STAT 650 - ASSIGNMENT 4

Name: Sneha Mishra

UIN: 733000826

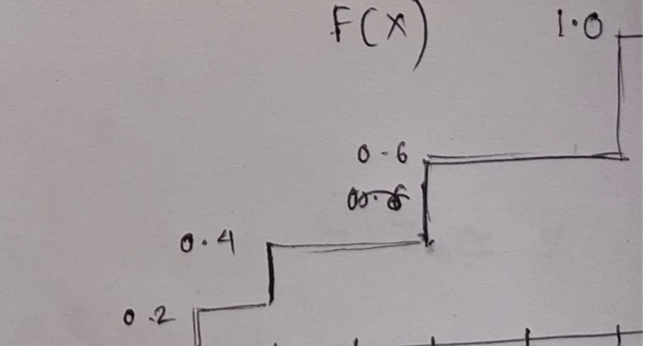
1) $P(X=x) = \begin{cases} 0.2 & \text{if } x=0 \\ 0.2 & \text{if } x=1 \\ 0.2 & \text{if } x=3 \\ 0.4 & \text{if } x=5 \\ 0 & \text{otherwise} \end{cases}$

a)

X	0	1	3	5	otherwise
$P(x)$	0.2	0.2	0.2	0.4	0
$F(X)$	0.2	0.4 $(0.2+0.2)$	0.6 $(0.4+0.2)$	1 $(0.6+0.4)$	1

$F(X) = P(X \le x)$

$$= \begin{cases} 0, & \text{if } x < 0 \\ 0.2, & \text{if } x \in [0,1) \\ 0.4, & \text{if } x \in [1,3) \\ 0.6, & \text{if } x \in [3,5) \\ 1, & \text{if } x \in [5,\infty) \end{cases}$$



b) $E(X) = \sum x P(x)$ where $x = 0, 1, 3, 5, \ldots$

$= [(0 \times 0.2) + (1 \times 0.2) + (3 \times 0.2) + (5 \times 0.4) + 0]$

$= 2.8$

$X =$	0	1	3	5	otherwise 0
$X^2 =$	0	1	9	25	0

$$E(X^2) = \sum_{x=0,1,3,5,\text{otherwise.}} x^2 \cdot P(X=x)$$

$$= [(0 \times 0.2) + (1 \times 0.2) + (9 \times 0.2) + (25 \times 0.4) + 0]$$

$$= [0.2 + 1.8 + 10]$$

$$= 12.$$

$$Var(X) = E(X^2) - (E(X))^2$$

$$= 12 - (2.8)^2$$

$$= 4.16$$

Q2 → Given: $n = 100$

p (% of drugs equivalently licensed for sale)

$= 0.2$

$\therefore$ q (% of drugs not licensed for sale)

$= 1 - p = 1 - 0.2 = 0.8$

a) $P(X \geq 15) = 1 - P(X \leq 14)$

$$= 1 - \left(\sum_{r=1}^{14} {}^{100}C_r (0.2)^r (0.8)^{100-r} \right)$$

$= 1 - 0.08044$

$= 0.91955$

$\therefore$ The probability that at least 15 of the 100 drugs are equivalently licesed is

0.91955.

b) Calculation of (a) using binomial-normal approximation:

$n = 100$

$p = 0.2$

$q = 1 - p = 1 - 0.2 = 0.8$

Condition for using binomial-normal approx: is $np > 5$ and $nq > 5$

Here, $np = 100 \times 0.2 = 20$ which is > 5

$nq = 100 \times 0.8 = 80$ which is > 5

$P(X \geq 15) \rightarrow$ Using continuity correction factor i.e $(15 - 0.5)$

$\Rightarrow P(X \geq 14.5)$

~~= 1 - P(X ≤~~

$= 1 - P(X \leq 14.5)$

$$= 1 - P\left(\frac{X-\mu}{\sigma} \le \frac{14.5-\mu}{\sigma}\right)$$

$$= 1 - P\left(z \le \frac{14.5-20}{4}\right)$$ $\left[\because \mu = np = 20,\ \sigma^2 = npq = 16,\ \sigma = 4\right]$

