MATH5905 Assignment 2

Peter Bradshaw z3463859

October 10, 2017

1. a) If $X = (X_1, X_2, \ldots, X_n)$, observe that

$$L(X, \theta) = \frac{2^n \theta^{2n}}{\prod_{i=1}^n X_i^3} \prod_{i=1}^n 1_{(0,X_i)}(\theta) = \frac{2^n \theta^{2n}}{\prod_{i=1}^n X_i^3} 1_{(0,X_{(1)})}(\theta).$$

If $Y = (Y_1, Y_2, \ldots, Y_n)$ is another set of data with the same density, then

$$\begin{aligned}\frac{L(Y,\theta)}{L(X,\theta)} &= \frac{\frac{2^n\theta^{2n}}{\prod_{i=1}^n Y_i^3} 1_{(0,Y_{(1)})}(\theta)}{\frac{2^n\theta^{2n}}{\prod_{i=1}^n X_i^3} 1_{(0,X_{(1)})}(\theta)} \\ &= \prod_{i=1}^n \left(\frac{X_i}{Y_i}\right)^3 \cdot \frac{1_{(0,Y_{(1)})}(\theta)}{1_{(0,X_{(1)})}(\theta)}\end{aligned}$$

which is independent of θ if an only if $X_{(1)} = Y_{(1)}$. By Lehmann-Scheffe, the induced partition $\{A(x)\}$ where $A(x) = \{y : \frac{L(y,\theta)}{L(x,\theta)} \text{ indep of } \theta\}$ is minimal sufficient. Since $A(x) = \{y : Y_{(1)} = X_{(1)}\}$, we conclude that $T(X) = X_{(1)}$ is a minimal sufficient statistic.

b) Firstly note that for $x < \theta$, we have $\mathbb{P}(X_1 \geq x) = 1$ and for $x \geq \theta$

$$\mathbb{P}(X_1 \geq x) = \int_x^\infty \frac{2\theta^2}{t^3} dt = \left(\frac{\theta}{x}\right)^2.$$

Thus

$$\begin{aligned}F_{X_{(1)}}(x;\theta) &= \mathbb{P}(X_{(1)} \leq x) \\ &= 1 - \mathbb{P}(X_1 \geq x, X_2 \geq x, \ldots, X_n \geq x) \\ &= 1 - \mathbb{P}(X_1 \geq x)^n \qquad \text{by independence} \\ &= \begin{cases} 1 - (\frac{\theta}{x})^{2n} & x \geq \theta \\ 0 & x < \theta \end{cases}.\end{aligned}$$

Differentiating gives

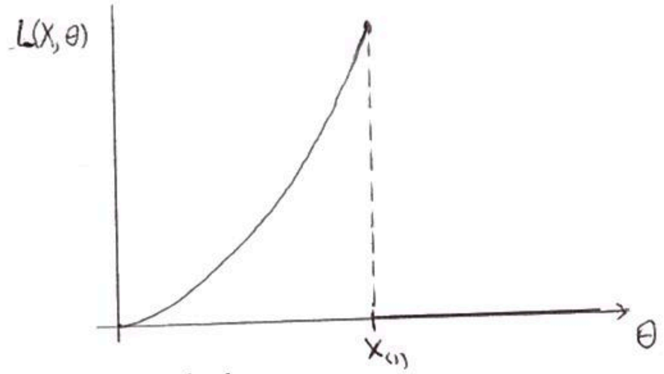
$$f_{X_{(1)}}(x;\theta) = \begin{cases} \frac{2n\theta^{2n}}{x^{2n+1}} & x \geq \theta \\ 0 & x < \theta \end{cases},$$

as required.

c) The MLE for θ is calculated by maximising over all θ values

$$L(X,\theta) = \frac{2^n \theta^{2n}}{\prod_{i=1}^n X_i^3} 1_{(0,X_{(1)})}(\theta).$$

The graph of this function is given below, and is clearly the MLE is $\hat{\theta}_{\text{MLE}} = X_{(1)}$.



