MTH5131 Actuarial Statistics

Coursework 2 — Solutions

Exercise 1. 1. The likelihood is

$$
\begin{aligned}
L(\pi; y_1, \dots, y_n) &= P(Y_1 = y_1, \dots, Y_n = y_n) \\
&= \prod_{i=1}^{n} P(Y_i = y_i) \\
&= \prod_{i=1}^{n} \binom{m}{y_i} \pi^{y_i} (1-\pi)^{m-y_i} \\
&= \pi^{\sum_{i=1}^{n} y_i} (1-\pi)^{mn - \sum_{i=1}^{n} y_i} \prod_{i=1}^{n} \binom{m}{y_i}.
\end{aligned}
$$

2. We may write

$$
\begin{aligned}
E\left(\frac{\overline{Y}}{m}\right) &= \frac{1}{mn} E\left(\sum_{i=1}^{n} Y_i\right) \\
&= \frac{1}{mn} mn\pi = \pi,
\end{aligned}
$$

and so $\overline{Y}/m$ is an unbiased estimator of π. Also, by independence,

$$
\begin{aligned}
\operatorname{var}\left(\frac{\overline{Y}}{m}\right) &= \frac{1}{(mn)^2} \operatorname{var}\left(\sum_{i=1}^{n} Y_i\right) \\
&= \frac{1}{(mn)^2} mn\pi(1-\pi) = \frac{\pi(1-\pi)}{mn}.
\end{aligned}
$$

Since

$$
\begin{aligned}
E\left(\frac{\overline{Y}(m-\overline{Y})}{m^2}\right) &= E\left(\frac{\overline{Y}}{m}\right) - E\left(\frac{\overline{Y}^2}{m^2}\right) \\
&= E\left(\frac{\overline{Y}}{m}\right) - E\left(\frac{\overline{Y}}{m}\right)^2 - \left(E\left(\frac{\overline{Y}^2}{m^2}\right) - E\left(\frac{\overline{Y}}{m}\right)^2\right) \\
&= E\left(\frac{\overline{Y}}{m}\right) - E\left(\frac{\overline{Y}}{m}\right)^2 - \operatorname{Var}\left(\frac{\overline{Y}}{m}\right) \\
&= \pi - \pi^2 - \frac{\pi(1-\pi)}{mn} = \frac{(mn-1)}{mn}\pi(1-\pi),
\end{aligned}
$$

it follows that

$$
E\left(\frac{1}{(mn-1)} \frac{\overline{Y}(m-\overline{Y})}{m^2}\right) = \frac{\pi(1-\pi)}{mn},
$$

and so $c = 1/(mn-1)$.

Exercise 2. 1. We may write

$$E(\overline{Y}) = \frac{1}{n}\left(E(Y_1) + \ldots + E(Y_n)\right) = \mu,$$

and so $\overline{Y}$ is an unbiased estimator of μ.

2. By independence, we have

$$\text{var}(\overline{Y}) = \frac{1}{n^2}\left(\text{var}(Y_1) + \ldots + \text{var}(Y_n)\right) = \frac{\sigma^2}{n}.$$

Thus, as $n \to \infty$, $\text{var}(\overline{Y}) \to 0$. Since $\text{bias}(\overline{Y}) = 0$, MSE$(\overline{Y}) = \text{var}(\overline{Y}) \to 0$ and it follows that $\overline{Y}$ is consistent for μ.

3. We may write

$$\begin{aligned}
E\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2\right) &= E\left(\frac{1}{n}\sum_{i=1}^{n}(Y_i^2 - 2Y_i\overline{Y} + \overline{Y}^2)\right) \\
&= \frac{1}{n}\sum_{i=1}^{n} E(Y_i^2) - 2E(\overline{Y}^2) + E(\overline{Y}^2) \\
&= \frac{1}{n}\sum_{i=1}^{n} E(Y_i^2) - E(\overline{Y}^2) \\
&= E(Y_1^2) - \mu^2 + \mu^2 - (E(\overline{Y}^2) - E(\overline{Y})^2) - E(\overline{Y})^2 \\
&= \text{Var}(Y_1) + \mu^2 - \text{Var}(\overline{Y}) - \mu^2 \\
&= \sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2 \\
&= \frac{n-1}{n}\sigma^2 \neq \sigma^2,
\end{aligned}$$

and so it is a biased estimator of σ^2.

Exercise 3.

