Part A Statistics 2016 – solutions

1. (a) [B] If X is from F then $F(X) \sim U(0,1)$, so $X \sim F^{-1}(U)$. Hence

$$\begin{aligned} E[X_{(k)}] &= E[F^{-1}(U_{(k)})] \\ &\approx F^{-1}(E[U_{(k)}]) \quad \text{by delta method} \\ &= F^{-1}(\tfrac{k}{n+1}). \end{aligned}$$

So if $x_1, \dots, x_n$ are from F then we should have $x_{(k)} \approx F^{-1}(\frac{k}{n+1})$.
A Q-Q plot is a plot of $x_{(k)}$ against $F^{-1}(\frac{k}{n+1})$ for $k = 1, \dots, n$. **[4]**
If the observations are really from F then the plotted points should lie approx on the line $y = x$. **[1]**

(b) [There are other suitable ways to answer (b)(i)–(iii). Any brief and correct answer is ok, less than is given below could be fine.]

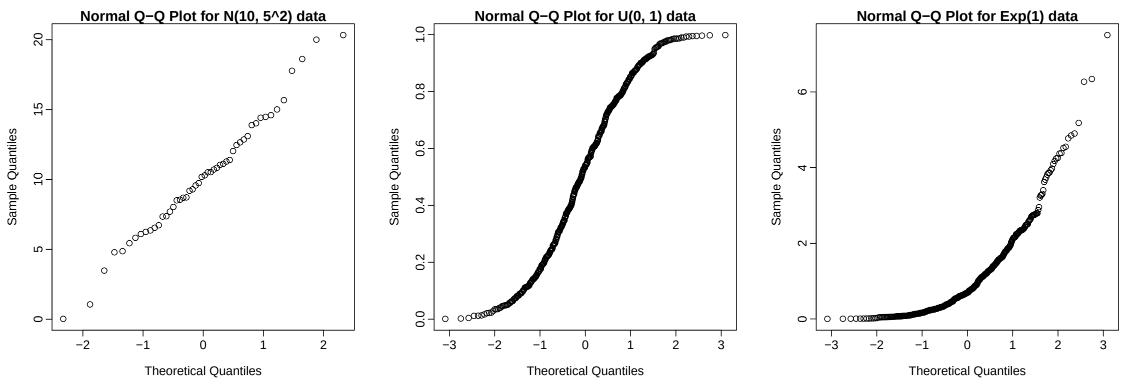
(i) [B] Let G be the cdf of $N(\mu, \sigma^2)$. Since the X_i actually have cdf G, we have from (a) that $x_{(k)} \approx G^{-1}(\frac{k}{n+1})$ or $G(x_{(k)}) \approx \frac{k}{n+1}$.

But $\frac{X_i - \mu}{\sigma} \sim N(0,1)$ and so $G(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$.

Hence $\Phi\left(\frac{x_{(k)} - \mu}{\sigma}\right) \approx \frac{k}{n+1}$ or $x_{(k)} \approx \sigma \Phi^{-1}(\frac{k}{n+1}) + \mu$. So, since $F = \Phi$, the Q-Q plot of (a) will result in an approx straight line with gradient σ, intercept μ. **[2]**

(ii) [B/S] In the tails of the distribution, the sample quantiles of $U(0,1)$ increase/decrease more slowly than those of $N(0,1)$, since $U(0,1)$ is restricted to $(0,1)$. Hence the Q-Q plot flattens out as $x \to \pm\infty$ and we get an S-type shape as below. **[2]**

(iii) [B/S] The exponential is like the uniform for small x, since the exponential is restricted to $x > 0$. But we get the opposite behaviour for large x as the Exp pdf decays more slowly than the Normal pdf. Hence we get a convex-type shape as below. **[2]**



(c) (i) [B] $\overline{X}$ and Q_X are independent and their marginal distributions are given by $\overline{X} \sim N(\mu_1, \frac{\sigma_1^2}{n_1})$ and $\frac{Q_X}{\sigma_1^2} \sim \chi^2_{n_1 - 1}$. **[2]**

