STAT743 FOUNDATIONS OF STATISTICS (PART II)

Winter 2019

Assignment 4 Solutions

Q. 1 **a)** To show that $X_nY_n \xrightarrow{p} 0$ we need to show that, for any $\delta > 0$ and $\varepsilon > 0$ there exists N such that

$$n > N \quad \Rightarrow \quad \mathrm{P}\left(|X_nY_n| > \varepsilon\right) \; < \; \delta$$

To do this we note that

$$\begin{aligned}
\mathrm{P}\left(|X_nY_n| > \varepsilon\right) &= \left(|X_n||Y_n| > \varepsilon\right) \\
&\leqslant \mathrm{P}\left(|X_n| > \sqrt{\varepsilon} \text{ OR } |Y_n| > \sqrt{\varepsilon}\right) \\
&\qquad (\text{since } |X_n||Y_n| > \varepsilon \;\Rightarrow\; |X_n| > \sqrt{\varepsilon} \;\cup\; |Y_n| > \sqrt{\varepsilon}) \\
&\leqslant \mathrm{P}\left(|X_n| > \sqrt{\varepsilon}\right) + \mathrm{P}\left(|Y_n| > \sqrt{\varepsilon}\right) \\
&\qquad (\text{since } \mathrm{P}(A \;\cup\; B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \bigcap B) \leqslant \mathrm{P}(A) + \mathrm{P}(B))
\end{aligned}$$

[4 marks]

Now $X_n \xrightarrow{p} 0$ implies that there exists N_1 such that

$$n > N_1 \quad \Rightarrow \quad \mathrm{P}\left(|X_n| > \sqrt{\varepsilon}\right) \; < \; \frac{\delta}{2}$$

And similarly $Y_n \xrightarrow{p} 0$ implies there exists N_2 such that

$$n > N_2 \quad \Rightarrow \quad \mathrm{P}\left(|Y_n| > \sqrt{\varepsilon}\right) \; < \; \frac{\delta}{2}$$

Therefore

$$n > \max\{N_1, N_2\} \quad \Rightarrow \quad \mathrm{P}\left(|X_nY_n| > \varepsilon\right) \; \leqslant \; \mathrm{P}\left(|X_n| > \sqrt{\varepsilon}\right) + \mathrm{P}\left(|Y_n| > \sqrt{\varepsilon}\right) \; < \; \delta$$

and so $X_nY_n \xrightarrow{p} 0$ as required. **[4 marks]**

b) This proceeds in a similar way to the first part of the question. First we note that the Triangle Inequality tells us $|X_n + Y_n - (X+Y)| < |X_n - X| + |Y_n - Y|$ and so

$$\begin{aligned}
&\mathrm{P}\left(|X_n + Y_n - (X+Y)| > \varepsilon\right) \\
&\leqslant\ \mathrm{P}\left(|X_n - X| + |Y_n - Y| > \varepsilon\right) \\
&\leqslant\ \mathrm{P}\left(|X_n - X| > \frac{\varepsilon}{2}\ \cup\ |Y_n - Y| > \frac{\varepsilon}{2}\right) \\
&\qquad\qquad (\text{since } |X_n - X| + |Y_n - Y| > \varepsilon\ \Rightarrow\ |X_n - X| > \varepsilon/2\ \cup\ |Y_n - Y| > \varepsilon/2) \\
&\leqslant\ \mathrm{P}\left(|X_n - X| > \frac{\varepsilon}{2}\right) + \mathrm{P}\left(|Y_n - Y| > \frac{\varepsilon}{2}\right) \\
&\qquad\qquad (\text{since } \mathrm{P}(A\ \cup\ B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \bigcap B) \leqslant \mathrm{P}(A) + \mathrm{P}(B))
\end{aligned}$$

[4 marks]

