MATH562 – Fall 2025

Problem Set: Week 8

1. Let $Y_1, \ldots, Y_n \overset{i.i.d.}{\sim} U(a, b)$. Find minimal sufficient statistics for θ when (a) $a = -\theta$, $b = \theta$, and (b) $a = \theta - 1$, $b = \theta + 1$. If the minimal sufficient statistic is not scalar, can you find an ancillary?

   Solution:

   The likelihood is

$$\prod_{j=1}^{n} \frac{1}{b-a} I(a < y_j < b) = (b-a)^{-n} I(a < u < v < b), \quad a < b,$$

   where $u = \min y_j < v = \max y_j$ are the sample maxima and minima. Hence a sufficient statistic is $S = (U, V) = (\min Y_j, \max Y_j)$, using the factorisation theorem; as usual we suppose that the sample size is fixed.

   In the first case $I(a < u < v < b)$ equals

$$I(v < \theta) I(u < v) I(-\theta < u) = I(v < \theta) I(-u < \theta) = I\{\max(-u, v) < \theta\},$$

   because $U < V$ with probability one, so $S = \max(-U, V)$ is a sufficient statistic. It is also minimal sufficient, since (in an obvious notation and setting $0/0 = 1$)

$$\frac{f(z;\theta)}{f(y;\theta)} = \frac{(2\theta)^{-n} I(s_z < \theta)}{(2\theta)^{-n} I(s_y < \theta)} = \frac{I(s_z < \theta)}{I(s_y < \theta)}$$

   does not depend on θ iff $s_z = s_y$.

   In the second case

$$\frac{f(z;\theta)}{f(y;\theta)} = \frac{2^{-n} I(u_z < \theta < v_z)}{2^{-n} I(u_y < \theta < v_y)} = \frac{I(u_z < \theta < v_z)}{I(u_y < \theta < v_y)},$$

   so, as we must have $u_z = u_y$ and $v_z = v_y$ for this not to depend on θ, $S = (U, V)$ is minimal sufficient.

   It seems clear that the distribution of $A = V - U$ does not depend on θ. To formalise this we argue as follows: $Y_j \overset{d}{=} \theta + W_j$, where $W_1, \ldots, W_n \overset{iid}{\sim} U(-1, 1)$. Hence

$$V \overset{d}{=} \max(Y_1, \ldots, Y_n) = \max(\theta + W_1, \ldots, \theta + W_n) = \theta + \max W_j,$$

   and likewise $U \overset{d}{=} \theta + \min W_j$. Thus $A = V - U \overset{d}{=} \max W_j - \max W_j$, whose distribution does not depend on θ. Thus A is ancillary.

   A clumsier approach uses results on the joint densities of order statistics, which give

$$f_{U,V}(u, v) = \frac{n!}{1!(n-2)!1!} 2^{-n} (u - v)^{n-2}, \quad \theta - 1 < u < v < \theta + 1.$$

   The Jacobian for the transformation $(u, v) \mapsto (u, a) = (u, v - u)$ is unity, so the joint density of U and A is

$$f_{U,A}(u, a) = n(n-1) 2^{-n} a^{n-2} \times 1, \quad -1 < u - \theta < u - \theta + a < 1,$$

   and as $-1 < u - \theta < 1$ we have $-1 < a < 1$, and the marginal density of A is therefore

$$f_A(a) = \int_{\theta-1}^{\theta+1-a} n(n-1) 2^{-n} a^{n-2} \, \mathrm{d}u = n(n-1) 2^{-n} (2-a) a^{n-2}, \quad 0 < a < 2.$$

   As this does not depend on θ, A is ancillary.

2. Find minimal sufficient statistics in the following settings:

   (a) $Y_1, \ldots, Y_N \stackrel{i.i.d.}{\sim} \text{Poiss}(\theta)$ with N a geometric random variable with success probability θ; **Solution**: The joint density $f(y, n; \theta)$ of $y_1, \ldots, y_n, n$ is

$$f(y \mid n; \theta) f(n; \theta) = \prod_{j=1}^{n} f(y_j; \theta) \times f(n; \theta) = \prod_{j=1}^{n} \frac{\theta^{y_j}}{y_j!} e^{-\theta} \times (1-\theta)^{n-1}\theta \propto \theta^{s+1}(1-\theta)^{n-1} e^{-n\theta}, \quad 0 < \theta < 1,$$

   so the factorization theorem implies that $(S, N) = (Y_1 + \cdots + Y_N, N)$ is sufficient, and it is easy to check that it is also minimal, as the model is an exponential family. Here N is part of the minimal sufficient statistic because its value is informative about θ (very large N would imply that θ is very low, which would also reduce the likely value of S).

