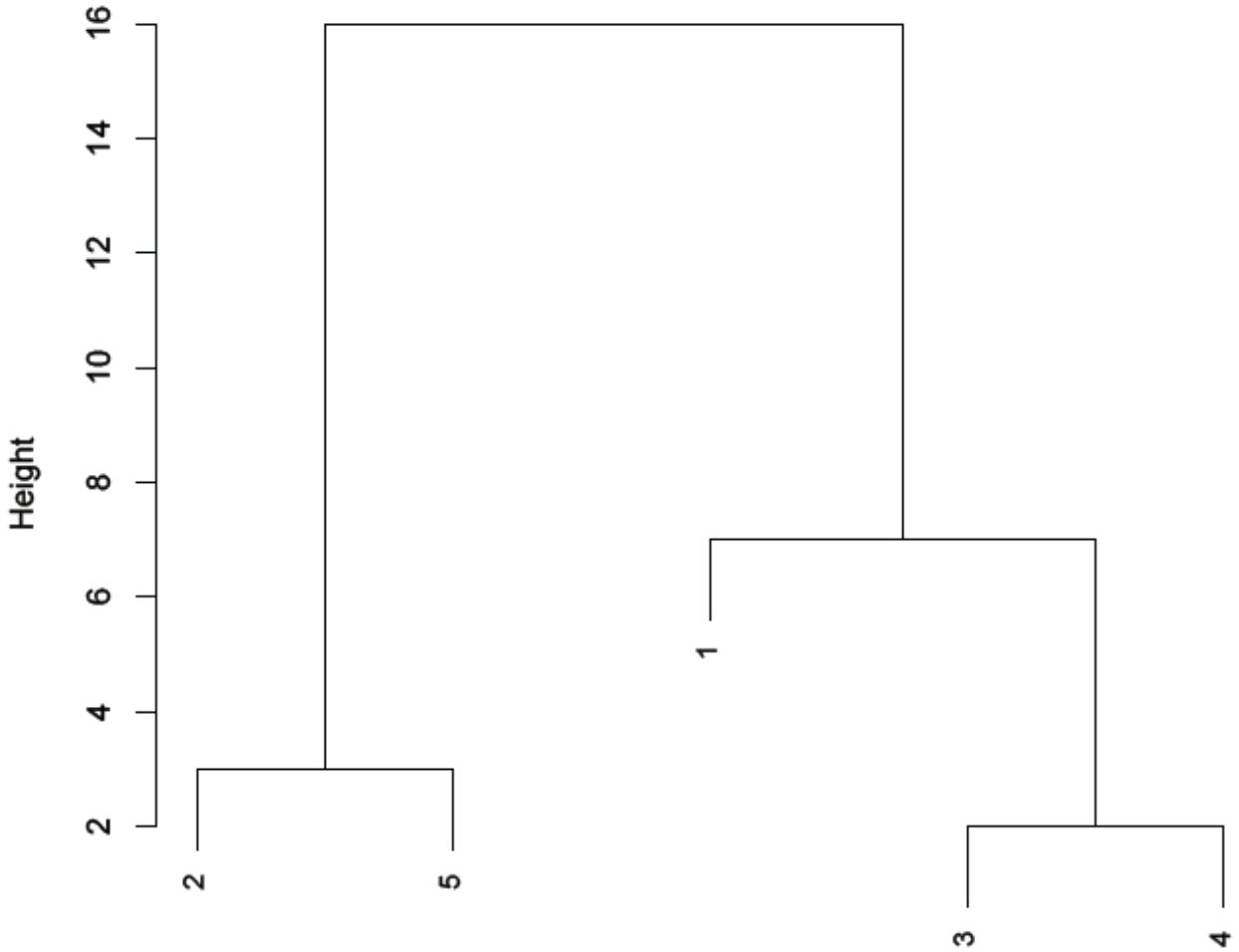


Since the earlier principal components explain more of the variance than the later ones, such a graph should be concave downwards; however, that is not visually obvious in this case.
(If the original variables were independent, then each principal component would explain an equal amount of the variance.)

