Part A Statistics 2015 – solutions

1. (a) [B: problem sheet]

(i) Since pdfs integrate to 1, we have

$$\int_0^1 u^{r-1}(1-u)^{n-r}\,du = \frac{(r-1)!\,(n-r)!}{n!}$$

for all r and n. Hence

$$\begin{aligned} E[U_{(r)}] &= \int_0^1 u f_{(r)}(u)\,du \\ &= \frac{n!}{(r-1)!\,(n-r)!}\int_0^1 u^{(r+1)-1}(1-u)^{(n+1)-(r+1)}\,du \\ &= \frac{n!}{(r-1)!\,(n-r)!}\frac{r!\,(n-r)!}{(n+1)!} \\ &= \frac{r}{n+1} \end{aligned}$$ **[3]**

and

$$\begin{aligned} E[U_{(r)}^2] &= \frac{n!}{(r-1)!\,(n-r)!}\int_0^1 u^{(r+2)-1}(1-u)^{(n+2)-(r+2)}\,du \\ &= \frac{n!}{(r-1)!\,(n-r)!}\frac{(r+1)!\,(n-r)!}{(n+2)!} \\ &= \frac{r(r+1)}{(n+1)(n+2)}. \end{aligned}$$

Then

$$\begin{aligned} \text{var}[U_{(r)}] &= E[U_{(r)}^2] - \{E[U_{(r)}]\}^2 \\ &= \frac{r(r+1)}{(n+1)(n+2)} - \frac{r^2}{(n+1)^2} \\ &= r\left[\frac{(r+1)(n+1) - r(n+2)}{(n+1)^2(n+2)}\right] \\ &= \frac{r(n-r+1)}{(n+1)^2(n+2)}. \end{aligned}$$ **[3]**

(ii) Now $M_U = X_{(m+1)}$ and so

$$\begin{aligned} E[M_U] &= \frac{m+1}{(2m+1)+1} = \frac{1}{2} \\ \text{var}[M_U] &= \frac{(m+1)(n-m)}{(n+1)^2(n+2)} = \frac{(m+1)^2}{(2m+2)^2(n+2)} = \frac{1}{4(n+2)}. \end{aligned}$$ **[2]**

(b) [B, then S/N: $F^{-1}(U)$ bit is bookwork, have seen similar applications of delta method, but this one is new.]

(i) The X_i are independent since the U_i are, and the cdf of X_i is

$$\begin{aligned} P(X_i \leqslant x) &= P(F^{-1}(U_i) \leqslant x) \\ &= P(U_i \leqslant F(x)) \\ &= F(x) \qquad \text{since } U_i \sim U[0,1] \end{aligned}$$

and so the X_i are a random sample from F. **[3]**

(ii) Delta method:

$$\begin{aligned} M_X &= F^{-1}(M_U) \\ &= h(M_U) \quad \text{say} \\ &= h(\mu_U) + (M_U - \mu_U)h'(\mu_U) + \cdots \end{aligned}$$

where $\mu_U = E[M_U] = \frac{1}{2}$, and for the asymptotic variance of M_X we want

$$\text{var}[(M_U - \mu_U)h'(\mu_U)] = \{h'(\mu_U)\}^2 \,\text{var}[M_U].$$

We have

$$\begin{aligned} \text{var}[M_U] &= \frac{1}{4(n+2)} \sim \frac{1}{4n} \\ h'(\mu_U) &= [F^{-1}]'(\tfrac{1}{2}) = \frac{1}{F'(F^{-1}(\frac{1}{2}))} = \frac{1}{F'(\theta)} = \frac{1}{f(\theta)}. \end{aligned}$$

Hence

$$\text{var}[M_X] \sim \left\{\frac{1}{f(\theta)}\right\}^2 \cdot \frac{1}{4n}.$$ **[6]**

(c) [S/N: have seen many similar calculations of this type, but not really comparing variances like this.]