   (b) $Y_1, \ldots, Y_n \perp\!\!\!\perp \text{Poiss}(\theta_1, \ldots, \theta_n)$ for fixed n, where $\log \theta_j = x_j^\top \beta$ depends on known $d \times 1$ vectors of covariates $x_1, \ldots, x_n$ and an unknown parameter $\beta \in \mathbb{R}^d$; and
   Solution: The joint density is

$$\prod_{j=1}^{n} \frac{\theta_j^{y_j}}{y_j!} e^{-\theta_j} = \prod_{j=1}^{n} \frac{\exp(y_j x_j^\top \beta)}{y_j!} e^{-\exp(x_j^\top \beta)} \propto \exp\left(\sum_{j=1}^{n} y_j x_j^\top \beta - \sum_{j=1}^{n} e^{x_j^\top \beta} \right), \quad 0 < \theta < 1,$$

   so the $d \times 1$ vector $S = \sum_{j=1}^{n} Y_j x_j$ is sufficient for β, using the factorisation theorem. This is a (d, d) exponential family, so S is also minimal sufficient.

   (c) $Y_1, \ldots, Y_n \stackrel{iid}{\sim} \exp(\lambda)$. In this last case show also that $(Y_1/\bar{Y}, \ldots, Y_n/\bar{Y})$ is distribution-constant, and without computing their joint density show that it is independent of $\bar{Y}$. **Solution**: We can use exponential family results from Section 2.2, or just note that

$$f(y_1, \ldots, y_n; \lambda) = \lambda^n \exp(-\lambda s), \quad s = y_1 + \cdots + y_n > 0,$$

   so $S = Y_1 + \cdots + Y_n$ is minimal sufficient. This is an exponential family, so S is also complete, and therefore it is independent of any distribution-constant statistics (by Basu's theorem). It is clear from its density function that $\lambda Y_j \stackrel{d}{=} E_j$, where $E_1, \ldots, E_n \stackrel{iid}{\sim} \exp(1)$, so

$$(Y_1/\bar{Y}, \ldots, Y_n/\bar{Y}) \stackrel{d}{=} (E_1/\bar{E}, \ldots, E_n/\bar{E}),$$

   which is distribution-constant, and therefore independent of $\bar{Y} = S/n$.

3. If Y_1/θ and $Y_2\theta$ are independent gamma variables with unit scale parameter and shape parameter n, check that their joint density function is

$$f(y_1, y_2; \theta) = \frac{(y_1 y_2)^{n-1}}{\Gamma(n)^2} \exp(-y_1/\theta - \theta y_2), \quad y_1, y_2 > 0, \theta > 0.$$

   Solution: They are independent, so

$$f(y_1, y_2; \theta) = f(y_1; \theta) f(y_2; \theta) = \frac{y_1^{n-1}\theta^{-n}}{\Gamma(n)} e^{-y_1/\theta} \times \frac{y_2^{n-1}\theta^{n}}{\Gamma(n)} e^{-y_2\theta}, \quad y_1, y_2 > 0, \theta > 0.$$

   (a) Show that this is a $(2, 1)$ exponential family with minimal sufficient statistic $S = (T, A)$, where $T = (Y_1/Y_2)^{1/2}$ and $A = (Y_1 Y_2)^{1/2}$. **Solution**: On setting $\varphi(\theta)^\top = (1/\theta, \theta)$, $s(Y)^\top = (-Y_1, -Y_2)$ and $k(\theta) = 0$ we see that this is an exponential family in which φ lies in the set $\{(x, y) : xy = 1\} \subset \mathbb{R}_+^2$, a one-dimensional subset of the positive quadrant, while $-s(Y) \in \mathbb{R}_+^2$. Hence this is a $(2, 1)$ exponential family. It is clear from Example 36 that (Y_1, Y_2) are minimal sufficient.
   There is a 1–1 mapping between $(Y_1, Y_2) = (TA, A/T)$ and (T, A), which is thus minimal sufficient.

