Distributions- Assignment

1. import pandas as pd

```
import pandas as pd
import matplotlib.pyplot as plt
titanic_df = pd.read_csv(" .. \Titanic.csv")
titanic_df
```

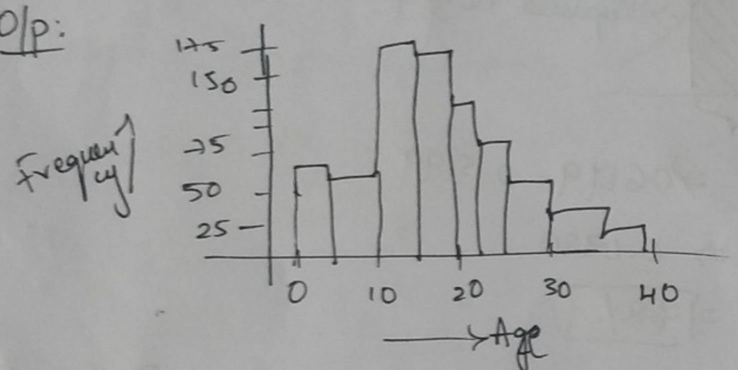
O/p: Data

Passenger Id Survived PClass Name Sex Age SibSp ...

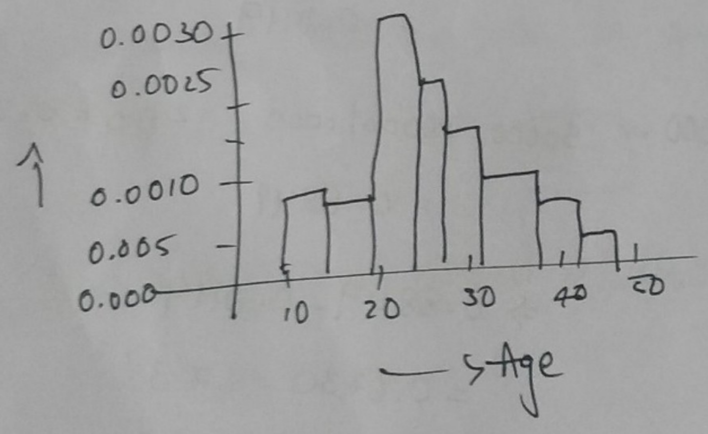
- - - - - - - - - - -

```
titanic_df.Age.plot.hist().set_title("Frequency Distribution")
```

O/p:




```
titanic_df.Age.hist(density=True).set_title("Probability Mass function")
```



2. Given, $\mu = 38000$, $\sigma = 10000$, $n = 2000$ firms

(a) Number of firms with sales of over 50,000

$P(X > 50K)$, Z-Scr $= X-\mu/\sigma = 50000-38000/10000 = 12/10 = 1.2 = 0.8849$

over 50,000 means $1-0.8849 = \boxed{0.11 = \boxed{11\%}} \Rightarrow 11/100 \times 2000 = \boxed{220}$

(b) 220 firms have sales greater than 50,000.

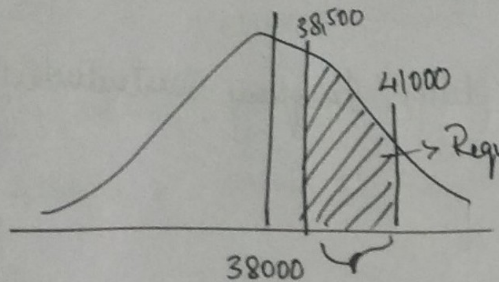
↳ Percentage of firms with sales b/n Rs. 38,500 and Rs. 41,000

Z-Score for 38,500 → $\frac{38500-38000}{10000} = 500/10000$

$= 0.05 = 0.5199$

Z-Score for 41,000 = $41000 - 38000/10000 = 3000/10000$

$= 0.3 = 0.6179$



→ Required no of firms b/n particular sales

$\Rightarrow 0.6179 - 0.5199$

$\Rightarrow 0.0980$

$= \boxed{9.8\%}$

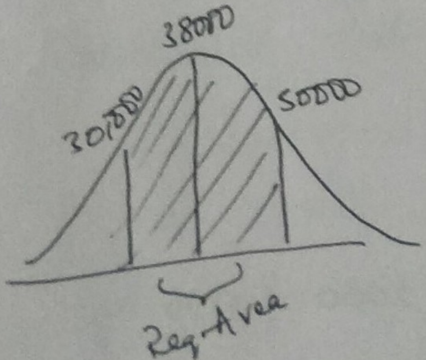
(c) Number of firms with sales b/n Rs. 30,000 and Rs. 50,000

Z-score = $30000 - 38000/10000 = -8000/10000 = -0.8$

$= 0.2119$

Z-Score for 50000 → $50000 - 38000/10000 = 12/10 = 0.2$

$= 0.8849$



$\Rightarrow 0.8849 - 0.2119$

$= 0.6730 = 67.3\%$

$673/1000 \times 2000$

$= \boxed{1346 \text{ firms}}$

3. Given $n = 20$ $\quad r(\text{success}) = 1/4$

$1-r(\text{failure}) = 3/4$

Exactly 5 wrong answers means $r = 15$ (success)