Then we calculate

$$\begin{aligned} E(\hat{\theta}_{\text{MLE}}) &= EX_{(1)} \\ &= \int_\theta^\infty \frac{2n\theta^{2n}}{x^{2n}} dx \\ &= \left[-\frac{2n}{2n-1} \theta^{2n} x^{-2n+1} \right]_\theta^\infty \\ &= \frac{2n}{2n-1}\theta \end{aligned}$$

which is not equal to θ and therefore $\hat{\theta}_{\text{MLE}}$ is biased.

d) Firstly notice that $\theta > 0$ and $T(X) = X_{(1)}$ takes values greater than θ. Consider a function g (with natural domain) and assume that $E_\theta g(T) = 0$. This means that

$$\int_\theta^\infty g(x) \frac{2n\theta^{2n}}{x^{2n+1}} dx = 0 \qquad \text{for all } \theta > 0.$$

Treating both sides as functions of θ and differentiating, we get

$$-g(\theta) \frac{2n\theta^{2n}}{\theta^{2n+1}} = 0,$$

or equivalently

$$g(\theta) \cdot \frac{2n}{\theta} = 0,$$

for all $\theta > 0$. This implies that $g(\theta) = 0$ for $\theta > 0$. Then since $g(T) = 0$ everywhere where T has positive density, we conclude that $\mathbb{P}_\theta(g(T) = 0) = 1$. Thus, $T = X_{(1)}$ is complete.

Now from (c), $W = \frac{2n-1}{2n} X_{(1)}$ is an unbiased estimator for θ. Thus, by Lehmann-Scheffe theorem, $\hat{\theta} = E(W|X_{(1)}) = \frac{2n-1}{2n} X_{(1)}$ is the unique UMVUE of θ.

e) We calculate for $\theta > \theta'$

$$
\begin{aligned}
\frac{L(X,\theta)}{L(X,\theta')} &= \frac{\frac{2^n\theta^{2n}}{\prod_{i=1}^n X_i^3} 1_{(0,X_{(1)})}(\theta)}{\frac{2^n\theta'^{2n}}{\prod_{i=1}^n X_i^3} 1_{(0,X_{(1)})}(\theta')} \\
&= \left(\frac{\theta}{\theta'}\right)^{2n} \cdot \frac{1_{(0,X_{(1)})}(\theta)}{1_{(0,X_{(1)})}(\theta')} \\
&= \begin{cases} \text{undefined} & X_{(1)} \le \theta' \\ 0 & \theta' < X_{(1)} \ge \theta \\ \left(\frac{\theta}{\theta'}\right)^{2n} & X_{(1)} > \theta \end{cases} .
\end{aligned}
$$

This is a non-decreasing function of $X_{(1)}$ so the family $\{L(X,\theta)\}$ has monotone likelihood ratio in $T = X_{(1)}$.

f) For a continuous distribution the UMP α-test has the form

$$
\varphi^*(X) = \begin{cases} 1 & \text{if } T(X) > k \\ 0 & \text{if } T(X) < k \end{cases} .
$$

It remains only to find k such that $E_{\theta=1}\varphi^*(X) = \alpha$, or equivalently, that $\mathbb{P}_{\theta=1}(X_{(1)} > k) = \alpha$. Noting that $\mathbb{P}_{\theta=1}(X_{(1)} > k) = \int_k^\infty \frac{2n}{x^{2k+1}}dx = (\frac{1}{k})^{2n}$, we conclude that $k = \alpha^{-1/2n}$. So the UMP α-test is

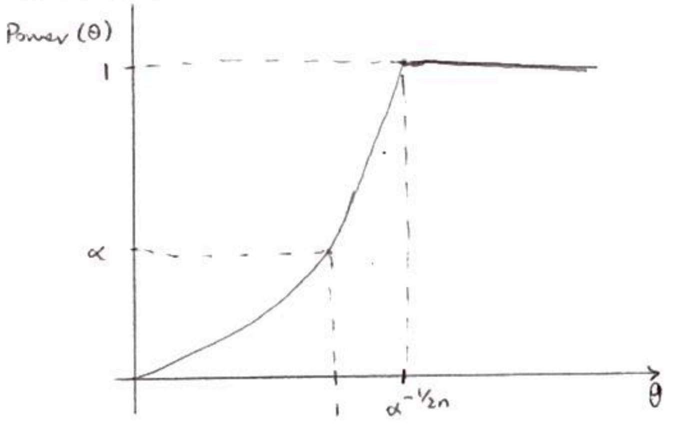
$$
\varphi^*(X) = \begin{cases} 1 & \text{if } T(X) > \alpha^{-1/2n} \\ 0 & \text{if } T(X) < \alpha^{-1/2n} \end{cases} .
$$

