Final Exam 2011

Part I

Answer questions 1 and 2 using the following information:
Let $X_1, \ldots, X_n$ be iid with pdf $f(x|\theta) = \theta e^{-x\theta}, \;\; 0 \le x \le \infty \;\&\; \theta \ge 0$.

1. Derive the MLE estimate of θ.

$$L(\theta|X_1, \ldots, X_n) = \prod_{i=1}^{n} \theta e^{-X_i\theta} \, 1_{\{X_{(1)} \ge 0\}}$$
$$\Longrightarrow \log L(\theta|X_1, \ldots, X_n) = \sum_{i=1}^{n} \log\left(\theta e^{-X_i\theta} 1_{\{X_{(1)} \ge 0\}}\right) = n\log(\theta) - \theta\sum_{i=1}^{n} X_i + n\log\left(1_{\{X_{(1)} \ge 0\}}\right)$$
$$\therefore, \frac{\partial \log L\left(\hat{\theta}|x_1, \ldots, x_n\right)}{\partial\hat{\theta}} = \frac{n}{\hat{\theta}} - \sum_{i=1}^{n} x_i = 0$$
$$\Longrightarrow \hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\overline{x_n}}$$

2. What is the Cramer-Rao lower bound for the variance of an unbiased estimator of θ?

This distribution satisfies all the necessary requirements to use the most simplified form of the denominator for the Cramer-Rao lower bound (it's worth checking, just for practice….). Since we are assuming our estimator for θ – let's call it $\hat{\theta}$ – is unbiased, the numerator reduces to 1.
$$\therefore, var\left(\hat{\theta}\right) \ge \frac{1}{-n\mathbb{E}\left[\frac{\partial^2 \log(\theta e^{-x\theta})}{\partial\theta}\right]} = \frac{1}{-n\mathbb{E}\left(-\frac{1}{\theta^2}\right)} = \frac{1}{\frac{n}{\theta^2}} = \frac{\theta^2}{n}$$

3. Suppose that the random variable X has exponential distribution with parameter 1 (i.e. $\lambda = 1$), and the random variable $Y|X$ has uniform distribution with parameters 0 and X (i.e. $a = 0$ and $b = X$). What is the function of the conditional mean of Y given X, $\mathbb{E}(Y|X)$?

The question asks for $\mathbb{E}(Y|X)$, and it already supplies us with the distribution for $Y|X$, which means that the information regarding the distribution of X is superfluous information! Don't let simple things like extraneous information throw you off your game….
$$\mathbb{E}(Y|X) = \int_0^X Yf(X|Y)dY = \int_0^X Y\frac{1}{X}dY = \frac{1}{X}\int_0^X YdY = \frac{1}{X}\frac{1}{2}(X^2 - 0^2) = \frac{X}{2}$$

4. Let $X_1, X_2, \ldots$ be a sequence of random variables that converge in probability to a constant a. Assume that $P(X_i > 0) = 1$ for all i. Verify that the sequences defined by $Y_i = \sqrt{X_i}$ and $Y_i' = \frac{a}{X_i}$ converge in probability, and find the limits (what do they converge to?).

$X_n \rightarrow_p a$ is given
By Theorem 5.5.4, $Y_n = \sqrt{X_n} \rightarrow_p \sqrt{a} \Longrightarrow Y_n \rightarrow_p \sqrt{a}$
By Theorem 5.5.4, $\frac{1}{X_n} \rightarrow_p \frac{1}{a}$
By Slutsky Theorem, $Y_n' = \frac{a}{X_n} = a\frac{1}{X_n} \rightarrow_p a\frac{1}{a} = 1 \Longrightarrow Y_n' \rightarrow_p 1$

5. For any two random variables X and Y with finite variances, prove that X and $Y - \mathbb{E}(Y|X)$ are uncorrelated.

$cov[X, Y - \mathbb{E}(Y|X)] = 0 \Longrightarrow X$ and Y are uncorrelated.

$\therefore, cov[X, Y - \mathbb{E}(Y|X)] = \mathbb{E}[X(Y - \mathbb{E}(Y|X))] - \mathbb{E}(X)\mathbb{E}[Y - \mathbb{E}(Y|X)]$	Def of Cov
$= \mathbb{E}[XY - X\mathbb{E}(Y|X)] - \mathbb{E}(X)[\mathbb{E}(Y) - \mathbb{E}(\mathbb{E}(Y|X))]$	Linearity
$= \mathbb{E}(XY) - \mathbb{E}[X\mathbb{E}(Y|X)] - \mathbb{E}(X)[\mathbb{E}(Y) - E(Y)]$	Linearity & LIE
$= \mathbb{E}(XY) - \mathbb{E}[\mathbb{E}(XY|X)]$	CE
$= \mathbb{E}(XY) - \mathbb{E}(XY)$	LIE
$= 0$	Algebra

"CE" stands for property of conditional expectation.