$1-r = 5$

follows Binomial Distribution,

$$P(x=r) = {}^{n}C_{r} (p)^{r} (q)^{n-r}$$

$$= \boxed{{}^{20}C_{15} \left(\frac{1}{4}\right)^{15} \left(\frac{3}{4}\right)^{5}}$$

4. Given, $\lambda = 4$ photon/second

Probability of no photons reaching telescope means $r=0$.

follows a Poisson Distribution,

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

$$P(x=0) = \frac{e^{-4} \cdot \lambda^{0}}{0!} = \boxed{e^{-4}}$$

5. Given, $\lambda = 3$ (calls per minute)

(a) Probability that no calls in given 1-minute means $r=0$

$$P(r=0) = \frac{e^{-3} \cdot (3)^{0}}{0!} = \boxed{e^{-3}}$$

(b) Probability that atleast 2 calls will arrive in 2-minute

$\lambda = 3$ per minute

for 2 minute $\lambda = 6$

(atleast 2 calls means) $P(x=2) + P(x=3) + \cdots$ for that

$1 - (P(x=0) + P(x=1))$

$$1 - e^{-6}\left[\frac{6^{0}}{0!} + \frac{6^{1}}{1!}\right] = \boxed{1 - 7e^{-6}} = \boxed{0.982}$$

6. Given defective rate = 20% = 0.2

$P = 0.2, q = 0.8$ (Non-defective)

Probability of 1st defective after 3 good parts, ..
means; after 3 good part, in any chance, you can find defective part

Follows Geometric Distribution,

$P(X=x) \Rightarrow q^{x-1} \cdot P$

$P(x \geq 3) \Rightarrow 1 - P(x < 3)$

$1 - [P(x=1) + P(x=2)]$

$P(x=1) = (0.2)^{1-1}(0.8) = 0.8$

$P(x=2) = (0.2)^{2-1}(0.8) = 0.16$

$\Rightarrow 1 - [0.8 + 0.16] = \boxed{0.04}$

$= \boxed{4\%}$

Avg no of inspections to obtain first defective part

$\mu = 1/p = 1/0.2 = \boxed{5}$

7. Given,

$r(\text{success}) = 0.3$ $n = 5$

$1 - r(\text{fail}) = 0.7$

Probability atmost 2 are accepted to college means
0 student, 1 student & 2 student

$P(x=0) + P(x=1) + P(x=2)$

Follows Binomial Distribution,

$\Rightarrow 5_{C_0}(0.3)^0(0.7)^5 + 5_{C_1}(0.3)^1(0.7)^4 + 5_{C_2}(0.3)^2(0.7)^3$

$\Rightarrow 1(1)(0.7)^5 + 5(0.3)(0.7)^4 + 10(0.09)(0.7)^3$

$\Rightarrow 0.168 + (1.5)(0.2401) + 10(0.09)(0.343)$

$\Rightarrow 0.168 + 0.36015 + 0.3087 = \boxed{0.83685 = 83.68\%}$

8. Given, maximum weight elevator can bare
$= 800kg$

Avg adult weight $\mu = 70kg$

Variance $= 200 \Rightarrow$ std(σ) $= \sqrt{200} = 14.14$

Probability that elevator reaches safely if 10 diff adults

→ All 10 adults can have any weight, but elevator cap $= 800kg$

So, we calculate new μ

$$\bar{\mu} = \frac{800}{10} = 80kg.$$

Now take 80kg. as x.

$$z = \frac{x-\mu}{\sigma} = \frac{80-70}{14.14} = 0.707$$

P-value $= 0.7794 = \boxed{77.94\%}$ lift reaches safely if there are 10 adults.

If '12' adults', we know elevator capacity $= 800kg$

New $\mu = 800/12 = 66.66kg$

So, on avg when there are 12 adults, if weight of each person is 66.66, then lift would reach the ground safely

So,

$$z\text{-score} = \frac{66.66-70}{14.14} = -3.34/14.14 = -0.2362$$

P-value $= 0.4090$

$= \boxed{40.9\%}$ elevator reaches safely if there are 12 adults; the avg wt of adult in that apt is itself 70kg;

9. Given, $n = 50$ if '2' choices for each question

r (success) $= 1/2$

$1-r$ (fail) $= 1/2$

Atleast 20; clear the exam that means $P(x=20)+P(x=21)+P(x=22)+\dots+P(x=50)$

Probability to clear the exam; even if 20 answers are right, he clears the test.