We accept H_0 if $X_{(1)} \le \alpha^{-1/2n}$ and reject H_0 if $X_{(1)} > \alpha^{-1/2n}$.

g) Finally,

$$
\begin{aligned}
Power(\theta) &= E_\theta\varphi^*(X) \\
&= \mathbb{P}(X_{(1)} > \alpha^{-1/2n}) \\
&= \left(\frac{\theta}{\alpha^{-1/2n}}\right)^{2n} \\
&= \theta^{2n}\alpha,
\end{aligned}
$$

for $\theta \le \alpha^{-1/2n}$. For $\theta > \alpha^{-1/(2n)}$, the power function is 1. A graph is sketched below.



2. a) If $T = \sum_{i=1}^n (\log X_i)^2$, the likelihood function of $X = (X_1, X_2, \ldots, X_n)$ is given by

$$\begin{aligned} L(X,\theta) &= \frac{1}{(2\pi)^{n/2}\theta^n \prod_{i=1}^n X_i} e^{-\frac{1}{2}\sum_{i=1}^n (\frac{\log X_i}{\theta})^2} \\ &= \frac{1}{(2\pi)^{n/2}\theta^n \prod_{i=1}^n X_i} e^{-\frac{1}{2\theta^2}T}. \end{aligned}$$

Then for $\theta > \theta'$,

$$\begin{aligned} \frac{L(X,\theta)}{L(X,\theta')} &= \frac{\frac{1}{(2\pi)^{n/2}\theta^n \prod_{i=1}^n X_i} e^{-\frac{1}{2\theta^2}T}}{\frac{1}{(2\pi)^{n/2}\theta'^n \prod_{i=1}^n X_i} e^{-\frac{1}{2\theta'^2}T}} \\ &= \left(\frac{\theta'}{\theta}\right)^n e^{\frac{T}{2}(\frac{1}{\theta'^2} - \frac{1}{\theta^2})}, \end{aligned}$$

is an increasing function in T. Thus $\{L(X,\theta)\}$ is a family with a monotone likelihood ratio in T.

b) Let θ be in the alternative set, or equivalently $\theta > \theta_0$. Since X_i came from a continuous distribution, the lemma of Neyman–Pearson gives an α-test of the form

$$\varphi^*(X) = \begin{cases} 1 & \text{if } \frac{L(X,\theta)}{L(X,\theta_0)} > K \\ 0 & \text{otherwise} \end{cases},$$

which, by the monotonicity of $\frac{L(X,\theta)}{L(X,\theta_0)}$ is equivalent to

$$\varphi^*(X) = \begin{cases} 1 & \text{if } T > k \\ 0 & \text{it } T \leq k \end{cases}.$$

Since $\alpha = E_{\theta_0}\varphi^* = \mathbb{P}_{\theta_0}(T > k)$, it is clear that k is independent of the choice of θ. Since the Neyman–Pearson lemma gives the most powerful test for $\Theta = \{\theta_0, \theta\}$, and this test doesn't depend on which element θ of the alternative set we take, we conclude that φ^* is uniformly most powerful, the UMP α-test.

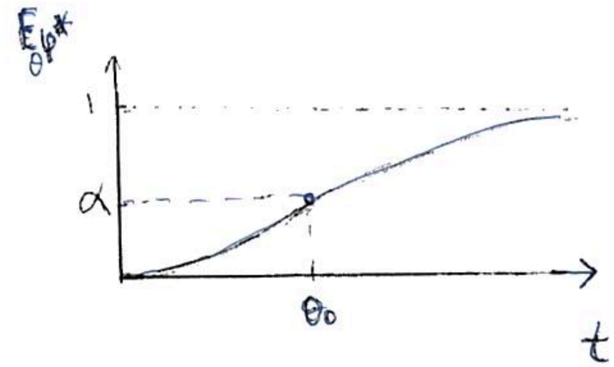
c) If we let $Y_i = \log X_i$ for all i, then $X_i = e^{Y_i}$. By the density transformation formula, this gives

$$\begin{aligned} f_{Y_i}(y) &= f(e^y) \cdot e^y \\ &= \frac{e^y}{\sqrt{2\pi}e^y\theta} e^{-\frac{1}{2}\left(\frac{\log e^y}{\theta}\right)^2} \\ &= \frac{1}{\sqrt{2\pi}\theta} e^{-\frac{y^2}{2\theta^2}}, \end{aligned}$$

the pfd of a $N(0,\theta^2)$ random variable. Thus $\log X_i = Y_i \sim N(0,\theta^2)$ for all i.