$$= 1 - P(Z \le -1.375)$$

$$= 1 - 0.08457 \quad \text{(from Z table)}$$

$$= 0.91543.$$

Q3 → a) To find: $P(7 \le Y \le 11)$

Sol^n: Poisson distribution : $\lambda = 9$

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

$$P(7 \le Y \le 11)$$

$$= P(Y \le 11) - P(Y \le 6)$$

$$= 0.8030 - \cancel{0.3239}\ 0.2068 \quad \text{(from poisson's table)}$$

$$= \cancel{0.4791} = 0.5962$$

b) Calculation of (a) using poisson-normal approximation:

Sol^n: $Var(Y) = \mu = \lambda = 9$ $\quad \sigma = \sqrt{9} = 3.$

$P(7 \le y \le 11)$ with using continuity correction factor →

$P(6.5 \leq y \leq 11.5)$

$= P\left[\frac{6.5-9}{3} \leq \frac{y-\mu}{\sigma} \leq \frac{11.5-9}{3}\right]$

$= P\left[-\frac{2.5}{3} \leq Z \leq \frac{2.5}{3}\right]$

$= P\left(Z \leq \frac{2.5}{3}\right) - P\left(Z \leq -\frac{2.5}{3}\right)$

$= 0.7976 - 0.20242$ (from z table)

$= 0.595$

Q4 → $X \rightarrow$ Normal distribution

$\mu = 3$

$\sigma^2 = 9$

a) $P(1 < x < 6)$

$= P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right)$

$= P\left(\frac{1-3}{3} < Z < \frac{6-3}{3}\right)$

$= P\left(-\frac{2}{3} < Z < 1\right)$

$= P(-0.667 < z < 1)$

$= F(1) - F(-0.667)$

$= 0.84134 - 0.25239$

$= 0.58895.$

b) $E((2+x)^2)$ Given : $\mu = 3$, $\sigma^2 = 9$

$= E(4 + x^2 + 4x)$

$= E(4) + E(x^2) + 4E(x)$

Now, $Var = E(x^2) - E(x)^2$

$\Rightarrow E(x^2) = Var + E(x)^2$

$= \sigma^2 + \mu^2$

$= 9 + 9$

$= 18$

$\therefore E((2+x)^2)$
$= E(4) + E(x^2) + 4E(x)$
$= 4 + 18 + 12$
$= 34.$

$Var(4+3x) = Var(3x)$

$= 3^2 \cdot var(x)$

$= 9 \times 9$

$= 81$

$\therefore E((2+x)^2) = 34$ & $Var(4+3x) = 81$

Q5 → (a) $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$.

Now, $F(x) = \int f(x)\, dx$.

$= \int \lambda e^{-\lambda x} dx$.

Let $-\lambda x = z$ $\therefore$ $F(x) = \int \cancel{\lambda} e^{z} \cdot \left(-\frac{1}{\cancel{\lambda}}\right) dz$

$-\lambda dx = dz$

$= -\int e^{z} dz$.

Applying limits to CDF $F(x) = -\int_0^x e^{-\lambda x}$.

$= 1 - e^{-\lambda x}$.

Now, for $g(x) = y = 2x + 4$,

$x = (y-4)/2$.

We know, $CDF(y) = CDF(x)$

$= CDF(g^{-1}(y))$

$CDF(y) = 1 - e^{-\lambda\left(\frac{y-4}{2}\right)} = 1 - e^{-(y-4)}$ $[\because \lambda = 2]$

Now, $PDF(y) = \frac{\partial}{\partial y} CDF(y) = -\frac{\lambda}{2} \times -e^{-\lambda\left(\frac{y-4}{2}\right)}$

$= \frac{\lambda}{2} e^{-\lambda\left(\frac{y-4}{2}\right)}$

$\therefore$ PDF of Y can be written as

putting $\lambda = 2$, $g(y) = \begin{cases} e^{-(y-4)} & , y > 4 \\ 0 & , \text{otherwise} \end{cases}$

b) $E(Y) = \int f(y)\, y\, dy$

$= \int \frac{\lambda}{2} e^{-\lambda\left(\frac{y-4}{2}\right)} \cdot y \cdot dy$.

