Stat 252.01 Winter 2007
Assignment #6 Solutions

1. Textbook

(9.73) If $Y_1, \ldots, Y_n$ are i.i.d. exponential(θ), then each Y_i has common density

$$f_Y(y|\theta) = \frac{1}{\theta} e^{-y/\theta}, \quad y > 0.$$

Therefore, the likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f_Y(y_i|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-y_i/\theta} = \theta^{-n} \exp\left\{-\frac{1}{\theta} \sum_{i=1}^{n} y_i\right\}.$$

The maximum likelihood estimator $\hat{\theta}_{\text{MLE}}$ is obtained by maximizing $L(\theta)$. In order to maximize $L(\theta)$ we will attempt to maximize the log-likelihood function $\ell(\theta)$ given by

$$\ell(\theta) = -n \log \theta - \frac{1}{\theta} \sum_{i=1}^{n} y_i.$$

Since

$$\ell'(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} y_i$$

we find $\ell'(\theta) = 0$ when

$$-\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} y_i = 0 \quad \text{or} \quad \theta = \frac{1}{n} \sum_{i=1}^{n} y_i.$$

Finally

$$\ell''(\theta) = \frac{n}{\theta^2} - \frac{2}{\theta^3} \sum_{i=1}^{n} y_i$$

so that

$$\ell''\left(\frac{1}{n} \sum y_i\right) = -\frac{n^3}{\left(\sum y_i\right)^2} < 0.$$

By the second derivative test we conclude,

$$\hat{\theta}_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^{n} Y_i = \overline{Y}.$$

Since $\theta > 0$, we see that the function $T(\theta) = \theta^2$ is one-to-one. Therefore, the maximum likelihood estimator of θ^2 is

$$\hat{\theta}^2_{\text{MLE}} = \overline{Y}^2.$$

(9.74) (b) If $Y_1, \ldots, Y_n$ are a random sample from the density function

$$f_Y(y|\theta) = \frac{1}{\theta} r y^{r-1} e^{-y^r/\theta}, \quad y > 0$$

where $\theta > 0$ is a parameter, then the likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f_Y(y_i|\theta) = \prod_{i=1}^{n} \frac{1}{\theta} r y_i^{r-1} e^{-y_i^r/\theta} = \theta^{-n} \cdot r^n \cdot \left(\prod_{i=1}^{n} y_i \right)^{r-1} \cdot \exp\left\{ -\frac{1}{\theta} \sum_{i=1}^{n} y_i^r \right\}.$$

The maximum likelihood estimator $\hat{\theta}_{\text{MLE}}$ is obtained by maximizing $L(\theta)$. In order to maximize $L(\theta)$ we will attempt to maximize the log-likelihood function $\ell(\theta)$ given by

$$\ell(\theta) = -n \log \theta + n \log r + (r-1) \sum_{i=1}^{n} \log y_i - \frac{1}{\theta} \sum_{i=1}^{n} y_i^r.$$

Since

$$\ell'(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} y_i^r$$

we find $\ell'(\theta) = 0$ when

$$-\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} y_i^r = 0 \quad \text{or} \quad \theta = \frac{1}{n} \sum_{i=1}^{n} y_i^r.$$

Finally

$$\ell''(\theta) = \frac{n}{\theta^2} - \frac{2}{\theta^3} \sum_{i=1}^{n} y_i^r$$

so that

$$\ell''\left(\frac{1}{n} \sum y_i^r \right) = -\frac{n^3}{\left(\sum y_i^r\right)^2} < 0.$$

By the second derivative test we conclude,

$$\hat{\theta}_{\text{MLE}} = \frac{1}{n} \sum_{i=1}^{n} Y_i^r.$$

(9.75) (a) Suppose that $Y_1, \ldots, Y_n$ are i.i.d. Uniform$(0, 2\theta + 1)$ so that each Y_i has density function

$$f_Y(y|\theta) = \frac{1}{2\theta + 1}, \quad 0 \le y \le 2\theta + 1.$$

The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f_Y(y_i|\theta) = (2\theta + 1)^{-n}$$

provided that $0 \le y_i \le 2\theta + 1$ for each $i = 1, \ldots, n$. In other words,

$$L(\theta) = \begin{cases} (2\theta + 1)^{-n}, & 0 \le \max\{y_1, \ldots, y_n\} \le 2\theta + 1, \\ 0, & \text{otherwise.} \end{cases}$$