Now $X_n \xrightarrow{p} X$ implies that there exists N_1 such that

$$n > N_1 \quad \Rightarrow \quad \mathrm{P}\left(|X_n - X| > \frac{\varepsilon}{2}\right) \ < \ \frac{\delta}{2}$$

And similarly $Y_n \xrightarrow{p} Y$ implies there exists N_2 such that

$$n > N_2 \quad \Rightarrow \quad \mathrm{P}\left(|Y_n - Y| > \frac{\varepsilon}{2}\right) \ < \ \frac{\delta}{2}$$

Therefore if we set $N = \max\{N_1, N_2\}$ we have that $n > N$ implies

$$\mathrm{P}\left(|X_n + Y_n - (X+Y)| > \varepsilon\right) \ \leqslant\ \mathrm{P}\left(|X_n - X| > \frac{\varepsilon}{2}\right) + \mathrm{P}\left(|Y_n - Y| > \frac{\varepsilon}{2}\right) \ < \ \delta$$

and so $X_n + Y_n \xrightarrow{p} X + Y$ as required. **[4 marks]**

c)

$$Z_n \ = \ \sqrt{n}(Y_n - \mu) \xrightarrow{d} Z \ \sim \ \text{normal}(0, \sigma^2)$$

Now for every fixed n and $\varepsilon > 0$ we have

$$\mathrm{P}\left(|Y_n - \mu| < \varepsilon\right) \ = \ \mathrm{P}\left(|Z_n| < \sqrt{n}\varepsilon\right)$$

so that

$$\begin{aligned}
\mathrm{P}\left(|Y_n - \mu| < \varepsilon\right) \ &= \ \mathrm{P}\left(|Z_n| < \sqrt{n}\varepsilon\right) \\
&= \ \mathrm{P}\left(-\sqrt{n}\varepsilon < Z_n < \sqrt{n}\varepsilon\right) \\
&= \ F_n(\sqrt{n}\varepsilon) - F_n(-\sqrt{n}\varepsilon)
\end{aligned}$$

where F_n is the cumulative distribution function of Z_n.

Now for any N, because of monotonicity of the F_n, we have that

$$n > N \quad \Rightarrow \quad F_n\left(\sqrt{n}\varepsilon\right) \geqslant F_n\left(\sqrt{N}\varepsilon\right) \quad \text{and} \quad F_n\left(-\sqrt{n}\varepsilon\right) \leqslant F_n\left(-\sqrt{N}\varepsilon\right)$$

so that for $n > N$ we have

$$\begin{aligned} \mathrm{P}\left(|Y_n - \mu| < \varepsilon\right) &\geqslant F_n\left(\sqrt{N}\varepsilon\right) - F_n\left(-\sqrt{N}\varepsilon\right) \\ &= 1 - 2\Phi\left(-\sqrt{N}\varepsilon\right) + F_n\left(\sqrt{N}\varepsilon\right) - \Phi\left(\sqrt{N}\varepsilon\right) - F_n\left(-\sqrt{N}\varepsilon\right) + \Phi\left(-\sqrt{N}\varepsilon\right) \end{aligned}$$

where Φ is the cdf of the normal$(0,1)$ and for any $a > 0$, $\Phi(a) - \Phi(-a) < 1 - 2\Phi(-a)$. For any $\varepsilon > 0, \delta > 0$ we can choose an N_0 such that $\Phi\left(-\sqrt{N_0}\varepsilon\right) < \delta/4$ and since $F_n(x) \to \Phi(x)$ at every $x \in \mathbb{R}$ we can also find $N > N_0$ such that

$$n > N \quad \Rightarrow \quad \left|F_n\left(\sqrt{N}\varepsilon\right) - \Phi\left(\sqrt{N}\varepsilon\right)\right| < \frac{\delta}{4} \quad \text{and} \quad \left|F_n\left(-\sqrt{N}\varepsilon\right) - \Phi\left(-\sqrt{N}\varepsilon\right)\right| < \frac{\delta}{4}$$

from which we see that

$$n > N \quad \Rightarrow \quad F_n\left(\sqrt{N}\varepsilon\right) - \Phi\left(\sqrt{N}\varepsilon\right) > -\frac{\delta}{4} \quad \text{and} \quad F_n\left(-\sqrt{N}\varepsilon\right) - \Phi\left(-\sqrt{N}\varepsilon\right) < \frac{\delta}{4}$$

