Student ID: 1492182

Subject Code: MAST20006

Subject Name: Probabilities for Statistics

Assignment 3

Q1: $f(x) = \begin{cases} \frac{cx}{2}, & 0 \leq x < 2 \\ c, & 2 \leq x \leq 3 \end{cases}$

c: constant, $f(x)=0$ elsewhere

(a) $\int_0^2 \frac{cx}{2}\,dx + \int_2^3 c\,dx = 1$

$\therefore \frac{c}{2}\left[\frac{x^2}{2}\right]_0^2 + c\left[x\right]_2^3 = 1$

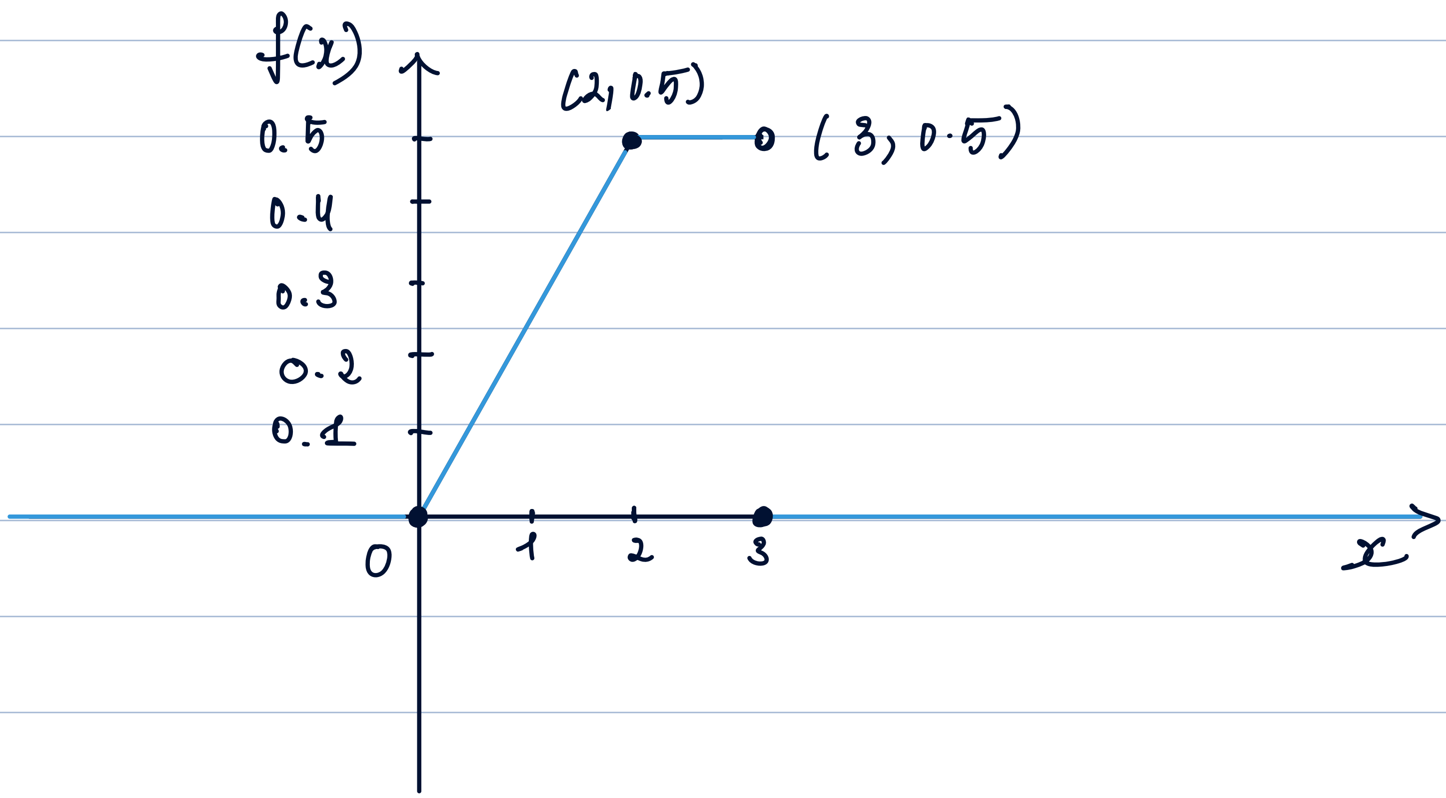
$\therefore \frac{c}{2}(2-0) + c(3-2) = 1$

