NATIONAL UNIVERSITY OF SINGAPORE
MATHEMATICS SOCIETY

PAST YEAR PAPER SOLUTIONS
with credits to Terry Lau Shue Chien, Zhuang Linjie

MA2216/ST2131 Probability
AY 2008/2009 Sem 1

Question 1

(a) $X \sim NB(r, p)$ with $p = \frac{\sqrt{5}-1}{2}$. Recall that X can be regarded as a sum of r identical independent geometric random variables X_i with parameter p.

$$X_i \sim Geom(p)$$

$$\mathbb{E}[X_i] = \frac{1}{p}$$

We need to find how large r needs to be so that

$$\mathbb{P}\left\{\left|\frac{X}{r} - \frac{1}{p} > 0.01\right|\right\} < 0.01$$

This is a "central limit theorem" type of problem. Define $\overline{X}_r = \frac{X}{r}$. Then $\overline{X}_r \sim N(\frac{1}{p}, \frac{\text{Var}(X_i)}{r})$ with $\text{Var}(X_i) = \frac{1-p}{p^2}$. We thus have

$$\begin{aligned}
\mathbb{P}\left\{\left|\frac{X}{r} - \frac{1}{p}\right| > 0.01\right\} &= \mathbb{P}\left\{\left|\overline{X}_i - \frac{1}{p}\right| > 0.01\right\} \\
&= \mathbb{P}\left\{|Z| > 0.01\sqrt{\frac{rp^2}{1-p}}\right\} \\
&= 2\mathbb{P}\left\{Z > 0.01\sqrt{\frac{rp^2}{1-p}}\right\}
\end{aligned}$$

by symmetry of $N(0, 1)$. Referring to the normal distribution table, we have

$$\mathbb{P}(Z > 2.5758) = 0.005 \Rightarrow 0.01\sqrt{\frac{rp^2}{1-p}} \geq 2.5758$$

Solving, we get $r \geq 66347.45 \Rightarrow r = 66348$.

(b) Let X_1, X_2 be the sales for the next two weeks. Then

$$f_{(X_1,X_2)}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{Q(x_1,x_2;\mu_1,\mu_2;\sigma_1,\sigma_2;\rho)}{2}}$$

where

$$Q(x_1, x_2; \mu_1, \mu_2; \sigma_1, \sigma_2; \rho) = \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$$

Let

$$U = \frac{X_1 - \mu_1}{\sigma_1} \sim N(0, 1)$$

$$V = \frac{-\rho}{\sqrt{1-\rho^2}}\frac{X_1-\mu_1}{\sigma_1} + \frac{1}{\sqrt{1-\rho^2}}\frac{X_2-\mu_2}{\sigma_2} \sim N(0,1)$$

$$X_1 + X_2 \geq 90 \Rightarrow (\sigma_1 u + \mu_1) + (\mu_2 + \rho\sigma_2 u + \sqrt{1-\rho^2}\sigma_2 v) \geq 90$$

Note that $\mu_1 = \mu_2 = 40$, $\sigma_1 = \sigma_2 = 6$ and $\rho = 0.6$, hence we have

$$(\sigma_1 u + \mu_1) + (\mu_2 + \rho\sigma_2 u + \sqrt{1-\rho^2}\sigma_2 v) \geq 90 \Rightarrow 9.6u + 4.8v + 80 \geq 90$$

Note that U, V are independent random variables, hence

$$X_1 + X_2 \equiv 9.6U + 4.8V + 80 \sim N(80, 9.6^2 + 4.8^2)$$

Thus $\mathbb{P}(X_1 + X_2 \geq 90) = \mathbb{P}(Z \geq \frac{90-80}{\sqrt{115.2}}) = \mathbb{P}(Z \geq 0.9317) = 1 - 0.3238 = 0.6762.$

Question 2

(i) For $0 < y < \infty$,

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)\,dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y} e^{-\frac{(x-y)^2}{2}}\,dx \\ &= e^{-y}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2}}\,dx \\ &= e^{-y}. \end{aligned}$$

(ii)

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f_{(X,Y)}(x,y)}{f_Y(y)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-y)^2}{2}}\,dx \end{aligned}$$

for $0 < y < \infty$, $-\infty < x < \infty$.

(iii)

$$\begin{aligned} \mathbf{E}[X|Y=y] &= \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx \\ &= \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}} e^{-\frac{(x-y)^2}{2}}\,dx \\ &= y. \end{aligned}$$

(iv) Since $Y \sim exp(1)$, $\mathbf{E}[X] = \mathbf{E}[\mathbf{E}[X|Y]] = \mathbf{E}[Y] = 1$.