Hence for such an N we have that

$$n > N \quad \Rightarrow \quad \mathrm{P}\left(|Y_n - \mu| < \varepsilon\right) \; > \; 1 - 2\frac{\delta}{4} - \frac{\delta}{4} - \frac{\delta}{4} \; = \; 1 - \delta$$

Hence we have that for any $\varepsilon > 0$, $\delta > 0$ we can find an N such that

$$\mathrm{P}\left(|Y_n - \mu| < \varepsilon\right) \; > \; 1 - \delta$$

and so $Y_n \xrightarrow{p} \mu$ as required. **[9 marks]**

Q. 2 **a)** For any $\varepsilon > 0, B > 0$ we have

$$\begin{aligned}
\mathrm{P}\big(|X_nY_n| > \varepsilon\big) &= \mathrm{P}\big(|X_n||Y_n| > \varepsilon\big) \\
&= \mathrm{P}(|X_n||Y_n| > \varepsilon, |X_n| \leqslant B\big) + \mathrm{P}(|X_n||Y_n| > \varepsilon, |X_n| > B\big) \\
&\leqslant \mathrm{P}\big(|Y_n| > \varepsilon/B\big) + \mathrm{P}(|X_n| > B\big)
\end{aligned}$$

Take an arbitrary $\delta > 0$ then, since X_n is bounded in probability we find N_1 and B such that

$$n > N_1 \quad \Rightarrow \quad \mathrm{P}\big(|X_n| > B\big) \;=\; 1 - \mathrm{P}\big(|X_n| \leqslant B\big) \;\leqslant\; \frac{\delta}{2}$$

Also since $Y_n \xrightarrow{p} 0$ we can find N_2 such that

$$n > N_2 \quad \Rightarrow \quad \mathrm{P}\big(|Y_n| > \varepsilon\big) \;<\; \frac{\delta}{2}$$

Hence, for any fixed $\delta > 0, \varepsilon > 0$, we can find $N > \max\{N_1, N_2\}$ such that

$$n > N \quad \Rightarrow \quad \mathrm{P}\big(|X_nY_n| > \varepsilon\big) \;\leqslant\; \mathrm{P}\big(|Y_n| > \varepsilon/B\big) + \mathrm{P}(|X_n| > B\big) \;<\; \delta$$

and so $X_nY_n \xrightarrow{p} 0$ as required. **[8 marks]**

b) Using two terms in the Taylors expansion we have

$$g(Y_n) - g(\theta) \;=\; g'(\theta)(Y_n - \theta) + \frac{g''(\theta)}{2}(Y_n - \theta)^2 + R_2(Y_n)$$

However, the statement of the Theorem tells us that $g'(\theta) = 0$ so we have

$$g(Y_n) - g(\theta) \;=\; \frac{g''(\theta)}{2}(Y_n - \theta)^2 + R_2(Y_n)$$

which we can rewrite as

$$\frac{n\big(g(Y_n) - g(\theta)\big)}{\sigma^2} \;=\; \left(\frac{\sqrt{n}(Y_n - \theta)}{\sigma}\right)^2 \left[\frac{g''(\theta)}{2} + \frac{R_2(Y_n)}{(Y_n - \theta)^2}\right]$$

Now Taylor's Theorem tells us that

$$\lim_{y \to \theta} \frac{R_2(y)}{(y - \theta)^2} \;=\; 0$$

which means that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|y - \theta| < \delta \;\Rightarrow\; \frac{|R_2(y)|}{(y - \theta)^2} < \varepsilon$$

