14.30 Problem Set #7 solutions
Due Tuesday, November 16, 2004
Guy Michaels

You must answer ALL questions in order to fulfill the course requirement. For each student, the answers will be ordered by score and the top four answers will then comprise the final problem set grade.

Question 1:
Let $X_1, X_2, ..., X_n$ be a random sample where $X_i \sim$exponential(β), that is: $f_{X_i}(x_i) = \frac{1}{\beta}e^{-x_i/\beta}, x_i > 0$ and 0 elsewhere.

a. Derive the MLE for β.

answer: The likelihood function is: $L(\beta; x) = \prod_{i=1}^{n} \frac{1}{\beta}e^{-x_i/\beta}$.

Therefore, the log likelihood function is: $\ln L(\beta; x) = \ln\left(\prod_{i=1}^{n} \frac{1}{\beta}e^{-x_i/\beta}\right) = \sum_{i=1}^{n}(-x_i/\beta) - n\ln(\beta)$.

Differentiating with respect to β and equating the result to zero gives: $\frac{\partial \ln L(\beta;x)}{\partial\beta} = \ln\sum_{i=1}^{n}\left(x_i/\beta^2\right) - \frac{n}{\beta} = 0$

So: $\widehat{\beta}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$.

b. Find the MLE for $\sqrt{\beta}$

answer: Using the invariance property, the MLE for $\sqrt{\beta}$ is:
$\widehat{\sqrt{\beta}}_{MLE} = \sqrt{\widehat{\beta}_{MLE}} = \sqrt{\overline{X}}$

c. Is your MLE in part a. unbiased? Formally prove or disprove.

answer: $\widehat{\beta}_{MLE}$ is unbiased, since: $E\left[\widehat{\beta}_{MLE}\right] = E\left[\overline{X}\right] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\cdot n\beta = \beta$

d. Is your MLE from part a. consistent?

answer: Yes, as your handout indicates a MLE is consistent.

Question 2:
Assume a sample of continuous random variables: $X_1, X_2, ..., X_n$, where $E[X_i] = \mu, Var[X_i] = \sigma^2 > 0$. Consider the following estimators: $\widehat{\mu}_{1,n} = X_n, \widehat{\mu}_{2,n} = \frac{1}{n+1}\sum_{i=1}^{n} X_i$.

a. Are $\widehat{\mu}_{1,n}$ and $\widehat{\mu}_{2,n}$ unbiased?

answer: $Bias\left[\widehat{\mu}_{1,n}\right] = E\left[\widehat{\mu}_{1,n}\right] - \mu = E\left[X_n\right] - \mu = \mu - \mu = 0$ so $\widehat{\mu}_{1,n}$ is unbiased for every n.

$Bias\left[\widehat{\mu}_{2,n}\right] = E\left[\widehat{\mu}_{2,n}\right] - \mu = E\left[\frac{1}{n+1}\sum_{i=1}^{n} X_i\right] - \mu = \frac{1}{n+1}\sum_{i=1}^{n} E\left[X_i\right] - \mu = \frac{n}{n+1}\mu - \mu = \frac{-\mu}{n+1} \neq 0$ so $\widehat{\mu}_{2,n}$ is biased for for every n.

b. Are $\widehat{\mu}_{1,n}$ and $\widehat{\mu}_{2,n}$ consistent?

answer: $\lim_{n\to\infty} P\left(\left|\widehat{\mu}_{1,n} - \mu\right| < \varepsilon\right) = \lim_{n\to\infty} P\left(\left|X_n - \mu\right| < \varepsilon\right) = P\left(\mu - \varepsilon < X_n < \mu + \varepsilon\right) = P\left(X_n \leq \mu + \varepsilon\right) - P\left(X_n \leq \mu - \varepsilon\right)$ (because the R.V. are continuous)$= F\left(\mu + \varepsilon\right) - F\left(\mu - \varepsilon\right) \neq 0$ for some $\varepsilon > 0$. Thus $\widehat{\mu}_{1,n}$ is not consistent.

