Math 448. Third Midterm. Friday, April 16, 2004.

Name: ....

Show all your work. No credit for lucky answers.

1. A sample of 150 brand A light bulbs showed a mean lifetime of 1500 hours and a standard deviation of 1300 hours. A sample of 200 brand B light bulbs showed a mean lifetime of 1200 hours and a standard deviation of 80 hours. Find 95% confidence limits for the difference of the mean lifetimes of the populations of brands A and B. Assume that the observations come from a normal distribution with possible different variances.

$$\bar{y}_1 - \bar{y}_2 \pm z_{\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$= 1500 - 1200 \pm (1.96) \sqrt{\frac{(1300)^2}{150} + \frac{(80)^2}{200}}$$

$$= 300 \pm 208.34 = [91.66, 508.34]$$

2. A business school researcher is interested in the mean amount of time that hotel managers have stayed with their current employers. She plans to take a random sample of hotel managers, and ask each how long he or she has stayed with the current employer. A reasonable guess for the population variance of staying time for all hotel managers is 4. If a margin of error of 1.5 and a confidence level of 86 percent are required, how large a random sample will be needed?

$$z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} = 1.5 = 1.475$$

$$n = \left(\frac{(1.475)(2)}{1.5}\right)^2 = 3.86$$

$$\boxed{n = 4}$$

3. From a recent AP story (reported in The Capital, 20 April 2003) titled "Salmonella Prevalence Decreases":

"Of the 58,085 samples of meat and poultry tested for salmonella last year, 4.3% [2498 samples] had the germ which can cause food poisoning, the report says. That's down from two years ago when 5% [2297 samples] of the 45,941 samples of meat and poultry tested positive."

Construct a 99% confidence interval for the difference in proportions which had germs between last year's samples and two years ago samples.

$$\hat{p}_1 - \hat{p}_2 \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$= 0.043 - 0.05 \pm (2.575)\sqrt{\frac{(0.043)(1-0.043)}{58085} + \frac{(0.05)(1-0.05)}{45941}}$$

$$= -0.007 \pm 0.0034 = [-0.0104, -0.0036]$$

4. Given a random sample of size n from the density

$$f(y|\theta,\alpha) = \begin{cases} \frac{\alpha y^{\alpha-1}}{\theta^\alpha} & \text{if } 0 < y < \theta, \\ 0 & \text{else.} \end{cases}$$

where $\theta > 0$ and $\alpha > 0$ are two unknown parameters. Find two jointly sufficient statistics of θ and α.

$$L(\theta,\alpha) = \prod_{i=1}^{n} \frac{\alpha y_i^{\alpha-1} I(0<Y_i<\theta)}{\theta^\alpha} = \frac{\alpha^n \left(\prod Y_i\right)^{\alpha-1} I(0<Y_{(1)}, Y_{(n)}<\theta)}{\theta^{n\alpha}}$$

$$= \underbrace{\frac{\alpha^n I(Y_{(n)}<\theta)\left(\prod Y_i\right)^{\alpha-1}}{\theta^{n\alpha}}}_{g(u,(\theta,\alpha))} \quad \underbrace{I(0<Y_{(1)})}_{h(Y_1,\ldots,Y_n)}$$

$$u = \left(Y_{(n)}, \prod_{i=1}^{n} Y_i\right)$$

5. Given a random sample of size n from the density

$$f(y|\theta) = \begin{cases} \frac{y^2 e^{-\frac{y}{\theta}}}{2\theta^3} & \text{if } 0 < y, \\ 0 & \text{else.} \end{cases}$$

where $\theta > 0$ is unknown parameter.
(a) Find a sufficient statistic of θ.
(b) Find the MVUE of θ (find a function of the sufficient statistic which an unbiased estimator of θ).

$$\text{(a) } L(\theta) = \prod_{j=1}^{n} \left(\frac{Y_j^2 e^{-Y_j/\theta}}{2\theta^3} \right) = \frac{(\prod Y_j)^2 e^{-\sum Y_j/\theta}}{2^n \theta^{3n}} = \underbrace{\frac{e^{-\sum Y_j/\theta}}{\theta^{3n}}}_{g(\sum Y_j, \theta)} \underbrace{\frac{\prod Y_j^2}{2^n}}_{h(Y_1 \dots Y_n)}$$

$U = \sum Y_j$

$U = \sum_{j=1}^{n} Y_j$ is a sufficient statistic for θ

$$\text{(b) } E[U] = nE[Y] = n\int_0^\infty y \frac{y^2 e^{-\frac{y}{\theta}}}{2\theta^3} dy = \frac{n\theta}{2} \int_0^\infty x^3 e^{-x} dx = \frac{n\theta}{2} \cdot 3! = 3n\theta$$

$\frac{y}{\theta} = x$

$\frac{1}{3n} \sum_{j=1}^{n} Y_j$ is the MVUE of θ.