(ii) [S] By the independence and the definition of χ^2 we have $\frac{Q_X + Q_Y}{\sigma^2} \sim \chi^2_r$ where $r = n_1 + n_2 - 2$. Hence

$$P\left(c_1 < \frac{Q_X + Q_Y}{\sigma^2} < c_2\right) = 1 - \alpha$$

where c_1, c_2 are such that $P(\chi^2_r < c_1) = P(\chi^2_r > c_2) = \frac{\alpha}{2}$. Rearranging gives

$$P\left(\frac{Q_X + Q_Y}{c_2} < \sigma^2 < \frac{Q_X + Q_Y}{c_1}\right) = 1 - \alpha$$

and so $\left(\frac{Q_X+Q_Y}{c_2}, \frac{Q_X+Q_Y}{c_1}\right)$ is a $1-\alpha$ confidence interval for σ^2. **[3]**

(iii) [N] We have $\overline{X} \sim N(\mu_1, \frac{\sigma_1^2}{n_1})$ and $\overline{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$. Hence $\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$ and so, as $\sigma_1^2 = c\sigma_2^2$,

$$\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{c\sigma_2^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1).$$

Also $\frac{Q_X}{c\sigma_2^2} = \frac{Q_X}{\sigma_1^2} \sim \chi^2_{n_1-1}$ and $\frac{Q_Y}{\sigma_2^2} \sim \chi^2_{n_2-1}$ and so

$$\frac{1}{\sigma_2^2}\left(\frac{1}{c}Q_X + Q_Y\right) \sim \chi^2_r$$

again with $r = n_1 + n_2 - 2$. Hence from the definition of a t-distribution ($t = \frac{N(0,1)}{\sqrt{\chi^2/df}}$),

$$\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{D} \sim t_r$$

where

$$D = \left[\left(\frac{c}{n_1} + \frac{1}{n_2}\right)\left(\frac{1}{c}Q_X + Q_Y\right)\Big/ r\right]^{1/2}$$

(as σ_2^2 cancels). Hence

$$P\left(-k < \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{D} < k\right) = 1 - \alpha$$

where k is such that $P(t_r > k) = \frac{\alpha}{2}$. Rearranging gives

$$P\left(\overline{X} - \overline{Y} - kD < \mu_1 - \mu_2 < \overline{X} - \overline{Y} + kD\right) = 1 - \alpha$$

and so $(\overline{X} - \overline{Y} \pm kD)$ is a $1-\alpha$ confidence interval for $\mu_1 - \mu_2$. **[5]**

(d) [N] In large samples the distribution of $\Lambda = -2\log\left(\frac{L(\theta)}{L(\widehat{\theta})}\right) = 2\{\ell(\widehat{\theta}) - \ell(\theta)\}$ is approximately χ^2_p where $p = 1$ here as we have a scalar parameter. Hence

$$\begin{aligned} 0.95 &\approx P(\Lambda < 3.84) \\ &= P(2\{\ell(\widehat{\theta}) - \ell(\theta)\} < 3.84) \\ &= P(\ell(\theta) > \ell(\widehat{\theta}) - 1.92). \end{aligned}$$

Hence $\{\theta \colon \ell(\theta) > \ell(\widehat{\theta}) - 1.92\}$ is an approx 95% for θ. So (with sufficient regularity, a unimodal likelihood) solve $\ell(\theta) = \ell(\widehat{\theta}) - 1.92$ to obtain the two roots θ_1 and θ_2, and then the approx confidence interval is (θ_1, θ_2). **[4]**

2. (a) (i) [B] Neyman–Pearson lemma: Consider testing simple null $H_0 : \theta = \theta_0$ against simple alternative $H_1 : \theta = \theta_1$. Define critical region C by

$$C = \left\{ \mathbf{x} : \frac{L(\theta_0; \mathbf{x})}{L(\theta_1; \mathbf{x})} \leqslant k \right\}$$

and suppose constants k and α are such that $P(\mathbf{X} \in C \,|\, H_0) = \alpha$. Then among all tests of size $\leqslant \alpha$, the test with critical region C has maximum power. **[2]**