- We have $\text{var}[\overline{X}] = \frac{\sigma^2}{n}$ and $f(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}}$. So

$$\text{var}[M_X] \sim \frac{2\pi\sigma^2}{4n}$$

and hence

$$\lim_{n\to\infty} \frac{\text{var}[\overline{X}]}{\text{var}[M_X]} = \lim_{n\to\infty} \frac{\sigma^2}{n} \cdot \frac{4n}{2\pi\sigma^2} = \frac{2}{\pi}.$$ **[3]**

- By symmetry, $E[X_i] = \theta$. Hence

$$\begin{aligned} \text{var}[X_i] &= \int_{-\infty}^{\infty} (x-\theta)^2 \frac{1}{2\sigma} \exp\left(-\frac{|x-\theta|}{\sigma}\right) dx \\ &= \int_{-\infty}^{\infty} y^2 \frac{1}{2\sigma} \exp\left(-\frac{|y|}{\sigma}\right) dy \\ &= \int_{0}^{\infty} y^2 \frac{1}{\sigma} e^{-y/\sigma}\, dy \\ &= E[Y^2] \qquad \text{for } Y \sim \text{Exp(mean } \sigma) \\ &= 2\sigma^2 \end{aligned}$$

and so $\text{var}[\overline{X}] = \frac{2\sigma^2}{n}$. Now $f(\theta) = \frac{1}{2\sigma}$, so

$$\text{var}[M_X] \sim \frac{(2\sigma)^2}{4n}.$$

Hence

$$\lim_{n\to\infty} \frac{\text{var}[\overline{X}]}{\text{var}[M_X]} = \lim_{n\to\infty} \frac{2\sigma^2}{n} \cdot \frac{n}{\sigma^2} = 2.$$ **[3]**

For the normal case prefer the sample mean (since relative efficiency of sample median is less than 1), and for the second case prefer the sample median (since relative efficiency is greater than 1). **[2]**

2. (a) [B]

The setup is that we want to test $H_0 : \theta \in \Theta_0$ against the general alternative $H_1 : \theta \in \Theta$, where $\Theta_0 \subset \Theta$.

The likelihood ratio statistic is

$$\Lambda(\mathbf{x}) = -2\log\left(\frac{\sup_{\theta\in\Theta_0} L(\theta;\mathbf{x})}{\sup_{\theta\in\Theta} L(\theta;\mathbf{x})}\right)$$

where $L(\theta;\mathbf{x})$ is the likelihood.

The test rejects H_0 for large values of $\Lambda(\mathbf{x})$.

Assuming large sample size, $\Lambda(\mathbf{X}) \approx \chi^2_p$ under H_0, where $p = \dim\Theta - \dim\Theta_0$.

So for a test of approx size α we reject H_0 iff $\Lambda(\mathbf{x}) \geqslant k_{p;1-\alpha}$, where $k_{p;1-\alpha}$ is such that $P(\chi^2_p \geqslant k_{p;1-\alpha}) = \alpha$. **[5]**

(b) [B/S: may not have seen $\pi_i = 1/k$ but have seen similar examples.]

(i) The likelihood is

$$L(\pi) \propto \pi_1^{n_1} \cdots \pi_k^{n_k}.$$

Write the LR statistic as

$$\Lambda = 2\Big(\sup_{H_1} l - \sup_{H_0} l\Big)$$

where $l = \sum_i n_i \log \pi_i$ is the log-likelihood (drop the additive constant in the log-likelihood since it cancels).

Under H_1, to maximise l subject to $\sum \pi_i = 1$, we maximise the Lagrangian $\mathcal{L}$:

$$\mathcal{L} = \sum_i n_i \log \pi_i + \lambda(1 - \sum \pi_i).$$

We have

$$\frac{\partial \mathcal{L}}{\partial \pi_j} = \frac{n_j}{\pi_j} - \lambda$$

so that $\frac{\partial \mathcal{L}}{\partial \pi_j} = 0$ when $\pi_j = \frac{n_j}{\lambda}$. Choosing λ so that the constraint is satisfied: $1 = \sum \pi_i = \sum \frac{n_i}{\lambda} = \frac{n}{\lambda}$, hence $\lambda = n$. So under H_1, $\widehat{\pi}_i = \frac{n_i}{n}$. **[5]**

Under H_0 there is no maximisation to do, $\pi_i = \frac{1}{k}$ for all i. Hence

$$\begin{aligned}\Lambda &= 2\Big\{\sum n_i \log \widehat{\pi}_i - \sum n_i \log\left(\frac{1}{k}\right)\Big\} \\ &= 2\sum n_i \log\left(\frac{k n_i}{n}\right) \\ &= 2\sum n_i \log\left(\frac{n_i}{e_i}\right)\end{aligned}$$

where $e_i = \frac{n}{k}$ (the expected number in category i under H_0). **[2]**

(ii) We have $\dim H_1 = 4 - 1 = 3$ ($k = 4$ probabilities, but 1 constraint) and $\dim H_0 = 0$ (probabilities completely fixed under H_0), so $p = \dim H_1 - \dim H_1 = 3 - 0 = 3$. **[2]**