6. Given a random sample of size n from the density

$$f(y|\theta) = \begin{cases} \frac{\alpha y^{\alpha-1}}{\theta^\alpha} & \text{if } 0 < y < \theta, \\ 0 & \text{else.} \end{cases}$$

where $\theta > 0$ is unknown parameter and $\alpha > 0$ is known.
(a) Find a sufficient statistic of θ.
(b) Find the MVUE of θ^2.

(a) $L(\theta) = \prod_{j=1}^{n} \frac{\alpha Y_j^{\alpha-1}}{\theta^\alpha} I(0 < Y_j < \theta) = \frac{\alpha^n (\prod Y_j)^{\alpha-1} I(0 < Y_{(1)}, Y_{(n)} < \theta)}{\theta^{n\alpha}}$

$= \underbrace{\frac{I(Y_{(n)} < \theta)}{\theta^{n\alpha}}}_{g(u,\theta)} \quad \underbrace{\alpha^n (\prod Y_j)^{\alpha-1} I(0 < Y_{(1)})}_{h(Y_1 \cdots Y_n)}$

$U = Y_{(n)}$ is a sufficient statistic for θ

$f_{Y_{(n)}}(y) = n\,(F(y))^{n-1} f(y) = n \left(\frac{y^\alpha}{\theta^\alpha}\right)^{n-1} \frac{\alpha y^{\alpha-1}}{\theta^\alpha} = \frac{n\alpha\, y^{n\alpha-1}}{\theta^{n\alpha}}, \text{ if } 0 < y < \theta$

$E[Y_{(n)}^2] = \int_0^\theta y^2 \frac{n\alpha\, y^{n\alpha-1}}{\theta^{n\alpha}} = \frac{n\alpha}{n\alpha+2}\theta^2$

$\frac{(n\alpha+2)}{n\alpha} Y_{(n)}^2$ is the MVUE of θ^2

7. Let $Y_1, \ldots, Y_n$ denote a random sample from the density

$$f(y|\theta) = \begin{cases} \frac{y^2 e^{-\frac{y}{\theta}}}{2\theta^3} & \text{if } 0 < y, \\ 0 & \text{else.} \end{cases}$$

(a) Find the method of moments estimator of θ. Is this estimator unbiased for θ? (prove it!) Is this estimator consistent for θ? (prove it!)
(b) Find the maximum likelihood estimator of θ. Is this estimator unbiased for θ? (prove it!) Is this estimator consistent for θ? (prove it!)

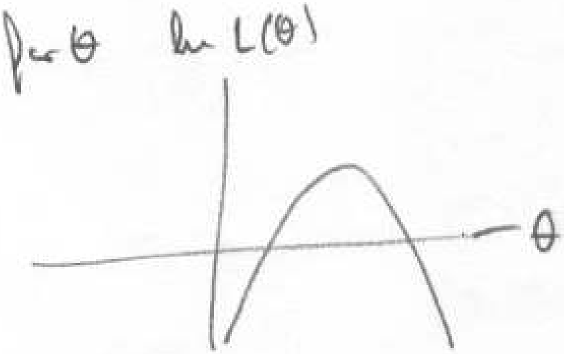
(a) $E[Y] = \int_0^\infty y \, \frac{y^2 e^{-y/\theta}}{2\theta^3} \, dy = \int_0^\infty \frac{(\theta x)^3}{2\theta^3} e^{-x} \, \theta \, dx = \frac{\theta}{2} \int_0^\infty x^3 e^{-x} \, dx = 3\theta$

$\frac{y}{\theta} = x$

$\bar{Y} = 3\theta \quad \theta = \frac{\bar{Y}}{3} \qquad \hat{\theta} = \frac{\bar{Y}}{3}$ is the MME of θ

$\hat{\theta} = \frac{\bar{Y}}{3} \xrightarrow{P} \frac{E[Y]}{3} = \theta$, $\quad \hat{\theta}$ is consistent for θ

$E[\hat{\theta}] = E[\frac{\bar{Y}}{3}] = \theta$, $\quad \hat{\theta}$ is unbiased for θ



(b) $L(\theta) = \prod_{j=1}^{n} \frac{Y_j^2 e^{-Y_j/\theta}}{2\theta^3} = \frac{\prod Y_j^2 \, e^{-\sum Y_j/\theta}}{2^n \theta^{3n}}$