6. (Bonus) Define $X_1, X_2, \dots, X_n$ as a random sample of exponentially distributed variables with parameter λ, $f_{X_i}(x) = \lambda e^{-\lambda x}$, $F_{X_i}(x) = 1 - e^{-\lambda x}$. Define the statistic $X_{\{1\}}$ as $\min\{X_1, \dots, X_n\}$. Derive the cdf of $X_{\{1\}}$.

$$
\begin{aligned}
F_{X_{\{1\}}}(x) &= P(X_{\{1\}} \leq x) \\
&= 1 - P(X_1 > x) * P(X_2 > x) * \dots * P(X_n > x) \\
&= 1 - [1 - P(X_1 \leq x)] * \dots * [1 - P(X_n \leq x)] \\
&= 1 - [1 - (1 - e^{-\lambda x})] * \dots * [1 - (1 - e^{-\lambda x})] \\
&= 1 - [1 - (1 - e^{-\lambda x})]^n \\
&= 1 - (e^{-\lambda x})^n \\
&= 1 - e^{-n\lambda x}
\end{aligned}
$$

$$
\therefore, F_{X_{\{1\}}}(x) = \begin{cases} 1 - e^{-n\lambda x}, & x \geq 0 \\ 0, & otherwise \end{cases}
$$

Part II

1.

a) Show that $f(x|\theta) = \theta x^{\theta-1},\ 0 < x \leq 1,\ \theta > 0$ is a pdf

1) For $\theta > 0$, $f(x|\theta) \geq 0\ \forall x$.	Satisfied
2) $\int_{-\infty}^{\infty} f(x|\theta)dx = \int_0^1 \theta x^{\theta-1}dx = \theta \frac{1}{\theta} x^{\theta} \Big|_0^1 = 1$	Satisfied

For parts (b) through (e), let $X_1, \dots, X_n$ be iid with pdf $f(x|\theta) = \theta x^{\theta-1},\ 0 < x \leq 1,\ \theta > 0$ and $\mathbb{E}(X_i) = \frac{\theta}{1+\theta}$.

b) Show that $\prod_{i=1}^n X_i$ is a sufficient statistic for θ.

$$
L(\theta|X_1, \dots, X_n) = \prod_{i=1}^{n} \theta X_i^{\theta-1} 1_{\{X_i \in (0,1]\}} = \theta^n \left(\prod_{i=1}^{n} X_i\right)^{\theta-1} \prod_{i=1}^{n} 1_{\{X_i \in (0,1]\}}
$$

Applying the Factorization Theorem, $g[T(\boldsymbol{X})|\theta] = \theta(\prod_{i=1}^n X_i)^{\theta-1}$ and $h(\boldsymbol{X}) = \prod_{i=1}^n 1_{\{X_i \in (0,1]\}}$.
$\therefore, T(\boldsymbol{X}) = \prod_{i=1}^n X_i$ is a sufficient statistic for θ.

c) Find the maximum likelihood estimator (MLE) of θ.

$$
l(\theta|X_1, \dots, X_n) = n\log(\theta) + (\theta - 1)\sum_i \log(X_i) + \sum_i \log(1_{\{X_i \in (0,1]\}})
$$

$$
\frac{\partial l(\hat{\theta}|X_1, \dots, X_n)}{\partial \hat{\theta}} = \frac{n}{\hat{\theta}} + \sum_i \log(X_i) = 0
$$

$$
\Longrightarrow \hat{\theta}_{MLE} = -\frac{n}{\sum_i \log(X_i)} = -\frac{1}{\frac{1}{n}\sum_i \log(X_i)}
$$

d) Find the method of moments estimator of θ.

$$\mathbb{E}(X_i) = \frac{\theta}{1+\theta}$$

$\therefore$, by the analogy principle, $\frac{1}{n}\sum_i X_i = \frac{\theta}{1+\theta}$

$\Longrightarrow \hat{\theta}_{MoM} = \frac{\bar{X}}{1-\bar{X}}$

e) Is the method of moments estimator in (d) biased? Explain why or why not.