1. We have

$$E(T_n) = \frac{E(X) + 1}{n+2} = \frac{n\theta + 1}{n+2} \neq \theta.$$

$$\text{bias}(T_n) = E(T_n) - \theta = \frac{n\theta + 1}{n+2} - \theta = \frac{1 - 2\theta}{n+2}.$$

Since $\text{bias}(T_n) \neq 0$, T_n is biased.

2.

$$\text{Var}(T_n) = \frac{1}{(n+2)^2}\text{Var}(X) = \frac{n\theta(1-\theta)}{(n+2)^2}$$

and so

$$\begin{aligned}
\text{MSE}(T_n) &= \text{Var}(T_n) + (\text{bias}(T_n))^2 \\
&= \frac{n\theta(1-\theta)}{(n+2)^2} + \frac{(1-2\theta)^2}{(n+2)^2} \\
&= \frac{n\theta(1-\theta) + (1-2\theta)^2}{(n+2)^2}.
\end{aligned}$$

3. As $\lim_{n\to\infty} \mathrm{MSE}(T_n) = 0$, the sequence of estimators T_n are consistent.

Exercise 4. 1. The likelihood is

$$\begin{aligned} L(\theta; \underline{y}) &= \prod_{i=1}^{n} f_{Y_i}(y_i) \\ &= \prod_{i=1}^{n} \frac{1}{\theta} y_i^{\frac{1}{\theta}-1} \\ &= \frac{1}{\theta^n} \left(\prod_{i=1}^{n} y_i \right)^{\frac{1}{\theta}-1}. \end{aligned}$$

2. Using integration by parts, we may write

$$\begin{aligned} E(\log Y) &= \frac{1}{\theta} \int_0^1 (\log y) y^{\frac{1}{\theta}-1} dy \\ &= \left[(\log y) y^{\frac{1}{\theta}} \right]_0^1 - \int_0^1 y^{\frac{1}{\theta}-1} dy \\ &= -\theta \left[y^{\frac{1}{\theta}} \right]_0^1 = -\theta. \end{aligned}$$

It follows that

$$\begin{aligned} E\left(-\frac{1}{n} \sum_{i=1}^{n} \log(Y_i) \right) &= -\frac{1}{n} E\left(\sum_{i=1}^{n} \log(Y_i) \right) \\ &= \frac{1}{n} n\theta = \theta, \end{aligned}$$

and so $-\sum_{i=1}^{n} \log(Y_i)/n$ is an unbiased estimator of θ.

3. Let $g(\theta) = \theta$. Then we have $dg/d\theta = 1$. The log-likelihood is

$$\log L(\theta; \underline{y}) = -n \log \theta + \left(\frac{1}{\theta} - 1 \right) \sum_{i=1}^{n} \log(y_i).$$

Thus, we have

$$\frac{d \log L(\theta; \underline{y})}{d\theta} = -\frac{n}{\theta} - \frac{1}{\theta^2} \sum_{i=1}^{n} \log(y_i)$$

and

$$\frac{d^2 \log L(\theta; \underline{y})}{d\theta^2} = \frac{n}{\theta^2} + \frac{2}{\theta^3} \sum_{i=1}^{n} \log(y_i).$$

It follows that

$$\begin{aligned} E\left(-\frac{d^2 \log L(\theta; \underline{Y})}{d\theta^2} \right) &= -\frac{n}{\theta^2} - \frac{2}{\theta^3} E\left(\sum_{i=1}^{n} \log(Y_i) \right) \\ &= -\frac{n}{\theta^2} + \frac{2}{\theta^3} n\theta = \frac{n}{\theta^2}, \end{aligned}$$

and hence $\mathrm{CRLB}(\theta) = \theta^2/n$.

4. By independence, we have

$$\begin{aligned} \operatorname{var}\left(-\frac{1}{n}\sum_{i=1}^{n}\log(Y_i)\right) &= \frac{1}{n^2}\operatorname{var}\left(\sum_{i=1}^{n}\log(Y_i)\right) \\ &= \frac{1}{n^2}n\theta^2 = \frac{\theta^2}{n}. \end{aligned}$$

Since this is equal to the CRLB, $-\sum_{i=1}^{n}\log(Y_i)/n$ is a minimum variance unbiased estimator of θ.