$= \frac{\lambda}{2} e^{2\lambda} \int e^{-\lambda y/2} \cdot y\, dy$.

Let $\frac{\lambda y}{2} = z$

$\frac{\lambda}{2} dy = dz$.

$\therefore E(Y) = \frac{\cancel{\lambda}}{\cancel{2}} e^{2\lambda} \int e^{-z} \cdot \frac{\cancel{2}z}{\cancel{\lambda}} \cdot \frac{2\, dz}{\lambda}$

$= \frac{2}{\lambda} \cdot e^{2\lambda} \int e^{-z} \cdot z\, dz$

Applying integration by parts $\longrightarrow$

$= \frac{2}{\lambda} e^{2\lambda} \left[\int -e^{-z} \cdot z - \int -e^{-z} dz \right]$

$x = 0, \infty$
$\Rightarrow y = 4, \infty$
$\Rightarrow z = 2\lambda, \infty$

$= \frac{2}{\lambda} e^{2\lambda} \left[-e^{-z} \cdot z - e^{-z} \right]_{2\lambda}^{\infty}$

$= \frac{2}{\lambda} e^{2\lambda} \left[e^{-2\lambda} \cdot 2\lambda + e^{-2\lambda} \right]$

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

$= 4 + \frac{2}{\lambda}$.

Now, $\lambda = 2$. So, $E(Y) = 4 + \frac{2}{2} = 5$

$E(Y^2)$

$$\cancel{Var(y)} = \int f(y)\, y^2\, dy$$

$$= \int \frac{\lambda}{2} \cdot e^{-\lambda\left(\frac{y-4}{2}\right)} y^2 \cdot dy$$

$$= \int \frac{\lambda}{2} \cdot e^{-\lambda y/2} \cdot e^{2\lambda} \cdot y^2 dy$$

$$= \frac{\lambda}{2} e^{2\lambda} \int e^{-\lambda y/2} \cdot y^2 dy. \quad \text{--- (1)}$$

Let $z = \frac{\lambda y}{2} \Rightarrow y = \frac{2z}{\lambda}$

$dz = \frac{\lambda}{2} dy$

when $x = 0, \infty$

$y = 4, \infty$

$z = 2\lambda, \infty$.

Now, (1) becomes,

$$\frac{\cancel{\lambda}}{\cancel{2}} e^{2\lambda} \int e^{-z} \times \frac{4z^2}{\lambda^2} \cdot \frac{\cancel{2}\, dz}{\cancel{\lambda}}$$

$$= \frac{4}{\lambda^2} e^{2\lambda} \int e^{-z} z^2 dz \quad \text{--- (2)}$$

Applying integration by parts:

$$eq^n\ (2) = \frac{4}{\lambda^2} e^{2\lambda} \left[-e^{-z} \cdot z^2 - \int -2 \cdot e^{-z} \cdot z\, dz \right]$$

$$= \frac{4}{\lambda^2} e^{2\lambda} \left[-e^{-z} \cdot z^2 - \left(-2\left(-e^{-z} \cdot z - e^{-z}\right)\right) \right]$$

$$= -\frac{4}{\lambda^2} e^{2\lambda} \left[e^{-z} \cdot z^2 + 2\left(e^{-z} \cdot z + e^{-z}\right) \right]_{2\lambda}^{\infty}$$