$$\varphi^* = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} ln(X_i)^2 \geq \theta_0^2 \chi^2_{n,\alpha} \\ 0 & \text{if } \sum_{i=1}^{n} ln(X_i)^2 < \theta_0^2 \chi^2_{n,\alpha} \end{cases}$$

$$\begin{aligned} Power(t) &= P(\sum_{i=1}^{n} ln(X_i)^2 \geq \theta_0^2 \chi^2_{n,\alpha} | \theta = t) \\ &= P(\sum_{i=1}^{n} Y_i^2 \geq \theta_0^2 \chi^2_{n,\alpha} | \theta = t) \\ &= P(\chi^2_n \geq (\frac{\theta_0}{t})^2 \chi^2_{n,\alpha}) \end{aligned}$$



4

d) Observe that $\frac{T}{\theta^2} = \sum_{i=1}^{n} \left(\frac{\log X_i}{\theta}\right)^2 \sim \chi_n^2$. We know that $\mathbb{P}_{\theta_0}(T > k) = \alpha$ so $\chi^2_{n,\alpha} = \frac{k}{\theta_0^2}$ where $\chi^2_{n,\alpha}$ is number giving the upper α-region of a χ^2_n distribution. Rearranging and using part (ii) gives the UMP α-test

$$\varphi^*(X) = \begin{cases} 1 & \text{if } T > \theta_0^2\chi^2_{n,\alpha} \\ 0 & \text{it } T \leq \theta_0^2\chi^2_{n,\alpha} \end{cases}.$$

4. i) Using the formula in the notes with $f(x) = e^{-x}, x > 0$ and therefore $F(x) = 1 - e^{-x}, x > 0$, and also $n = 4$ and $r = 3$, we get

$$\begin{aligned} f_{X_{(3)}}(x) &= \frac{4!}{2!1!}(1-e^{-x})^2(1-(1-e^{-x})e^{-x} \\ &= 12(1-e^{-x})^2e^{-2x}. \end{aligned}$$

Then $E(X_3) = \int_0^\infty 12x(1-e^{-x})^2e^{-2x}dx = \frac{13}{12}$, using computer program for integration.

ii) The joint distribution of $X_{(1)}$ and $X_{(4)}$ is

$$f_{(X_{(1)},X_{(4)})}(x,y) = 12(e^{-x} - e^{-y})^2e^{-(x+y)},$$

for $0 < x < y < \infty$. If we apply the transformation,

$$U = \frac{1}{2}(X_{(1)} + X_{(4)}) \qquad V = X_{(1)},$$

then we can equivalently write this as

$$X_{(1)} = V \qquad X_{(4)} = 2U - V.$$

The absolute determinant of the Jacobian evaluates to 2 and we can apply a density transformation formula to obtain

$$\begin{aligned} f_{(U,V)}(u,v) &= f_{(X_{(1)},X_{(4)})}(v, 2u - v) \cdot 2 \\ &= 24(e^{-v} - e^{-2u+v})^2e^{-2u}, \end{aligned}$$

for $0 < v < u < \infty$. Then if $B = \frac{1}{2}(X_{(1)} + X_{(4)})$, then the marginal density is given by

$$\begin{aligned} f_B(u) &= 24\int_0^u (e^{-v} - e^{-2u+v})^2e^{-2u}dv \\ &= 24e^{-2u}\int_0^u e^{-2v} - 2e^{-2u} + e^{-4u+2v}dv \\ &= 24e^{-2u}\left[-\frac{e^{-2v}}{2} - 2e^{-2u}v + \frac{e^{-4u+2v}}{2}\right]_0^u \\ &= 12e^{-2u}(1 - e^{-2u} - 4ue^{-2u} + e^{-2u} - e^{-4u}) \\ &= 12e^{-2u}(1 - e^{-4u} - 4ue^{-2u}), \end{aligned}$$

for $0 < u < \infty$. Then by integration on a computer

$$\mathbb{P}(B > 1) = \int_1^\infty 12e^{-2x}(1-e^{-2x}-4xe^{-2x})dx = \frac{6e^4 - 15e^2 - 2}{e^6} \approx 0.532320.$$