Exercise 5. We have $\mu_1' = E(Y) = \exp(\theta + 1/2)$ and $m_1' = \overline{Y}$. So the method of moments estimator of θ satisfies the equation

$$\exp\left(\tilde{\theta} + \frac{1}{2}\right) = \overline{Y}.$$

Thus, we obtain $\tilde{\theta} = \log(\overline{Y}) - 1/2$.§

Exercise 6. The method of moments estimators of μ and σ^2 satisfy the equations

$$e^{\tilde{\mu}+\frac{\tilde{\sigma}^2}{2}} = \overline{Y} \quad \text{and} \quad e^{2(\tilde{\mu}+\tilde{\sigma}^2)} = \frac{1}{n}\sum_{i=1}^{n}Y_i^2.$$

By dividing the second equation by the square of the first, we obtain

$$e^{\tilde{\sigma}^2} = \frac{\sum_{i=1}^{n}Y_i^2}{n\overline{Y}^2} \Rightarrow e^{2\tilde{\mu}} = \frac{n\overline{Y}^4}{\sum_{i=1}^{n}Y_i^2}$$

So the method of moments estimators are given by

$$\tilde{\mu} = \frac{1}{2}\log\left(\frac{n\overline{Y}^4}{\sum_{i=1}^{n}Y_i^2}\right) \quad \text{and} \quad \tilde{\sigma}^2 = \log\left(\frac{\sum_{i=1}^{n}Y_i^2}{n\overline{Y}^2}\right).$$

Exercise 7.

1. We have $\mu_1' = E(Y) = m\pi$ and $m_1' = \overline{Y}$. Hence, the method of moments estimator of π satisfies the equation $m\tilde{\pi} = \overline{Y}$, so that $\tilde{\pi} = \overline{Y}/m$.

2. The likelihood is

$$\begin{aligned} L(\pi;\underline{y}) &= \prod_{i=1}^{n}\binom{m}{y_i}\pi^{y_i}(1-\pi)^{m-y_i} \\ &= \pi^{\sum_{i=1}^{n}y_i}(1-\pi)^{mn-\sum_{i=1}^{n}y_i}\prod_{i=1}^{n}\binom{m}{y_i}, \end{aligned}$$

and so the log-likelihood is

$$\ell(\pi;\underline{y}) = \sum_{i=1}^{n}\log\binom{m}{y_i} + \sum_{i=1}^{n}y_i\log\pi + \left(mn - \sum_{i=1}^{n}y_i\right)\log(1-\pi).$$

Thus, solving the equation

$$\frac{d\ell}{d\pi} = \frac{\sum_{i=1}^{n}y_i}{\pi} - \frac{mn - \sum_{i=1}^{n}y_i}{1-\pi} = 0,$$

we obtain the maximum likelihood estimator of π as $\hat{\pi} = \overline{Y}/m$.

3. The Maximum Likelihood estimator is asymptotically normally distributed, with mean π and variance given by $I^{-1}(\pi) = \frac{\pi(1-\pi)}{nm}$.
(Note: $E(\hat{\pi}) = \pi$ and $\mathrm{Var}(\hat{\pi}) = \frac{\pi(1-\pi)}{nm}$, which is consistent with the parameters of the normal distribution just given.)

Exercise 8. 1. The likelihood is

$$L(\theta; \underline{y}) = \prod_{i=1}^{n} e^{-(y_i - \theta)} = e^{-\sum_{i=1}^{n} y_i + n\theta}, \quad \min_i y_i \geq \theta.$$

Since $L(\theta; \underline{y})$ is an increasing function of θ, it is maximised at the largest value of θ, which is $\min_i y_i$. So the maximum likelihood estimator of θ is $\hat{\theta} = \min_i Y_i$.

2. We have $\mu_1' = \theta + 1$ and $m_1' = \overline{Y}$. Thus, the method of moments estimator of θ satisfies the equation $\tilde{\theta} + 1 = \overline{Y}$, so that $\tilde{\theta} = \overline{Y} - 1$.

Exercise 9. 1. $\overline{x} = \frac{1}{n} x_i = 2.95$. $\hat{\lambda} = \frac{1}{2.95} = 0.33898$

2. $X \sim \chi^2_\nu \Rightarrow E(X) = \nu$, so $\hat{\nu} = 2.95$

3. We have

$$E(X^2) = \mathrm{var}(X) + (E(X))^2 = \frac{k(1-p)}{p^2} + \left(\frac{k(1-p)}{p}\right)^2$$

and $\frac{1}{n} \sum x_i^2 = 13.635$. Set

$$\frac{\hat{k}(1-\hat{p})}{\hat{p}} = 2.95 \quad \text{and} \quad \frac{\hat{k}(1-\hat{p})}{\hat{p}^2} + \left(\frac{\hat{k}(1-\hat{p})}{\hat{p}}\right)^2 = 13.635$$

Substituting the first equation into the second gives

$$\frac{2.95}{\hat{p}} + 2.95^2 = 13.635 \Rightarrow \frac{2.95}{\hat{p}} = 4.9325 \Rightarrow \hat{p} = 0.59807$$

Hence, substituting this back into the first equation gives

$$\hat{k} = 4.3896$$

Note that k must be an integer but $\hat{k}$ is not an integer.

