STATISTICS 630 - Solution to Test II **November 7, 2012**

1. Suppose that $X_1 \sim N(2, 2^2)$ and $X_2 \sim N(-1, 3^2)$ are independent random variables.

   (a) Let $U = 4X_1 - X_2$. Find the distribution of U.

   $E(U) = 4(2) - (-1) = 9$ and $\text{Var}(U) = 4^2(4) + (-1)^2(9) = 73$. Since U is a linear function of independent normal rvs, $U \sim N(9, 73)$.

   (b) Find values of $C_1,\ C_2,\ C_3,\ C_4$, and C_5 (where $C_1 \neq 0$ and $C_3 \neq 0$) so that

   $$C_1(X_1 + C_2)^2 + C_3(X_2 + C_4)^2 \sim \chi^2(C_5).$$

   Choose $C_2 = -E(X_1) = -2$ and $C_4 = -E(X_2) = -(-1) = 1$ so that each term inside the square has zero mean. Choose $C_1 = 1/\text{Var}(X_1) = 1/4$ and $C_3 = 1/\text{Var}(X_2) = 1/9$ so that the two terms are squares of standard normal rvs. Thus, $C_5 = 2$ since the expression is the sum of two independent squared standard normal rvs.

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

$$p_\theta(x) = \begin{cases} (x+1)\theta^2(1-\theta)^x, & x = 0, 1, 2, 3, \ldots, \quad (0 \leq \theta \leq 1) \\ 0 & \text{otherwise,} \end{cases}$$

and mean $E(X_i) = 2(1-\theta)/\theta$. Find the maximum likelihood estimator estimator of θ and also the method of moments estimator of θ. Are they the same?

The log-likelihood is

$$\begin{aligned} \ell(\theta) = \log(L(\theta)) &= \log\left(\prod_{i=1}^n (x_i + 1)\theta^2(1-\theta)_i^x\right) \\ &= \log\left(\prod_{i=1}^n (x_i + 1)\right) + 2n\log(\theta) + \sum_{i=1}^n x_i \log(1-\theta). \end{aligned}$$

The score equation is

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

Solve to get $\hat{\theta} = \frac{2n}{2n + \sum_{i=1}^n x_i}$. To check for maximum,

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

when $\theta = \hat{\theta}$ since both terms are negative.

Solve $\bar{X} = 2(1-\theta)/\theta$ to obtain the MOM estimator, $\tilde{\theta} = 2/(2 + \bar{X})$ which equals the mle.

3. Suppose that X and Y are jointly distributed random variables with means, $E(X) = 0$, $E(Y) = 0$, variances, $\text{Var}(X) = 6$, $\text{Var}(Y) = 5$, and covariance, $\text{Cov}(X, Y) = 2$. Let $U = 3X - 2Y$ and $W = 2X + Y$. Obtain the following expectations:

   (a) $E(U) = 3(0) - 2(0) = 0$

   (b) $\text{Var}(U) = 3^2\text{Var}(X) + (-2)^2\text{Var}(Y) + 2(3)(-2)\text{Cov}(X, Y) = 54 + 20 - 24 = 50$

   (c) $E(W) = 2(0) + 0 = 0$

   (d) $\text{Var}(W) = 2^2\text{Var}(X) + \text{Var}(Y) + 2(2)(1)\text{Cov}(X, Y) = 24 + 5 + 8 = 37$

   (e) $\text{Cov}(U, W) = E(UW) = E[(3X - 2Y)(2X + Y)] = 6\text{Var}(X) + 3\text{Cov}(X, Y) - 4\text{Cov}(X, Y) - 2\text{Var}(Y) = (6)(6) - 2 - (2)(5) = 24.$

4. Suppose that undergraduate statistics students take a multiple choice exam with 20 questions and that each question has 5 possible answers. Since all the students neglected to study, each student guesses at random on each question. We assume that all the students take the test independently.

   (a) Let $X_i = 1$ be the score of the i^{th} student taking the exam. Find $E(X_i)$ and $\text{Var}(X_i)$.

   Let $W_j = 1$ if the i^{th} student gets the j^{th} problem correct and $= 0$, otherwise. Then $P(W_j = 1) = 1/5$ and $P(W_j = 0) = 4/5$. Thus, $E(W_j) = 1/5$, $E(W_j^2) = 1/5$, and $\text{Var}(W_j) = 1/5 - (1/5)^2 = 4/25$. Then $E(X_i) = E(\sum_{j=1}^{20} W_j) = 20(1/5) = 4$, and $\text{Var}(X_i) = \text{Var}(\sum_{i=1}^{20} W_j) = 20(4/25) = 16/5$.

   Alternatively, you could state that $X_i \sim \text{binomial}(20, 1/5)$ since the assumptions for a binomial experiment hold. Then $E(X_i) = 20(1/5) = 4$ and $\text{Var}(X_i) = 20(1/5)(4/5) = 16/5$.

   (b) Suppose that a class of n students take the exam independently, and let their scores be $X_1, \ldots, X_n$. Find with a proof a number m such that the average score of the class converges in probability to that number as $n \longrightarrow \infty$; i.e., find a number m such that $\frac{1}{n}(X_1 + \cdots + X_n) \xrightarrow{P} m$.

   Since the variance of X_i is finite, the assumptions of the Weak Law of Large Numbers hold and $\frac{1}{n}(X_1 + \cdots + X_n) \xrightarrow{P} E(X_i) = 4$.

5. Suppose that T is a random variable such that $E(T) = 4\theta$ and $\mathrm{Var}(T) = 8\theta^2$. Consider the following estimators of θ:

$$\hat{\theta}_1 = \frac{T}{4}, \qquad \hat{\theta}_2 = \frac{T}{5}.$$

Find the mean, variance, and mean squared error of each of these estimators. Then determine which one has smaller mean squared error.

$$E(\hat{\theta}_1) = \frac{E(T)}{4} = \frac{4\theta}{4} = \theta, \quad \text{and} \quad E(\hat{\theta}_2) = \frac{E(T)}{5} = \frac{4\theta}{5}.$$

$$\mathrm{Var}(\hat{\theta}_1) = \frac{\mathrm{Var}(T)}{4^2} = \frac{8\theta^2}{4^2} = \frac{\theta^2}{2}, \quad \text{and} \quad \mathrm{Var}(\hat{\theta}_2) = \frac{\mathrm{Var}(T)}{5^2} = \frac{8\theta^2}{5^2} = \frac{8\theta^2}{25}.$$

$$\mathrm{MSE}(\hat{\theta}_1) = \frac{\theta^2}{2} \quad \text{and} \quad \mathrm{MSE}(\hat{\theta}_2) = \frac{8\theta^2}{25} + \left(\frac{4\theta}{5} - \theta\right)^2 = \frac{9\theta^2}{25} < \frac{\theta^2}{2}.$$

Thus, $\hat{\theta}_2$ has smaller mean squared error.