(ii) [B] The likelihood is

$$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda} \lambda^{n\overline{x}}}{\prod x_i!}.$$

By NP lemma a test of the following form is most powerful:

$$\begin{aligned} \text{reject } H_0 &\iff \frac{L(\lambda_0)}{L(\lambda_1)} \leqslant k_1 \\ &\iff \frac{e^{-n\lambda_0} \lambda_0^{n\overline{x}}}{e^{-n\lambda_1} \lambda_1^{n\overline{x}}} \\ &\iff \left(\frac{\lambda_0}{\lambda_1} \right)^{n\overline{x}} \leqslant k_2 \\ &\iff \overline{x} \geqslant c \qquad \text{since } \tfrac{\lambda_0}{\lambda_1} < 1. \end{aligned}$$

[2]

[S] The size is

$$\begin{aligned} P(\overline{X} \geqslant c | H_0) &= P\left(\sum X_i \geqslant nc | H_0 \right) \\ &= P(\text{Poisson}(n\lambda_0) \geqslant nc) \\ &= \sum_{k=nc}^{\infty} \frac{e^{-n\lambda_0} (n\lambda_0)^k}{k!}. \end{aligned}$$

[1]

[B] Yes, it is UMP because it does not depend on the value of $\lambda_1 \in (\lambda_0, \infty)$. **[1]**

(iii) [B/S] For size α we want

$$\begin{aligned} \alpha = P\left(\overline{X} \geqslant c \,|\, H_0 \right) &= P\left(\frac{\overline{X} - \lambda_0}{\sqrt{\lambda_0/n}} \geqslant \frac{c - \lambda_0}{\sqrt{\lambda_0/n}} \,\middle|\, H_0 \right) \\ &\approx P\left(N(0,1) \geqslant \frac{c - \lambda_0}{\sqrt{\lambda_0/n}} \right) \end{aligned}$$

and so we want

$$\frac{c - \lambda_0}{\sqrt{\lambda_0/n}} = z_\alpha, \quad \text{that is } c = \lambda_0 + z_\alpha \sqrt{\lambda_0/n}$$

where z_α is such that $P(N(0,1) > z_\alpha) = \alpha$. **[3]**

[S/N] So the power function is

$$\begin{aligned} w(\lambda) = P(\text{reject } H_0 \,|\, \lambda) &= P\left(\overline{X} \geqslant \lambda_0 + z_\alpha \sqrt{\lambda_0/n} \right) \\ &= P\left(\frac{\overline{X} - \lambda}{\sqrt{\lambda/n}} \geqslant \frac{\lambda_0 - \lambda}{\sqrt{\lambda/n}} + z_\alpha \sqrt{\lambda_0/\lambda} \,\middle|\, H_0 \right) \\ &\approx P\left(N(0,1) \geqslant (\lambda_0 - \lambda)\sqrt{n/\lambda} + z_\alpha \sqrt{\lambda_0/\lambda} \right) \\ &= 1 - \Phi\left((\lambda_0 - \lambda)\sqrt{n/\lambda} + z_\alpha \sqrt{\lambda_0/\lambda} \right). \end{aligned}$$

[3]

(b) [S]

(i) The likelihood ratio is

$$\lambda = \frac{\sup_{H_0} L(\lambda)}{\sup_{H_1} L(\lambda)} = \frac{L(\lambda_0)}{L(\widehat{\lambda})} = \frac{e^{-n\lambda_0}\lambda_0^{n\overline{x}}}{e^{-n\widehat{\lambda}}\widehat{\lambda}^{n\overline{x}}} = e^{n(\widehat{\lambda}-\lambda_0)}\left(\frac{\lambda_0}{\widehat{\lambda}}\right)^{n\overline{x}}.$$

Now $\widehat{\lambda} = \overline{x}$ (ok just to state this), so the LR statistic Λ is

$$\begin{aligned}\Lambda = -2\log\lambda &= -2\left[n(\overline{x}-\lambda_0) + n\overline{x}\log\frac{\lambda_0}{\overline{x}}\right] \\ &= 2n\left[\lambda_0 - \overline{x} + \overline{x}\log\frac{\overline{x}}{\lambda_0}\right].\end{aligned}$$ **[4]**