Exercise 10. 1. Since $0 \leq P(X = x) \leq 1$, using this for each of the probabilities gives lower bounds for α of $-1/16$, $-1/6$, and $-3/8$. Hence $\alpha \geq -1/16$. We also obtain upper bounds for α of $7/16$, $1/6$, and $5/8$. Hence, $\alpha \leq 1/6$.

(a) We have one unknown, so we will use $E(Y) = \bar{y}$:

$$E(Y) = 2(1/8 + 2\alpha) + 4(1/2 - 3\alpha) + 5(3/8 + \alpha) = 33/8 - 3\alpha.$$

From the data we have

$$\bar{y} = \frac{7 \times 2 + 6 \times 4 + 17 \times 5}{30} = 123/30 = 4.1$$

Therefore,

$$33/8 - 3\hat{\alpha} = 4.1 \Rightarrow \hat{\alpha} = 0.0083$$

This value lies between the limits derived in part (i).

(b) The likelihood of obtaining the observed results is:

$$L(\alpha) = C \times (1/8 + 2\alpha)^7 (1/2 - 3\alpha)^6 (3/8 + \alpha)^{17}$$

where C is a normalising constant. Taking logs and differentiating gives

$$\ln L(\alpha) = \ln C + 7\ln(1/8 + 2\alpha) + 6\ln(1/2 - 3\alpha) + 17\ln(3/8 + \alpha)$$

$$\Rightarrow \frac{d}{d\alpha} \ln L(\alpha) = \frac{14}{1/8 + 2\alpha} - \frac{18}{1/2 - 3\alpha} + \frac{17}{3/8 + \alpha}$$

Equating this to zero to find the maximum value of α gives

$$\frac{14}{1/8 + 2\hat{\alpha}} - \frac{18}{1/2 - 3\hat{\alpha}} + \frac{17}{3/8 + \hat{\alpha}} = 0$$

$$\Rightarrow 14(1/2 - 3\hat{\alpha})(3/8 + \hat{\alpha}) - 18(1/8 + 2\hat{\alpha})(3/8 + \hat{\alpha}) + 17(1/8 + 2\hat{\alpha})(1/2 - 3\hat{\alpha}) = 0$$

$$\Rightarrow 180\hat{\alpha}^2 + \frac{111}{8}\hat{\alpha} - \frac{91}{32} = 0$$

(c) Solving the quadratic equation gives:

$$\hat{\alpha} = \frac{-\frac{111}{8} \pm \sqrt{\left(\frac{111}{8}\right)^2 - 4(180)(-91/32)}}{360} = -0.170, 0.0929$$

The maximum likelihood estimate is 0.0929.

The other solution of -0.170 oes not lie between the bounds calculated in (i). It is not feasible as it is less than the smallest possible value for α of -0.0625.

Exercise 11. The likelihood is given by

$$L(\lambda) = \left(\prod_{i=1}^{10} f_X(x_i)\right) \times (P(X > 4))^6 = \lambda^{10} e^{-\lambda \sum_{i=1}^{10} x_i} \times (e^{-4\lambda})^6.$$

Taking logarithms gives

$$\ell(\lambda) = 10\ln\lambda - \lambda\sum_{i=1}^{10} x_i - 24\lambda.$$

Since $\sum_{i=1}^{10} x_i = 18.3$, we get

$$\ell(\lambda) = 10\ln\lambda - 42.3\lambda.$$

Differentiating gives

$$\frac{d}{d\lambda}\ell(\lambda) = \frac{10}{\lambda} - 42.3.$$

This equals 0 when

$$\lambda = \frac{10}{42.3} = 0.2364.$$

Differentiating again gives

$$\frac{d}{d\lambda}\ell(\lambda) = -\frac{10}{\lambda^2} < 0$$

So the maximum likelihood estimate is

$$\hat{\lambda} = \frac{10}{42.3} = 0.2364.$$