As for $\widehat{\mu}_{2,n}$: $Var\left[\widehat{\mu}_{2,n}\right] = Var\left[\frac{1}{n+1}\sum_{i=1}^{n} X_i\right] = \left(\frac{1}{n+1}\right)^2 \sum_{i=1}^{n} Var\left[X_i\right] = \left(\frac{1}{n+1}\right)^2 n\sigma^2$ so:

$\lim_{n\to\infty}\left(MSE\left[\widehat{\mu}_{2,n}\right]\right) = \lim_{n\to\infty}\left(Var\left[\widehat{\mu}_{2,n}\right]\right) + \lim_{n\to\infty}\left(Bias\left[\widehat{\mu}_{2,n}\right]\right)^2 = 0$, proving that $\widehat{\mu}_{2,n}$ is consistent.

c. What do you conclude about the relation between unbiased and consistent estimators?

answer: An unbiased estimator is not necessarily consistent; a consistent estimator is not necessarily unbiased.

Question 3:

Consider a sample of random variables: $X_1, X_2, ..., X_n$, where $n > 10$, $E\left[X_i\right] = \mu$, $Var\left[X_i\right] = \sigma^2 > 0$ and the estimator: $\widehat{\mu}_n = \frac{1}{n-10}\sum_{i=11}^{n} X_i$. [Hint: note that the sum is over $(n-10)$ random variables].

a. Calculate the bias of $\widehat{\mu}_n$

answer: $Bias\left[\widehat{\mu}_n\right] = E\left[\widehat{\mu}_n\right] - \mu = E\left[\frac{1}{n-10}\sum_{i=11}^{n} X_i\right] - \mu = \frac{1}{n-10}\sum_{i=11}^{n} E\left[X_i\right] - \mu = \frac{1}{n-10}\left(n-10\right)E\left[X_1\right] - \mu = 0$

b. Calculate the variance of $\widehat{\mu}_n$.

answer: $Var\left[\widehat{\mu}_n\right] = Var\left[\frac{1}{n-10}\sum_{i=11}^{n} X_i\right] = \left(\frac{1}{n-10}\right)^2 \sum_{i=11}^{n} Var\left[X_i\right] = \left(\frac{1}{n-10}\right)^2 \left(n-10\right)\sigma^2 = \frac{\sigma^2}{n-10}$

c. Calculate the MSE of $\widehat{\mu}_n$.

answer: $MSE\left[\widehat{\mu}_n\right] = Var\left[\widehat{\mu}_n\right] + \left(Bias\left[\widehat{\mu}_n\right]\right)^2 = \frac{\sigma^2}{n-10} + 0^2 = \frac{\sigma^2}{n-10}$

d. Is $\widehat{\mu}_n$ efficient in a finite sample?

answer: No. Consider the estimator: $\widetilde{\mu}_n = \overline{X}_n$. Clearly $\widetilde{\mu}_n$ is unbiased and $Var\left[\widetilde{\mu}_n\right] = \frac{\sigma^2}{n} < \frac{\sigma^2}{n-10} = Var\left[\widehat{\mu}_n\right]$ so we have found an unbiased estimator with a strictly lower variance than $\widehat{\mu}_n$. Thus $\widehat{\mu}_n$ is not efficient.

e. Can you think of a scenario where you might want to use $\widehat{\mu}_n$?

answer: If we suspect that the first ten observations may be coming from a different distribution we might choose to use $\widehat{\mu}_n$.

Question 4:

Suppose all MIT undergraduates have one of the four following living arrangements: (i) fraternity/sorority, (ii) independent living group, (iii) a dormitory, (iv) an off-campus apartment. Assume this list is exhaustive and mutually exclusive. You are interested in estimating the probability that an incoming freshman will choose each of the four options. Assume that there is no individual heterogeneity, so each freshman chooses any option with the same (independent) probability as all other freshmen. Assume also that the probabilities do not differ by year. You conduct a survey of students to determine the probability of each alternative.

a. Using the results from your survey as your i.i.d. random sample $(X_1, X_2, ..., X_n)$ derive the MLE for each of the four probabilities.