$\bar{X}$ is an unbiased estimator of $\mathbb{E}(X_i)$; however, $\frac{\bar{X}}{1-\bar{X}}$ is a convex function of $\bar{X}$

$\therefore$, Jensen's Inequality for convex functions $\Longrightarrow \mathbb{E}[g(\bar{X})] > g[E(\bar{X})]$, where $g(\alpha) = \frac{\alpha}{1-\alpha}$

$$\therefore, \mathbb{E}\left(\hat{\theta}_{MoM}\right) = \mathbb{E}[g(\bar{X})] > g[E(\bar{X})] = \frac{\mathbb{E}(\bar{X})}{1-\mathbb{E}(\bar{X})} = \frac{\frac{\theta}{1+\theta}}{1-\frac{\theta}{1+\theta}} = \theta$$

$\therefore$, $\hat{\theta}_{MoM}$ is biased upwards.

f) Find the MLE estimator for μ_X, $\hat{\mu}_X$, where $\mu_X = \mathbb{E}(X_i)$.

$$\mu = \mathbb{E}(X_i) = h(\theta) = \frac{\theta}{1+\theta}$$

By invariance property of MLE,

$$\hat{\mu}_{MLE} = h\left(\hat{\theta}_{MLE}\right) = \frac{-\frac{n}{\sum_i \log(X_i)}}{1-\frac{n}{\sum_i \log(X_i)}} = -\frac{n}{\sum_i \log(X_i)} \frac{\sum_i \log(X_i)}{\sum_i \log(X_i) - n} = \frac{n}{n - \sum_i \log(X_i)}$$

g) Derive the asymptotic distribution for $\sqrt{n}\left(\hat{\theta}_{MLE} - \theta\right)$.

Let $Y_i = g(X_i) = \log(X_i)$

$\therefore, X_i = g^{-1}(Y_i) = e^{Y_i}$

Since this transformation satisfies all the necessary requirements,

$$f_Y(y) = \begin{cases} f_X[g^{-1}(y_i)]\left|\left(\frac{dg^{-1}(y_i)}{d\theta}\right)^2\right|, & -\infty < y \le 0 \\ 0, & otherwise \end{cases} = \begin{cases} \theta e^{\theta y_i}, & -\infty < y \le 0 \\ 0, & otherwise \end{cases}$$

$$= \begin{cases} \theta e^{-\theta y_i}, & 0 \le y < \infty \\ 0, & otherwise \end{cases}$$

This is the exponential distribution!

$\Longrightarrow E(Y_i) = \frac{1}{\theta}$ & $var(Y_i) = \frac{1}{\theta^2}$

$CLT \Longrightarrow \sqrt{n}(\bar{Y} - \mu) \to_d N(0, \mu^2)$, where $\bar{Y} = \frac{1}{n}\sum_i Y_i$ & $\mu = \frac{1}{\theta}$

Let $h(\alpha) = -\frac{1}{\alpha} \Longrightarrow [h'(\alpha)]^2 = \frac{1}{\alpha^4}$

By the Delta Method, $\sqrt{n}[h(\bar{Y}) - h(\mu)] \to_d N\left(0, \mu^2\frac{1}{\mu^4}\right) = N\left(0, \frac{1}{\mu^2}\right)$

Substituting back in for the original arguments and functions,

$$\sqrt{n}\left(-\frac{n}{\sum_i \log(X_i)} - \theta\right) \to_d N(0, \theta^2)$$

$\therefore, \sqrt{n}\left(\hat{\theta}_{MLE} - \theta\right) \to_d N(0, \theta^2)$

h) Construct the likelihood ratio test for $H_0: \theta = 1, H_1: \theta \neq 1$.

$$LRT = \frac{\prod_{i=1}^{n} 1X_i^{1-1}}{\prod_{i=1}^{n} \hat{\theta}_{MLE} X_i^{\hat{\theta}_{MLE}-1}} = \frac{1}{\prod_{i=1}^{n}\left[-\frac{n}{\sum_{j=1}^{n}\log(X_j)} X_i^{-\frac{n}{\sum_i \log(X_i)}-1}\right]}$$
$$= \left[-\frac{1}{n}\sum_{j=1}^{n}\log(X_j)\right]^n \left(\prod_i X_i\right)^{1+\frac{n}{\sum_j \log(X_j)}}$$

2. A study is interested in finding what the **mean and variance of the opportunity cost of tree planting** is for farmers. The study randomly assigns n farmers to 5 different compensation levels, c_k $(k = 1,2,\dots,5)$, for planting 50 trees. Hence, there are 5 equally sized groups of farmers, each with $m_k = m = \frac{n}{5}$ members.

