STATISTICS 630 - Solution to Final Exam **December 11, 2012**

1. Suppose that X and Y are independent random variables with respective moment generating functions

$$M_X(s) = \exp(2s + 8s^2) \quad \text{and} \quad M_Y(s) = \exp(-s + s^2).$$

(a) Obtain the moment generating functions of (i) $W = X + Y$ and (ii) $V = 5X$.

$$M_W(s) = M_{X+Y}(s) = E[e^{sX}]E[e^{sY}] = \exp(2s + 8s^2)\exp(-s + s^2) = \exp(s + 9s^2)$$
$$M_V(s) = M_{5X}(s) = E[e^{5sX}] = M_X(5s) = \exp(10s + 200s^2)$$

(b) Obtain $E(W)$ and $E(V)$.

$$E(W) = \left.\frac{dM_W(s)}{ds}\right|_{s=0} = (1 + 18s)e^{s+9s^2}\big|_{s=0} = 1$$
$$E(V) = \left.\frac{dM_V(s)}{ds}\right|_{s=0} = (10 + 400s)e^{10s+200s^2}\big|_{s=0} = 10$$

2. Let $X_1, \ldots, X_n$ be a random sample from the Weibull distribution with density

$$f(x|\theta) = 3\theta x^2 e^{-\theta x^3}, \qquad x > 0, \quad 0 < \theta < \infty.$$

Suppose that θ has the prior density

$$\pi(\theta) = \theta e^{-\theta}, \quad \theta > 0.$$

Obtain the posterior distribution of θ given $X_1 = x_1, \ldots, X_n = x_n$. Then obtain the mean of the posterior distribution.

$$\pi(\theta|x_1, \ldots, x_n) \sim \theta e^{-\theta} \times \theta^n e^{-\theta \sum_{i=1}^n x_i^3} = \theta^{n+1} e^{-\theta(1+\sum_{i=1}^n x_i^3)}$$

This is the kernel of a gamma$(n + 2, 1 + \sum_{i=1}^n x_i^3)$ distribution. Thus, the posterior pdf is

$$\pi(\theta|x_1, \ldots, x_n) = \frac{1}{\Gamma(n+2)} \left((1 + \sum_{i=1}^n x_i^3\right)^{n+2} \theta^{n+1} e^{-\theta(1+\sum_{i=1}^n x_i^3)}, \quad \theta > 0,$$

and

$$E(\theta|x_1, \ldots, x_n) = \frac{n+2}{1 + \sum_{i=1}^n x_i^3}.$$

3. The weekly CPU time used by an accounting firm has probability density function (measured in hours) given by

$$f(x) = \begin{cases} \frac{3}{64}x^2(4-x), & 0 < x < 4, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Obtain $E(X)$ and $\mathrm{Var}(X)$. You may use without proof the fact that $E(X^2) = 32/5$.

$$E(X) = \int_0^4 x\frac{3}{64}x^2(4-x)dx = \frac{3}{64}\left(\frac{4x^4}{4} - \frac{x^5}{5}\right)\Big|_0^4 = \frac{12}{5}$$

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{32}{5} - \left(\frac{12}{5}\right)^2 = \frac{16}{25}.$$

(b) Suppose that the weekly CPU times are recorded for 50 weeks during the year. Assuming that the times are independent and all have the above distribution, obtain an expression in terms of $E(X)$ and $\mathrm{Var}(X)$ that approximates the probability that the total CPU time used during the 50 weeks exceeds 130 hours.

$$\begin{aligned} P[\sum_{i=1}^{50} X_i > 130] &= P\left[\frac{\sum_{i=1}^{50} X_i - 50(12/5)}{\sqrt{50(16/25)}} > \frac{130 - 50(12/5)}{\sqrt{50(16/25)}}\right] \\ &= P\left[Z > \frac{10}{\sqrt{32}}\right] = 1 - \Phi\left(\frac{10}{\sqrt{32}}\right) \end{aligned}$$

The second equality is an approximation justified by the Central Limit Theorem.