$\therefore c + 3c - 2c = 1$

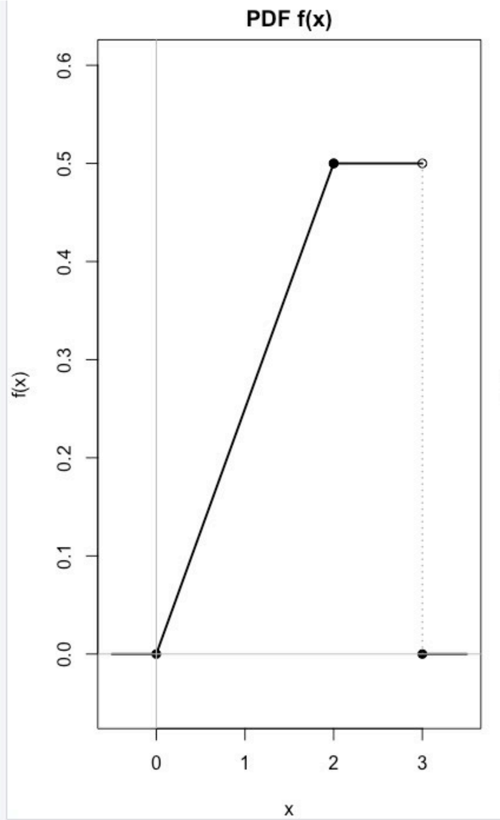
$\therefore 2c = 1$

$\therefore c = \frac{1}{2} = 0.5$

graph of pmf $f(x)$:



R graph:



(b)

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$c = 1/2$

For $x < 0$: $F(x) = 0$

For $0 \le x < 2$: $F(x) = \int_0^x \frac{ct}{2}\,dt = \frac{c}{2}\left[\frac{t^2}{2}\right]_0^x$

