Christopher Marais

gmarais@ufl.edu

Problem 1

$$n = 5$$

$$S_n = \sum_{i=1}^{n=5} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}$$

Define the likelihood function:

$$L(\lambda|x) = \prod_{i=1}^{n=5} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}$$

Take the derivative of the log likelihood function with regards to lambda:

$$0 = \frac{d}{d\lambda}\left(\ln\left(\prod_{i=1}^{n=5} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}\right)\right)$$

Set it equal to 0:

$$0 = \frac{\lambda(-x_1 - x_2 - x_3 - x_4 - x_5) + 15}{\lambda}$$

Solve for lambda:

$$\lambda = -\frac{15}{-x_1 - x_2 - x_3 - x_4 - x_5}; \ -x_1 - x_2 - x_3 - x_4 - x_5 \neq 0$$

$$\lambda = \frac{\sum_{i=1}^{n} n}{\sum_{i=1}^{n} x_n} \ ; \ \sum_{i=1}^{n} -x_n \neq 0$$

$$\lambda = \frac{1}{\beta} \ ; \ \beta = \frac{1}{\lambda}$$

$$\hat{\beta} = \frac{S_n}{n} \ ; \hat{\lambda} = \frac{n}{S_n}$$

$$\frac{\sum_{i=1}^{n} n}{\sum_{i=1}^{n} x_n} \neq \frac{n}{\sum_{i=1}^{n=5} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!}}$$

The estimate is biased. The size of the bias is defined by the difference in the two equations above.

1.2 I do not expect it to be unbiased, but I do expect the bias to reduce. This is in fact what we see.

Problem 2

2.1 The value of x is [2, -2, 1]. This was done using the solve command in R.

2.2 The A matrix is not invertible and therefore the system of equations can not be solved. It is likely because of the linear dependence between columns 2 and 3 of the matrix.

2.3 Because matrix A from 2.2 is unavertable it either has no solution or multiple solutions. By calculating the determinant of the matrix we see that it is a very small non-zero value. This tells us that the matrix does in fact have at least one solution for some values of b. By calculating the eigenvectors of the matrix we find a solution to a general b value.

Problem 3

3.1 the null hypothesis of X1 = 0 can be rejected as the P-value of the t-test is 0.0255 which is smaller than the generally used 0.05. The t value is also larger than the critical value so we are able to reject the null hypothesis.

3.2 It would be possible to do with a student's t-test. We did this with the information provided by using the t-distribution. First we find the critical value for our degrees of freedom at an assumed alpha of 0.05. Then we computed the t-value by dividing the difference between 1 and our coefficient by the standard error of our coefficient. Thereafter, we compared the t-value and the critical value. the t-value is smaller than the critical value so we fail to reject the null hypothesis that the slope is equal to 1. Thereafter we calculated a p-value for this test which was much larger than 0.05 (0.34). This means that there is a large probability of getting the given data if both populations had the same mean. This p-value is thus not significant.

3.3 We can see the results of testing the null hypothesis that X2 = 0 from looking at the given output. From there we can see that the t-value is smaller than the critical value and we can not reject the null hypothesis. Furthermore, the p-value is 0.1992 which says there is a reasonably high probability of having the data if X2 were 0. We cannot definitively say that there is no association between Y and X2, but we can say that the association is weaker than what is seen between X1 and Y. We can in fact say that the association between X2 and Y is negative rather than positive, but we cannot conclude that it is significant.

3.4 The summary above gives us the F-statistic which should suffice. By using the F statistic, we can test this. From the F-statistic we can almost confidently reject the null-hypothesis that X1=X2=0. The p-value for the f-statistic is not entirely smaller than 0.05, but it is close. Therefore, we cannot reject the null hypothesis of no association, but it would be unwise to assume no association as well.

3.5 Yes, the summary can be used. From the linear model we can see how well the model is able to predict the new observation by using the coefficients and including the Residual error. We can see that the model fails to describe the mean with high confidence. As the difference in the estimated mean and the observed mean is quite large.

Problem 4

To generate a subset from a value of p and m as described in the problem I created a function that firstly calculates the maximum decimal number that can be represented by a binary vector of p length. Then I make a list of all positive numbers up to the maximum number. Thereafter, I converted the decimal numbers in this list to their decimal representations. This was stored in a matrix. A subset from the matrix is then selected where the sum of each row equals m. This then produces a binary matrix that is already sorted lexicographically.

Problem 5

5.1 iii) With a higher correlation (p) between variables the variance increases. The model turns to be more sensitive, and this results in a less accurate prediction model.

I only got to spend approximately 4-5 hours on this. The extra time on Saturday did not help me too much as I already had other responsibilities planned for the day. I do appreciate the extra time though and I am sure I will be on top of things next year for this class.