The maximum likelihood estimator $\hat{\theta}_{\text{MLE}}$ is obtained by maximizing $L(\theta)$. Since $(2\theta + 1)^{-n}$ is strictly decreasing in θ provided that $0 \le \max\{y_1, \ldots, y_n\} \le 2\theta + 1$, we see that the maximum value is obtained when θ is chosen as small as possible subject to the constraint $\max\{y_1, \ldots, y_n\} \le 2\theta + 1$. Thus, the maximum likelihood estimator is

$$\hat{\theta}_{\text{MLE}} = \frac{\max\{Y_1, \ldots, Y_n\} - 1}{2}.$$

(9.76) **(a)** Suppose that Y_1, Y_2, Y_3 are i.i.d. Gamma$(2, \theta)$ random variables so that each has density

$$f_Y(y|\theta) = \frac{1}{\theta^2} y e^{-y/\theta}, \quad y > 0$$

Therefore, the likelihood function is

$$L(\theta) = \prod_{i=1}^{3} f_Y(y_i|\theta) = \prod_{i=1}^{3} \frac{1}{\theta^2} y_i e^{-y_i/\theta} = \theta^{-6} \cdot \prod_{i=1}^{3} y_i \cdot \exp\left\{ -\frac{1}{\theta} \sum_{i=1}^{3} y_i \right\}.$$

The maximum likelihood estimator $\hat{\theta}_{\text{MLE}}$ is obtained by maximizing $L(\theta)$. In order to maximize $L(\theta)$ we will attempt to maximize the log-likelihood function $\ell(\theta)$ given by

$$\ell(\theta) = -6 \log \theta + \sum_{i=1}^{3} \log y_i - \frac{1}{\theta} \sum_{i=1}^{3} y_i.$$

Since

$$\ell'(\theta) = -\frac{6}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{3} y_i$$

we find $\ell'(\theta) = 0$ when

$$-\frac{6}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{3} y_i = 0 \quad \text{or} \quad \theta = \frac{1}{6} \sum_{i=1}^{3} y_i = \frac{\overline{y}}{2}.$$

Finally

$$\ell''(\theta) = \frac{6}{\theta^2} - \frac{2}{\theta^3} \sum_{i=1}^{3} y_i$$

so that

$$\ell''\left(\frac{1}{6} \sum y_i \right) = -\frac{6^3}{\left(\sum y_i\right)^2} < 0.$$

By the second derivative test we conclude,

$$\hat{\theta}_{\text{MLE}} = \frac{1}{6} \sum_{i=1}^{3} Y_i = \frac{\overline{Y}}{2}.$$

Based on the observed data, we find that the maximum likelihood estimate of θ is

$$\hat{\theta}_{\text{MLE}} = \frac{120 + 130 + 128}{6} = 63.$$

(b) Since each Y_1, Y_2, Y_3 are i.i.d. Gamma$(2, \theta)$, and since

$$\hat{\theta}_{\text{MLE}} = \frac{1}{6} \sum_{i=1}^{3} Y_i = \frac{\overline{Y}}{2}$$

we conclude

$$E(\hat{\theta}_{\text{MLE}}) = \frac{1}{6} \sum_{i=1}^{3} E(Y_i) = \frac{3 \cdot 2\theta}{6} = \theta$$

and

$$\mathrm{Var}(\hat{\theta}_{\mathrm{MLE}}) = \frac{1}{6^2}\sum_{i=1}^{3}\mathrm{Var}(Y_i) = \frac{3\cdot 2\theta^2}{6^2} = \frac{\theta^2}{6}.$$

(c) If $\theta = 130$, then an approximate bound for the error of estimation is given by

$$2\sqrt{\mathrm{Var}(\hat{\theta}_{\mathrm{MLE}})} = 2\sqrt{\frac{\theta^2}{6}} = 2\sqrt{\frac{130^2}{6}} \approx 106.14.$$

(d) Since the variance of Y is $\mathrm{Var}(Y) = 2\theta^2$, we conclude that the MLE of $\mathrm{Var}(Y)$ is

$$\hat{\mathrm{Var}(Y)}_{\mathrm{MLE}} = 2\hat{\theta}^2_{\mathrm{MLE}} = 2(63)^2 = 7938.$$