$$= \frac{c}{2}\left(\frac{x^2}{2}\right) = \frac{cx^2}{4} = \frac{x^2}{8}$$

for $2 \le x \le 3$: $F(x) = \int_0^2 \frac{ct}{2}\,dt + \int_2^x c\,dt$

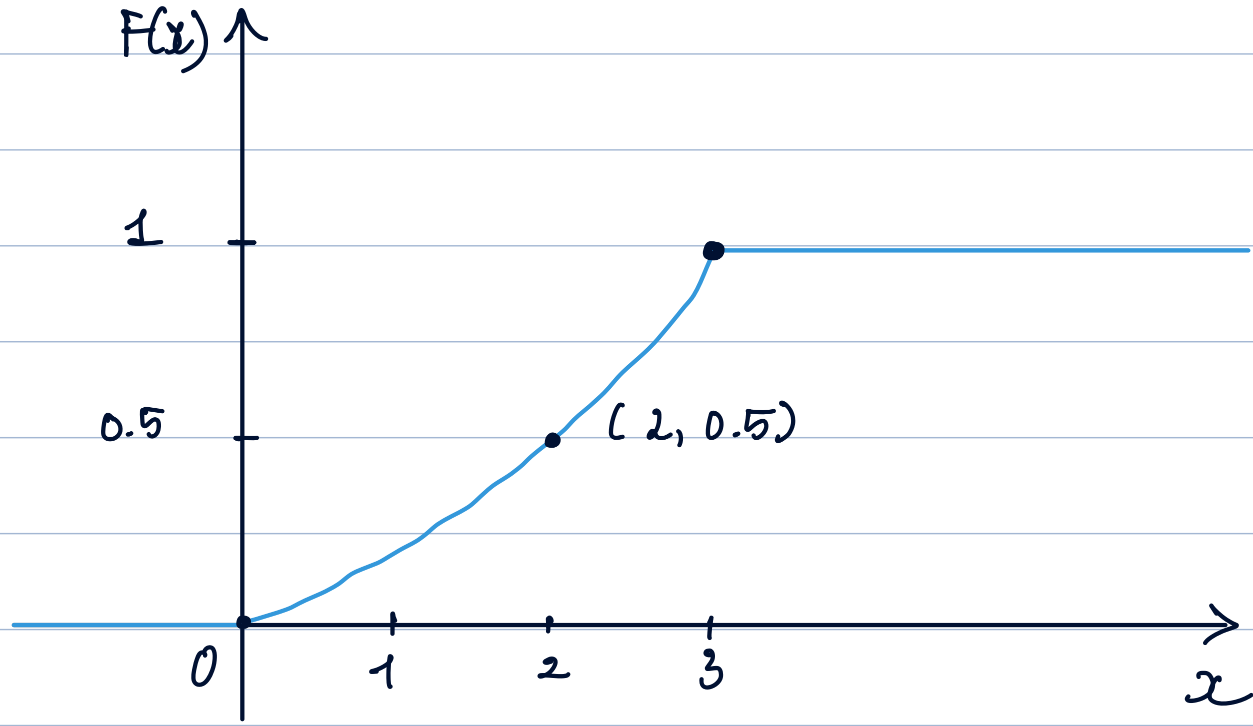
$$= \frac{1}{4}\left[\frac{t^2}{2}\right]_0^2 + \frac{1}{2}\left[dt\right]_2^x = \frac{1}{4}(2-0) + \frac{1}{2}(x-2)$$

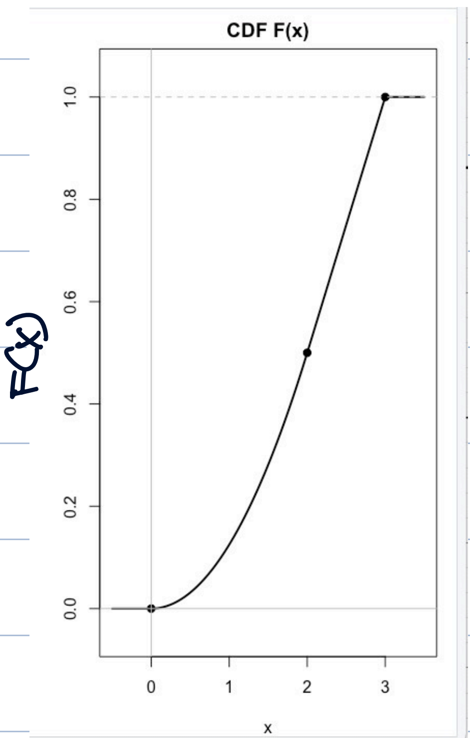
$$= \frac{x-1}{2}$$

For $x \geq 3$: $F(x) = 1$

So $F(x) = \begin{cases} 0 & , x < 0 \\ x^2/8 & , 0 \leq x < 2 \\ \dfrac{x-1}{2} & , 2 \leq x \leq 3 \\ 1 & , x \geq 3 \end{cases}$

graph of CDF $F(x)$:





(In R)

(c) $E(X) = \int_0^2 x \cdot \frac{cx}{2} dx + \int_2^3 cx \cdot dx$

$$= \frac{c}{2}\int_0^2 x^2 dx + c\int_2^3 x dx = \frac{c}{2}\left[\frac{x^3}{3}\right]_0^2 + c\left[\frac{x^2}{2}\right]_2^3$$

$$= \frac{c}{2}\left(\frac{8}{3} - 0\right) + c\left(\frac{9}{2} - 2\right)$$

$$= \frac{1}{4} \cdot \frac{8}{3} + \frac{1}{2} \cdot \frac{5}{2} = \frac{2}{3} + \frac{5}{4} = \frac{23}{12} = 1.6429$$
(4 d.p.)

