## STATISTICS 630 - Solutions to Test II
## July 12, 2013

1. Suppose that the random variable $V$ has probability density function

$$f_V(v) = \begin{cases} 2v & 0 \leq v \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Find $E(V)$ and $\text{Var}(V)$.

$$E(V) = \int_0^1 2v^2 dv = \left.\frac{2v^3}{3}\right|_0^1 = \frac{2}{3}, \quad E(V^2) = \int_0^1 2v^3 dv = \left.\frac{2v^4}{4}\right|_0^1 = \frac{2}{4},$$
$$\text{Var}(V) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

   (b) Use Chebychev's inequality to give a bound on $P[|V - E(V)| \geq 1/3]$. Compare this bound to the actual probability.

$$P[|V - E(V)| \geq 1/3] \leq \frac{1/18}{(1/3)^2} = \frac{1}{2}.$$

   The actual value is

$$\int_0^{1/3} 2v dv = \left.\frac{2v^2}{2}\right|_0^{1/3} = \frac{1}{9}.$$

2. We now consider inference for the exponential distribution using a mean parameter instead of a rate parameter. Let $X_1, \ldots, X_n$ be a random sample from the exponential distribution with mean parameter $\theta > 0$ and probability density function,

$$f_\theta(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & 0 < x < \infty \\ 0 & \text{otherwise.} \end{cases}$$

   One can show that $E(X_i) = \theta$, $\text{Var}(X_i) = \theta^2$, and that the maximum likelihood estimator of $\theta$ is given by $\hat{\theta} = \frac{1}{n}\sum_{i=1}^n X_i$. (You do not have to derive any of these results.) Consider also the alternative estimator

$$W = \frac{1}{n+1}\sum_{i=1}^n X_i.$$

   Obtain the bias, variance, and mean squared error for each estimator, $\hat{\theta}$ and $W$. Which estimator has smaller mean squared error as an estimator of $\theta$, $\hat{\theta}$ or $W$?

$$E(\bar{X}) = E(X_i) = \theta, \ \text{bias}(\bar{X}) = 0, \ \text{Var}(\bar{X}) = \text{Var}(X_i)/n = \theta^2/n, \ \text{MSE}(\bar{X}) = \theta^2/n.$$

$$E(W) = \frac{nE(X_i)}{n+1} = \frac{n\theta}{n+1}, \text{ bias}(W) = \frac{n\theta}{n+1} - \theta = \frac{-\theta}{n+1},$$

$$\text{Var}(W) = \frac{n\text{Var}(X_i)}{(n+1)^2} = \frac{n\theta^2}{(n+1)^2}, \text{ MSE}(W) = \frac{n\theta^2}{(n+1)^2} + \left(\frac{-\theta}{n+1}\right)^2 = \frac{\theta^2}{n+1}.$$

Thus, $W$ has smaller MSE.

3. Let $X$ and $Y$ be independent random variables where $X$ has a normal distribution with mean 0 and variance 4 and $Y$ has a Poisson distribution with $\lambda = 3$. Define $Z = X+Y$ and $W = X - Y$. Compute $E(Z)$, $E(W)$, $\text{Var}(Z)$, $\text{Var}(W)$, and $\text{Cov}(W, Z)$.

   $E(Z) = 0 + 3 = 3$
   $E(W) = 0 - 3 = -3$
   $\text{Var}(Z) = 4 + 3 = 7 = \text{Var}(W)$
   $\text{Cov}(W, Z) = \text{Cov}(X+Y, X-Y) = \text{Var}(X) - \text{Cov}(X, Y) + \text{Cov}(X, Y) - \text{Var}(Y) = 4 - 3 = 1$

4. Suppose that $X_1, \ldots, X_n$ are a random sample from a distribution with probability mass function,

$$f(x|\theta) = \begin{cases} \frac{e^{-\theta^2}\theta^{2x}}{x!}, & x = 0, 1, 2, \ldots, \ 0 < \theta < \infty \\ 0 & \text{otherwise,} \end{cases}$$

   and mean $E(X_i) = \theta^2$. Find the maximum likelihood estimator estimator of $\theta$ and also the method of moments estimator of $\theta$. Are they the same?

   The log-likelihood is

$$\ell(\theta) = \log(L(\theta)) = \log\left(\prod_{i=1}^{n} \frac{e^{-\theta^2}\theta^{2x_i}}{x_i!}\right)$$
$$= -n\theta^2 + \sum_{i=1}^{n} 2x_i \log(\theta) - \log\left(\prod_{i=1}^{n} x_i!\right).$$

   The score equation is

$$\frac{\partial \ell(\theta)}{\partial \theta} = -2n\theta + \frac{2\sum_{i=1}^{n} x_i}{\theta} = 0.$$

   Solve to get $\hat{\theta} = \sqrt{\frac{\sum_{i=1}^{n} x_i}{n}}$. To check for maximum,

$$\frac{\partial^2 \ell(\theta)}{\partial \theta^2} = -2n - \frac{2\sum_{i=1}^{n} x_i}{\theta^2} < 0$$

   To obtain the method of moments estimate, solve $\bar{X} = E(X_i) = \theta^2$ for $\theta$. Obtain the same estimator, $\hat{\theta} = \sqrt{\frac{\sum_{i=1}^{n} x_i}{n}}$.

5. Suppose that the weights $W_1, \ldots, W_{10}$ of checked luggage for 10 first class customers on a certain airline are normally distributed with a mean of 40 pounds and a standard deviation of 10 pounds. Also, suppose that the weights of checked luggage $V_1, \ldots, V_{25}$ for 25 coach customers on this airline are normally distributed with a mean of 30 pounds and a standard deviation of 5 pounds. You may assume that all the weights are independent.

   Identify the distributions (including the values of the parameters) of

$$\bar{W} = (W_1 + \cdots + W_{10})/10 \qquad \text{and} \qquad \bar{V} = (V_1 + \cdots + V_{25})/25.$$

   Then express $P[\bar{W} > \bar{V}]$ in terms of the cdf of the standard normal distribution.

   The sample mean of a random sample from the normal distribution is normally distributed. Thus, $\bar{W} \sim N(40, (10^2/10))$ and $\bar{V} \sim N(30, (5^2/25))$.

   Next since $\bar{W}$ and $\bar{V}$ are independent normal rvs, $\bar{W} - \bar{V} \sim N(10, 11)$ since $E(\bar{W} - \bar{V}) = 40 - 30 = 10$ and $\mathrm{Var}(\bar{W} - \bar{V}) = \mathrm{Var}(\bar{W}) + \mathrm{Var}(\bar{V}) = 10 + 1 = 11$. Thus,

$$P[\bar{W} > \bar{V}] = P[\bar{W} - \bar{V} > 0] = P\left[\frac{\bar{W} - \bar{V} - 10}{\sqrt{11}} > \frac{0 - 10}{\sqrt{11}}\right] = P\left[Z > \frac{-10}{\sqrt{11}}\right] = 1 - \Phi\left(\frac{-10}{\sqrt{11}}\right).$$