DS 203 - Assignment 2

Name - Rohan Rajesh Kalbag
Roll - 20D170033

Exercise 1: Let F denote CDF of random variables
Let X, Y are independent exponential random variables with λ_1, λ_2 are parameters

$$F_X(x : \lambda_1) = \begin{cases} 1-e^{-\lambda_1 x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$F_Y(x : \lambda_2) = \begin{cases} 1-e^{-\lambda_2 x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

(i) Define $Z = \min(X, Y)$, $a \in \mathbb{R}$

$$P(z > t) = P(x > t, y > t)$$
$$= P(x > t) \cdot P(y > t) \quad \{\text{Independence}\}$$
$$1 - P(z \leq t) = (1 - P(x \leq t)) \cdot (1 - P(y \leq t))$$
$$1 - F_Z(x) = (e^{-\lambda_1 x}) \cdot e^{-\lambda_2 x}$$
$$\boxed{F_Z(t) = 1 - e^{-(\lambda_1 + \lambda_2) t}}$$

Therefore Min(X, Y) also has an exponential distribution with parameter $\lambda' = \lambda_1 + \lambda_2$.

(ii) Define $H = \max(X, Y)$

$$P(H \le t) = P(X \le t \cap Y \le t)$$
$$= P(X \le t) \cdot P(Y \le t) \quad \{\text{Independence}\}$$
$$F_H(t) = (1-e^{-\lambda_1 t})(1-e^{-\lambda_2 t})$$
$$F_H(t) = 1 - e^{-\lambda_1 t} - e^{-\lambda_2 t} - e^{-(\lambda_1+\lambda_2)t}$$

$\therefore$ The CDF of the random variable $H = \max(X, Y)$ is given by

$$F_H(t) = 1 - e^{-\lambda_1 t} - e^{-\lambda_2 t} - e^{-(\lambda_1+\lambda_2)t}.$$

Exercise 2 : 3 white, 6 red, 5 blue

X = the no of white balls selected
Y = the no of blue balls selected

given $Y=3 \Rightarrow$ 3 blue balls present

$$P(Y=3) = {}^6C_3 \left(\frac{5}{14}\right)^3 \left(\frac{9}{14}\right)^3$$

$$P(X=0 \mid Y=3) = \frac{P(X=0 \cap Y=3)}{P(Y=3)} = \frac{{}^6C_3 \left(\frac{5}{14}\right)^3 \left(\frac{6}{14}\right)^3}{{}^6C_3 \left(\frac{5}{14}\right)^3 \left(\frac{9}{14}\right)^3}$$

$$P(X=1 \mid Y=3) = \frac{P(X=1 \cap Y=3)}{P(Y=3)} = \frac{{}^6C_3 \left(\frac{5}{14}\right)^3 \, {}^3C_1 \left(\frac{3}{14}\right)\left(\frac{6}{14}\right)^2}{{}^6C_3 \left(\frac{5}{14}\right)^3 \left(\frac{9}{14}\right)^3}$$

Similarly, $P(X=2 \mid Y=3) = \dfrac{{}^6C_3\left(\frac{5}{14}\right)^3 \, {}^3C_2\left(\frac{3}{14}\right)^2\left(\frac{6}{14}\right)}{{}^6C_3\left(\frac{5}{14}\right)^3\left(\frac{9}{14}\right)^3}$

, $P(X=3 \mid Y=3) = \dfrac{{}^6C_3\left(\frac{5}{14}\right)^3\left(\frac{3}{14}\right)^3}{{}^6C_3\left(\frac{5}{14}\right)^3\left(\frac{9}{14}\right)^3}$.