(d) $M(t) = E[e^{tX}] = \int_0^2 e^{tx} \frac{cx}{2} dx + \int_2^3 e^{tx} c dx$

$$= M_1 + M_2$$

$$M_1 = \frac{c}{2}\int_0^2 xe^{tx} dx$$

Using integration by parts: Let $u = x$, $dv = e^{tx} dx$

$du = 1 dx$, $v = \int e^{tx} dx = \frac{e^{tx}}{t}$

Then $\int xe^{tx} dx = \frac{xe^{tx}}{t} - \int \frac{e^{tx}}{t} dx = \frac{xe^{tx}}{t} - \frac{1}{t}\int \frac{e^u}{t} du$

By substituting $u = tx \Rightarrow x = \frac{u}{t} \Rightarrow dx = \frac{1}{t} du$

$$= \frac{xe^{tx}}{t} - \frac{1}{t^2} e^u = \frac{xe^{tx}}{t} - \frac{e^{tx}}{t^2} = e^{tx}\left(\frac{x}{t} - \frac{1}{t^2}\right)$$

So $M_1 = \frac{1}{4}\left[e^{tx}\left(\frac{x}{t} - \frac{1}{t^2}\right)\right]_0^2 = \frac{1}{4}\left[e^{2t}\left(\frac{2}{t} - \frac{1}{t^2}\right) - \left(0 - \frac{1}{t^2}\right)\right]$

$= \frac{1}{4}\left[e^{2t}\left(\frac{2}{t} - \frac{1}{t^2}\right) + \frac{1}{t^2}\right]$

$M_2 = c\int_2^3 e^{tx}\,dx = \frac{1}{2}\left[\frac{e^{tx}}{t}\right]_2^3 = \frac{e^{3t} - e^{2t}}{2t}$

Hence, $M(t) = \frac{1}{4}\left[e^{2t}\left(\frac{2}{t} - \frac{1}{t^2}\right) + \frac{1}{t^2}\right] + \frac{e^{3t} - e^{2t}}{2t}$

Q2: $X \overset{d}{=} U(0,1)$

(a) Since the only integers near $[0,1]$ are 0 and 1:

$$Y = \min(|X-0|, |X-1|) = \min(X, 1-X)$$

As X ranges over $(0,1)$, the smallest $\min(X, 1-X)$ is 0 at $X=0$ or 1, the largest is at $X=\frac{1}{2}$ where $\min(\frac{1}{2},\frac{1}{2}) = \frac{1}{2}$

$\therefore$ Support $0 \leq Y \leq \frac{1}{2}$ or $Y \in [0, \frac{1}{2}]$

(b)

To find the CDF $F(y) = P(Y \leq y)$ for $0 \leq y \leq \frac{1}{2}$:

$\{Y > y\} = \{\min(X, 1-X) > y\} = \{X > y \text{ and } 1-X > y\}$
$= \{y < X < 1-y\}$

Since $X \overset{d}{=} U(0,1)$:

$$P(y < X < 1-y) = (1-y) - y = 1-2y$$

So,

$$F(y) = 1 - P(Y > y) = 1 - (1-2y) = 2y$$

and outside $[0, \frac{1}{2}]$, we have $F(y) = 0$ for $y < 0$
and $F(y) = 1$ for $y > \frac{1}{2}$.

(c)

From (c), $F(y) = 2$

By the distribution-function technique, the PDF is the derivative of the CDF at points where the derivatives exist, so:

$$f(y) = \frac{d}{dy} F(y) = \lim_{h \to 0} \frac{F(y+h) - F(y)}{h}$$

For any fixed y with $0 < y < \frac{1}{2}$:

$F(y+h) = 2(y+h)$

$F(y) = 2y$

so, $$\frac{F(y+h) - F(y)}{h} = \frac{2(y+h) - 2y}{h} = \frac{2h}{h} = 2$$

$$f(y) = \lim_{h \to 0} 2 = 2, \quad 0 < y < \frac{1}{2}$$

outside $(0, \frac{1}{2})$, the CDF is 0 for $y < 0$, 1 for $y > \frac{1}{2}$, so the derivative $= 0$ there.