answer: Define the following parameters:
$\alpha =$fraction of students in group (i)
$\beta =$fraction of students in group (ii)
$\gamma =$fraction of students in group (iii)
$1-\alpha-\beta-\gamma =$fraction of students in group (iv)
The number of observations in groups $(i)-(iv)$ in your sample is $X_\alpha, X_\beta, X_\gamma, X_{1-\alpha-\beta-\gamma}$, where: $X_\alpha + X_\beta + X_\gamma + X_{1-\alpha-\beta-\gamma} = n$.
The likelihood function is: $L(\alpha,\beta,\gamma;X) = \frac{n!}{X_\alpha! X_\beta! X_\gamma! X_{1-\alpha-\beta-\gamma}!}\alpha^{X_\alpha}\beta^{X_\beta}\gamma^{X_\gamma}(1-\alpha-\beta-\gamma)^{X_{1-\alpha-\beta-\gamma}}$
The log-likelihood function is: $\ln(L(\alpha,\beta,\gamma;X)) =$constant$+X_\alpha\ln(\alpha)+X_\beta\ln(\beta)+X_\gamma\ln(\gamma)+X_{1-\alpha-\beta-\gamma}\ln(1-\alpha-\beta-\gamma)$
Taking first order conditions with respect to α, β and γ:
$\frac{\partial\ln(L)}{\partial\alpha} = \frac{X_\alpha}{\alpha} - \frac{X_{1-\alpha-\beta-\gamma}}{1-\alpha-\beta-\gamma} = 0$
$\frac{\partial\ln(L)}{\partial\beta} = \frac{X_\beta}{\beta} - \frac{X_{1-\alpha-\beta-\gamma}}{1-\alpha-\beta-\gamma} = 0$
$\frac{\partial\ln(L)}{\partial\gamma} = \frac{X_\gamma}{\gamma} - \frac{X_{1-\alpha-\beta-\gamma}}{1-\alpha-\beta-\gamma} = 0$
Solving for α, β and γ we get:
$\widehat{\alpha}_{MLE} = \frac{X_\alpha}{n}, \widehat{\beta}_{MLE} = \frac{X_\beta}{n}, \widehat{\gamma}_{MLE} = \frac{X_\gamma}{n}$.

b. What happens if one of the categories in your survey had zero respondents?

answer: If you have zero respondents for a certain group, your estimate for that groups probability is also zero.

Question 5:

You observe a random sample $(X_1, X_2, ..., X_n)$ from a distribution $f_X(x) = \begin{cases} e^{-(x-\theta)} \text{ if } x > \theta \\ 0 \text{ otherwise} \end{cases}$

You do not know θ and you wish to estimate it.

a. Calculate a method of moments (MM) estimator for θ.

answer: The MM estimator is derived by solving the following moment equation: $E\left[X|\theta = \widehat{\theta}\right] = \overline{X}$. To find $E[X]$ note that: $E[X] = \int_{-\infty}^{\infty} xf(x)\,dx = \int_{\theta}^{\infty} xe^{-(x-\theta)}dx = \theta + 1$. This means that $E\left[X|\theta = \widehat{\theta}\right] = \overline{X} = \widehat{\theta} + 1$. Thus: $\widehat{\theta} = \overline{X} - 1$.

b. Is the MM estimator unbiased?

answer: Yes, since: $E\left[\widehat{\theta}\right] = E\left[\overline{X} - 1\right] = E\left[\overline{X}\right] - 1 = E[X] - 1 = (\theta + 1) - 1 = \theta$

c. Is the MM estimator consistent?

answer: Yes, since by the law of large numbers $\overline{X}$ converges to $E[X]$, which equals $\theta + 1$, so: $\widehat{\theta} = \overline{X} - 1$ converges to θ.