$$\therefore P(X=t \mid Y=3) = \begin{cases} {}^3C_0 \cdot \dfrac{6^3}{9^3}, & t=0 \\ {}^3C_1 \dfrac{\cdot 3 \cdot 6^2}{9^3}, & t=1 \\ {}^3C_2 \dfrac{\cdot 3^2 \cdot 6}{9^3}, & t=1 \\ {}^3C_3 \cdot \dfrac{3^3}{9^3}, & t=2 \end{cases}$$

$$\therefore E[X \mid Y=3] = 0\, P(X=0 \mid Y=3) + 1\, P(X=1 \mid Y=3) + 2\, P(X=2 \mid Y=3) + 3\, P(X=3 \mid Y=3)$$

$$= \frac{{}^3C_1(3\cdot 6^2)}{9^3} + \frac{2\left({}^3C_2 \cdot 3^2 \cdot 6\right)}{9^3} + \frac{3\left({}^3C_3 3^3\right)}{9^3}$$

$$= \frac{3^4 \cdot 2^2}{3^6} + \frac{2^2 \cdot 3^4}{3^6} + \frac{3^4}{3^6}$$

$$= \frac{4}{9} + \frac{4}{9} + \frac{1}{9}$$

$$= \frac{9}{9} = 1$$

$$\boxed{E[X \mid Y=3] = 1}$$

Exercise 3 :

x_1 and x_2 are independent binomial random with (n_1, p) and (n_2, p)

$$P(x_1 = i) = {}^{n_1}C_i \; p^i (1-p)^{n_1 - i}$$

$$P(x_2 = i) = {}^{n_2}C_i \; p^i (1-p)^{n_2 - i}$$

Conditional PMF of x_1 given $x_1 + x_2 = m$

$$P(x_1 = t \mid x_1 + x_2 = m) = \frac{P(x_1 = t \cap x_2 = m-t)}{P(x_1 + x_2 = m)}$$

$$= \frac{P(x_1 = t) \cdot P(x_2 = m-t)}{P(x_1 + x_2 = m)}$$

$$= \frac{{}^{n_1}C_t \; p^t (1-p)^{n_1 - t} \cdot {}^{n_2}C_{m-t} \; p^{m-t} (1-p)^{n_2 - m + t}}{P(x_1 + x_2 = m)}$$

we can treat $x_1 + x_2 = X$ as another binomial random variable as $P_1 = P_2 = p$

$$P(X = m) = {}^{n_1 + n_2}C_m \; p^m (1-p)^{n_1 + n_2 - m}$$

$$P(x_1 = t \mid x_1 + x_2 = m) = \frac{{}^{n_1}C_t \; {}^{n_2}C_{m-t} \; \cancel{p^m (1-p)^{n_1 + n_2 - m}}}{{}^{n_1 + n_2}C_m \; \cancel{p^m (1-p)^{n_1 + n_2 - m}}}$$

$\therefore$ The conditional probability mass function of X_1

$$P(X_1 = t \mid X_1 + X_2 = m) = \frac{{}^{n_1}C_t \; {}^{n_2}C_{m-t}}{{}^{n_1+n_2}C_m}$$

Exercise 4:

Example is $X = \{-1, 0, 1\}$ with probability $\frac{1}{3}$ each. Let $Y = X^4$.

X	P
-1	$\frac{1}{3}$
0	$\frac{1}{3}$
1	$\frac{1}{3}$

Y	P
1	$2/3$
0	$\frac{1}{3}$

$E(X) = E(X^5)$ as $X \in \{-1, 0, 1\}$

$$\begin{aligned}
Cov(X,Y) &= E(XY) - E(X)E(Y) \\
&= E(X \cdot X^4) - E(X)E(X^4) \\
&= E(X^5) - E(X)E(X^4) \\
&= E(X)\left[1 - E(X^4)\right] \\
&= 0
\end{aligned}$$

$\therefore Cov(X,Y) = 0$ {Hence they are uncorelated}

$P(X=0, Y=0) = P(X=0) \cdot P(Y=0 \mid X=0) = \frac{1}{3} \times 1 = \frac{1}{3}$

$P(X=0) = \frac{1}{3}$ $\quad P(Y=0) = \frac{1}{3}$ $\quad P(X=0) \cdot P(Y=0) = \frac{1}{9}$

$\therefore P(X=0, Y=0) \neq P(X=0) \cdot P(Y=0)$ {Hence they are not independent}

