DEPARTMENT OF MATHEMATICS AND STATISTICS
UMASS - AMHERST
ADVANCED EXAM - STATISTICS (I)
August 29, 2016

Work all problems. Show all work. Explain your answers. State the theorems used whenever possible. 70 points are required to pass, including at least 35 points from questions 1-4 and 20 points from question 5.

1. A family $\mathcal{G}$ of distributions is said to be *closed under convolution* if whenever independent random variables X and Y have distributions in $\mathcal{G}$, he same is true for random variable $X + Y$.

   Consider the family of Gamma distributions with density given by:

   $$f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} \qquad x \geq 0.$$

   (a) (4 points) Consider the family $\mathcal{G}$ of distributions including all Gamma distributions with a fixed value of $\alpha = 1$. Is this family closed under convolution?

   (b) (4 points) Consider the family $\mathcal{G}$ of distributions including all Gamma distributions with a fixed value of $\beta = 1$. Is this family closed under convolution?

   (c) (4 points) Consider the family $\mathcal{G}$ of distributions including all Gamma distributions. Is this family closed under convolution?

2. Suppose that in a certain society, the distribution of the number of children in a family is Poisson with mean 4.

   (a) (7 points) What is the average number of children in a family in this society?

   (b) (7 points) What is the average number of siblings of a child in this society?
   hint: this is not the same answer as the previous part, and it is not just 1 less than part (a)!

3. (12 points) n draws are taken from a uniform distribution on $[0, \theta]$. Observe U_n, the number of draws which are less than 3. Find the MLE of θ. *hint: Be sure to specify the MLE for all possible values of* U_n.

4. Let X_1, X_2 be independent random variables, each with a uniform distribution on $(0, 1)$. Find and sketch the density functions of each of the following quantities:

   (a) (7 points) $X_1 - X_2$

   (b) (7 points) $|X_1 - X_2|$

   (c) (7 points) $max(X_1, X_2)$

   (d) (7 points) $X_1 X_2$

5. A researcher conducted an experiment to compare the growth rates of three varieties of corn. She grew three plants of each variety, measured their growth, and now wants to compare the mean growth of the three varieties. She consults three statisticians — Anna, Krista, and Michael — who all tell her to use a linear model of the form $Y = X\beta + \epsilon$ where the vector Y contains the measured growth rate of the nine plants, β is a vector of parameters and ϵ is a vector of errors. But the three statisticians differ on what X should be and how long β should be.

- Anna says X should be

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

  and that β will be a 3-vector: $\beta^A = (\beta_1^A, \beta_2^A, \beta_3^A)^t$.

- Krista says X should be

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

  and that β will be a 3-vector: $\beta^K = (\beta_1^K, \beta_2^K, \beta_3^K)^t$.

- Michael says X should be

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

  and that β will be a 4-vector: $(\beta_1^M, \beta_2^M, \beta_3^M, \beta_4^M)$.

(a) (7 points) Describe in words the meaning of $(\beta_1^A, \beta_2^A, \beta_3^A)$.

(b) (7 points) Describe in words the meaning of $(\beta_1^K, \beta_2^K, \beta_3^K)$.

(c) (5 points) Are there any data sets $(y_1, \ldots, y_9)$ for which Anna's model fits better than Krista's? Explain.

(d) (8 points) Anna fits her model to the data and finds the covariance matrix Σ^A of $\hat{\beta}^A$. Is there any way other than fitting Krista's model to the data to use Σ^A to find Σ^K, the covariance matrix of $\hat{\beta}^K$? Explain.

(e) (7 points) Describe in words why Michael is wrong.