Question 6:

a. Find a Method of Moments estimator for $\theta = (\theta_1, \theta_2)$ given a random sample from the uniform distribution $U[\theta_1, \theta_2]$.

answer: The first two moments are:

$E[X] = \int_{\theta_1}^{\theta_2} xf(x)\,dx = \int_{\theta_1}^{\theta_2} x\frac{1}{\theta_2 - \theta_1}dx = \frac{1}{\theta_2 - \theta_1}\left[\frac{x^2}{2}\right]_{\theta_1}^{\theta_2} = \frac{\theta_2 + \theta_1}{2}$

$E[X^2] = \int_{\theta_1}^{\theta_2} x^2 f(x)\,dx = \int_{\theta_1}^{\theta_2} x^2\frac{1}{\theta_2 - \theta_1}dx = \frac{1}{\theta_2 - \theta_1}\left[\frac{x^3}{3}\right]_{\theta_1}^{\theta_2} = \frac{\theta_2^2 + \theta_1^2 + \theta_2\theta_1}{3}$

So MM estimators can be obtained by solving:

$$\frac{\widehat{\theta}_2 + \widehat{\theta}_1}{2} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

$$\frac{\widehat{\theta}_2^2 + \widehat{\theta}_1^2 + \widehat{\theta}_2\widehat{\theta}_1}{3} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \overline{X^2}$$

Note that: $\overline{X^2} - \overline{X}^2 = \frac{1}{12}\left(4\widehat{\theta}_2^2 + 4\widehat{\theta}_1^2 + 4\widehat{\theta}_2\widehat{\theta}_1 - 3\widehat{\theta}_2^2 - 3\widehat{\theta}_1^2 - 6\widehat{\theta}_2\widehat{\theta}_1\right) = \frac{1}{12}\left(\widehat{\theta}_2 - \widehat{\theta}_1\right)^2$

$\Rightarrow \sqrt{3\left(\overline{X^2} - \overline{X}^2\right)} = \frac{1}{2}\left(\widehat{\theta}_2 - \widehat{\theta}_1\right)$

Thus the MM estimators are:

$$\widehat{\theta}_1 = \overline{X} - \sqrt{3\left(\overline{X^2} - \overline{X}^2\right)}$$

$$\widehat{\theta}_2 = \overline{X} + \sqrt{3\left(\overline{X^2} - \overline{X}^2\right)}$$

b. Assume $(X_1, X_2, ..., X_n)$ are a random sample from the distribution with a cdf: $F_X(x) = \begin{cases} 1 - (\theta_1/x)^{\theta_2} \text{ if } x \geq \theta_1 \\ 0 \text{ otherwise} \end{cases}$, where $\theta_1 > 0, \theta_2 > 0$. Find the Maximum Likelihood estimator of $\theta = (\theta_1, \theta_2)$

answer: The pdf is:
$f_X(x) = \frac{\partial}{\partial x} F_X(x) = \theta_1^{\theta_2} \theta_2 x^{-(\theta_2+1)}, x \geq \theta_1.$
The likelihood function is therefore:
$L(\theta_1, \theta_2; x) = \frac{\theta_1^{n\theta_2} \theta_2^n}{(x_1 \cdot x_2 \cdot ... \cdot x_n)^{(\theta_2+1)}}$
And the log-likelihood function is:
$\ln(L(\theta_1, \theta_2; x)) = n\theta_2 \ln(\theta_1) + n \ln(\theta_2) - (\theta_2 + 1) \sum_{i=1}^{n} \ln(x_i)$
To maximize this function with respect to θ_1 we would want to choose the highest possible value of θ_1. Bearing in mind the constraint: $x \geq \theta_1$ we choose: $\widehat{\theta_1} = \min_{i \in \{1,2,...,n\}}(x_i)$. This is true for any θ_2, so we can take the derivative with respect to θ_2 :
$\frac{\partial}{\partial \theta_2} \ln\left(L\left(\widehat{\theta_1}, \theta_2; x\right)\right) = n \ln\left(\widehat{\theta_1}\right) + \frac{n}{\theta_2} - \sum_{i=1}^{n} \ln(x_i) = 0$
So: $\widehat{\theta_2} = \frac{n}{\sum_{i=1}^{n} \ln\left(x_i / \widehat{\theta_1}\right)}$