$\boxed{Q3}$

```
In[112]:= f = 1/Sqrt[2 π] Exp[-x^2/2]; domain[f] = {x, -∞, ∞};

          g = OrderStat[r, f, 3]
          domain[g] = OrderStatDomain[r, f, 3]

          g1 = OrderStat[1, f, 3]
          domain[g1] = OrderStatDomain[1, f, 3]
          g2 = OrderStat[2, f, 3]
          domain[g2] = OrderStatDomain[2, f, 3]
          g3 = OrderStat[3, f, 3]
          domain[g3] = OrderStatDomain[3, f, 3]

          PlotDensity[g /. {r → {1, 2, 3}}]

          Expect[x, g1]
          Expect[x, g2]
          Expect[x, g3]
```

Out[113]= $$\frac{3\, e^{-\frac{x^2}{2}} \left(1 - \mathrm{Erf}\left[\frac{x}{\sqrt{2}}\right]\right)^{3-r} \left(1 + \mathrm{Erf}\left[\frac{x}{\sqrt{2}}\right]\right)^{-1+r}}{2\sqrt{2\pi}\ (3-r)!\ (-1+r)!}$$

Out[114]= {x, -∞, ∞} && {r ∈ Integers, 1 ≤ r ≤ 3}

Out[115]= $$\frac{3\, e^{-\frac{x^2}{2}} \left(-1 + \mathrm{Erf}\left[\frac{x}{\sqrt{2}}\right]\right)^{2}}{4\sqrt{2\pi}}$$

Out[116]= {x, -∞, ∞}

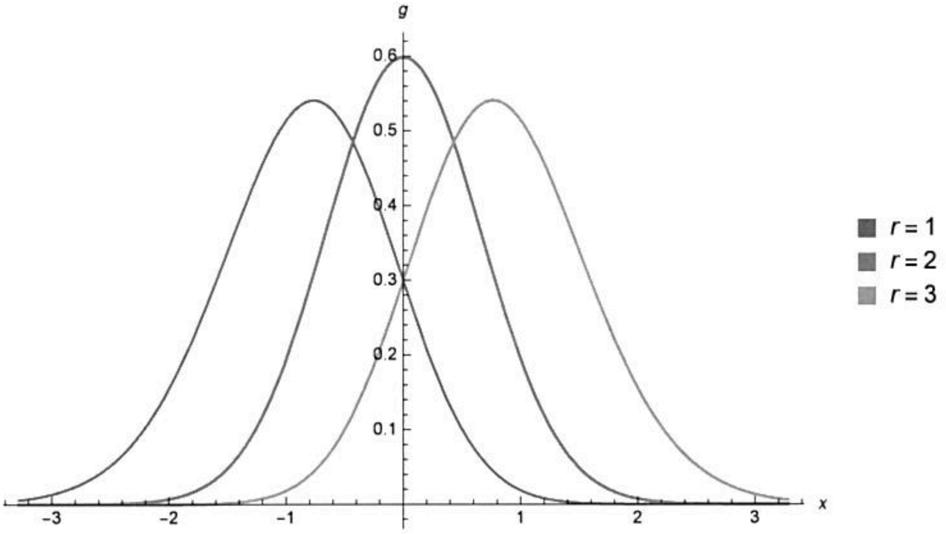
Out[117]= $$-\frac{3\, e^{-\frac{x^2}{2}} \left(-1 + \mathrm{Erf}\left[\frac{x}{\sqrt{2}}\right]^{2}\right)}{2\sqrt{2\pi}}$$

Out[118]= {x, -∞, ∞}

Out[119]= $$\frac{3\, e^{-\frac{x^2}{2}} \left(1 + \mathrm{Erf}\left[\frac{x}{\sqrt{2}}\right]\right)^{2}}{4\sqrt{2\pi}}$$