4. Let X be a single observation from one of the two following distributions:

x	0	1	2	3	4	5	6	7
$p_0(x)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$p_1(x)$	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{5}{32}$	$\frac{8}{32}$	$\frac{8}{32}$	$\frac{5}{32}$	$\frac{2}{32}$	$\frac{1}{32}$
LR	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{5}{4}$	$\frac{8}{4}$	$\frac{8}{4}$	$\frac{5}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Compute the likelihood ratio for testing $H_0 : \theta = 0$ versus $H_a : \theta = 1$. Then obtain the most powerful test with size $\alpha = \frac{1}{4}$ and compute its power.

By the Neyman-Pearson Lemma, the MP size $\alpha = 1/4$ test has rejection region $\{3, 4\}$ since this corresponds to $LR \geq 2$. The power of the test equals $p_1(3) + p_1(4) = 8/32 + 8/32 = 1/2$.

5. The reading on a voltage meter connected to a test circuit is uniformly distributed over the interval $[\theta, \theta+1]$ where $\theta > 0$ is the true but unknown voltage of the circuit. Suppose that $Y_1, \ldots, Y_n$ is a random sample of such readings.

   (a) Obtain the method of moments estimator $\hat{\theta}$ of θ based on $E(Y_i)$.

   Since $Y_1, \ldots, Y_n$ have a uniform $[\theta, \theta+1]$ distribution, $E(Y_i) = (\theta+\theta+1)/2 = \theta+1/2$. Set $\bar{Y} = \theta + 1/2$ and solve for θ to obtain the moment estimator, $\hat{\theta} = \bar{Y} - 1/2$.

   (b) Obtain the mean squared errors of the two estimators, $\hat{\theta}$ and the sample mean, $\bar{Y}$. Which is the better estimator?

   Using basic properties of the sample mean, we have $E(\bar{Y}) = E(Y_i) = \theta + 1/2$ and $\mathrm{Var}(\bar{Y}) = \mathrm{Var}(Y_i)/n = 1/(12n)$ since $Y_i \sim$ uniform$[\theta, \theta+1]$. Thus,

$$E(\hat{\theta}) = E(\bar{Y}) - 1/2 = \theta,\ MSE(\hat{\theta}) = \mathrm{Var}(\hat{\theta}) = \mathrm{Var}(\bar{Y}) = \frac{1}{12n}, \text{ and}$$

$$E(\bar{Y}) = \theta + 1/2,\ \mathrm{Var}(\bar{Y}) = \frac{1}{12n}, \text{ and } MSE(\bar{Y}) = \frac{1}{4} + \frac{1}{12n}.$$

   Thus, the moment estimator has smaller MSE and is the better estimator.

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

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

and mean $E(X_i) = 2(1-\theta)/\theta$. In Test 2, you found that the maximum likelihood estimator is $\hat{\theta} = \frac{2n}{2n+\sum_{i=1}^n X_i}$.

   (a) Obtain Fisher's information for θ.

   Since

$$\frac{\partial \log(p_\theta(x))}{\partial \theta} = \frac{2}{\theta} - \frac{x}{1-\theta} \quad \text{and} \quad \frac{\partial^2 \log(p_\theta(x))}{\partial \theta^2} = -\frac{2}{\theta^2} - \frac{x}{(1-\theta)^2},$$

   Fisher's information in a single observation is

$$I(\theta) = -E\left[\frac{\partial^2 \log(p_\theta(X))}{\partial \theta^2}\right] = E\left[\frac{2}{\theta^2} + \frac{X}{(1-\theta)^2}\right] = \frac{2}{\theta^2} + \frac{2(1-\theta)}{\theta(1-\theta)^2} = \frac{2}{\theta^2(1-\theta)}.$$

   (b) Use Fisher's information to construct an approximate level γ confidence interval for θ based on the maximum likelihood estimator.

.

The approximate $\gamma-$confidence interval for θ is given by

$$\hat{\theta} \pm Z_{(1+\gamma)/2} \frac{1}{\sqrt{nI(\hat{\theta})}} = \hat{\theta} \pm Z_{(1+\gamma)/2} \frac{\hat{\theta}\sqrt{1-\hat{\theta}}}{\sqrt{2n}}.$$