Hence we have that

$$\mathrm{P}\left(\frac{|R_2(Y_n)|}{(Y_n - \theta)^2} < \varepsilon\right) \;\geqslant\; \mathrm{P}\big(|Y_n - \theta| < \delta\big)$$

Since we showed in Question 1(c) above that $Y_n \xrightarrow{p} \theta$ we can choose N such that

$$n > N \quad \Rightarrow \quad \mathrm{P}\left(\frac{|R_2(Y_n)|}{(Y_n-\theta)^2} < \varepsilon\right) \;\geqslant\; \mathrm{P}\big(|Y_n - \theta| < \delta\big) \;\geqslant\; 1-\delta$$

and so we have that

$$\frac{R_2(Y_n)}{(Y_n-\theta)^2} \xrightarrow{p} 0 \quad \Rightarrow \quad \frac{g''(\theta)}{2} + \frac{R_2(Y_n)}{(Y_n-\theta)^2} \xrightarrow{p} \frac{g''(\theta)}{2}$$

[8 marks]

Now let us consider the random variable

$$Z_n \;=\; \frac{\sqrt{n}(Y_n-\theta)}{\sigma^2} \;\xrightarrow{d}\; Z \;\sim\; \mathrm{N}(0,1)$$

and let $X_n = Z_n^2$ then we have that the cumulative distribution function of X_n is

$$\begin{aligned} F_{X_n}(x) &= \mathrm{P}\left(X_n \;\leqslant\; x\right) \\ &= \mathrm{P}\left(Z_n^2 \;\leqslant\; x\right) \\ &= \mathrm{P}\left(-\sqrt{x} \;\leqslant\; Z_n \;\leqslant\; \sqrt{x}\right) \\ &= F_{Z_n}\big(\sqrt{x}\big) - F_{Z_n}\big(-\sqrt{x}\big) \end{aligned}$$

Hence, since $F_{Z_n}(z) \to \Phi(z)$ as $n \to \infty$ for all $z \in \mathbb{R}$ we have

$$\lim_{n\to\infty} F_{X_n}(x) \;=\; \Phi\big(\sqrt{x}\big) - \Phi\big(-\sqrt{x}\big) \;=\; \mathrm{P}\left(Z^2 < x\right)$$

where Z^2 is the square of a standard normal random variable and, from Example 2.1.9 in the text book, $Z^2 \sim \chi^2_1$ so we have that $X_n \xrightarrow{d} X \sim \chi^2_1$. **[5 marks]**

We can now apply Slutsky's Theorem to

$$\frac{n\big(g(Y_n) - g(\theta)\big)}{\sigma^2} \;=\; X_n W_n$$

where

$$\begin{aligned} X_n &= \frac{n(Y_n-\theta)^2}{\sigma^2} &\xrightarrow{d}\; & X \\ W_n &= \frac{g''(\theta)}{2} + \frac{R_2(Y_n)}{(Y_n-\theta)^2} &\xrightarrow{p}\; & \frac{g''(\theta)}{2} \end{aligned}$$

to see that

$$n\big(g(Y_n) - \theta\big) \xrightarrow{d} \frac{\sigma^2 g''(\theta)}{2} X \qquad \text{where } X \sim \chi^2_1$$

[4 marks]

Q. 3 **a)** **(i)** From Theorem 6.15 in my notes the cdf of $X_{(n)}$ is

$$\mathrm{P}(X_{(n)} \leqslant x) = \big(\mathrm{P}(X_1 \leqslant x)\big)^n = \big(1 - \mathrm{e}^{-x/\mu}\big)^n$$

For any $B_\varepsilon > 0$ we have

$$\begin{aligned} \mathrm{P}(X_{(n)} \leqslant B_\varepsilon) &= \big(1 - \mathrm{e}^{-B_\varepsilon/\mu}\big)^n \\ \Rightarrow \quad \lim_{n\to\infty} \mathrm{P}(X_{(n)} \leqslant B_\varepsilon) &= \lim_{n\to\infty} \big(1 - \mathrm{e}^{-B_\varepsilon/\mu}\big)^n \\ &= 0 \end{aligned}$$