$\ln L(\theta) = \sum_{j=1}^{n} \ln Y_j^2 - \frac{\sum Y_j}{\theta} - n \ln 2 - 3n \ln \theta$

$0 = \frac{\partial}{\partial \theta} \ln L(\theta) = \frac{\sum Y_j}{\theta^2} - \frac{3n}{\theta} \qquad \hat{\theta} = \frac{1}{3n} \sum_{j=1}^{n} Y_j = \frac{\bar{Y}}{3}$

Since this is the same estimator as before, the MLE is unbiased and consistent

Note that $\lim_{\theta \to 0} \ln L(\theta) = -\infty$ and $\lim_{\theta \to \infty} \ln L(\theta) = -\infty$

So $\hat{\theta}$ maximizes $\ln L(\theta)$

8. Given a random sample $Y_1, \ldots, Y_n$ of size n from the a normal density with unknown mean μ and unknown variance σ^2. Find the MSE of $\hat{\sigma}_1^2 = \frac{1}{n}\sum_{j=1}^{n}(Y_j - \bar{Y})^2$ and $\hat{\sigma}_2^2 = \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \bar{Y})^2$ as estimators of σ^2. Which estimator has a smaller MSE?
Hint: The mean and the variance of a r.v. with a $\chi^2(v)$ distribution are v and $2v$ respectively.

Let $U = \sum_{j=1}^{n} \frac{(Y_j - \bar{Y})^2}{\sigma^2} \sim \chi^2(n-1)$

$$MSE(\hat{\sigma}_1^2) = Var(\hat{\sigma}_1^2) + \left(E[\hat{\sigma}_1^2 - \sigma^2]\right)^2$$

$$= Var\left(\frac{\sigma^2 U}{n}\right) + \left(E\left[\frac{\sigma^2 U}{n} - \sigma^2\right]\right)^2$$

$$= \frac{\sigma^4}{n^2} 2(n-1) + \left(\frac{\sigma^2}{n}(n-1) - \sigma^2\right)^2 = \frac{\sigma^4 2(n-1)}{n^2} + \frac{\sigma^4}{n^2} = \frac{\sigma^4(2n-1)}{n^2}$$

$$MSE(\hat{\sigma}_2^2) = Var(\hat{\sigma}_2^2) + \left(E[\hat{\sigma}_2^2 - \sigma^2]\right)^2$$

$$= Var\left(\frac{\sigma^2 U}{n-1}\right) + \left(E\left[\frac{\sigma^2 U}{n-1} - \sigma^2\right]\right)^2 = \frac{\sigma^4}{(n-1)^2} 2(n-1) = \frac{2\sigma^4}{n-1}$$

Now

$$\sigma^4\frac{(2n-1)}{n^2} < \frac{2\sigma^4}{n-1} \iff (2n-1)(n-1) < 2n^2 \iff 2n^2 - 3n + 1 < 2n^2$$

$$\iff 1 < 3n.$$

So $\hat{\sigma}_1^2$ has smaller MSE than $\hat{\sigma}_2^2$ for each $n \geq 1$

9. Given a random sample of size n from the density

$$f(y|\theta) = \begin{cases} \frac{2ye^{-\frac{y^2}{\theta^2}}}{\theta^2} & \text{if } 0 < y, \\ 0 & \text{else.} \end{cases}$$

where $\theta > 0$ is an unknown parameter.

(a) Find the method of moments estimator of θ. Is this estimator consistent of θ? (prove it!)

(b) Find the maximum likelihood estimator of θ. Is this estimator consistent of θ? (prove it!)

(a) $E[Y] = \int_0^\infty y \frac{2y e^{-y^2/\theta^2}}{\theta^2} dy = \frac{2}{\theta^2} \int_0^\infty y^2 e^{-y^2/\theta^2} dy = \frac{2}{\theta^2} \int_0^\infty \theta^2 x e^{-x} \frac{\theta}{2\sqrt{x}} dx$

$\frac{y^2}{\theta^2} = x \quad y = \theta\sqrt{x}$

$dy = \frac{\theta}{2\sqrt{x}} dx$

$= \theta \int_0^\infty x^{1/2} e^{-x} dx = \theta\, \Gamma(\frac{3}{2}) = \theta \frac{1}{2}\sqrt{\pi}$

$\bar{Y} = \frac{\theta\sqrt{\pi}}{2} \quad \hat{\theta} = \frac{2\bar{Y}}{\sqrt{\pi}}$ is the MME of θ