   The study allows farmers to voluntarily participate in the program. According to economics theory, a farmer will participate if her opportunity cost, Y_i, is below the compensation level offered, c_k, where k denotes the compensation group they were assigned to. The researcher does not observe data on Y_i, but observes data on participation decisions:

$$X_{ik} = g(Y_i|c_k) = \begin{cases} 1, & if\ Y_i \le c_k \\ 0, & otherwise \end{cases}$$

   Note that the first subscript in X_{ik} indexes the number of observation from 1 to n and the second subscript indexes the compensation group from 1 to 5.

a) Use the first m observations of the sample $X_{11}, X_{21}, \dots, X_{m1}, X_{(m+1)2}, \dots, X_{n5}$ to derive the method of moments estimator for $p_1 = P(Y_i \le c_1)$?

   Since X_{i1} is a binary variable that depends on the value of Y_i, the expected value of X_{i1} corresponds to the probability that $Y_i \le c_1$.
   $\therefore, p_1 = P(Y_i \le c_1) = \mathbb{E}(X_{i1})$
   By the analogy principle, $\hat{p}_1^{MoM} = \frac{1}{m}\sum_{i=1}^{m} X_{i1}$.

b) Write the variance estimator you derived in (a) as a function of p_1 and m.

$$var(\hat{p}_1^{MoM}) = \frac{1}{m^2} var\left(\sum_{i=1}^{m} X_{i1}\right)$$
   Since X_{i1} is independent, $cov(X_{i1}, X_{-i}1) = 0$.
$$\therefore, var\left(\sum_{i=1}^{m} X_{i1}\right) = \sum_{i=1}^{m} var(X_{i1})$$
   Since X_{i1} is identically distributed $\forall i$, $\sum_{i=1}^{m} var(X_{i1}) = m\, var(X_{i1})$
   X_{i1} is a Bernoulli random variable $\Longrightarrow var(X_{i1}) = p_1(1-p_1)$
$$\therefore, var(\hat{p}_1^{MoM}) = \frac{1}{m^2} m\, p_1(1-p_1) = \frac{1}{m} p_1(1-p_1)$$

c) The researcher assumes that the opportunity cost for each farmer, Y_i, has an identical and independent normal distribution with unknown mean, μ, and unknown variance, σ^2. Hence, the cdf of $\frac{Y_i-\mu}{\sigma}$ evaluated at z can be written as $F_Z(z)$, where $F_Z(z)$ is the standard normal cdf. Write p_1 as a function of parameters μ, σ^2, and a constant c_1.

$$p_1 = P(Y_i \le c_1) = P\left(\frac{Y_i-\mu}{\sigma} \le \frac{c_1-\mu}{\sigma}\right) = F_Z\left(\frac{c_1-\mu}{\sigma}\right)$$

d) Write the 5 equations that match the m-sized sample means of X_{ik} with their theoretical counterparts as a function of unknown parameters μ, σ^2, and known constants c_k, $k = 1, \dots, 5$. Hint: each equation should involve one of the expressions: $\sum_{i=1}^{m} X_{i1}, \sum_{i=m+1}^{2m} X_{i2}, \dots, \sum_{i=4m+1}^{n} X_{i5}$.

$$\frac{1}{m}\sum_{i=1}^{m} X_{i1} = F_Z\left(\frac{c_1-\mu}{\sigma}\right)$$
$$\frac{1}{m}\sum_{i=m+1}^{2m} X_{i2} = F_Z\left(\frac{c_2-\mu}{\sigma}\right)$$
$$\frac{1}{m}\sum_{i=2m+1}^{3m} X_{i3} = F_Z\left(\frac{c_3-\mu}{\sigma}\right)$$
$$\frac{1}{m}\sum_{i=3m+1}^{4m} X_{i4} = F_Z\left(\frac{c_4-\mu}{\sigma}\right)$$
$$\frac{1}{m}\sum_{i=4m+1}^{n} X_{i5} = F_Z\left(\frac{c_5-\mu}{\sigma}\right)$$

e) Based on the number of unknown parameters of the distribution of Y_i, what is the minimum number of equations in (d) that you would need in a method of moments estimation? What is the minimum number of farmer groups with different compensations that the study should have?