Hence we have that for any $B_\varepsilon > 0$ we can find N_ε such that

$$n \geqslant N_\varepsilon \;\Rightarrow\; \mathrm{P}(X_{(n)} \leqslant B_\varepsilon) \leqslant \varepsilon$$

and so we see that the sequence $X_{(n)}$ is not bounded in probability **[6 marks]**

(ii) For the above we know that, for any $z \in \mathbb{R}$

$$\begin{aligned} \mathrm{P}(Z_n \leqslant z) &= \mathrm{P}(X_{(n)} - \mu \log \leqslant z) \\ &= \mathrm{P}(X_{(n)} \leqslant z + \mu \log n) \\ &= \begin{cases} \left(1 - \exp\left\{\dfrac{z}{\mu} - \log(n)\right\}\right)^n & \text{for } z > -mu \log n \\ 0 & \text{for } z < -\mu \log n \end{cases} \\ &= \begin{cases} \left(1 - \dfrac{\mathrm{e}^{-z/\mu}}{n}\right)^n & \text{for } z > -mu \log n \\ 0 & \text{for } z < -\mu \log n \end{cases} \end{aligned}$$

Hence for any $z \in \mathbb{R}$ $\lim\limits_{n\to\infty} \mathrm{P}(Z_n \leqslant z) = \exp\left\{-\mathrm{e}^{-z/\mu}\right\}$. **[4 marks]**

Now we need to show that this is a valid cumulative distribution function. Let $G(z) = \exp\{-\mathrm{e}^{z/\mu}\}$ the we see that

$$\frac{dG(z)}{dz} = \frac{1}{\mu} \exp\left\{-\frac{z}{\mu} - \mathrm{e}^{-z/\mu}\right\} > 0$$

Hence $G(z)$ is monotone increasing.
Furthermore we have

$$\lim_{z\to-\infty} G(z) = \lim_{z\to-\infty} \exp\left\{-\mathrm{e}^{z/\mu}\right\} = \exp\left\{-\lim_{z\to-\infty} \mathrm{e}^{-z/\mu}\right\} = 0$$

$$\lim_{z\to\infty} G(z) = \lim_{z\to\infty} \exp\left\{-\mathrm{e}^{z/\mu}\right\} = \exp\left\{-\lim_{z\to\infty} \mathrm{e}^{-z/\mu}\right\} = \exp\{0\} = 1$$

Therefore G is indeed a valid cumulative distribution function and so there is a random variable Z such that $Z_n \xrightarrow{d} Z$. **[4 marks]**

Finally the probability density function of the limiting random variable Z is

$$f_Z(x) = \frac{dG(z)}{dz} = \frac{1}{\mu}\exp\left\{-\frac{z}{\mu} - \mathrm{e}^{-z/\mu}\right\} \qquad \text{for } z \in \mathbb{R}$$

[2 marks]

b) Casella & Berger 5.44

(i) Since $X_1, \ldots, X_n$ are *iid* Bernoulli(p) random variables we have

$$\mathrm{E}(X_i) = p \qquad \mathrm{Var}(X_i) - p(1-p) < \infty \qquad i = 1, \ldots, n$$

Hence the Central Limit Theorem applies and

$$\frac{\sqrt{n}(Y_n - p)}{\sqrt{p(1-p)}} \xrightarrow{d} \mathrm{N}(0,1) \quad \Rightarrow \quad \sqrt{n}(Y_n - p) \xrightarrow{d} \mathrm{N}\big(0, p(1-p)\big)$$

[3 marks]

