Part IB — Statistics Example Sheet 1

Supervised by Mr. Higson
Examples worked through by Christopher Turnbull

Lent 2018

QUESTION 2

If $X \sim \text{Exp}(\lambda), Y \sim \text{Exp}(\mu)$, X, Y independent, we can derive the standard result that the minimum of exponentials is exponential:

$$\begin{aligned}\mathbb{P}(\min[X,Y] < t) &= 1 - \mathbb{P}(\min[X,Y] \geq t)\\ &= 1 - \int_0^\infty \int_0^\infty I(\lambda e^{-\lambda x_1} \geq t, \mu e^{-\mu x_2} \geq t)\ \mathrm{d}x_2 \mathrm{d}x_1\\ &= 1 - \int_t^\infty \lambda e^{-\lambda x_1}\ \mathrm{d}x_1 \int_t^\infty \mu e^{-\mu x_2}\ \mathrm{d}x_2\\ &= 1 - e^{-(\lambda+\mu)t},\ \text{i.e.} \min[X,Y] \sim \text{Exp}(\lambda+\mu)\end{aligned}$$

Next, suppose $X \sim \Gamma(\alpha, \lambda)$, $Y \sim \Gamma(\beta, \lambda)$. We want to find the joint PDF of

$$U = X + Y, \quad \text{and} \quad V = X/(X+Y)$$

Consider the map

$$T : (x,y) \mapsto (u,v), \quad \text{where } \ u = x + y,\ v = \frac{x}{x+y}$$

where $x, y, u \geq 0$, $0 \leq v \leq 1$ The inverse map T^{-1} acts by

$$T^{-1} : (u,v) \mapsto (x,y),\ \text{where } x = uv,\ y = u(1-v)$$

and has the Jacobian

$$\begin{aligned}J(u,v) &= \det \begin{pmatrix} v & u \\ 1-v & -u \end{pmatrix}\\ &= -u\end{aligned}$$

Then the joint PDF

$$f_{U,V}(u,v) = f_{X,Y}(uv, u(1-v))\,|-u|$$

Substituting in $f_{X,Y}(x,y) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \frac{\lambda^\beta y^{\beta-1} e^{-\lambda y}}{\Gamma(\beta)}$, $x, y \geq 0$, yields

$$\begin{aligned}f_{U,V}(u,v) &= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} (uv)^{\alpha-1} (u(1-v))^{\beta-1} u e^{-\lambda u},\ u \geq 0,\ 0 \leq v \leq 1\\ &= \frac{\lambda^{\alpha+\beta}}{\Gamma(\alpha)\Gamma(\beta)} u^{\alpha+\beta-1} v^{\alpha-1} (1-v)^{\beta-1} e^{-\lambda u}\\ &= \text{Beta}(v; \alpha, \beta) \frac{\lambda^{\alpha+\beta}}{\Gamma(a+b)} u^{\alpha+\beta-1} e^{-\lambda u}\\ &= \text{Beta}(v; \alpha, \beta)\text{Gamma}(u; \alpha+\beta)\end{aligned}$$

This factorises, so the respective marginal PDFs are

$$f_U(u) = \text{Gamma}(u; \alpha+\beta), \quad f_V(v) = \text{Beta}(v; \alpha, \beta)$$

QUESTION 3

The factorization criterion states that a statistic $T = t(\mathbf{x})$ is sufficient for θ iff

$$f_{\mathbf{X}}(\mathbf{x}\,;\theta) = g(t(\mathbf{x}),\theta)h(\mathbf{x})$$

We have proved the discrete case in lectures. The continuous case is similar:

Proof. Suppose we are given the factorization $f_{\mathbf{X}}(\mathbf{x}\,;\theta) = g(t(\mathbf{x}),\theta)h(\mathbf{x})$. If $T = u$, then

$$\begin{aligned}
f_{\mathbf{X}|T=u}(\mathbf{x}\,;u) &= \frac{g(t(\mathbf{x}),\theta)h(\mathbf{x})}{\int_{\mathbf{y}\,;T(\mathbf{y})=u} g(t(\mathbf{y}),\theta)h(\mathbf{y})\,\mathrm{d}\mathbf{y}} \\
&= \frac{g(u,\theta)h(\mathbf{x})}{g(u,\theta)\int_{\mathbf{y}\,;T(\mathbf{y})=u} h(\mathbf{y})\,\mathrm{d}\mathbf{y}} \\
&= \frac{h(\mathbf{x})}{\int_{\mathbf{y}} h(\mathbf{y})\,\mathrm{d}\mathbf{y}}
\end{aligned}$$

which does not depend on θ; thus T is sufficient for θ.