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

$$= \frac{4}{\lambda^2} [4\lambda^2 + 4\lambda + 2]$$

As $\lambda = 2$,

$$E(Y^2) = \frac{4}{2^2} [4 \times (2)^2 + 4(2) + 2]$$

$$= 26$$

$$\therefore \quad Var(y) = E(Y^2) - E(Y)^2$$

$$= 26 - 5^2$$

$$= 1$$

c) 95 percentile of Y

To find Y such that CDF of Y = 0.95

$$\Rightarrow \quad 1 - e^{-\lambda/2\,(y-4)} = 0.95$$

$$\Rightarrow \quad 1 - e^{-(y-4)} = 0.95$$

$$\Rightarrow \quad 0.05 = e^{-y} \cdot e^{+4}$$

$$\Rightarrow \quad e^{Y} = \frac{e^4}{0.05}$$

$$\Rightarrow \quad y = \ln \frac{e^4}{0.05} \doteq 6.9957$$

$$\simeq 7$$

d) $P[2 \le y \le 36]$

$Y = 2x + 4$

$P[2 \le 2x + 4 \le 36]$

$= P\left(\frac{2-4}{2} \le x \le \frac{36-4}{2}\right)$

$= P(-1 \le x \le 16)$

$= F(16) - F(-1)$

$= 0.999 - 0$

$= 0.999.$

(attached python code at the end)

Q6 →

First 3 months purchases	Heavy	Medium	Light	
0	5	15	60	80
1	10	30	20	60
2	30	40	15	85
3 +	55	15	5	75
	100	100	100	300

P(L) = probability of light

P(0) = probability of zero books purchased

P(L/0) = probability of light given zero books purchased

$$P(L|0) = \frac{P(0|L)\cdot P(L)}{P(0)}$$

$$= \frac{0.6 \times 0.5}{P(0|H)\cdot P(H) + P(0|M)\cdot P(M) + P(0|L)\cdot P(L)}$$

$$= \frac{0.6 \times 0.5}{(0.05 \times 0.2) + (0.15 \times 0.3) + (0.6 \times 0.5)}$$

$$= \frac{0.6 \times 0.5}{0.355} = 0.8450.$$

Q 7 → a) Given: $N(\mu, a=3)$

$$\bar{x} \pm Z_{\alpha/2} \frac{a}{\sqrt{2}}$$

$$= \bar{x} \pm 1.96 \frac{a}{\sqrt{n}}$$ (for 95% of confidence interval)

So, for 95% confidence interval of μ, we have

$$45 \pm 1.96 \times \frac{3}{10}$$

$$= 45 \pm 1.96 \times 0.3$$

$$= [44.412, 45.588]$$

b) Given that

μ_1 of $x_1, x_2, \ldots, x_{50} = 42$

We know,

$$\frac{\mu_1(x_1, x_2, \ldots, x_{50}) + \mu_2(x_{51}, \ldots, x_{100})}{2} = \mu \text{ (Sample mean)}$$

Given: Sample mean = 45

So, $\frac{42 + \mu_2}{2} = 45$

$\mu_2 = 2 \times 45 - 42$
$= 48$

As per question,
if we assume μ is calculated on the last 50 observations $(x_{51}, \ldots, x_{100})$

then let $\mu = \mu_2$

Confidence interval for $\mu = \mu_2$

$\Rightarrow \mu_2 \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{50}}$ ($n = 50$ as last 50 observations are considered)

So, new confidence interval

$= 48 \pm 1.96 \times \frac{3}{\sqrt{50}}$

$= 48 \pm 0.8315$

$= [47.1685, 48.8315]$

(c) For (a) the width of the interval

$= 2 \times 5.88$

$= 1.176$

For (b), the width of the interval

$= 2 \times 0.8315$

$= 1.663$

i.e as the no. of samples decreased, the width of the interval with same 95% confidence increased.

Since, we have a narrower confidence interval for N in case (a), it is considered better.

—x—

```
#Question 5-(d)
# As per question,
from scipy.stats import expon
expon.cdf(x=16)-expon.cdf(x=-1)
```

```
0.9999998874648253
```

```
#Question-2(a) Bionomial part calculation verification
import numpy as np
s=0
for i in range(1,15):
    s =  s +
np.math.factorial(100)/(np.math.factorial(i)*np.math.factorial(100-
i))*pow(0.2,i)*pow(0.8,(100-i))
print(s)
```

```
0.08044372093434754
```