(ii) We can define the function

$$g(x) = x(1-x) \quad \Rightarrow \quad g'(x) = 1 - 2x$$

Hence $g'(p) \neq 0$ provided $p \neq 0.5$ and we can apply the first order delta method to get

$$\begin{aligned}\sqrt{n}\big(g(Y_n) - g(p)\big) &= \sqrt{n}\big(Y_n(1-Y_n) - p(1-p)\big)\\ &\xrightarrow{d} \mathrm{N}\big(0, \big(g'(p)\big)^2 p(1-p)\big)\\ &\stackrel{d}{=} \mathrm{N}\big(0, (1-2p)^2 p(1-p)\big)\end{aligned}$$

[3 marks]

(iii) When $p = 0.5$ we have $g'(p) = 0$ and so the first order delta method is not applicable but the second order delta method is applicable as long as $g''(0.5)$ exists and is not 0. In our case we have

$$g''(p) = -2 \qquad \text{for any } p \in (0,1)$$

Hence we have

$$\begin{aligned}n\big(g(Y_n) - g(0.5)\big) &= \sqrt{n}\big(Y_n(1-Y_n) - 0.25\big)\\ &\xrightarrow{d} \frac{g''(0.5) \times 0.5 \times (1-0.5)}{2}\chi_1^2\\ &\stackrel{d}{=} -0.25\chi_1^2\end{aligned}$$

[3 marks]

Q. 4 **a)** The probability mass function of the Binomial($n = 4, p = 1/3$) distribution can be written as

x	0	1	2	3	4
$f(x)$	$\frac{1}{81}$	$\frac{8}{81}$	$\frac{24}{81}$	$\frac{32}{81}$	$\frac{16}{81}$

The cumulative distribution function is then

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{81} & 0 \leqslant x < 1 \\ \frac{9}{81} & 1 \leqslant x < 2 \\ \frac{33}{81} & 2 \leqslant x < 3 \\ \frac{65}{81} & 3 \leqslant x < 4 \\ 1 & x \geqslant 4 \end{cases}$$

Hence an algorithm to generate such random variables is

1. Generate $U \sim \text{Uniform}(0, 1)$.

2. If $U < \frac{1}{81}$ then return $X = 0$

Else if $\frac{1}{81} < U < \frac{9}{81}$ then return $X = 1$

Else if $\frac{9}{81} < U < \frac{33}{81}$ then return $X = 2$

Else if $\frac{33}{81} < U < \frac{65}{81}$ then return $X = 3$

Else return $X = 4$

[6 marks]

it was not required but a more second step in this algorithm is

If $U < \frac{32}{81}$ *then return* $X = 3$

Else if $\frac{32}{81} < U < \frac{56}{81}$ *then return* $X = 2$

Else if $\frac{56}{81} < U < \frac{72}{81}$ *then return* $X = 4$

Else if $\frac{72}{81} < U < \frac{80}{81}$ *then return* $X = 1$

Else return $X = 0$

b) **(i)** Even though the cdf of the logistic is given in the textbook, you were required to derive it as stated in the question.

The cdf is given by

$$F(x) \;=\; \int_{-\infty}^{x} \frac{e^{-y}}{(1+e^{-y})^2}\,dy$$

Now consider the change of variables

$$u \;=\; 1+e^{-y} \quad\Rightarrow\quad du \;=\; -\,e^{-y}dy$$

and the limits transform to

$$y \to -\infty \;\Rightarrow\; u \to \infty \qquad y = x \;\Rightarrow\; u = 1+e^{-x}$$

Hence we get

$$\begin{aligned} F(x) &= \int_{\infty}^{1+e^{-x}} -\frac{1}{u^2}\,du \\ &= \int_{1+e^{-x}}^{\infty} u^{-2}\,du \\ &= -\frac{1}{u}\bigg|_{1+e^{-x}}^{\infty} \\ &= \frac{1}{1+e^{-x}} \end{aligned}$$

[2 marks]