(b) Find the joint density of T and A, show that A is ancillary, and find the conditional density of T given A. **Solution**: The Jacobian of the transformation $(y_1, y_2) \mapsto (t, a)$ is $2a/t > 0$. Hence

$$f(t, a; \theta) = \frac{2a^{2n-1}}{t\Gamma(n)^2} \exp\{-a(t/\theta + \theta/t)\}, \quad a, t > 0, \theta > 0,$$

and the marginal density of A is

$$f(a; \theta) = C(a) \int_0^\infty t^{-1} \exp\{-a(t/\theta + \theta/t)\} \, \mathrm{d}t = C(a) \int_{-\infty}^\infty \exp(-2a \cosh w) \, \mathrm{d}w = C(a) I(a),$$

say, where $C(a) = 2a^{2n-1}/\Gamma(n)^2$ and we have changed variables from t to $w = \log(t/\theta)$, so $\mathrm{d}t/\mathrm{d}w = t$. As $f(a; \theta)$ does not depend on θ, A is ancillary. This was obvious from the problem statement, because if we write the two gamma variables as $X_1 \stackrel{d}{=} Y_1/\theta$ and $X_2 \stackrel{d}{=} Y_2\theta$, then the distribution of $A = (Y_1Y_2)^{1/2} \stackrel{d}{=} (\theta X_1 \times X_2/\theta)^{1/2} = (X_1X_2)^{1/2}$ does not depend on θ. Hence

$$f(t \mid a; \theta) = \{tI(a)\}^{-1} \exp\{-a(t/\theta + \theta/t)\}, \quad t > 0, \theta > 0.$$

(c) Show that the observed information for θ is proportional to a, compute the unconditional Fisher information, and hence discuss the role of A. **Solution**: The log likelihood is $\ell(\theta) \equiv -a(t/\theta + \theta/t)$, with first and second derivatives $a(t/\theta^2 - 1/t)$ and $-at/\theta^3$, so the unconditional Fisher information is $\mathrm{E}(AT/\theta^3) = \mathrm{E}(Y_1/\theta^3) = n\theta/\theta^3 = n/\theta^2$. This implies that a can be seen as an *observed sample size*, replacing n in the observed information.

The maximum likelihood estimate is $\hat{\theta} = t$ and $\jmath(\hat{\theta}) = a/\hat{\theta}^2$. The standard error $\jmath(\hat{\theta})^{-1/2} = \hat{\theta}/a^{1/2}$ is decreasing in a, which as we just saw plays the role of a sample size.

4. Independent exponential random variables Y_1 and Y_2 have respective densities $\theta_1 e^{-\theta_1 y_1}$ and $\theta_2 e^{-\theta_2 y_2}$, where $\theta_1, \theta_2 > 0$, and $\lambda, \psi > 0$ below.

(a) Find the joint density of Y_1 and Y_2 when $\theta_1 = \lambda$ and $\theta_2 = \lambda + \psi$. Inspect this and hence eliminate λ and thus obtain a $1 - 2\alpha$ confidence interval for ψ. **Solution**: The joint density is

$$f(y_1, y_2; \psi, \lambda) = \lambda(\lambda + \psi) \exp\{-(y_1 + y_2)\lambda - y_2\psi\}, \quad y_1, y_2 > 0, \psi, \lambda > 0.$$

This is a $(2, 2)$ exponential family with $\varphi = (\psi, \lambda)$ and $s(y) = (y_2, y_1 + y_2)$, so we can eliminate λ using the conditional density of $T = Y_2$ given $W = Y_1 + Y_2$. The Jacobian of the transformation $(y_1, y_2) \mapsto (t = y_2, w = y_1 + y_2)$ is 1, so the joint density of (T, W) is

$$f(t, w; \lambda, \psi) = f(y_1, y_2; \psi, \lambda) \times 1|_{y_1 = w - t, y_2 = t} = \lambda(\lambda + \psi) \exp(-w\lambda - t\psi), \quad 0 < t < w.$$

Hence the marginal density of W is

$$f(w; \lambda, \psi) = \int_0^w f(t, w; \lambda, \psi) \, \mathrm{d}t = \lambda(\lambda + \psi) \exp(-w\lambda)\psi^{-1}(1 - e^{-w\psi}),$$

and the required conditional density and distribution are of truncated exponential form

$$f(t \mid w; \psi) = \frac{\psi e^{-t\psi}}{1 - e^{-w\psi}}, \quad F(t \mid w; \psi) = \frac{1 - e^{-t\psi}}{1 - e^{-w\psi}}, \quad 0 < t < w.$$

Given observed values w^o and t^o, the limits of the $1 - 2\alpha$ confidence interval are the values of ψ that satisfy $F(t^o \mid w^o; \psi) = \alpha, 1 - \alpha$.