Hence, the PDF is: $f(y) = \begin{cases} 2, & 0 < y < \frac{1}{2} \\ 0, & \text{elsewhere} \end{cases}$

A constant density 2 over an interval of length $\frac{1}{2}$ is the $U(0, \frac{1}{2})$ distribution since $1/(\frac{1}{2} - 0) = 2$, so:

$Y \sim U(0, \frac{1}{2})$, $f(y) = 2$ for $0 \le y \le \frac{1}{2}$

Q3 :

(a)

Let X be the waiting time (minutes) from 7am until the first customer arrives

In a Poisson process with rate λ events per unit time, the waiting time until the first event follows an Exponential distribution

$$\lambda_{min} = \frac{5 \text{ customers}}{1 \text{ hour}} \times \frac{1 \text{ hour}}{60 \text{ minutes}} = \frac{1}{12} \text{ customer/min}$$

So, $X \stackrel{d}{=} \text{Exponential}\left(\theta = \frac{1}{\lambda_{min}} = 12 \text{ min}\right)$

$$f(t) = \frac{1}{12} e^{-t/12}, \quad t \geq 0$$

(b) $P(X > 15) = e^{-15/12} = e^{-5/4} = 0.2865$ (4 d.p.)

(c) By Memoryless property of the Exponential distribution

$P(X > 20 \mid X > 15) = P(X > 15 + 5 \mid X > 15)$

$= P(X > 5) = e^{-5/12} = 0.6592$ (4 d.p.)

(d) Let T_2 be the time (in hours) from 7am to the 2^{nd} arrival

$$P(1 < T_2 \leq 2)$$

since 8am is $t=1$ and 9am is $t=2$

<u>Method 1</u>: Gamma distribution

In a Poisson process with rate λ, the waiting time until α-th arrival, T_2, follows a Gamma distribution

Here, $\alpha = 2$ and the rate (in hours) is $\lambda = 5$ so:

$$T_2 \overset{d}{=} \text{Gamma}(\alpha = 2, \theta = \tfrac{1}{5})$$

$P(1 < T_2 \leq 2) = F_{T_2}(2) - F_{T_2}(1)$, where $F_{T_2}(t)$ is the CDF of $T_2 = 1 - \sum_{k=0}^{1} \frac{(\lambda x)^k e^{-\lambda x}}{k!}$ so:

$P(1 < T_2 \leq 2)$

$$= \left[1 - \sum_{k=0}^{1} \frac{(5\cdot 2)^k e^{-10}}{k!}\right] - \left[1 - \sum_{k=0}^{1} \frac{(5\cdot 1)^k e^{-5}}{k!}\right]$$

$$= \left(1 - (1 + 5\cdot 2)e^{-10}\right) - \left(1 - (1+5)e^{-5}\right)$$

$$= 1 - (1+10)e^{-10} - 1 + (1+5)e^{-5}$$

$$= 6e^{-5} - 11e^{-10}$$

$$\simeq 0.0399 \text{ (4 d.p.)}$$

Method 2 : Poisson distribution:

Let $N(t)$ be the number of arrivals by time t

Then $P(1 < T_2 \leq 2) = P(N(2) \geq 2) - P(N(1) \geq 2)$

Since $N(t) \overset{d}{=} \text{Pois}(5t)$

and $$P(N = k) = \frac{(5t)^k e^{-5t}}{k!}$$

Then, $P(1 < T_2 \leq 2)$

$$= \left[1 - (e^{-10} + 10e^{-10})\right] - \left[1 - (e^{-5} + 5e^{-5})\right]$$

$$= 1 - e^{-10} - 10e^{-10} - 1 + e^{-5} + 5e^{-5}$$

$$= -11e^{-10} + 6e^{-5} \approx 0.0399 \text{ (4 d.p.)}$$

(Same as Method 1's result)