Math 341, Exam II, Spring 2010, Name_____________

Student # _____________

Thursday, April 8. *Please show the complete solution (with all steps) to each problem to receive full credit.*

1. Suppose that $Y_1, Y_2, \ldots, Y_n$ is a random sample of size n > 3 from normal distribution with mean μ and variance σ^2. Let $\overline{Y}_4 = \frac{\sum_{i=1}^{4} Y_i}{4}$ and $\overline{Y}_3 = \frac{\sum_{i=1}^{3} Y_i}{3}$. Show that $\hat{\sigma}_1^2 = \frac{\sum_{i=1}^{4}(Y_i - \overline{Y}_4)^2}{3}$ and $\hat{\sigma}_2^2 = \frac{\sum_{i=1}^{3}(Y_i - \overline{Y}_3)^2}{2}$ are unbiased estimators of σ^2. (6 points) Find the efficiency of $\hat{\sigma}_1^2$ relative to $\hat{\sigma}_2^2$. Which one of the two estimators is better? Why? (12 points)

Note that $\frac{3\hat{\sigma}_1^2}{\sigma^2} \sim \chi^2_3$ & $\frac{2\hat{\sigma}_2^2}{\sigma^2} \sim \chi^2_2$,

Hence, $E\frac{\cancel{3}\hat{\sigma}_1^2}{\sigma^2} = \cancel{3}$ & $Var\left(\frac{3\hat{\sigma}_1^2}{\sigma^2}\right) = 6$ and

$E\frac{\cancel{2}\hat{\sigma}_2^2}{\sigma^2} = \cancel{2}$ & $Var\left(\frac{2\hat{\sigma}_2^2}{\sigma^2}\right) = 4$. Thus,

$E\hat{\sigma}_1^2 = E\hat{\sigma}_2^2 = \sigma^2$. $Var(\hat{\sigma}_1^2) = \frac{6}{9}\sigma^4$ & $Var(\hat{\sigma}_2^2) = \sigma^4$ Therefore, $eff(\hat{\sigma}_1^2, \hat{\sigma}_2^2) = \sigma^4 = 1.5$

$\hat{\sigma}_1^2$ is better because its variance is 2/3 the variance $\boxed{\frac{2}{3}\sigma^4}$ of $\hat{\sigma}_2^2$.

2. a. – b. The number of persons coming through a blood bank until the first person with type A blood is found is a random variable Y with a geometric distribution. If p denotes the probability that any one randomly selected person will posses type A blood, the E(*Y*) $= \frac{1}{p}$ and V(*Y*) $= \frac{1-p}{p^2}$. $= \frac{1}{p^2} - \frac{1}{p}$

   a. Find a function of Y that is an unbiased estimator of V(*Y*). (8 points)

Note that $EY^2 = Var(Y) + (EY)^2 = \frac{1-p}{p^2} + \frac{1}{p^2}$

$= \frac{2-p}{p^2} = \frac{2}{p^2} - \frac{1}{p}$

$= \frac{2}{p^2} - EY \Rightarrow *$

$* \; EY^2 + EY = \frac{2}{p^2} \Rightarrow E\left(\frac{Y^2+Y}{2}\right) = \frac{1}{p^2}$

Therefore $\frac{Y^2+Y}{2} - Y = \frac{Y^2-Y}{2}$ is an unbiased estimator of $V(Y)$.

b. Form a reasonable 2-standard–error bound on the error of estimation when Y is used to estimate $\frac{1}{p}$. (7 points)

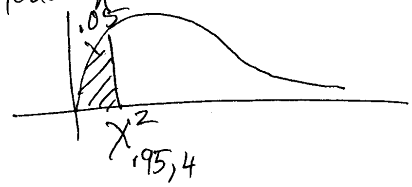
bound is given by $2\sigma_Y = 2\sqrt{Var(Y)}$

$$= 2\sqrt{\frac{Y^2 - Y}{2}}$$

3. The EPA (Environmental Protection Agency) has set a maximum noise level for heavy trucks at 83 decibels (dB). The manner in which this limit is applied will greatly affect the trucking industry and the public. One way to apply the limit is to require all trucks to conform to the noise limit. A second but less satisfactory method is to require the truck fleet's mean noise level to be less than the limit. If the latter rule is adopted, variation in the noise level from truck to truck becomes important because a large value of σ would imply that many trucks exceed the limit, even if the mean fleet level were 83 decibels. A random sample of five heavy trucks produced the following noise levels (in decibels): 85.4 86.8 86.1 85.3 84.8. Use these data to construct a 95% upper confidence interval for σ. (15 points)

Assume that the data follows d r.s. from normal distribution.