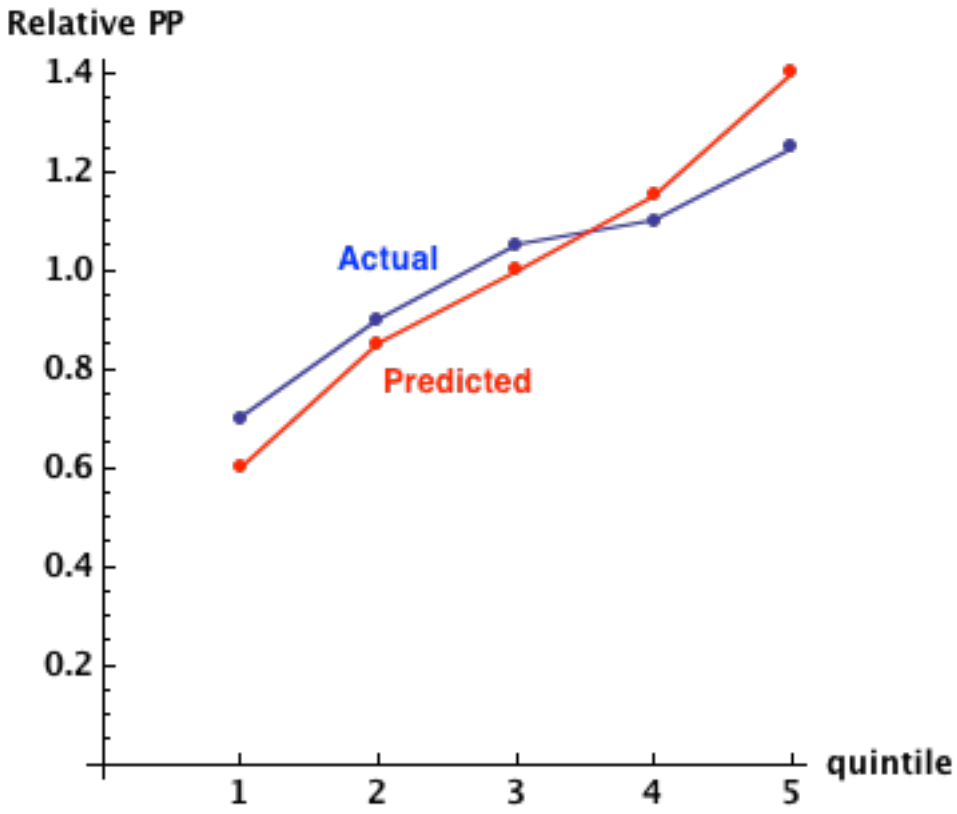
5, revise solution **12.21: C.** The distances between the each pair of points:

	15	4	2	18
9	6	5	7	9
15		11	13	3
4			2	14
2				16

At the first step we group the two closest points: points 2 and points 4.
Then for each of the remaining points we calculate the maximum of the distances from the group of these two points: {2, 4}.
#1 to {2, 4} is Max[9 - 4, 9 - 2] = 7.
#2 to {2, 4} is Max[15 - 4, 15 - 2] = 11.
#5 to {2, 4} is Max[18 - 4, 18 - 2] = 14.
However, the closest ungrouped pair of points is 15 and 18 with a distance of 3.
Thus we group {15, 18}.
The distance from #1 to {15, 18} is Max[15 - 9, 18 - 9] = 9.
The distance from {2, 4} to {15, 18} is Max[15 - 4, 15 - 2, 18 - 4, 18 - 2] = 16.
Thus we next group X = 9 with {2, 4} with distance **7**.
Comment: A dendrogram for this clustering, #1 = 9, #2 = 15, #3 = 4, #4 = 2, #5 = 15:



6, Add to the end of the solution 3.4:
"A model that could be used in practice", would have the actuals increase monotonically, have good but not perfect predictive accuracy, and a reasonably large vertical distance between the actuals in the first and last quintiles. A simple quintile plot:



Comment: Similar to 8, 11/07, Q. 5.
There are many possible examples of the last plot.
Since the records are ordered by predicted values, the records in each bucket change for each graph. Thus, actuals are not the same between the graphs.
Quintile plots are sorted by predicted values from smallest to largest value. Thus the predicted values must be monotonically increasing (or in the case of the null model equal). Actuals need not be monotonically increasing, although that is desirable.
In every graph, the average of the actuals should be the grand mean.
In the final plot, the average of the predicteds should be close to if not equal to the grand mean; the GLM may have a small bias.
In the final plot, the predicted and actuals for the final quintile should each be less than in the saturated model. In the final plot, the predicted and actuals for the final quintile should each be more than in the null model.

8, page 52: MAS-2, 5/19, Q.9 should be labeled 4.**16**.

8, Solution 4.21: The Linear Mixed Model includes **random** effects, while the Implied Marginal Model does not.

8, Q.6.2:

	Restricted Maximum **Log**likelihood	Maximum **Log**likelihood
Reference Model	-1626.517	-1623.517
Nested Model	-1629.219	-1627.823

8, Q.6.16:

	Restricted Maximum **Log**likelihood	Maximum **Log**likelihood
Reference Model	-2657.715	-2646.751
Nested Model	-2659.912	-2648.382

8, **Sol. 6.18. A or B.** The likelihood ratio test statistic is: (2){(-1063.290) - (1065.453)} = 4.326.

For the more complicated model, $D = \begin{pmatrix} \sigma^2_{int} & \sigma_{int:slope} \\ \sigma_{int:slope} & \sigma^2_{slope} \end{pmatrix}$, rather than just σ^2_{int}.

There is a difference of two parameters between the two models. Thus we compare to a 50%-50% mixture of Chi-Square Distributions with 1 and 2 degree of freedom.
For one degree of freedom, the 5% critical value is 3.84 while the 2.5% critical value is 5.02.
$3.84 < 4.326 < 5.02$. Thus comparing to the Chi-Square Distributions with 1 degree of freedom the p-value would be between 2.5% and 5%.
For two degrees of freedom, the 5% critical value is 5.99 while the 2.5% critical value is 5.02.
$4.326 < 5.99$. Thus comparing to the Chi-Square Distributions with 2 degree of freedom the p-value is greater than 5%. Averaging the two p-values we get either something greater than 5% or something between 2.5% and 5%.
Either do not reject at 5%, or reject H_0 at 5% but not at 2.5%.

Comment: The null hypothesis is the model without the random slope, while the alternative hypothesis is the model with the random slope.
Using a computer the p-value is: (3.75% + 11.50%)/2 = 7.6%.

8, page 185:

$$a_i = \frac{\sigma_i^2}{\sigma^2_{\text{random factor}} + \sigma_i^2} = \frac{\sigma^2_{error} / n_i}{\sigma^2_{\text{random factor}} + \sigma^2_{error} / n_i} = \frac{\sigma^2_{error}/\sigma^2_{\text{random factor}}}{n_i + \sigma^2_{error}/\sigma^2_{\text{random factor}}}$$

$$= 1 - \frac{n_i}{n_i + \sigma^2_{error}/\sigma^2_{\text{random factor}}}.$$

$u_i = a_i \mu + (1 - a_i) \mu_i$. $\Leftrightarrow$ Estimate = (1 - Z_i) (overall mean) + Z_i (observed individual mean).

Thus $a_i \Leftrightarrow 1 - Z_i$. $\Rightarrow$ $Z_i \Leftrightarrow \frac{n_i}{n_i + \sigma^2_{error}/\sigma^2_{\text{random factor}}}$.