(v)

$$\begin{aligned}
\mathbf{E}[XY] &= \int_0^\infty \int_{-\infty}^\infty xy \frac{1}{\sqrt{2\pi}} e^{-y} e^{-\frac{(x-y)^2}{2}}\, dx\, dy \\
&= \int_0^\infty ye^{-y} \int_{-\infty}^\infty x \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-y)^2}{2}}\, dx\, dy \\
&= \int_0^\infty ye^{-y}(y)\, dy \\
&= \int_0^\infty y^2 e^{-y}\, dy \\
&= \left[-y^2e^{-y}\right]_0^\infty + 2\int_0^\infty ye^{-y}\, dy \\
&= 0 + 2\left[\left[-ye^{-y}\right]_0^\infty + \int_0^\infty e^{-y}\, dy\right] \\
&= 2(1-0) \\
&= 2.
\end{aligned}$$

(vi) $\mathrm{Cov}(X,Y) = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = 2 - 1 = 1.$

(vii) $(t+y)^2 = t^2 + 2ty + y^2$, so

$$\begin{aligned}
-(x-(t+y))^2 &= -(x^2 - 2x(t+y) + t^2 + 2ty + y^2) \\
&= -x^2 + 2x(t+y) - t^2 - 2ty - y^2
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}[e^{tX}|Y=y] &= \int_{-\infty}^\infty e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-y)^2}{2}}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{\frac{-x^2+2xy+2tx-y^2}{2}}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{\frac{-x^2+2x(t+y)-y^2-2ty-t^2+2ty+t^2}{2}}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{(x-(t+y))^2}{2}} e^{\frac{2ty+t^2}{2}}\, dx \\
&= e^{\frac{2ty+t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{(x-(t+y))^2}{2}}\, dx \\
&= e^{\frac{t(2y+t)}{2}}.
\end{aligned}$$

(viii) $t < 1$,

$$\begin{aligned}
\mathbf{E}[e^{tX}] &= \mathbf{E}[\mathbf{E}[e^{tX}|Y]] \\
&= \mathbf{E}[e^{\frac{t(2y+t)}{2}}] \\
&= e^{\frac{t^2}{2}} \mathbf{E}[e^{ty}]
\end{aligned}$$

Note that $Y \sim exp(1)$, hence $\mathbf{E}[e^{ty}] = M_Y(t) = \frac{1}{1-t}$.

$$\begin{aligned}
\mathbf{E}[e^{tX}] &= e^{\frac{t^2}{2}} \mathbf{E}[e^{tY}] \\
&= \frac{e^{\frac{t^2}{2}}}{1-t}.
\end{aligned}$$

Question 3

(a) (i) $X = g_1(U,V) = \frac{1}{2}(U+V)$, $Y = g_2(U,V) = \frac{1}{2}(U-V)$. We thus have

$$\begin{aligned} J(U,V) &= \begin{vmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{vmatrix} \\ &= \frac{\partial g_1}{\partial u}\frac{\partial g_2}{\partial v} - \frac{\partial g_2}{\partial u}\frac{\partial g_1}{\partial v} \\ &= \left(\frac{1}{2}\right)\left(-\frac{1}{2}\right) - \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) \\ &= -\frac{1}{2} \end{aligned}$$

Then $f_{(X,Y)}(x,y) = \frac{1}{|J(U,V)|} f_{(U,V)}(u,v) = 2f_{(U,V)}(u,v)$.
Since $U \sim N(0,1)$ and $V \sim N(0,1)$, and U, V are independent,

$$\begin{aligned} f_{(U,V)}(u,v) &= f_U(u) \cdot f_V(v) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \\ &= \frac{1}{2\pi} e^{-\frac{u^2+v^2}{2}} \end{aligned}$$

Therefore, for $-\infty < x < \infty$ and $-\infty < y < \infty$,

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

(ii)

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} \frac{1}{\pi} e^{-(x^2+y^2)}\, dy \\ &= \frac{e^{-x^2}}{\pi} \int_{-\infty}^{\infty} e^{-y^2}\, dy \end{aligned}$$

Now use the substitution $y = \frac{z}{\sqrt{2}} \Rightarrow \frac{dy}{dz} = \frac{1}{\sqrt{2}}$. Then we have

$$\begin{aligned} f_X(x) &= \frac{e^{-x^2}}{\pi} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} \frac{1}{\sqrt{2}}\, dz \\ &= \frac{e^{-x^2}}{\pi\sqrt{2}} \sqrt{2\pi} \\ &= \frac{e^{-x^2}}{\sqrt{\pi}} \end{aligned}$$

Thus $X \sim n(x; 0; \frac{1}{\sqrt{2}})$.
Similarly, $f_Y(y) = \frac{e^{-y^2}}{\sqrt{\pi}}$ and $Y \sim n(y; 0; \frac{1}{\sqrt{2}})$.