Exercise 5:

$$P(X=n|\lambda) = e^{-\lambda} \lambda^n / n! \qquad \left\{ \begin{array}{l} \text{Poisson's Distribution} \\ \text{with mean } \lambda \end{array} \right\}$$

$$f_\lambda(\lambda=t) = \begin{cases} \mu e^{-\mu t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Since mean is given as 1 $\Rightarrow$

$$\frac{1}{\mu} = 1 \Rightarrow \mu = 1.$$

using the theorem of total probability

$$P(X=n) = \int_{-\infty}^{\infty} P(X=n|\lambda) f_\lambda(\lambda=t)\,dt$$

$$P(X=n) = \int_{-\infty}^{\infty} \frac{e^{-t} t^n}{n!} f_\lambda(\lambda=t)\,dt$$

$$= \int_0^{\infty} \frac{e^{-2t} t^n\,dt}{n!} \qquad \text{put } 2t = v$$

$$= \int_0^{\infty} \frac{e^{-v} \left(\frac{v}{2}\right)^n}{n!} \frac{dv}{2}$$

$$= \frac{1}{2^{n+1}} \frac{\Gamma(n+1)}{n!} = \frac{1}{2^{n+1}}$$

as $\Gamma(n+1) = n!$ as $n \in \mathbb{N}$

$$\therefore \boxed{P(X=n) = \left(\frac{1}{2}\right)^{n+1}}$$

Exercise 6:

$$f_{X,Y}(x,y) = c(1+xy) \qquad \begin{array}{l} 2 \le x \le 3 \\ 1 \le y \le 2 \end{array}$$

(i) To find c we use

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$$

$$= c\int_{1}^{2}\int_{2}^{3} (1+xy)\,dx\,dy = 1$$

$$= c\int_{1}^{2}\left(1 + y\frac{(9-4)}{2}\right)dy$$

$$= c\left[1 + \left(\frac{4-1}{2}\right)\left(\frac{5}{2}\right)\right] = 1$$

$$= c\left[1 + \frac{15}{4}\right] = 1 \implies \boxed{c = \frac{4}{19}}$$

(ii)

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_{1}^{2} f_{X,Y}(x,y)\,dy$$

$$f_X(x) = \frac{4}{19}\int_{1}^{2}(1+xy)\,dy$$

$$= \frac{4}{19}\left[1 + \frac{3}{2}x\right]$$

$$\boxed{f_X(x) = \frac{2}{19}(2+3x)}$$

$$f_Y(y) = \frac{4}{19}\int_{2}^{3}(1+xy)\,dx$$

$$\boxed{f_Y(y) = \frac{4}{19}\left[1 + \frac{5}{2}y\right]}$$

Exercise 7:

X be the random variable, the no of accidents a random policyholder has in a year.

$$P(X=n|\lambda) = \frac{e^{-\lambda}\lambda^{n}}{n!}, \quad \text{mean} = \lambda$$

The distribution of λ, $g(\lambda) = \lambda e^{-\lambda}$

$$P(X=n) = \int_{-\infty}^{\infty} P(X=n|\lambda)\, g(\lambda)\, d\lambda$$

$$\therefore P(X=n) = \int_{-\infty}^{\infty} \frac{e^{-\lambda}\lambda^{n}}{n!}\, g(\lambda)\, d\lambda$$

$$= \int_{-\infty}^{\infty} \frac{e^{-\lambda}\lambda^{n}}{n!} \cdot \lambda e^{-\lambda}\, d\lambda$$

$$= \int_{-\infty}^{\infty} \frac{e^{-2\lambda}\lambda^{n+1}}{n!}\, d\lambda$$

$2\lambda = u$

$$P(X=n) = \frac{1}{n!}\int_{0}^{\infty} e^{-2\lambda}\lambda^{n+1}\, d\lambda$$

$$= \frac{1}{n!}\int_{0}^{\infty} e^{-u}\left(\frac{u}{2}\right)^{n+1} \cdot \frac{du}{2}$$

$$= \frac{1}{n!}\frac{1}{2^{n+2}}\left(\Gamma(n+2)\right) \qquad \Gamma n = (n-1)!$$

$$\boxed{P(X=n) = \frac{1}{n!}\frac{1}{2^{n+2}}(n+1)! = \frac{n+1}{2^{n+2}}}$$

Exercise 8:

$X \rightarrow$ no of people who visit yoga studio

$$P(X=n) = \frac{e^{-\lambda}\lambda^{n}}{n!}$$

$A \rightarrow$ That m men and n women visit the yoga studio today

$$P(A) = P(X=m+n) \cdot P(A/X=m+n)$$

$$= \frac{e^{-\lambda}\lambda^{m+n}}{(m+n)!} \cdot \left({}^{m+n}C_n \, p^{n}(1-p)^{m} \right)$$

$$P(A) = \frac{e^{-\lambda}\lambda^{m+n}}{(m+n)!} \left({}^{m+n}C_n \cdot p^{n}(1-p)^{m} \right)$$

Exercise 9:

X_1, X_2 and X_3 are random variables

(i) $Cov(aX_1+b, cX_2+d) = ac\, Cov(X_1, X_2)$

Result: $E(aX_1+b) = aE(X_1)+b$

Proof: $$\frac{\sum_{i=1}^{n}(ax_i+b)}{n} = a\left(\frac{\sum x_i}{n}\right)+b$$

$$= a\,E(x_i)+b$$

$$\mathrm{Cov}(aX_1+b,\ cX_2+d) = E\left[(aX_1+b-E(aX_1+b))(cX_2+d-E(cX_2+d))\right]$$

$$= E\left[(aX_1+b-aE(X_1)-b)(cX_2+d-cE(X_2)-d)\right]$$

$$= E\left[ac(X_1-E(X_1))(X_2-E(X_2))\right]$$

$$= ac\,E\left[(X_1-E(X_1))(X_2-E(X_2))\right]$$

$$= ac\,\mathrm{Cov}(X_1, X_2)$$

Hence we are done.

(ii) $\mathrm{Cov}(X_1+X_2,\ X_3) = \mathrm{Cov}(X_1, X_3) + \mathrm{Cov}(X_2, X_3)$

$$E(aX_1+bX_2) = aE(X_1) + bE(X_2)$$

$a, b \in \mathbb{R}$ where X_1 and X_2 are Random Variables.

$$= E\left[((X_1+X_2)-E(X_1+X_2))(X_3-E(X_3))\right]$$

$$= E\left[((X_1-E(X_1)) + (X_2-E(X_2)))(X_3-E(X_3))\right]$$

$$= E\left[(X_1-E(X_1))(X_3-E(X_3))\right] + E\left[(X_2-E(X_2))(X_3-E(X_3))\right]$$

$$= \mathrm{Cov}(X_1, X_3) + \mathrm{Cov}(X_2, X_3)$$

Hence we are done.

Exercise 10:

No of samples $n = 100$, let $\epsilon > 0$
y be the actual average.

$$P(|\hat{y} - y| \leq \epsilon) \geq 0.95$$

$$P(|\hat{y} - y| > \epsilon) < 0.05$$

(i) Using $\boxed{P(|\hat{y} - y| > \epsilon) \leq 2\exp(-n\epsilon^2)}$

$$\Rightarrow \quad 0.05 \leq 2\exp(-n\epsilon^2)$$

$$\Rightarrow \quad 0.025 \leq \exp(-n\epsilon^2)$$

$$\ln(0.025) \leq -n\epsilon^2$$

$$\epsilon^2 < \frac{-1}{100}\ln(0.025)$$

$$\epsilon < 0.192$$

$$\therefore \quad \epsilon_{max} = 0.192$$

$|\mu - \hat{\mu}| \leq \epsilon_{max}$ will be the confidence interval

$$|\mu - 0.45| \leq 0.192$$

$$0.45 - 0.192 \leq \mu \leq 0.45 + 0.192$$

$$\boxed{0.26 \leq \mu \leq 0.642}$$

(ii) as $\delta = \exp(-n\epsilon^2) \Rightarrow \epsilon^2 n = \text{const}$

which means $\epsilon \propto \frac{1}{\sqrt{n}}$.

if we want ϵ to become $\frac{\epsilon}{2}$

then $n' = 4n$, hence

the minimum number of samples needed then would be $\underline{\underline{n = 400}}$.