(9.77) (a) Suppose that $Y_1, \ldots, Y_n$ are a random sample from the density function

$$f_Y(y|\theta) = \frac{1}{\Gamma(\alpha)\theta^\alpha}y^{\alpha-1}e^{-y/\theta}, \quad y > 0$$

where $\alpha > 0$ is known. Therefore, the likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f_Y(y_i|\theta) = \prod_{i=1}^{n}\frac{1}{\Gamma(\alpha)\theta^\alpha}y_i^{\alpha-1}e^{-y_i/\theta} = \theta^{-n\alpha}\cdot\Gamma(\alpha)^{-n}\cdot\left(\prod_{i=1}^{n} y_i\right)^{\alpha-1}\cdot\exp\left\{-\frac{1}{\theta}\sum_{i=1}^{n} y_i\right\}.$$

The maximum likelihood estimator $\hat{\theta}_{\mathrm{MLE}}$ is obtained by maximizing $L(\theta)$. In order to maximize $L(\theta)$ we will attempt to maximize the log-likelihood function $\ell(\theta)$ given by

$$\ell(\theta) = -n\alpha\log\theta - n\log\Gamma(\alpha) + (\alpha-1)\sum_{i=1}^{n}\log y_i - \frac{1}{\theta}\sum_{i=1}^{n} y_i.$$

Since

$$\ell'(\theta) = -\frac{n\alpha}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} y_i$$

we find $\ell'(\theta) = 0$ when

$$-\frac{n\alpha}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} y_i = 0 \quad \text{or} \quad \theta = \frac{1}{n\alpha}\sum_{i=1}^{n} y_i = \frac{\overline{y}}{\alpha}.$$

Finally

$$\ell''(\theta) = \frac{n\alpha}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} y_i$$

so that

$$\ell''\left(\frac{1}{n\alpha}\sum y_i\right) = -\frac{n^3\alpha^3}{(\sum y_i)^2} < 0$$

since $\alpha > 0$. By the second derivative test we conclude,

$$\hat{\theta}_{\mathrm{MLE}} = \frac{1}{n\alpha}\sum_{i=1}^{n} Y_i = \frac{\overline{Y}}{\alpha}.$$

3. (a) It is highly unlikely that the i.i.d. assumption is reasonable. In order to postulate i.i.d. $\text{Bin}(k, p)$, she is assuming that each animal has the same probability of being trapped. This is doubtful both within a species and between species. (Are some animals "dumber" and others "smarter"? What about different species? Are some more cautious than others?) This is also doubtful because animals are likely to get "smarter" after being trapped once. (Think of any Pavlovian experiment.) The independent trials assumption is also dubious. Is it reasonable to assume that animals do not warn others of the danger of the trap? Probably not.

(b) For a $\text{Bin}(k, p)$ random variable Y, we have $\mathbb{E}(Y) = kp$ and $\mathbb{E}(Y^2) = \text{Var}(Y) + [\mathbb{E}(Y)]^2 = kp(1-p)+k^2p^2$. The method of moments system implies that $\hat{\mu}_1 = kp$ and $\hat{\mu}_2 = kp(1-p)+k^2p^2$. Solving gives

$$\hat{p}_{\text{MOM}} = 1 - \frac{\hat{\mu}_2 - (\hat{\mu}_1)^2}{\hat{\mu}_1}$$

and

$$\hat{k}_{\text{MOM}} = \frac{\hat{\mu}_1}{\hat{p}_{\text{MOM}}}.$$

The data yield $\hat{\mu}_1 = 12.6$ and $\hat{\mu}_2 = 163$. Thus,

$$\hat{p}_{\text{MOM}} = \frac{209}{315} \approx 0.663 \quad \text{and} \quad \hat{k}_{\text{MOM}} = \frac{3969}{209} \approx 19.$$

(c) In this case the data yield $\hat{\mu}_1 = 11.2$ and $\hat{\mu}_2 = 139.2$ which give

$$\hat{p}_{\text{MOM}} = \frac{-8}{35} \approx -0.229 \quad \text{and} \quad \hat{k}_{\text{MOM}} = -49.$$

These are nonsensical estimates since we require $p \in [0, 1]$ and $k > 0$. Clearly if these were the data observed, the postulate of a binomial distribution would definitely be cast into serious doubt!