Out[120]= {x, -∞, ∞}

Out[121]=



Out[122]= $-\frac{3}{2\sqrt{\pi}}$

Out[123]= 0

Out[124]= $\frac{3}{2\sqrt{\pi}}$

In[136]:=
```
f = 1/2 Exp[-Abs[x]]; domain[f] = {x, -∞, ∞};

g = OrderStat[r, f, 3]
domain[g] = OrderStatDomain[r, f, 3]

g1 = OrderStat[1, f, 3]
domain[g1] = OrderStatDomain[1, f, 3]
g2 = OrderStat[2, f, 3]
domain[g2] = OrderStatDomain[2, f, 3]
g3 = OrderStat[3, f, 3]
domain[g3] = OrderStatDomain[3, f, 3]

PlotDensity[g /. {r → {1, 2, 3}}]

Expect[x, g1]
Expect[x, g2]
Expect[x, g3]
```

Out[137]= $\begin{cases} \frac{3\, e^{r x} (2-e^{x})^{3-r}}{4\,(3-r)!\,(-1+r)!} & x \le 0 \\ \frac{3\, e^{(-3+r) x} (2-e^{-x})^{r}}{4\,(-1+2 e^{x})\,(3-r)!\,(-1+r)!} & \text{True} \end{cases}$

Out[138]= {x, -∞, ∞} && {r ∈ Integers, 1 ≤ r ≤ 3}

Out[139]= $\begin{cases} \frac{3\, e^{-3 x}}{8} & x \ge 0 \\ \frac{3}{8} e^{x} (-2 + e^{x})^{2} & \text{True} \end{cases}$

Out[140]= $\{x, -\infty, \infty\}$

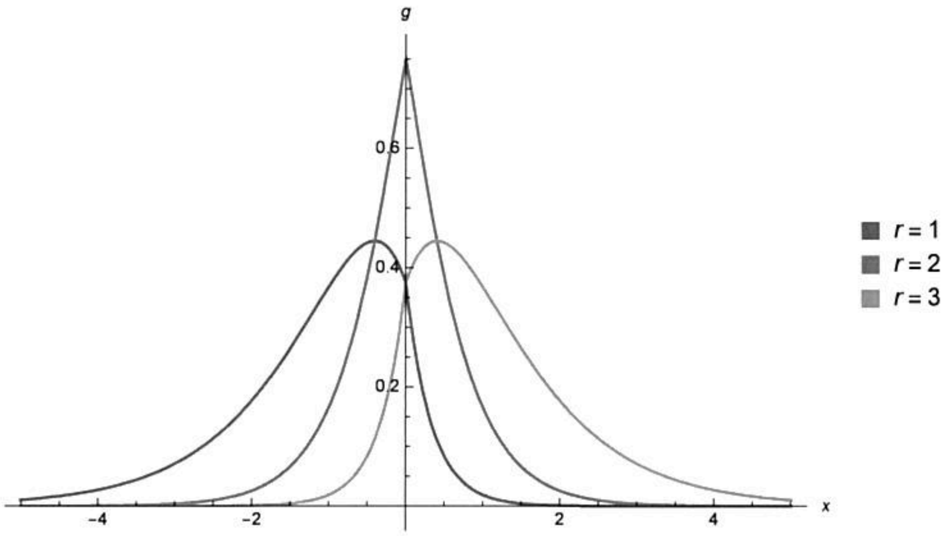
Out[141]= $\begin{cases} -\frac{3}{4} e^{2x} (-2 + e^{x}) & x \le 0 \\ \frac{3}{4} e^{-3x} (-1 + 2 e^{x}) & \text{True} \end{cases}$

Out[142]= $\{x, -\infty, \infty\}$

Out[143]= $\begin{cases} \frac{3 e^{3x}}{8} & x \le 0 \\ \frac{3}{8} e^{-3x} (1 - 2 e^{x})^2 & \text{True} \end{cases}$

Out[144]= $\{x, -\infty, \infty\}$

Out[145]=



Out[146]= $-\frac{9}{8}$

Out[147]= 0

Out[148]= $\frac{9}{8}$