The other direction is the same as the discrete case: Suppose T is sufficient for θ, ie. the conditional distribution of $\mathbf{X}\,|T = u$ does not depend on θ. Then

$$\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}) = \mathbb{P}_\theta(\mathbf{X} = \mathbf{x}, T = T(\mathbf{x})) = \mathbb{P}_\theta(\mathbf{X} = \mathbf{x}\,|T = T(\mathbf{x}))\mathbb{P}_\theta(T = T(\mathbf{x}))$$

The first factor does not depend on θ by assumption; call it $h(\mathbf{x})$. Let the second factor be $g(t,\theta)$, and so we have the required factorisation.

□

QUESTION 4

(a) Let $X_1, \cdots, X_n$ be independent Po($i\theta$). So

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x} \mid \theta) &= \prod_{i=1}^{n} \frac{e^{-i\theta}(i\theta)^{x_i}}{x_i!} \\ &= \underbrace{\exp\left(-\frac{n(n+1)}{2}\theta\right)\theta^{\sum x_i}}_{g(t(\mathbf{x}),\theta)} \cdot \underbrace{\prod_{i=1}^{n} \frac{i^{x_i}}{x_i!}}_{h(\mathbf{x})} \end{aligned}$$

Using the factorization criterion, $T = t(\mathbf{x}) = \sum_{i=1}^n x_i$ is a sufficient statistic, and $T \sim$ Po $(n(n+1)\theta/2)$

The log-likelihood is

$$l(\theta) = -\frac{n(n+1)}{2}\theta + \sum x_i \log\theta + \log\left(\prod_{i=1}^{n} \frac{i^{x_i}}{x_i!}\right)$$

and this is maximised when $\frac{\mathrm{d}l}{\mathrm{d}\theta} = 0$;

$$-\frac{n(n+1)}{2} + \frac{1}{\theta}\sum x_i = 0 \Rightarrow \hat{\theta} = \frac{2\sum x_i}{n(n+1)}$$

Thus the MLE $\hat{\theta}$ is a function of $T = \sum x_i$, and is unbiased:

$$\mathbb{E}_\theta(\hat{\theta}) = \frac{2}{n(n+1)} \cdot \frac{n(n+1)}{2}\theta = \theta$$

(b) Let $X_1, \cdots, X_n \sim$ iid Exp(θ). Then

$$\begin{aligned} \frac{f_{\mathbf{X}}(\mathbf{x} \mid \theta)}{f_{\mathbf{X}}(\mathbf{y} \mid \theta)} &= \frac{\lambda^n e^{-\lambda \sum x_i}}{\lambda^n e^{-\lambda \sum y_i}} \\ &= \exp\left\{-\lambda\left(\sum x_i - \sum y_i\right)\right\} \end{aligned}$$

This is constant as a function of λ iff $\sum x_i = \sum y_i$. Hence $T = \sum_{i=1}^n x_i$ is minimal sufficient, with $T \sim \Gamma(n, \lambda)$.

The log-likelihood is

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

and this is maximised when $\frac{\mathrm{d}l}{\mathrm{d}\lambda} = 0$;

$$\frac{n}{\lambda} - \sum x_i = 0 \Rightarrow \hat{\lambda} = \frac{n}{\sum x_i}$$

Thus the MLE $\hat{\lambda}$ is a function of $T = \sum x_i$, and

$$\mathbb{E}_\lambda(\hat{\lambda}) = n \cdot \left(\frac{n}{\lambda}\right)^{-1} = \lambda$$

so it is unbiased?

QUESTION 5

Given $\tilde{\theta} = \frac{2}{3}X_1$, we have

$$\mathbb{E}_\theta(\tilde{\theta}) = \frac{2}{3}\frac{1}{2}(\theta + 2\theta) = \theta$$

so $\tilde{\theta}$ is unbiased.
We have

$$\frac{f_{\mathbf{X}}(\mathbf{x} \mid \theta)}{f_{\mathbf{X}}(\mathbf{y} \mid \theta)} = \frac{\frac{1}{\theta^n}\mathbb{I}_{\{\max x_i < 2\theta\}}\mathbb{I}_{\{\min x_i > \theta\}}}{\frac{1}{\theta^n}\mathbb{I}_{\{\max y_i < 2\theta\}}\mathbb{I}_{\{\min y_i > \theta\}}}$$