(iii) X, Y are independent. By definition, for all $x, y \in \mathbb{R}$,

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

(b) Let I_i be the indicator random variables with

$$I_i = \begin{cases} 1 & \text{if } i\text{th guesses are correct} \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, 2, \ldots, n$.

$$N = I_1 + I_2 + \ldots + I_n$$

Then

$$\begin{aligned} \mathbf{E}[N] &= E[I_1 + I_2 + \ldots + I_n] \\ &= \mathbb{P}[I_1] + \mathbb{P}[I_2] + \ldots + \mathbb{P}[I_n] \\ &= \frac{1}{n} + \frac{1}{n} + \ldots + \frac{1}{n} \\ &= 1. \end{aligned}$$

Question 4

(a) Let X denote the number of A_i that occur. For the left hand side of the equation,

$$\begin{aligned} \mathbb{E}[X] &= \sum_{i=0}^{n} x\mathbb{P}(X = x) \\ &= \mathbb{P}(X = 1) + 2\mathbb{P}(X = 2) + \ldots + n\mathbb{P}(X = n) \\ &= \sum_{i=1}^{n} \mathbb{P}(X = i) + \sum_{i=2}^{n} \mathbb{P}(X = i) + \ldots + \sum_{i=n}^{n} \mathbb{P}(X = i) \\ &= \mathbb{P}(C_1) + \mathbb{P}(C_2) + \ldots + \mathbb{P}(C_k) \\ &= \sum_{k=0}^{n} \mathbb{P}(C_k) \end{aligned}$$

On the other hand, define

$$I_j = \begin{cases} 1 & \text{if } A_j \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

Note that $X = I_1 + I_2 + \ldots + I_n$.

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[I_1 + I_2 + \ldots + I_n] \\ &= \mathbb{E}[I_1] + \mathbb{E}[I_2] + \ldots + \mathbb{E}[I_n] \\ &= \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots + \mathbb{P}(A_n) \\ &= \sum_{i=1}^{n} \mathbb{P}(A_i) \end{aligned}$$

Therefore, $\sum_{k=0}^{n} \mathbb{P}(C_k) = \sum_{i=1}^{n} \mathbb{P}(A_i)$.

(b) $X \sim Po(10)$. Define

$$I_j = \begin{cases} 1 & \text{if elevator stops at floor } j \\ 0 & \text{otherwise} \end{cases}$$

for $j = 1, 2, \ldots, 10$.

(i) Note that $\mathbb{P}(I_j = 0|X = k)$ is the probability that no person gets off at floor j given that k people enter the elevator on the ground floor. We have

$$\mathbb{P}(I_j = 0|X = k) = \left(1 - \frac{1}{10}\right)^k = \left(\frac{9}{10}\right)^k$$

(ii)

$$\begin{aligned}\mathbb{E}[I_j|X = k] &= 1 \cdot \mathbb{P}(I_j = 1|X = k) + 0 \cdot \mathbb{P}(I_j = 0|X = k) \\ &= 1 - \left(\frac{9}{10}\right)^k\end{aligned}$$

$$\begin{aligned}\mathbb{E}[\text{number of stops}] &= \mathbb{E}[I_1 + I_2 + \ldots + I_{10}] \\ &= \mathbb{E}[I_1] + \mathbb{E}[I_2] + \ldots + \mathbb{E}[I_{10}] \\ &= 10\mathbb{E}[I_j] \\ &= 10\mathbb{E}[\mathbb{E}[I_j|X]] \\ &= 10\mathbb{E}[1 - (\frac{9}{10})^X] \\ &= 10 - 10\mathbb{E}[(\frac{9}{10})^X]\end{aligned}$$

Note that

$$\begin{aligned}\mathbb{E}[(\frac{9}{10})^X] &= \sum_{i=0}^{\infty} (\frac{9}{10})^i e^{-10} \frac{10^i}{i!} \\ &= \sum_{i=0}^{\infty} \frac{e^{-10} 9^i}{i!} \\ &= e^{-1} \sum_{i=0}^{\infty} \frac{e^{-9} 9^i}{i!} \\ &= e^{-1}\end{aligned}$$

Since $\mathbb{E}[(\frac{9}{10})^X] = e^{-1}$, we have $\mathbb{E}[\text{number of stops}] = 10 - 10e^{-1}$.