STAT 501/502 Comprehensive Exam
Monday August 22, 2016

Note: Start each problem on a new page, write on only one side of a page, and label each page with your name or initials.

1. Some probability problems.

   (a) Show that if $P(A|B) > P(A)$, then $P(B|A) > P(B)$. You may assume that none of the events has 0 probability.

   (b) Show that if A and B are independent then A^c and B^c are independent.

2. Let X be a binomial random variable with parameters n and p. Let $Y = n - X$.

   (a) Find the $E(Y)$ and the $Var(Y)$.

   (b) Use M_X (the mgf of X) to find M_Y and, based on that result, deduce the probability distribution of Y.

3. Suppose that the rv T has pdf

$$f_T(t) = \begin{cases} \frac{1}{1.5}\exp\left(-t/1.5\right) & t > 0 \\ 0 & \text{otherwise} \end{cases}$$

   Find the pdf of $V = 1/T$.

4. An urn contains balls numbered 1 and 2. First a ball is drawn randomly from the urn, and then a fair coin is tossed the number of times as the number shown on the drawn ball. Let X equal the number of heads.

   (a) Find the probability mass function for X.

   (b) Find the mean and variance of X.

   (c) Given that one head was observed, what is the probability that the drawn ball had a 1 on it?

5. Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} Expon(\theta)$.

   (a) Find the UMVUE of the median of the distribution, $\tau(\theta) = \theta \log 2$.

   (b) Derive an approximate $100(1-\alpha)\%$ CI for the median.

   (c) Find the UMVUE of the variance of the distribution.

6. An ecologist has conducted a study of the demography of a large land mammal using data from a 20 year period of time. One of the goals of the study was to obtain an estimate of the finite rate of population growth. After much analysis of data collected on 300 animals his estimate of the growth rate turns out to be 1.01 (an annual growth rate of 1%) with a standard error of 0.04. The resulting 95% confidence interval is given as 0.93 to 1.09. (You can assume that this is a frequentist confidence interval). The ecologist compares his estimate of 1.01 with an estimate of 1.05 from another study conducted on the same population, an estimate he thought was too high. He notes that his confidence interval contains the estimate of 1.05 but then claims that there is still a "real" difference between the two estimated growth rates writing:

"The differences between 1.05 and 1.01 are real, even though the estimate of 1.05 falls within my 95% confidence interval. There is no issue of statistical significance. My interval quantifies the amount I would expect my estimate of the growth rate to change if I were to gather data from an entirely different sample of 300 animals in my study area and were to analyze the new data using exactly the same method I employed here. They (the other researchers) got a different answer because they used a different method of analysis"

Comment on the validity of this statement.

7. A few years ago I was on a graduate committee for a student in another department. During his final defense I was asked the following question by another member of the committee:

   In our field we often report a sample mean based on a random sample from a population. Which should we report along with this: the sample standard deviation or the standard error of the mean?

   Answer the question.

8. Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution with pdf

$$f_X(x;\theta) = \frac{3x^2}{\theta}\exp(-x^3/\theta)I_{(0,\infty)}(x)$$

   for $\theta > 0$. Show that a UMP level α test of $\theta \leq \theta_0$ versus $\theta > \theta_0$ can be based on the sufficient statistic $T = \sum X_i^3$.

9. Let $Y_i = \beta X_i + \epsilon_i$ where $\epsilon_1, \epsilon_2, \cdots, \epsilon_n \overset{iid}{\sim} N(0,\theta)$. The X_is are known constants.

   (a) Find a joint minimal sufficient statistic for (β, θ).

   (b) Find the MLE of β and θ. You do not need to confirm maximization. Confirm that the MLE is a function of the statistic you found in part a.