(b) If $\theta_1 = \lambda$ and $\theta_2 = \lambda\psi$ show that $\psi Y_2/Y_1$ is a pivot and find a $1-2\alpha$ confidence interval for ψ.

Solution: We let $Q = \psi Y_2/Y_1$ and note that

$$\begin{aligned}
\Pr(Q \le q) &= \Pr(\psi Y_2/q \le Y_1) \\
&= \int_0^\infty f(y_2;\psi,\lambda)\Pr(Y_1 > \psi y_2/q) \\
&= \int_0^\infty \lambda\psi e^{-\lambda\psi y_2}\, e^{-\lambda y_1}\big|_{y_1=\psi y_2/q}\ \mathrm{d}y_2 \\
&= \frac{\lambda\psi}{\lambda\psi(1+1/q)} \\
&= \frac{q}{1+q}, \quad q > 0.
\end{aligned}$$

Hence Q is a pivot on which inference for ψ can be based. The α quantile q_α of Q satisfies $\Pr(Q \le q_\alpha)$ and hence $q_\alpha = \alpha/(1-\alpha)$. Thus

$$1 - 2\alpha = \Pr(q_\alpha < Q \le q_{1-\alpha}) = \Pr(q_\alpha < \psi Y_2/Y_1 \le q_{1-\alpha}),$$

and hence the $1-2\alpha$ confidence interval based on observed data y_1^o, y_2^o has limits

$$\frac{\alpha}{1-\alpha}\frac{y_1^o}{y_2^o}, \quad \frac{1-\alpha}{\alpha}\frac{y_1^o}{y_2^o}.$$

(c) If $\theta_1 = \lambda$ and $\theta_2 = \lambda\psi$, show that λ can be eliminated by conditioning on $W_\psi = Y_1 + \psi Y_2$, and that the conditional distribution of $T = Y_1$ given W_ψ is

$$\Pr(T \le t \mid W_\psi = w_\psi;\psi) = \frac{t}{w_\psi}, \quad 0 < t < w_\psi.$$

Deduce that the resulting confidence interval for ψ is the same as in (b). **Solution**: This model is not a linear exponential family, but if ψ is fixed then the density is

$$f(y_1, y_2;\psi,\lambda) = \lambda^2\psi\exp\{-\lambda(y_1 + \psi y_2)\}, \quad y_1, y_2 > 0,$$

and λ can be eliminated by conditioning on $w_\psi = y_1 + \psi y_2$. If we set $t = y_1$, then $|\partial(w_\psi, t)/\partial(y_1, y_2)| = \psi > 0$, so

$$f(t, w_\psi;\psi,\lambda) = f(y_1, y_2;\psi,\lambda)\left|\left|\frac{\partial(y_1,y_2)}{\partial(w_\psi,t)}\right|\right|_{y_1=t, y_2=(w_\psi-t)/\psi} = \lambda^2\exp(-\lambda w_\psi), \quad 0 < t < w_\psi,$$

and thus the marginal density of W_ψ is

$$f(w_\psi;\psi,\lambda) = \int_0^{w_\psi} f(t, w_\psi;\psi,\lambda)\ \mathrm{d}t = \lambda^2 w_\psi\exp(-\lambda w_\psi), \quad w_\psi > 0,$$

giving

$$f(t \mid w_\psi;\psi) = \frac{f(t, w_\psi;\psi,\lambda)}{f(w_\psi;\psi,\lambda)} = w_\psi^{-1}, \quad 0 < t < w_\psi,$$

i.e., $T \mid W_\psi = w_\psi \sim U(0, w_\psi)$.

The limits of a $1-2\alpha$ confidence interval for ψ solve the equations $F(t^o \mid w_\psi^o;\psi) = \alpha, 1-\alpha$, and setting $t^o = y_1^o$ and $w_\psi^o = y_1^o + \psi y_2^o$ leads to the interval in (b).