MAS programmes - Statistical Inference

Solutions 3

1.

$$L(\theta; \mathbf{x}) = \frac{\prod x_i^{\nu-1} \exp(-\sum x_i/\theta)}{\theta^{\nu n}[\Gamma(\nu)]^n} \qquad x > 0,$$

(i)

$$\ln L(\theta; \mathbf{x}) = \text{const} - \frac{n\bar{x}}{\theta} - \nu n \ln \theta$$

and hence

$$\frac{\partial \ln L(\theta; \mathbf{x})}{\partial \theta} = \frac{n\bar{x}}{\theta^2} - \frac{\nu n}{\theta}.$$

The MLE $\hat{\theta}$ is given by the solution of the likelihood equation,

$$\frac{n\bar{x}}{\theta^2} - \frac{\nu n}{\theta} = 0.$$

Thus $\hat{\theta} = \bar{X}/\nu$.

(ii) Note that the mean and variance of the gamma distribution are $\nu\theta$ and $\nu\theta^2$, respectively. Hence

$$E[\hat{\theta}] = E\left[\frac{\bar{X}}{\nu}\right] = \frac{\nu\theta}{\nu} = \theta.$$

Thus $\hat{\theta}$ is an unbiased estimator of θ.

$$\text{var}(\hat{\theta}) = \text{var}\left(\frac{\bar{X}}{\nu}\right) = \frac{1}{\nu^2}\frac{\nu\theta^2}{n} = \frac{\theta^2}{\nu n}.$$

(iii)

$$-\frac{\partial^2 \ln L(\theta; \mathbf{x})}{\partial \theta^2} = \frac{2n\bar{x}}{\theta^3} - \frac{\nu n}{\theta^2}.$$

$$I(\theta) = E\left[-\frac{\partial^2 \ln L(\theta; \mathbf{x})}{\partial \theta^2}\right] = \frac{2n\nu\theta}{\theta^3} - \frac{\nu n}{\theta^2} = \frac{\nu n}{\theta^2}.$$

(iv) We apply the Cramer-Rao inequality: under suitable regularity conditions, satisfied in the present case, if $\hat{\theta}$ is an unbiased estimator of θ then $\text{var}(\hat{\theta}|\theta) \geq 1/I(\theta)$, where $I(\theta)$ is the Fisher information.

In our case

$$\text{var}(\hat{\theta}) = \frac{\theta^2}{\nu n} = \frac{1}{I(\theta)},$$

so that the lower bound is attained and hence $\hat{\theta}$ is the MVUE.

2. (i) The likelihood function is given by

$$L(\theta;\mathbf{x}) = f(\mathbf{x};\theta) = \left(\frac{\theta^2}{1+\theta}\right)^n e^{-\theta n\bar{x}} \prod_{i=1}^{n}(1+x_i) \qquad \theta > 0.$$

Use the factorization criterion that a statistic T is sufficient for θ if and only if there exist functions $g(t;\theta)$ and $h(\mathbf{x})$ such that $f(\mathbf{x};\theta) = g(T(\mathbf{x});\theta)h(\mathbf{x})$. In the present case take $g(\bar{x};\theta) = [\theta^2/(1+\theta)]^n e^{-\theta n\bar{x}}$ and $h(\mathbf{x}) = \prod_{i=1}^{n}(1+x_i)$ to show that $\bar{x}$ is a sufficient statistic.

(ii) The mean of the distribution specified in the question is given by

$$\int_0^\infty x f(x;\theta)dx = \frac{\theta^2}{1+\theta}\int_0^\infty (x+x^2)e^{-\theta x}dx = \frac{\theta^2}{1+\theta}\left(\frac{1}{\theta^2}+\frac{2}{\theta^3}\right) = \frac{2+\theta}{\theta(1+\theta)} \ .$$

Hence the method of moments estimate is given by

$$\frac{2+\hat{\theta}}{\hat{\theta}(1+\hat{\theta})} = \bar{x} \ ,$$

i.e.,

$$\hat{\theta}^2 + \hat{\theta}\left(1-\frac{1}{\bar{x}}\right) - \frac{2}{\bar{x}} = 0 \ .$$

This equation has a unique positive root,

$$\hat{\theta} = -\frac{1}{2}\left(1-\frac{1}{\bar{x}}\right) + \frac{1}{2}\sqrt{1+\frac{6}{\bar{x}}+\frac{1}{\bar{x}^2}} \ ,$$

which specifies the method of moments estimator.

