University of Technology Sydney
School of Mathematical and Physical Sciences

Mathematical Statistics (37262) –
Class 4 Preparation Work
SOLUTIONS

1.

Consider the variables defined as $P = X^2$ and $Q = \frac{Y^2}{X^2}$.

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

i) $\boldsymbol{J} = \begin{pmatrix} \frac{\partial X}{\partial P} & \frac{\partial X}{\partial Q} \\ \frac{\partial Y}{\partial P} & \frac{\partial Y}{\partial Q} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\frac{1}{\sqrt{P}} & 0 \\ \frac{1}{2}\sqrt{\frac{Q}{P}} & \frac{1}{2}\sqrt{\frac{P}{Q}} \end{pmatrix}$ hence $|\det \boldsymbol{J}| = \frac{1}{4\sqrt{Q}}$.

ii) X and Y both have range $(-\infty, \infty)$. Considering the substitutions $X = \sqrt{P}$ and $Y = \sqrt{PQ}$ limits X and Y both to only non-negative values. This is only 1/4 of the full 2-dimensional range, since we could have $X \geq 0, Y \geq 0$ or $X \geq 0, Y < 0$ or $X < 0, Y \geq 0$ or $X < 0, Y < 0$, each with equal probability.

iii) $f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ for $(x,y) \in (-\infty,\infty)^2$ hence

$$f_{P,Q}(p,q) = \frac{1}{\sqrt{2\pi}} e^{-\frac{p}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{pq}{2}} 4|\det \boldsymbol{J}|$$

$$f_{P,Q}(p,q) = \begin{cases} \frac{1}{2\pi\sqrt{q}} e^{-\frac{p(1+q)}{2}} & (p,q) \in [0,\infty) \\ 0 & \text{otherwise} \end{cases}.$$

iv) $$f_Q(q) = \int_0^\infty f_{P,Q}(p,q)dp = \frac{1}{2\pi\sqrt{q}} \int_0^\infty e^{-\frac{p(1+q)}{2}} dp = \frac{1}{2\pi\sqrt{q}} \left[\frac{-2e^{-\frac{p(1+q)}{2}}}{1+q} \right]_0^\infty = \frac{1}{\pi(1+q)\sqrt{q}}$$

Hence $f_Q(q) = \begin{cases} \frac{1}{\pi(1+q)\sqrt{q}} & q \in [0,\infty) \\ 0 & \text{otherwise} \end{cases}$

We can note that this is $Q \sim F(1,1)$.

v) $$f_P(p) = \int_0^\infty f_{P,Q}(p,q)dq = \frac{1}{2\pi} e^{-\frac{p}{2}} \int_0^\infty \frac{1}{\sqrt{q}} e^{-\frac{pq}{2}} dq .$$

We can do this via integration by substitution with $r = \frac{pq}{2}$ hence. $\frac{dr}{dq} = \frac{p}{2}$.

$$\int_0^\infty \frac{1}{\sqrt{q}} e^{-\frac{pq}{2}} dq = \int_0^\infty \frac{1}{\sqrt{q}} \frac{2}{p} e^{-r} dr = \int_0^\infty \frac{\sqrt{p}}{\sqrt{2r}} \frac{2}{p} e^{-r} dr = \frac{\sqrt{2}}{\sqrt{p}} \int_0^\infty \frac{1}{\sqrt{r}} e^{-r} dr .$$

$\int_0^\infty \frac{1}{\sqrt{r}} e^{-\frac{r}{2}} dr = \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ since this is a gamma function hence

$$f_P(p) = \int_0^\infty f_{P,Q}(p,q)dq = \frac{1}{2\pi} e^{-\frac{p}{2}} \int_0^\infty \frac{1}{\sqrt{q}} e^{-\frac{pq}{2}} dq = \frac{1}{2\pi} e^{-\frac{p}{2}} \left(\frac{\sqrt{2}}{\sqrt{p}}\right)\left(\sqrt{\pi}\right)$$

$$f_P(p) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{p}} e^{-\frac{p}{2}} & p \in [0,\infty) \\ 0 & \text{otherwise} \end{cases}$$

We can note that this is $P \sim \chi^2(1)$.

vi) The conditional distribution of $Q \mid P = 1$ is simply the distribution of $Q = \frac{Y^2}{1^2} = Y^2$. Since Y is a standard normal variable, we know that $(Q \mid P = 1) \sim \chi^2(1)$ hence $f_{Q|P=1}(q) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{q}} e^{-\frac{q}{2}} & q \in [0,\infty) \\ 0 & \text{otherwise} \end{cases}$.

2.

Let $E(S) = \mu = 50$ and $Var(S) = \sigma^2 = 64$.

i) The central limit theorem tells us that $Z = \left(\frac{\bar{S} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \right) \sim N(0,1)$.

Here, we want $P(Z > z)$ where $z = \left(\frac{53 - 50}{\sqrt{\frac{64}{100}}} \right) = 3.75.$

This gives $P(\bar{S} > 53) \approx 0.00009$.

ii) Let the sample size be *n*.

We then want $P(Z > z)$ where $z = \left(\frac{51 - 50}{\sqrt{\frac{64}{n}}} \right) > 1.645.$

This gives $\frac{\sqrt{n}}{8} > 1.645$ so $n > (8 \times 1.645)^2$, giving a smallest (integer) sample size of $n = 174$.

3.

Under the null hypothesis, the expected frequencies are:

Outcome	1	2	3	4	5	4
Frequency	10	20	30	40	50	60

The Pearson statistic is therefore

$$\frac{(11-10)^2}{10} + \frac{(20-20)^2}{20} + \frac{(37-30)^2}{30} + \frac{(44-40)^2}{40} + \frac{(42-50)^2}{50} + \frac{(56-60)^2}{60} = 3.68$$

For $Y \sim \chi^2(5)$, $P(Y > 3.68) \approx 0.404$ hence with 0.05 significance, we do not reject the hypothesis that outcomes of the die rolls are draws from the stated distribution.