University of Technology Sydney
School of Mathematical and Physical Sciences

Mathematical Statistics (37262) –
Quiz 1
SOLUTIONS

Question 1.

i) One possible rule would be:

If $u_i < 0.1$ then $x_i = 0.1$
If $0.1 \le u_i < 0.3$ then $x_i = 0.2$
If $0.3 \le u_i < 0.6$ then $x_i = 0.3$
If $0.6 \le u_i$ then $x_i = 0.4$

Here, this gives $\{x_1, x_2, x_3\} = \{0.4,\ 0.2,\ 0.3\}$.

ii) $$F(y) = P(Y \le y) = \int_0^y 10e^{-10w}dw = \left[-e^{-10w}\right]_0^{10} = 1 - e^{-10y}.$$

iii) Inverting $F(y) = 1 - e^{-10y}$ gives $F^{-1}(y) = -\frac{\ln(1-y)}{10}$ hence we generate $\{y_1, y_2, y_3\}$ by setting $y_i = -\frac{\ln(1-u_i)}{10}$.

Here, $\{y_1, y_2, y_3\} \approx \{0.136,\ 0.012,\ 0.073\}$.

(6 Marks)

Question 2.

There are various different answers to these, but the basic form (e.g. summing squared normal variables etc.) must be the same. The index of which realisation is used is not important, provided the same is not used multiple times in the same solution e.g. in ii), if x_1 is in the numerator, it cannot be in the denominator as well.

i) $Q \sim \chi^2(3)$; ii) $T \sim t_2$; iii) $R \sim N(4,1)$;

i) $x_1^2 + x_2^2 + x_3^2$; ii) $\dfrac{x_1}{\sqrt{\dfrac{x_2^2 + x_3^2}{2}}}$; iii) $x_1 + 4$

(6 Marks)

Question 3.

i) $P(X=0)=P((X,Y)=(0,0))+P((X,Y)=(0,1))=0.3+0.4=0.7$.
$P(X=1)=P((X,Y)=(1,0))+P((X,Y)=(1,1))=0.2+0.1=0.3$.
Hence $X\sim Bern(0.3)$.

ii) $$P(X=0\,|\,Y=1)=\frac{P((X,Y)=(0,1))}{P((X,Y)=(0,1))+P((X,Y)=(1,1))}=\frac{0.4}{0.4+0.1}=0.8$$
$$P(X=1\,|\,Y=1)=\frac{P((X,Y)=(1,1))}{P((X,Y)=(0,1))+P((X,Y)=(1,1))}=\frac{0.1}{0.4+0.1}=0.2.$$
Hence $(X\,|\,Y=1)\sim Bern(0.2)$.

iii)
$$P((X,Z)=(a,c))=\begin{cases}0.21 & (a,c)=(0,-1)\\ 0.28 & (a,c)=(0,0)\\ 0.21 & (a,c)=(0,1)\\ 0.09 & (a,c)=(1,-1)\\ 0.12 & (a,c)=(1,0)\\ 0.09 & (a,c)=(1,1)\\ 0 & \text{otherwise}\end{cases}$$

(8 Marks)

Question 4.

i) $f_{X,Y}(x,y) = f_X(x)f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ for $-\infty < x,y < \infty$.

ii) $X = \frac{S+D}{2}$ and $Y = \frac{S-D}{2}$.

iii) As both X and Y can take any real number (positive or negative) both their difference and their sum can, so both S and D can take any real number, positive or negative.

iv) $$|\det \boldsymbol{J}| = \left|\det\begin{pmatrix} \frac{\partial X}{\partial S} & \frac{\partial X}{\partial D} \\ \frac{\partial Y}{\partial S} & \frac{\partial X}{\partial D} \end{pmatrix}\right| = \left|\det\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} \end{pmatrix}\right| = \frac{1}{2}.$$

v) $f_{S,D}(s,d) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{1}{2}$ for $-\infty < s,d < \infty$.

vi)

$$\begin{aligned} f_{S,D}(s,d) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\frac{s+d}{2}\right)^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\frac{s-d}{2}\right)^2}{2}} \frac{1}{2} \\ &= \left(\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2}} e^{-\frac{\left(s^2+d^2+2ds\right)}{8}}\right)\left(\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2}} e^{-\frac{\left(s^2+d^2-2ds\right)}{8}}\right) \\ &= \left(\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2}} e^{-\frac{s^2}{2(\sqrt{2})^2}}\right)\left(\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2}} e^{-\frac{d^2}{2(\sqrt{2})^2}}\right) \end{aligned}$$

The first term of this shows that is $S \sim N(0,2)$.

vii) As $f_{S,D}(s,d) = f_S(s)f_D(d)$ (where each marginal density is $N(0,2)$, we see that S and D are independent.

(10 Marks)

Question 5.

i) For this hypothesis, the Pearson statistic is

$$\frac{(10-20)^2}{20}+\frac{(10-20)^2}{20}+\frac{(20-20)^2}{20}+$$
$$\frac{(10-20)^2}{20}+\frac{(10-20)^2}{20}+\frac{(40-20)^2}{20}+$$
$$\frac{(25-20)^2}{20}+\frac{(25-20)^2}{20}+\frac{(30-20)^2}{20}$$
$$=5+5+0+5+5+20+\frac{5}{4}+\frac{5}{4}+5=47.5$$

ii) We have hypothesised 9 proportions, hence have 8 degrees of freedom. The 5% critical value is approximately 15.507, which our test statistic is larger than, so we reject the null hypotheses of these proportions.

iii) Under this new hypothesis, the expected counts would be approximately:

	Activity A	**Activity B**	**Activity C**
Adult	10	10	20
Senior	15	15	30
Child	20	20	40

For this hypothesis, the Pearson statistic is

$$\frac{(10-10)^2}{10}+\frac{(10-10)^2}{10}+\frac{(20-20)^2}{20}+$$
$$\frac{(10-15)^2}{15}+\frac{(10-15)^2}{15}+\frac{(40-30)^2}{30}+$$
$$\frac{(25-20)^2}{20}+\frac{(25-20)^2}{20}+\frac{(30-40)^2}{40}$$
$$=0+0+0+\frac{5}{3}+\frac{5}{3}+\frac{10}{3}+\frac{5}{4}+\frac{5}{4}+\frac{5}{2}\approx 11.67$$

iv) For a contingency table with 3 rows and 3 columns, we have $(3-1)(3-1)=4$ degrees of freedom. The 5% critical value is approximately 9.488, which our test statistic is larger than, so we reject the null hypotheses of independence between the column and the row factors.

(10 Marks)