(iii)

$$\ln L(\theta;\mathbf{x}) = 2n\ln\theta - n\ln(1+\theta) - \theta n\bar{x} + \sum_{i=1}^{n}\ln(1+x_i) \ .$$

Hence

$$\frac{\partial \ln L}{\partial\theta} = \frac{2n}{\theta} - \frac{n}{1+\theta} - n\bar{x}$$

and the likelihood equation is given by

$$\frac{2n}{\theta} - \frac{n}{1+\theta} - n\bar{x} = 0 \ ,$$

i.e.,

$$\theta^2 + \theta\left(1-\frac{1}{\bar{x}}\right) - \frac{2}{\bar{x}} = 0 \ ,$$

which is exactly the same equation as the one satisfied by the method of moments estimator.

To check that the positive solution of the likelihood equation does indeed maximize the likelihood, we note that

$$\frac{\partial^2 \ln L}{\partial\theta^2} = -\frac{2n}{\theta^2} + \frac{n}{(1+\theta)^2} = -\frac{n(\theta^2+4\theta+2)}{\theta^2(1+\theta)^2} < 0 \qquad \theta > 0.$$

Hence the log-likelihood function is concave on $(0,\infty)$ and its unique maximum is given by the positive solution of the likelihood equation.

(iv) The Fisher information is given by

$$I(\theta) = E\left[-\frac{\partial^2 \ln L}{\partial \theta^2}\right] = \frac{n(\theta^2 + 4\theta + 2)}{\theta^2(1+\theta)^2} \,.$$

Maximum likelihood estimators are consistent, asymptotically efficient and asymptotically normally distributed. Hence, approximately, for large n,

$$\hat{\theta} \sim N\left(\theta, \frac{1}{I(\theta)}\right) \sim N\left(\theta, \frac{\theta^2(1+\theta)^2}{n(\theta^2 + 4\theta + 2)}\right) \,.$$

3. (i)

$$\begin{aligned} f(y;p) &= p^y(1-p)^{1-y} = \exp\{y \log p + (1-y)\log(1-p)\} \\ &= \exp\left\{y \log\left(\frac{p}{1-p}\right) + \log(1-p)\right\}, \end{aligned}$$

which has the form of the exponential family with $h(x) = 1$, $A(p) = -\log(1-p)$, $\eta_1(p) = \log(p/(1-p))$ and $t_1(y) = y$. We deduce that the canonical statistic is $\sum_{i=1}^n t_1(Y_i) = \sum_{i=1}^n Y_i$.

(ii) The likelihood is

$$L(p; \mathbf{y}) = \prod_{i-1}^{n} p^{y_i}(1-p)^{1-y_i} = p^{\sum y_i}(1-p)^{n-\sum y_i},$$

and so the log-likelihood is

$$\log L(p; \mathbf{y}) = \sum y_i \log p + \left(n - \sum y_i\right)\log(1-p).$$

Differentiating with respect to p and setting to 0 gives the MLE as $\hat{p} = \frac{1}{n}\sum Y_i$ (and can check is indeed maximum by seeing that the 2nd derivative is negative). This estimator has variance

$$\text{Var}(\frac{1}{n}\sum Y_i) = \frac{1}{n}p(1-p).$$

The Cramér-Rao bound is (in this case $\tau(p) = p$) $1/I(p)$ where

$$\begin{aligned} I(p) = \mathbb{E}\left[-\frac{\partial^2}{\partial p^2}\log f(\mathbf{Y};p)\right] &= \mathbb{E}\left[-\left(-\frac{\sum Y_i}{p^2} - \frac{(n - \sum Y_i)}{(1-p)^2}\right)\right] \\ &= \frac{1}{p^2}n\mathbb{E}[Y_1] + \frac{1}{(1-p)^2}(n - n\mathbb{E}[Y_1]) \\ &= \frac{n}{p} + \frac{n(1-p)}{(1-p)^2} \\ &= \frac{n}{p(1-p)}. \end{aligned}$$

So the C-R bound is $p(1-p)/n$ but this is the variance of the MLE of p, so the MLE attains the bound.

(iii) By independence $\mathbb{E}[Y_1Y_2Y_3] = \mathbb{E}[Y_1]\mathbb{E}[Y_2]\mathbb{E}[Y_3] = p^3$, and so is unbiased.