2 equations
2 farmer groups

f) Write the likelihood function for the full sample, $X_{11}, X_{21}, \dots, X_{m1}, X_{(m+1)2}, \dots, X_{n5}$ as a function of sample realizations, $x_{11}, x_{21}, \dots, x_{m1}, x_{(m+1)2}, \dots, x_{n5}$, known constants $c_1, \dots, c_k$ and unknown parameters μ and σ^2. Note that although the Y_i's are iid, the X_{ik}'s are not identically distributed, since their distribution depends on the same unknown parameters, μ and σ^2, but different known constants $c_1, \dots, c_k$. Note also that X_{ik} are independently distributed since farmers were assigned randomly to different compensation groups.

$$L(\mu,\sigma|X_{11},\dots,X_{nk}) = \prod_{i=1}^{m}\left[p_1^{X_{i1}}(1-p_1)^{1-X_{i1}}\right]\dots\prod_{i=4m+1}^{n}\left[p_5^{X_{i5}}(1-p_5)^{1-X_{i5}}\right]$$
$$= \prod_{i=1}^{m} F_Z\left(\frac{c_1-\mu}{\sigma}\right)^{X_{i1}}\left[1-F_Z\left(\frac{c_1-\mu}{\sigma}\right)\right]^{1-X_{i1}} * \dots$$
$$* \prod_{i=4m+1}^{n} F_Z\left(\frac{c_5-\mu}{\sigma}\right)^{X_{i5}}\left[1-F_Z\left(\frac{c_5-\mu}{\sigma}\right)\right]^{1-X_{i5}}$$

g) Write the first order conditions of the log –likelihood maximization problem with respect to parameters μ and σ^2. You do not need to solve for the derivatives of the standard normal cdf, just leave them indicated.

$$l(\mu,\sigma|X_{11},\dots,X_{nk}) = \sum_{i=1}^{m}\left\{X_{i1}\log\left[F_z\left(\frac{c_1-\mu}{\sigma}\right)\right] + (1-X_{i1})\log\left[1-F_z\left(\frac{c_1-\mu}{\sigma}\right)\right]\right\} + \dots +$$
$$\sum_{i=4m+1}^{n}\left\{X_{i5}\log\left[F_z\left(\frac{c_5-\mu}{\sigma}\right)\right] + (1-X_{i5})\log\left[1-F_z\left(\frac{c_5-\mu}{\sigma}\right)\right]\right\}$$

$$\frac{\partial l}{\partial \mu} = \sum_{i=1}^{m}\left\{-X_{i1}\frac{1}{F_z\left(\frac{c_1-\mu}{\sigma}\right)} f_z\left(\frac{c_1-\mu}{\sigma}\right)\frac{1}{\sigma} + (1-X_{i1})\frac{1}{1-F_z\left(\frac{c_1-\mu}{\sigma}\right)} f_z\left(\frac{c_1-\mu}{\sigma}\right)\frac{1}{\sigma}\right\} + \dots +$$
$$\sum_{i=4m+1}^{n}\left\{-X_{i5}\frac{1}{F_z\left(\frac{c_5-\mu}{\sigma}\right)} f_z\left(\frac{c_5-\mu}{\sigma}\right)\frac{1}{\sigma} + (1-X_{i5})\frac{1}{1-F_z\left(\frac{c_5-\mu}{\sigma}\right)} f_z\left(\frac{c_5-\mu}{\sigma}\right)\frac{1}{\sigma}\right\}$$

$$\frac{\partial l}{\partial \sigma^2} = \sum_{i=1}^{m} \left\{ -\mathrm{X}_{i1} \frac{1}{F_Z\left(\frac{c_1-\mu}{\sigma}\right)} f_Z\left(\frac{c_1-\mu}{\sigma}\right) \frac{1}{2} \frac{c_1-\mu}{\sigma^3} + (1-X_{i1}) \frac{1}{1-F_Z\left(\frac{c_1-\mu}{\sigma}\right)} f_Z\left(\frac{c_1-\mu}{\sigma}\right) \frac{1}{2} \frac{c_1-\mu}{\sigma^3} \right\} + \cdots +$$

$$\sum_{i=4m+1}^{n} \left\{ -\mathrm{X}_{i5} \frac{1}{F_Z\left(\frac{c_5-\mu}{\sigma}\right)} f_Z\left(\frac{c_5-\mu}{\sigma}\right) \frac{1}{2} \frac{c_5-\mu}{\sigma^3} + (1-X_{i5}) \frac{1}{1-F_Z\left(\frac{c_5-\mu}{\sigma}\right)} f\left(\frac{c_5-\mu}{\sigma}\right) \frac{1}{2} \frac{c_5-\mu}{\sigma^3} \right\}$$