$$4\frac{S^2}{\sigma^2} \sim \chi^2_4$$



$$P\left(4\frac{S^2}{\sigma^2} > \chi^2_{.95,4}\right) = .95$$

$$= P\left(\sigma^2 < \frac{4S^2}{\chi^2_{.95,4}}\right) = .95$$

$$= P\left(\sigma < \frac{2S}{\sqrt{\chi^2_{.95,4}}}\right) = .95$$

$S = .779102047$

$$\left(-\infty, \frac{1.558204094}{\sqrt{.710721}} = 1.848309804\right)$$

4. Suppose that $Y_1, Y_2, ..., Y_n$ is a random sample of size n from Poisson distributed population with mean λ. Again, assume that n = 2k for some integer k. Consider $\hat{\lambda} = \frac{1}{2k}\sum_{i=1}^{k}(Y_{2i} - Y_{2i-1})^2$.

   a. Show that $\hat{\lambda}$ is an unbiased estimate for λ. (6 points)

   b Show that $\hat{\lambda}$ is a consistent estimator of λ. (12 points)

$\hat{\lambda}$ is the average of $\frac{(Y_{2i} - Y_{2i-1})^2}{2}$, $i = 1, \dots k$. Hence,

$E\hat{\lambda} = \frac{E(Y_2 - Y_1)^2}{2}$, also because $\left\{\frac{(Y_{2i} - Y_{2i-1})^2}{2}, i = 1, 2 \dots n\right\}$ is a random sample.

$E(Y_2 - Y_1) = EY_2 - EY_1 = \lambda - \lambda = 0 \Rightarrow$

$E(Y_2 - Y_1)^2 = Var(Y_2 - Y_1) \underset{\substack{Y_1 \& Y_2 \\ \text{are independent}}}{=} Var(Y_2) + Var(Y_1) = \lambda + \lambda = 2\lambda.$

$\Rightarrow \frac{E(Y_2 - Y_1)^2}{2} = \lambda$. Hence $E\hat{\lambda} = \lambda$.

$\hat{\lambda}$ being an average $\xrightarrow{P}$ by L.L.N. to $\frac{E(Y_2 - Y_1)^2}{2} = \lambda$ and because $Var\left[\frac{(Y_2 - Y_1)^2}{2}\right] < \infty$, which itself is true because $EY_1^4 < \infty$. $\left[Var(W) \le E(W^2)\right]$.

5. a. – c. Assume that $Y_1, \dots, Y_n$ be a random sample of observations from a gamma-distributed population with $\alpha = 3$ and unknown β.

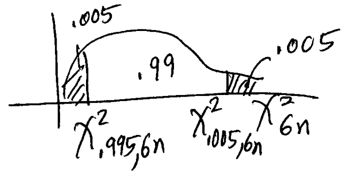
   a. Show that $2\sum_{i=1}^{n} Y_i \big/ \beta$ is a pivotal quantity and has χ^2 distribution with 6n degrees of freedom. (7 points)

$2\sum_{i=1}^{n} Y_i / \beta$ is function of data & β only, and it is distributed χ^2_{6n} because $2Y_i/\beta$ by transformation $\sim \chi^2_6$ & Sum of indep. χ^2 is again χ^2 with df's added.

b. Use the pivotal quantity in 5 a. above to derive a 99% confidence interval for β. (10 points)

c. If a sample of size n = 3 yields $\sum_{i=1}^{3} Y_i = 27$, use the result in 5 b. above to give a 99% confidence interval for β. (7 points)

b.



$$P\left(\chi^2_{.995,6n} \le \frac{2\sum Y_i}{\beta} \le \chi^2_{.005,6n}\right) = .99$$

$$.99 = P\left(\frac{2\sum Y_i}{\chi^2_{.005,\underbrace{6n}_{18}}} \le \beta \le \frac{2\sum Y_i}{\chi^2_{.995,\underbrace{6n}_{18}}}\right)$$

(c) 99% C.I. for β

$$\left(\frac{54}{37.1564}, \frac{54}{6.26481}\right) = (1.453316252, 8.619576053)$$

6. Let $X_1,\ldots,X_n$ be a random sample from Binomial (1, p) find a sufficient statistics for p. (10 points)

$$L(\theta=p) = \binom{1}{x_1} p^{x_1}(1-p)^{1-x_1} \cdots \binom{1}{x_n} p^{x_n}(1-p)^{1-x_n}$$

$$= \binom{1}{x_1}\cdots\binom{1}{x_n} p^{x_1+\cdots+x_n}(1-p)^{n-x_1-\cdots-x_n}$$

$$= \binom{1}{x_1}\cdots\binom{1}{x_n} p^{\sum_{i=1}^{n} x_i}(1-p)^{n-(x_1+\cdots+x_n)}$$

$$= \binom{1}{x_1}\cdots\binom{1}{x_n} p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

Let $h(x_1\cdots x_n) = \binom{1}{x_1}\cdots\binom{1}{x_n}$ &

$g\left(U=\sum_{i=1}^{n} x_i, p\right) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$, $h \& g \geq 0$

and U is sufficient for p by factorization criterion.