(iv) If $y \in \{0, 1, 2\}$ and we condition on $\sum_{i=1}^n Y_i = y$ then at least one of the Y_i for $i \in \{1, 2, 3\}$ must be 0. So in this case the conditional expectation is 0.
If $y \geq 3$ then

$$\begin{aligned}
\mathbb{E}\left[Y_1Y_2Y_3 \mid \sum_{i=1}^n Y_i = y\right] &= 1 \times \mathbb{P}\left(Y_1Y_2Y_3 = 1 \mid \sum_{i=1}^n Y_i = y\right) + 0 \\
&= \frac{\mathbb{P}\left(Y_1Y_2Y_3 = 1,\, \sum_{i=1}^n Y_i = y\right)}{\mathbb{P}\left(\sum_{i=1}^n Y_i = y\right)} \\
&= \frac{\mathbb{P}\left(Y_1Y_2Y_3 = 1,\, \sum_{i=4}^n Y_i = y-3\right)}{\mathbb{P}\left(\sum_{i=1}^n Y_i = y\right)} \\
&= p^3 \frac{\mathbb{P}\left(\sum_{i=4}^n Y_i = y-3\right)}{\mathbb{P}\left(\sum_{i=1}^n Y_i = y\right)} \\
&= p^3 \frac{\binom{n-3}{y-3} p^{y-3}(1-p)^{n-y}}{\binom{n}{y} p^y (1-p)^{n-y}} \\
&= \frac{\binom{n-3}{y-3}}{\binom{n}{y}},
\end{aligned}$$

where we have used that $\sum_{i=1}^n Y_i \sim \text{Bin}(n, p)$ and $\sum_{i=4}^n Y_i \sim \text{Bin}(n-3, p)$.

(v) The Bernoulli distribution is from the exponential family and so any unbiased estimator that is a function of the canonical sufficient statistic is the unique MVUE. So we calculate $\hat{\phi} = \mathbb{E}[\hat{p}|\, T]$ where $\hat{p} = Y_1Y_2Y_3$ and $T = \sum_{i=1}^n Y_i$ and use the Rao-Blackwell theorem. By part (iv) this is

$$\frac{\binom{n-3}{T-3}}{\binom{n}{T}} I_{[3,\infty)}(T) = \frac{T(T-1)(T-2)}{n(n-1)(n-2)}.$$

4. (i)

$$f(y; p) = \binom{k}{y} p^y (1-p)^{k-y} = \binom{k}{y} \exp\left\{ y \log\left(\frac{p}{1-p}\right) + k\log(1-p)\right\},$$

which has the form of the exponential family with $h(y) = \binom{k}{y}$, $A(p) = -k\log(1-p)$, $\eta_1(p) = \log(p/(1-p))$ and $t_1(y) = y$. We deduce that the canonical statistic is $T = \sum_{i=1}^n Y_i$.

(ii) $\hat{\tau}(\mathbf{Y}) = I_{\{Y_1=1\}}$ is unbiased since $\mathbb{E}[\hat{\tau}(\mathbf{Y})] = \mathbb{P}(Y_1 = 1) = \tau(p)$.

(iii) For $t \in \{0, \ldots, nk\}$ we have

$$\begin{aligned}
\mathbb{E}[\hat{\tau}(\mathbf{Y}) \mid T = t] &= \mathbb{E}\left[I_{\{Y_1=1\}} \mid \sum_{i=1}^{n} Y_i = t \right] \\
&= \mathbb{P}\left(Y_1 = 1 \mid \sum_{i=1}^{n} Y_i = t \right) \\
&= \frac{\mathbb{P}(Y_1 = 1,\ \sum_{i=1}^{n} Y_i = t)}{\mathbb{P}(\sum_{i=1}^{n} Y_i = t)} \\
&= \frac{\mathbb{P}(Y_1 = 1,\ \sum_{i=2}^{n} Y_i = t - 1)}{\mathbb{P}(\sum_{i=1}^{n} Y_i = t)} \\
&= \frac{\mathbb{P}(Y_1 = 1)\mathbb{P}(\sum_{i=2}^{n} Y_i = t - 1)}{\mathbb{P}(\sum_{i=1}^{n} Y_i = t)} \\
&= \frac{kp(1-p)^{k-1}\binom{k(n-1)}{t-1}p^{t-1}(1-p)^{k(n-1)-(t-1)}}{\binom{kn}{t}p^t(1-p)^{kn-t}} \\
&= k\frac{\binom{k(n-1)}{t-1}}{\binom{kn}{t}},
\end{aligned}$$

since $\sum_{i=1}^{n} Y_i \sim \mathrm{Bin}(kn, p)$ and $\sum_{i=2}^{n} Y_i \sim \mathrm{Bin}(k(n-1), p)$.

(iv) The Binomial distribution is from the exponential family and so any unbiased estimator that is a function of the canonical sufficient statistic is the unique MVUE. So we calculate $\hat{\phi} = \mathbb{E}[\hat{\tau} | T]$ and use the Rao-Blackwell theorem. By part (iii) this is

$$k\frac{\binom{k(n-1)}{T-1}}{\binom{kn}{T}}.$$