Under the rule of thumb of all $e_i \geqslant 5$, we compare against a χ^2_3: $\Lambda = 7.3$ is less than $q_3 = 7.82$, but is fairly close (the p-value is 0.063), so at the 5% level we don't reject the hypothesis that all categories are equally likely, but as the p-value is only a bit greater than 0.05 we have weak evidence that all categories aren't equally likely. **[2]**

(c) [S/N]
The likelihood is

$$L(\theta,\beta) = \frac{1}{\beta^n} e^{-\sum(x_i-\theta)/\beta} I\{x_i \geqslant \theta \text{ for all } i\}$$
$$= \frac{1}{\beta^n} e^{n(\theta-\overline{x})/\beta} I\{x_{(1)} \geqslant \theta\}. \qquad \textbf{[1]}$$

Maximising under H_1: for fixed β, $L(\theta,\beta)$ is increasing in θ for $\theta \leqslant x_{(1)}$, and zero for $\theta > x_{(1)}$, so $\widehat{\theta} = x_{(1)}$. And to find the MLE of β:

$$l = \log L = -n\log\beta + n(\theta - \overline{x})/\beta$$
$$\frac{\partial l}{\partial \beta} = -n/\beta + n(\overline{x}-\theta)/\beta^2$$

and so $\partial l/\partial\beta = 0$ when $\beta = \overline{x} - \theta$ and this is a max. Hence $\widehat{\theta} = x_{(1)}$ and $\widehat{\beta} = \overline{x} - x_{(1)}$. So

$$\sup_{H_1} L = L(\widehat{\theta},\widehat{\beta}) = \frac{1}{(\overline{x}-x_{(1)})^n} e^{n(x_{(1)}-\overline{x})/(\overline{x}-x_{(1)})} = \frac{e^{-n}}{(\overline{x}-x_{(1)})^n}. \qquad \textbf{[4]}$$

Maximising under H_0: from above the MLE of β is $\widehat{\beta}_0 = \overline{x} - \theta|_{\theta=0} = \overline{x}$, and so

$$\sup_{H_0} L = L(0,\widehat{\beta}_0) = \frac{1}{\overline{x}^n} e^{-n\overline{x}/\overline{x}} = \frac{e^{-n}}{\overline{x}^n}. \qquad \textbf{[2]}$$

The critical region required is of the form

$$\frac{\sup_{H_0} L}{\sup_{H_1} L} \leqslant k$$

that is

$$\left(1 - \frac{x_{(1)}}{\overline{x}}\right)^n \leqslant k$$

so $g(x_1,\ldots,x_n) = x_{(1)}/\overline{x}$. **[2]**

3. (a) [B definitions, then easy N.]

The prior odds of H_0 relative to H_1 are

$$\text{prior odds} = \frac{P(H_0)}{P(H_1)} = \frac{P(H_0)}{1 - P(H_0)}$$

and the posterior odds of H_0 relative to H_1 are

$$\text{posterior odds} = \frac{P(H_0 \,|\, \mathbf{x})}{P(H_1 \,|\, \mathbf{x})} = \frac{P(H_0 \,|\, \mathbf{x})}{1 - P(H_0 \,|\, \mathbf{x})}. \qquad \textbf{[2]}$$

These are related via

$$\text{posterior odds} = \text{Bayes factor} \times \text{prior odds}$$

and so the Bayes factor B_{01} of H_0 relative to H_1 is given by

$$B_{01} = \frac{\text{posterior odds}}{\text{prior odds}} = \left\{ \frac{P(H_0 \,|\, \mathbf{x})}{P(H_1 \,|\, \mathbf{x})} \right\} \Big/ \left\{ \frac{P(H_0)}{P(H_1)} \right\} \qquad \textbf{[2]}$$

[or $B_{01} = P(\mathbf{x} \,|\, H_0)/P(\mathbf{x} \,|\, H_1)$].