Hence we can see $T = \Delta x := \lfloor \frac{\min x_i + \max x_i}{2} \rfloor$ is minimal sufficient, and

$$\begin{aligned}
\mathbb{E}_\theta(\tilde{\theta} | T = u) &= \frac{2}{3}\mathbb{E}_\theta(X_1 | \lfloor \frac{\min x_i + \max x_i}{2} \rfloor = u) \\
&= \frac{2}{3}\mathbb{E}_\theta(X_1 | \Delta x = u, X_1 = \Delta x)\mathbb{P}_\theta(X_1 = \Delta x \mid \Delta x = u) \\
&+ \frac{2}{3}\mathbb{E}_\theta(X_1 | \Delta x = u, X_1 \neq \Delta x)\mathbb{P}_\theta(X_1 \neq \Delta x \mid \Delta x = u) \\
&= \frac{2}{3}\left(u \times \frac{1}{n} + \frac{u}{2} \times \frac{n-1}{n}\right) \\
&= \frac{1}{3}\frac{n+1}{n}u
\end{aligned}$$

So the Rao-Blackwell estimator is $\frac{1}{3}\frac{n+1}{n}\lfloor \frac{\min x_i + \max x_i}{2} \rfloor$

Note sure about this as I don't think the probability that X_1 takes the value Δx is $\frac{1}{n}$. Not sure how to incorporate both $\min x_i$ and $\max x_i$ into a sufficient statistic.

QUESTION 6

Have

$$L(\theta) = f_{\mathbf{X}}(\mathbf{x}\,;\,\theta) = \frac{1}{\theta^n}\mathbb{I}_{\{\max x_i < \theta\}}\mathbb{I}_{\{\min x_i > 0\}}$$

So for $\theta \geq \max x_i$, $L(\theta) = \frac{1}{\theta^n}$ and is decreasing as θ increases, while for $\theta < \max x_i$, $L(\theta) = 0$. Hence the value $\hat{\theta} = \max x_i$ maximizes the likelihood.

Now, $\mathbb{P}_\theta(\theta \geq \max x_i) = 1$. So we can construct a one-sided $100(1-\alpha)$ % confidence interval with lower bound $\hat{\theta}$, and upper bound $b(\hat{\theta})$, such that $\mathbb{P}(\theta \leq b(\hat{\theta})) = 1 - \alpha$, for some function b to be determined.

Have that

$$1 - \alpha = \mathbb{P}(\theta \leq b(\hat{\theta})) = 1 - \mathbb{P}(b(\hat{\theta}) < \theta) = 1 - \mathbb{P}(\hat{\theta} < b^{-1}(\theta))$$

Hence $\mathbb{P}(\hat{\theta} < b^{-1}(\theta)) = \alpha$. For $0 \leq t \leq \theta$ the cumulative distribution function of $\hat{\theta}$ is

$$F_{\hat{\theta}}(t) = \mathbb{P}(\hat{\theta} \leq t) = \mathbb{P}(X_i \leq t \text{ for all } i) = (\mathbb{P}(X_i \leq t))^n = \left(\frac{t}{n}\right)^n$$

We obtain

$$\frac{(b^{-1}(\theta))^n}{\theta^n} = \alpha, \quad \text{whence } b^{-1}(\theta) = \theta\alpha^{1/n}$$

This gives inverse function $b(\hat{\theta}) = \hat{\theta}/\alpha^{1/n}$.

Hence, the $(1-\alpha)$ % CI for θ is $(\hat{\theta}, \hat{\theta}/\alpha^{1/n})$.

QUESTION 7

If we take $z_- < z_+$ such that $\Phi(z_+) - \Phi(z_-) = \sqrt{0.95}$, then the equation

$$\mathbb{P}\left(z_- < X_1 - \theta_1 < z_+, z_- < X_2 - \theta_2 < z_+\right) = 0.95$$

determines a 95 % confidence set for (θ_1, θ_2).
Owing to independence we can model this as a square:

$$\mathbb{P}\left(z_- < X_1 - \theta_1 < z_+\right) = \sqrt{0.95}$$

which can be written as

$$\mathbb{P}\left(X_1 - z_+ < \theta_1 < X_1 - z_-\right) = \sqrt{0.95}$$

ie. gives the interval

$$(X_1 - z_+, X_1 - z_-)$$

centred at $X_1 - (z_+ + z_-)/2$ with width $(z_+ - z_-)$.
Choosing $z_+ = -z_- := z$ gives the interval

$$(X_1 - z, X_1 - z)$$

and z will be the upper $(1 - \sqrt{0.95})/2$ point of the standard normal distribution, ie. $\Phi(a) = 1 - (1 - \sqrt{0.95})/2 = (1 + \sqrt{0.95})/a$. By the hint, we see $a = 2.236$; thus the confidence set can be written as

$$S = \{(\theta_1, \theta_2) : |\theta_1 - X_1| \leq 2.236, |\theta_2 - X_2| \leq 2.236\}$$

Similarly, we try model this as a circle:

$$\mathbb{P}\left((\theta_1 - X_1)^2 + (\theta_2 - X_2)^2 < z_+^2\right) 0.95$$

Not sure how to finish.

