Math 362, Problem set 1

Due 1/31/10

1. (4.1.8) Determine the mean and variance of the mean $\bar{X}$ of a random sample of size 9 from a distribution having pdf $f(x) = 4x^3$, $0 < x < 1$, zero elsewhere.

2. (4.2.25) Let $S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$d denote the sample variance of a random sample from a distribution with variance $\sigma^2 > 0$. Since $\mathbb{E}[S^2] = \sigma^2$, why isn't $\mathbb{E}[S] = \sigma$? *Note:* There is a hint in the book that gives it away, but maybe think about it for a second before looking there.

3. (5.1.4) Let $X_1, \ldots, X_n$ be a random sample from the $\Gamma(2, \theta)$ distribution, where θ is unknown. (Recall, look back in chapter 3 if you forget the specifics of the Γ distribution). Let $Y = \sum_{i=1}^n X_i$.

   (a) Find the distribution of Y and determine c so that cY is anunbiased estimator of θ.

   (b) If $n = 5$ show that
   $$\mathbb{P}\left(9.59 < \frac{2Y}{\theta} < 34.2\right) = 0.95$$

   (c) Using (b), show that if y is the value of Y once the sample is drawn then the interval
   $$\left(\frac{2y}{34.2}, \frac{2y}{9.59}\right)$$
   is a 95% confidence interval for Θ.

   (d) Suppose the sample results in the values,

   44.8079   1.5215   12.1929   12.5734   43.2305

   Based on these data, obtain the point estimate of θ as described in Part (a) and the computed 95% confidence interval in Part (c). What does the confidence interval mean?

4. (5.1.5) Suppose the number of customers X that enter a store between the hours of 9AM and 10AM follows a Poisson distribution with parameter θ. Suppose a random sample of the number of customers for 10 days results in the values

$$9 \quad 7 \quad 9 \quad 15 \quad 10 \quad 13 \quad 11 \quad 7 \quad 2 \quad 12$$

   Based on these data obtain an unbiased point estimate of θ. Explain the meaning of this estimate in terms of the number of customers.

5. (5.2.2) Obtain the probability that an observationis a potential outlier for the following distributions

   (a) The underlying distribution is normal

   (b) The underlying distribution is *logistic*, in other words it has pdf:

$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \qquad -\infty < x < \infty$$

   (c) The underlying distribution is Laplace, the pdf given by:

$$f(x) = \frac{1}{2}e^{-|x|}, \qquad -\infty < x < \infty.$$

6. (5.2.5) Let $Y_1 < Y_2 < Y_3 < Y_4$ be the order statistics of a random sample of size 4 from the distribution having pdf $f(x) = e^{-x}$, $0 < x < \infty$, zero elsewhere. Find $\mathbb{P}(3 \leq Y_4)$.

7. (5.2.12) Let $Y_1 < Y_2 < Y_3$ be the order statistics of a random sample of size 3 from a distribution having the pdf $f(x) = 2x$, $0 < x < 1$, zero elsewhere. Show that $Z_1 = Y_1/Y_2$, $Z_2 = Y_2/Y_3$ and $Z_3 = Y_3$ are mutually independent.

8. (5.2.21) Let $X_1, X_2, \ldots, X_n$ be a random sample. A measure of spread is Gini's mean difference

$$G = \sum_{j=2}^{n}\sum_{i=1}^{j-1} |X_i - X_j| / \binom{n}{2}$$

   (a) If $n = 10$, find $a_1, \ldots, a_10$ so that $G = \sum_{i=1}^{10} a_i Y_i$, where $Y_1, Y_2, \ldots, Y_10$ are the order statistics of the sample.

   (b) Show that $\mathbb{E}[G] = 2\sigma/\sqrt{\pi}$ if the sample arises from the normal distribution $N(\mu, \sigma^2)$.