With prior odds α and Bayes factor B, using posterior odds = Bayes factor × prior odds, we have

$$\frac{P(H_0 \,|\, \mathbf{x})}{1 - P(H_0 \,|\, \mathbf{x})} = B \times \alpha. \qquad (1)$$

Hence

$$P(H_0 \,|\, \mathbf{x}) = \frac{B\alpha}{1 + B\alpha}. \qquad \textbf{[3]}$$

(b) [Normal posterior is B, then S/N.]

(i) We have

$$\begin{aligned}
\pi(\theta \,|\, \mathbf{x}) &\propto f(\mathbf{x} \,|\, \theta)\pi(\theta) \\
&= \left[\prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} \exp\left\{ -\frac{(x_i - \theta)^2}{2\sigma^2} \right\} \right] (2\pi\tau^2)^{-1/2} \exp\left\{ -\frac{\theta^2}{2\tau^2} \right\} \\
&\propto \exp\left[-\frac{1}{2} \left\{ \frac{\theta^2}{\tau^2} + \sum \frac{(x_i - \theta)^2}{\sigma^2} \right\} \right].
\end{aligned}$$

Completing the square,

$$\begin{aligned}
\frac{\theta^2}{\tau^2} + \sum \frac{(x_i - \theta)^2}{\sigma^2} &= \theta^2 \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2} \right) - 2\theta \frac{n\overline{x}}{\sigma^2} + \text{constant} \\
&= \frac{1}{\sigma_1^2} (\theta - \mu_1)^2 + \text{constant} \qquad \textbf{[4]}
\end{aligned}$$

where

$$\mu_1 = \frac{\frac{n}{\sigma^2}\overline{x}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}$$

$$\frac{1}{\sigma_1^2} = \frac{1}{\tau^2} + \frac{n}{\sigma^2}, \quad \sigma_1^2 = \frac{1}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}}.$$

Hence

$$\pi(\theta \,|\, \mathbf{x}) \propto \exp\left\{ -\frac{1}{2\sigma_1^2} (\theta - \mu_1)^2 \right\}$$

and so $\theta \,|\, \mathbf{x} \sim N(\mu_1, \sigma_1^2)$. **[4]**

(ii) Since the prior is $\theta \sim N(0, \tau^2)$, we have prior probabilities $P(H_0) = P(N(0, \tau^2) < 0) = \frac{1}{2}$ and similarly $P(H_1) = \frac{1}{2}$, so the prior odds $\alpha = 1$. So by (1) we have

$$B = \frac{P(H_0 \,|\, \mathbf{x})}{1 - P(H_0 \,|\, \mathbf{x})}$$

Since posterior is $\theta \,|\, \mathbf{x} \sim N(\mu_1, \sigma_1^2)$, we have

$$P(H_0 \,|\, \mathbf{x}) = P(N(\mu_1, \sigma_1^2) < 0) = P(N(0,1) < -\tfrac{\mu_1}{\sigma_1}) = \Phi\big(-\tfrac{\mu_1}{\sigma_1}\big).$$

Hence

$$B = \frac{\Phi\big(-\frac{\mu_1}{\sigma_1}\big)}{1 - \Phi\big(-\frac{\mu_1}{\sigma_1}\big)}.$$ **[5]**

(iii) We reject H_0 when

$$\begin{aligned} B < \tfrac{1}{10} \iff & \frac{\Phi\big(-\frac{\mu_1}{\sigma_1}\big)}{1 - \Phi\big(-\frac{\mu_1}{\sigma_1}\big)} < \tfrac{1}{10} \\ \iff & \Phi\big(-\tfrac{\mu_1}{\sigma_1}\big) < \tfrac{1}{11} \\ \iff & -\frac{\mu_1}{\sigma_1} < \Phi^{-1}\big(\tfrac{1}{11}\big) \\ \iff & \mu_1 > -\sigma_1 \Phi^{-1}\big(\tfrac{1}{11}\big) \\ \iff & \frac{\frac{n}{\sigma^2}\overline{x}}{\frac{1}{\tau^2} + \frac{n}{\sigma^2}} > -\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1/2} \Phi^{-1}\big(\tfrac{1}{11}\big) \\ \iff & \overline{x} > c \end{aligned}$$

where

$$c = -\frac{\sigma^2}{n}\left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{1/2} \Phi^{-1}\big(\tfrac{1}{11}\big).$$ **[5]**