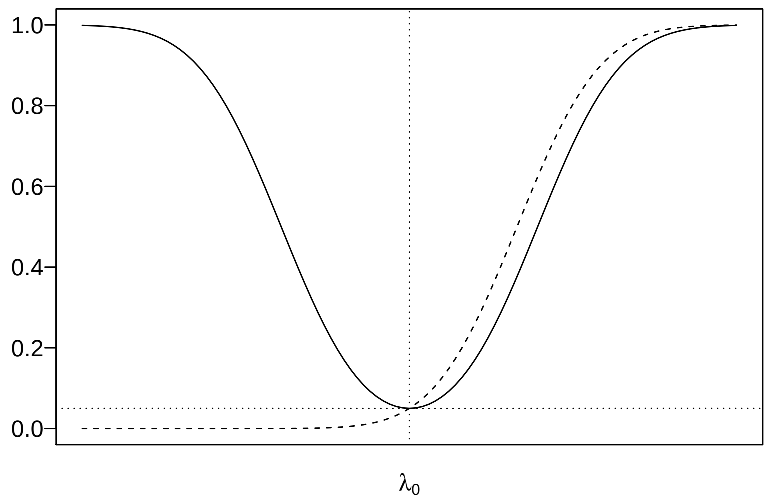
(ii) We compare Λ to a χ_p^2 where $p = \dim H_1 - \dim H_0 = 1 - 0 = 1$. So the critical region is $\Lambda \geqslant c_\alpha$ where c_α is such that $P(\chi_1^2 \geqslant c_\alpha) = \alpha$. **[2]**

(iii) $\Lambda = 3.2$ is less than $3.84 = F^{-1}(0.05)$ so a test of size 0.05 doesn't reject $\lambda = \lambda_0$. On the other hand 3.2 is not that much less than 3.84 so the the p-value won't be too much larger than 0.05 (in fact $p = P(\chi_1^2 \geqslant 3.2) = 0.074$). Conclusion: there is weak evidence that $\lambda \neq \lambda_0$. **[2]**

(c) [N]

Something like the plot below: $w_a(\lambda)$ is the dashed curve, $w_b(\lambda)$ is the solid curve.

- $w_a(\lambda_0) = w_b(\lambda_0) = \alpha$
- $w_a(\lambda)$ increases from 0 to 1 as λ increases
- $w_b(\lambda)$ must be below $w_a(\lambda)$ for $\lambda > \lambda_0$ because the test in (a) is UMP
- on the other hand the test in (b) is a good two-sided test so has power increasing towards 1 as λ decreases from λ_0



[5]

3. (a) (i) [B] Denoting the posterior by $\pi(\theta|\mathbf{x})$, a $100(1-\alpha)\%$ credible interval is an interval (θ_1, θ_2) such that

$$\int_{\theta_1}^{\theta_2} \pi(\theta|\mathbf{x})\, d\theta = 1-\alpha. \qquad \textbf{[1]}$$

(ii) [B] The interval if equal-tailed if $\int_{-\infty}^{\theta_1} \pi(\theta|\mathbf{x})\, d\theta = \int_{\theta_2}^{\infty} \pi(\theta|\mathbf{x})\, d\theta = \frac{\alpha}{2}$. **[1]**

The interval is HPD if $\pi(\theta|\mathbf{x}) \geqslant \pi(\theta'|\mathbf{x})$ for all $\theta \in (\theta_1, \theta_2)$, $\theta' \notin (\theta_1, \theta_2)$. **[1]**

(b) (i) [B]

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto f(\mathbf{x}|\theta)\pi(\theta) \\ &= \left[\prod_{i=1}^{n} \frac{\theta^2}{\Gamma(2)} x_i e^{-\theta x_i}\right] \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \\ &\propto \theta^{a+2n-1} e^{-(b+\sum x_i)\theta}\end{aligned}$$