$\hat{\theta}$ is consistent for θ, because

$\hat{\theta} = \frac{2\bar{Y}}{\sqrt{\pi}} \xrightarrow{P} \frac{2}{\sqrt{\pi}} E[Y] = \theta$

(b) $L(\theta) = \prod_{j=1}^n \frac{2Y_j e^{-Y_j^2/\theta^2}}{\theta^2} = \frac{2^n \prod Y_j\, e^{-\sum Y_j^2/\theta^2}}{\theta^{2n}}$

$\ln L(\theta) = n\ln 2 + \sum_{j=1}^n \ln Y_j - \frac{\sum Y_j^2}{\theta^2} - 2n \ln\theta$

$0 = \frac{\partial}{\partial\theta} \ln L(\theta) = \frac{2\sum Y_j^2}{\theta^3} - \frac{2n}{\theta}$: $\quad \hat{\theta} = \sqrt{\frac{1}{n}\sum_{j=1}^n Y_j^2}$ is the MLE of θ

$\sqrt{\frac{1}{n}\sum_{i=1}^{n} Y_i^2}$ is a maximum of $\ln L(\theta)$, because

$$\lim_{\theta \to 0} \ln L(\theta) = \lim_{\theta \to \infty} \ln L(\theta) = -\infty.$$

$$E[Y^2] = \int_0^\infty y^2 \frac{2y\, e^{-y^2/\theta^2}}{\theta^2}\, dy = \int_0^\infty \frac{2}{\theta^2} (\theta\sqrt{x})^3 e^{-x} \frac{\theta}{2\sqrt{x}}\, dx$$

$$\frac{y^2}{\theta^2} = x \quad y = \theta\sqrt{x} \quad dy = \frac{\theta}{2\sqrt{x}}\, dx$$

$$= \theta^2 \int_0^\infty x e^{-x}\, dx = \theta^2$$

$$\frac{1}{n}\sum_{i=1}^{n} Y_i^2 \xrightarrow{P} E[Y^2] = \theta^2$$

So, $\sqrt{\frac{1}{n}\sum_{i=1}^{n} Y_i^2} \xrightarrow{P} \sqrt{E[Y^2]} = \theta$

$\frac{1}{n}\sum_{i=1}^{n} Y_i^2$ is consistent for θ.

10. Let $Y_1, \ldots, Y_n$ denote a random sample of size n from an exponential distribution with mean θ. Find an asymptotic $100(1-\alpha)$ % confidence interval of $t(\theta) = \mathrm{Var}_\theta(Y)$.

The $100(1-\alpha)$ % confidence interval for $t(\theta)$ is

$$t(\hat\theta) \pm z_{\frac{\alpha}{2}} \frac{|t'(\hat\theta)|}{\sqrt{n E[-\frac{\partial^2}{\partial\theta^2} \ln f(x,\theta)]}\Big|_{\theta=\hat\theta}}$$

$$t(\theta) = \mathrm{Var}_\theta(Y) = \theta^2 \qquad t'(\theta) = 2\theta$$

$$f(y|\theta) = \frac{e^{-y/\theta}}{\theta}, \text{ if } y>0$$

$$L(\theta) = \prod_{i=1}^{n} \frac{e^{-y_i/\theta}}{\theta} = \frac{e^{-\sum y_i/\theta}}{\theta^n}$$

$$\ln L(\theta) = -\frac{\sum y_i}{\theta} - n\ln\theta$$

$$0 = \frac{\partial}{\partial\theta} \ln L(\theta) = \frac{\sum y_i}{\theta^2} - \frac{n}{\theta}$$

$\hat\theta = \bar y$ is the MLE of θ

$$\ln f(y,\theta) = \frac{-y}{\theta} - \ln\theta$$

$$\frac{\partial}{\partial\theta} \ln f(y,\theta) = \frac{y}{\theta^2} - \frac{1}{\theta} \qquad \frac{\partial^2}{\partial\theta^2} \ln f(y,\theta) = \frac{-2y}{\theta^3} + \frac{1}{\theta^2}$$

$$E[-\frac{\partial}{\partial\theta} \ln f(y,\theta)] = E[\frac{2y}{\theta^3} - \frac{1}{\theta^2}] = \frac{1}{\theta^2}$$

$$(\bar y)^2 \pm z_{\frac{\alpha}{2}} \frac{2\bar y}{\sqrt{n \frac{1}{(\bar y)^2}}} = (\bar y)^2 \pm z_{\frac{\alpha}{2}} \frac{2(\bar y)^2}{\sqrt{n}}$$