QUESTION 8

The likelihood is written as

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x} \mid \lambda) &= \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} \\ &= \exp(-n\lambda)\lambda^{\sum x_i} \cdot \prod_{i=1}^{n} \frac{1}{x_i!} \end{aligned}$$

Here $n = 5$, $\sum x_i = 16$, $\prod_{i=1}^{n} \frac{1}{x_i!} = 207,360$. Calculating,

$$\begin{aligned} f_X(x) &= f_X(x|1)\pi_\lambda(1) + f_X(x|1.5)\pi_\lambda(1.5) \\ &= \exp(-5) \cdot \frac{1}{207,360} \cdot 0.4 + \exp(-7.5)1.5^{16} \cdot \frac{1}{207,360} \cdot 0.6 \\ &= 3.249396 \times 10^{-8} \cdot 0.4 + 1.75197 \times 10^{-6} \times 0.6 \\ &= 7.20284 \times 10^{-7} \end{aligned}$$

Whence

$$\begin{aligned} \pi_{\lambda|X}(1|x) &= \frac{f_X(x|1) \cdot \pi_\lambda(1)}{f_X(x)} \\ &= \frac{3.249396 \times 10^{-8} \cdot 0.4}{7.20284 \times 10^{-7}} \\ &= 0.0270676 \end{aligned}$$

and

$$\begin{aligned} \pi_{\lambda|X}(1.5|x) &= \frac{f_X(x|1.5) \cdot \pi_\lambda(1.5)}{f_X(x)} \\ &= \frac{1.75197 \times 10^{-6} \times 0.6}{7.20284 \times 10^{-7}} \\ &= 0.972932 \end{aligned}$$

Getting different answers than those on the sheet but not sure where the error is.

QUESTION 9

By independence, $f_{X|\theta}(x|\theta) = \theta^n (x_1 x_2 \cdots x_n)^{\theta-1}, 0 < x < 1$, and so given a Gamma prior, the posterior is

$$\begin{aligned}\pi_{\theta|X}(\theta|x) &\propto f_{X|\theta}(x|\theta)\pi_\theta(\theta) \\ &= \theta^n (x_1 x_2 \cdots x_n)^{\theta-1} \cdot \frac{\lambda^\alpha \theta^{\alpha-1} e^{-\lambda\theta}}{\Gamma(\alpha)}, \quad 0 < x < 1\end{aligned}$$

which is $\Gamma(n+\alpha, \lambda)$ with the appropriate proportionality constant.

For quadratic loss, the Bayesian point estimator of θ is just the posterior mean, which is given as $\frac{n+\alpha}{\lambda}$.

QUESTION 10

We have that $S_n \sim \text{Bin}\ (n, p_n)$, so as $np_n \to \lambda$, $n \to \infty$

$$\begin{aligned}
\mathbb{P}(S_n = x) &= \binom{n}{x} p_n^x (1-p_n)^{n-x} \\
&= \frac{1}{x!}\frac{n(n-1)\cdots(n-x+1)}{n^x}(np_n)^x\left(1-\frac{np_n}{x}\right)^{n-x} \\
&\to \frac{1}{x!}\lambda^x e^{-\lambda} \\
&= \mathbb{P}(Y = x) \text{ where } Y \sim \text{Po}\ (\lambda)
\end{aligned}$$

since $(1-a/n)^n \to e^{-a}$

QUESTION 11

$f_{X_1}(x_1|\theta) \sim N(0,1)$, so $X_2 = \theta X_1 + (1-\theta^2)^{1/2}\varepsilon_2$, and

$$\begin{aligned} f_{X_2}(x_2|\theta) &\sim N(0,\theta^2) + N(0,1-\theta^2) \\ &\sim N(0,1) \end{aligned}$$

Similarly $f_{X_i}(x_i|\theta) \sim N(0,1)$ for $i = 1, \cdots, n$.
Hence

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}|\theta) &= (2\pi)^{-n/2} \exp\left(-\sum x_i^2\right) \\ &= \end{aligned}$$

Shouldn't be independent of θ?

QUESTION 12