Hence $\theta|\mathbf{x} \sim \text{Gamma}(a+2n, b+\sum x_i)$. **[5]**

(ii) [S] The prior mean is a/b so the posterior mean is $(a+2n)/(b+\sum x_i)$.
From $[\ldots]$ above, the likelihood is given by

$$L(\theta) \propto \theta^{2n} e^{-\theta \sum x_i}$$

so the log-likelihood is given by

$$l(\theta) = 2n \log \theta - \theta \sum x_i$$

and

$$l'(\theta) = \frac{2n}{\theta} - \sum x_i.$$

Setting $l'(\theta) = 0$ gives the MLE $\widehat{\theta} = \frac{2n}{\sum x_i}$.
So we want to show

$$\frac{a+2n}{b+\sum x_i} = \lambda\left(\frac{a}{b}\right) + (1-\lambda)\frac{2n}{\sum x_i}$$

for some $\lambda \in (0,1)$. Solving the above equation for λ gives $\lambda = \frac{b}{b+\sum x_i}$ which is $(0,1)$, hence the posterior mean is between the prior mean and the MLE. **[5]**

(iii) [S/N] For large n the value of λ is approx zero (as $\sum x_i$ will be large), so the posterior mean will be approx the same as the MLE. This is to be expected as we expect the likelihood to dominate the prior for large n. **[2]**

(c) [N] The density of each Y_i is

$$f(y) = \left(\frac{\psi}{2\pi}\right)^{1/2} e^{-\frac{\psi}{2}(y-\mu)^2}.$$

Hence the joint posterior is

$$\begin{aligned}\pi(\mu,\psi) &\propto f(\mathbf{y}|\mu,\psi)\pi(\mu,\psi) \\ &\propto \left[\prod_{i=1}^{n} \left(\frac{\psi}{2\pi}\right)^{1/2} e^{-\frac{\psi}{2}(y_i-\mu)^2}\right] \frac{1}{\psi} \\ &\propto \psi^{\frac{n}{2}-1} e^{-\frac{\psi}{2}\sum(y_i-\mu)^2} \\ &= \psi^{\frac{n}{2}-1} e^{-\frac{\psi}{2}s} e^{-\frac{\psi}{2}n(\mu-\overline{y})^2} \quad \text{by the hint}\end{aligned}$$

 Turn Over

where $s = \sum_{i=1}^{n}(y_i - \overline{y})^2$. **[2]**

The density for Y_i integrates to 1, so

$$\int_{-\infty}^{\infty} e^{-\frac{\psi}{2}(y-\mu)^2}\, dy \propto \psi^{-\frac{1}{2}}. \tag{1}$$

Hence

$$\begin{aligned}
\pi(\psi|\mathbf{y}) &= \int_{-\infty}^{\infty} \pi(\mu, \psi|\mathbf{y})\, d\mu \\
&\propto \psi^{\frac{n}{2}-1} e^{-\frac{\psi}{2}s} \int_{-\infty}^{\infty} e^{-\frac{n\psi}{2}(\mu-\overline{y})^2}\, d\mu \\
&\propto \psi^{\frac{n}{2}-1} e^{-\frac{\psi}{2}s} \psi^{-\frac{1}{2}} \quad \text{by (1)} \\
&= \psi^{(\frac{n-1}{2})-1} e^{-\frac{s}{2}\psi}.
\end{aligned}$$

This is a Gamma$(\frac{n-1}{2}, \frac{s}{2})$. **[6]**

So let F be the cdf of this Gamma.

Take ψ_1, ψ_2 such that $P(\text{Gamma}(\frac{n-1}{2}, \frac{s}{2}) < \psi_1) = P(\text{Gamma}(\frac{n-1}{2}, \frac{s}{2}) > \psi_2) = \frac{\alpha}{2}$ Then $(\psi_1, \psi_2) = (F^{-1}(\frac{\alpha}{2}), F^{-1}(1 - \frac{\alpha}{2}))$ is the required equal-tailed interval for ψ. **[2]**

End of Last Page