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 have

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

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 $\mathrm{Po}(i\theta)$. So

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x} \mid \theta) &= \prod_{i=1}^{n} \frac{e^{-i\theta}(i\theta)^{x_i}}{x_i!} \\ &= \underbrace{\exp\left(-\frac{n(n+1)}{2}\theta\right)\theta^{\sum x_i}}_{g(t(\mathbf{x}),\theta)} \cdot \underbrace{\frac{1^{x_1} \cdot 2^{x_2} \cdot \cdots \cdot n^{x_n}}{x_1!x_2!\cdots x_n!}}_{h(\mathbf{x})} \end{aligned}$$

Using the factorization criterion, $t(\mathbf{x}) = \sum_{i=1}^{n} x_i$ is a sufficient statistic, with distribution:

The maximum likelihood estimator $\hat{\theta}$ is given by

(b) Let $X_1, \cdots, X_n$ be independent $\mathrm{Exp}(\theta)$. So

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

Choosing $t(\mathbf{x}) = \sum_{i=1}^{n} x_i$, we can use the factorization criterion with $g(t(\mathbf{x}), \lambda) = \lambda^n e^{-\lambda \sum x_i}$, $h(\mathbf{x}) = 1$, to show that $t(\mathbf{x})$ is a sufficient statistic for λ.

QUESTION 5

(a) Let $X_1, \cdots, X_n$ be $\sim$ iid $\text{Bin}(1, p)$. (this is $\text{Ber}(p)$)

$$
\begin{aligned}
f_{\mathbf{X}}(\mathbf{x} \mid p) &= \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} \\
&= p^{\sum x_i}(1-p)^{n-\sum x_i}
\end{aligned}
$$

QUESTION 6

QUESTION 7

QUESTION 8

QUESTION 9

QUESTION 10

QUESTION 11

QUESTION 12