Statistical Inference 453/553 Spring 2011 Instructor: G. Huerta

Some review exercises for Test 1

1. Let $X_1, X_2, \ldots, X_n$ be a random sample where the pdf of each observ ation is:

$$f(x) = 2xI_{(0,1)}(x)$$

Define $T_n = \bar{X}_n = \sum_{i=1}^n X_i/n$. Find an associated limit distribution for T_n and find the corresponding values of parameters that define the limit distribution.

2. Let X_1, X_2 and X_3 be independent Bernoulli random variables of p arameter θ and let $T = \sum_{i=1}^3 X_i$, $T_1 = X_1$ and $T_2 = (T, T_1)$.

   (a) Obtain the partitions induced over the sample space by T, T_1 and T_2 respectively. The sample space is given by the points:

$$\chi = \{(0,0,0),(1,0,0),(0,1,0),(0,0,1),(1,1,0),(1,0,1),(0,1,1),(1,1,1)\}$$

   The statistic T can only take 4 values; $0, 1, 2, 3$. The partition induced by T is: $A_0 = \{(0,0,0)\}; A_1 = \{(1,0,0),(0,1,0),(0,0,1)\}; A_2 = \{(1,1,0),(1,0,1),(0,1,1)\}; A_1 = \{(1,1,1)\}$.
   T_1 only takes two values 0 and 1. Then, the partition induced by T_1 is:
   $A_0 = \{(0,0,0),(0,1,0),(0,0,1),(0,1,1)\}$ and $A_1 = \{(1,0,0),(1,1,0),(1,0,1),(1,1,1)\}$.
   Finally, the partition induced by T_2 is: $A_{(0,0)} = \{(0,0,0)\}$; $A_{(1,0)} = \{(0,1,0),(0,0,1)\}$; $A_{(2,0)} = \{(0,1,1)\}$; $A_{(1,1)} = \{(1,0,0)\}$; $A_{(2,1)} = \{(1,1,0),(1,0,1)\}$ $A_{(3,1)} = \{(1,1,1)\}$.

   (b) Show that T is a minimal sufficient statistic for θ, but T_2 is not, by showing first that T leads to the minimal sufficient partition over the sample space, but T_2 does not lead to such partition.

Notice from (a) that the partitions for T and T_2 are different. In fact, each subset of the partition for T_2 is contained in a subset of the partition for T, but the reverse is not true. Then T is function of T_2 but T_2 is not exclusively a function of T. Finally, notice that the partition for T is the partition that has the fewest elements, such that each of its elements contains points with the same pdf. This makes it the *minimal sufficient partition*.

3. Consider an urn that contains one white ball and two black. Suppose that the balls are extracted with replacement from the urn. Let $X = 0$, if the extracted ball is white and $X = 1$ when the ball results to be black.

   (a) For samples $X_1, X_2, X_3, \ldots X_9$ of size 9, what is the joint distribution of the obsevations?

   Here have Bernoulli sampling with $\theta = 2/3$. The joint pmf of the observations is:

   $$(2/3)^{\sum_{i=1}^{n} X_i}(1/3)^{9-\sum_{i=1}^{n} X_i}$$

   What is the distribution of the sum of the observations?

   A Binomial$(9, 2/3)$.

   (b) In relation to the previous point, find the expected values of the sample mean and the sample variance.

   Remember that $E(\bar{X}) = \mu$ and $Var(\bar{X}) = \sigma^2/n$. In this setting, $\mu = 2/3$ and $\sigma^2 = (2/3)(1/3)/9 = 2/81$

4. Let $X_1, X_2, \ldots, X_n$ be a random sample where $X_i \sim U(0, \theta)$, i.e., $f(x_i; \theta) = (1/\theta)I_{(0,\theta)}(x)$. Find the marginal densities of $X_{(1)}$ and $X_{(n)}$ respectively. Calculate $E(X_{(1)})$ and $E(X_{(n)})$. What is the minimal sufficient statistic for θ? Justify your answer.

5. For the following pdf let $X_1, X_2, \ldots, X_n$ be iid observations. Find a complete sufficient statistic, or show that one does not exist.

   $$f(x|\theta) = e^{-(x-\theta)}exp(-e^{-(x-\theta)}), \quad -\infty < x < \infty, \quad -\infty < \theta < \infty$$

   This density belongs to an exponential family model with $h(x) = e^{-x}$; $c(\theta) = e^{\theta}$; $t_1(x) = e^{-x}$; $w(\theta) = -e^{\theta}$. Then by Theorem 6.2.5:

   $$T_1(\mathbf{x}) = \sum_{i=1}^{n} e^{-x_i}$$

   is a sufficient and complete statistic.

6. Let T be an ancillary statistic for θ with pdf $f_T(t)$. Let $g(\cdot)$ be a one-to-one differentiable function that does not depend on θ. Show that $T^* = g(T)$ is also an ancillary statistic for θ.

   Recall that,

   $$f_{T^*}(t^*) = f_T(g^{-1}(t^*))|\frac{dg^{-1}}{dt^*}|$$

   where $f_{T^*}(\cdot)$ is the pdf of T^*. Since T is ancillary and $g(\cdot)$ does not depend on θ, both terms that define $f_{T^*}(\cdot)$ do not depend on θ. This implies that T^* is also ancillary.

7. Let X be a random variable with pdf

$$f(x|\theta) = \frac{1}{2\theta} I_{(-\theta,\theta)}(x);\ \theta > 0$$

This a case with $n = 1$ where X is sufficient statistic for θ.

(a) Is X a minimal sufficient statistic for θ? Justify your answer.
$-\theta < x < \theta$ iff $-\theta < x$ and $x < \theta$ iff $\theta > |x|$.
Then

$$f(x|\theta) = \frac{1}{2\theta} I_{(|x|,\infty)}(\theta);\ \theta > 0$$

and so the minimal sufficient statistic is $T = |X|$. X alone is not a function of its absolute value so X is not a minimal sufficient statistic.

(b) Is X a complete statistic? Justify your answer.
Take $g(X) = X$. Then $E(X) = \int_{-\theta}^{\theta} x/(2\theta)dx = (1/2\theta)(\theta^2 - \theta^2)/2 = 0$, but X is not identically zero because it has a uniform $(-\theta, \theta)$ pdf.

(c) What is the likelihood function for θ?
The likelihood function is:

$$L(\theta|x) = \frac{1}{2\theta} I_{(|x|,\infty)}(\theta);\ \theta > 0$$

as a function of θ this is equal to zero when $\theta < |x|$. For $\theta > |x|$, this a decreasing function of the form $1/2\theta$.