The inverse of this cumulative distribution can be found by setting $u = F(x)$ and solving for x. This gives

$$\begin{aligned} u \;=\; \frac{1}{1+e^{-x}} \;\Rightarrow\; 1+e^{-x} &= \frac{1}{u} \\ \Rightarrow\; e^{-x} &= \frac{1-u}{u} \\ \Rightarrow\; x &= \log(u) - \log(1-u) \end{aligned}$$

[2 marks]

Hence the algorithm to generate standard logistic random variables is

1. Generate $U \sim \text{Uniform}(0,1)$.

2. Return $X = \log(U) - \log(1-U)$. **[2 marks]**

(ii) Suppose that $Z \sim \text{logistic}(0,1)$ then consider the location scale transformation $X = \mu + \beta Z$. The cdf of X is then

$$F_X(x) \;=\; \mathrm{P}(X \leqslant x) \;=\; \mathrm{P}\left(Z \leqslant \frac{x-\mu}{\beta}\right) \;=\; \left(1+\exp\left\{\frac{x-\mu}{\beta}\right\}\right)^{-1}$$

and so the pdf of X is

$$f_X(x) \;=\; \frac{dF(x)}{dx} \;=\; \frac{1}{\beta}\,\frac{\exp\{-(x-\mu)/\beta\}}{\left(1+\exp\{-(x-\mu)/\beta\}\right)^2}$$

which from Page 624 in the textbook is the pdf of the logistic(μ, β) distribution.

[2 marks]

Hence the algorithm to generate general logistic random variables is

1. Generate $U \sim \text{Uniform}(0,1)$.

2. Return $X = \mu + \beta\big(\log(U) - \log(1-U)\big)$. **[2 marks]**

c) Casella and Berger 5.50

Suppose that U_1 and U_2 are independent Unif$[0,1]$ random variables and

$$X_1 \;=\; \cos(2\pi U_1)\sqrt{-2\log U_2} \qquad X_2 \;=\; \sin(2\pi U_1)\sqrt{-2\log U_2}$$

then we have that

$$X_1^2 + X_2^2 = -2\log U_2 \;\;\Rightarrow\;\; U_2 = \exp\left\{-\frac{1}{2}(X_1^2+X_2^2)\right\}$$
$$\frac{X_2}{X_1} = \tan(2\pi U_2) \;\;\Rightarrow\;\; U_2 = \frac{1}{2\pi}\tan^{-1}\left(\frac{X_2}{X_1}\right)$$

[3 marks]

The Jacobian of the transformation $(U_1, U_2) \to (X_1, X_2)$ is

$$|J| \;=\; \left|\begin{matrix} \dfrac{-x_2/x_1^2}{\dfrac{1}{2\pi}\left(1+\dfrac{x_2^2}{x_1^2}\right)} & \dfrac{1/x_1^2}{\dfrac{1}{2\pi}\left(1+\dfrac{x_2^2}{x_1^2}\right)} \\ -x_1\exp\left\{-\dfrac{1}{2}(x_1^2+x_2^2)\right\} & -x_2\exp\left\{-\dfrac{1}{2}(x_1^2+x_2^2)\right\} \end{matrix}\right|$$

$$=\; \frac{1}{2\pi}\exp\left\{-\frac{1}{2}(x_1^2+x_2^2)\right\}$$

[2 marks]

Now note that $\sqrt{-2\log U_2} \in (0, \infty)$ and in each case this is multiplied by something which lies in the interval $[-1, 1]$ so X_1 and X_2 can both take on any real values.

[2 marks]

Hence we have

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) &= \frac{1}{2\pi}\exp\left\{-\frac{1}{2}(x_1^2 + x_2^2)\right\} \qquad (x_1, x_2) \in \mathbb{R}^2 \\ &= \left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x_1^2/2}\right)\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x_2^2/2}\right) \qquad (x_1, x_2) \in \mathbb{R}^2 \end{aligned}$$

So X_1 and X_2 are independent $\mathrm{N}(0, 1)$ random variables. **[2 marks]**