Follows Binomial Distribution,

$$= 50_{C_{20}} \left(\frac{1}{2}\right)^{20} \left(\frac{1}{2}\right)^{30}$$

If there are '4' choices ($r \to$ (success) $= 1/4$

$1-r \to$ (fail) $= 3/4$

Then, probability to clear the exam (some exactly 20 answers to be correct)

$$\Rightarrow 50_{C_{20}} \left(\frac{1}{4}\right)^{20} \left(\frac{3}{4}\right)^{30}$$

10. Given $(1-r)$ (faulty) $= 30\% = 0.3$

r (success) $= 0.7$

$n=6$, $r=2$ (exactly faulty)

Probability of exactly 2 faulty bulbs being selected (means '4' good bulbs)

$$P(x=r) = n_{C_r} (P)^r (q)^{n-r}$$

$$P(x=\cancel{2}) = 6_{C_4} (0.7)^4 (0.3)^2 = 2.5 (0.7)^4 (0.09)$$

$P(x=4)$

$$= 2.5 \times 0.2401 \times 0.09$$

$$= \boxed{0.054}$$

Given,

11. Efficiency of writer, 6 error per hr

77 words per minute

$77 \times 60 = 4620$ words per hour with 6 errors is writer's efficiency

Probability of 2 errors in 322 word report

6 error → 4620 words → (Divide by '3' to equate)

2 error → 1540 words (as per writer efficiency)

But, for 322 word report, we need 'λ' value

(Mean typing speed) $\lambda = 1540/322 = 4.78$
to commit an error

Req. probability → Follows Poisson Distributions

2 error in 322 word report with avg typing speed in order to commit error = 4.78

$$P(X=2) = \frac{e^{-4.78}(4.78)^2}{2!} = \boxed{0.0959 \simeq 9.59\%}$$

12. (a) Probability less than 1 site exceeds the recommended level of dioxin

Given, mills with higher than Recommended level = 5%.

Sawmill sites with Recommended level = 95%.

Assume claim is correct,

$P(X<1)$ $(x) = 0.05$

$(1-x) = 0.95$

means no site (0 site exceeds recommend level)

$\Rightarrow P(X=0) =$ n is 20, x = 0

$$^{20}C_0 (0.05)^0 (0.95)^{20}$$

$$= (0.95)^{20} = 0.3584$$

$$= \boxed{35.84\%}$$

(b) Probability that less than or equal to 1 site exceeds the recommended level of ~~pollution~~ dioxin.

$X = 0, 1 \Rightarrow P(X=0) + P(X=1)$

$n = 20$

$$^{20}C_0 (0.05)^0 (0.95)^{20} + {}^{20}C_1 (0.05)^1 (0.95)^{19}$$

$$= 0.3584 + 20(0.05)(0.3773)$$

$$= 0.3584 + 0.3773 = \boxed{0.7357}$$

$$\boxed{73.57\%}$$

(c) Calculate probability that at most (i.e) maximum of 2 sites exceed the recommended level of dioxin.

$P(X=0) + P(X=1) + P(X=2) \rightarrow$ from above $P(X=0) + P(X=1) = 0.7357$

$$P(X=2) \rightarrow {}^{20}C_2 (0.05)^2 (0.95)^{18}$$

$$= 190(0.0025)(0.3972) = \boxed{0.1886}$$

Req. Probability $= 0.7357 + 0.1886 = \boxed{0.9243 = 92.43\%}$

13. a) Probability that company "Rose Waste Disposal" will be selected for auditing exactly twice in the next 5 years.

Follows Poisson distribution

Audit done $\rightarrow$ 5% every year of all companies, 'λ' value for '5' years $\rightarrow 25\% = 0.25$

for exactly twice $\rightarrow X = 2$

$$P(X=2) = \frac{e^{-0.25}(0.25)^2}{2!} = \boxed{0.0243}$$

13. (b) Probability that company will be audited exactly twice in the next 2 years.

5% → per year

per 2year → 10% of all companies

$\lambda = 0.1$

exactly twice to be audited $x=2$

$$P(x=2) = \frac{e^{-0.1}(0.1)^2}{2!} = \boxed{0.0045}$$

(c) Probability that the company will be audited atleast once in the next 4 years.

$x = 1,2,3,..4$ (any no. of times to be audited) but exactly probability for atleast once in next 4 years

next 4 years $\lambda = 20\% = 0.2$

(C.M.F) for exact probability of one audit for (4) years

$$= \frac{e^{-0.2}(0.2)^1}{1!}$$

If for any no. of audits more than once, in next 4 years it would be (4) audits

$$= e^{-0.2}\left(\frac{0.2}{1!} + \frac{(0.2)^2}{2!} + \frac{(0.2)^3}{3!} + \frac{(0.2)^4}{4!}\right)$$

$$= e^{-0.2}\left(\frac{0.2}{1} + \frac{0.04}{2} + \frac{0.008}{6} + \frac{0.0016}{24}\right)$$

14. Probability to stop = 0.2 ∈ (r)

No need to stop = 0.8 ∈ (1-r)

Given, n = 15

a) Probability to stop at exactly 2 of 15 sets of lights

$r = 2$

$P(x=2) = {}^{15}C_{2}(0.2)^{2}(0.8)^{13}$

$= 105(0.04)(0.0549)$

$= 0.23058 = \boxed{23.05\%}$

b) Probability that student will be stopped at 1 or more of 15 sets of traffic lights

$r = 1$ & $>1 \Rightarrow P(r=1) + P(r=2) + \ldots + P(r=15)$

$1 - P(r=0) = {}^{15}C_{0}(0.2)^{0}(0.8)^{15} = 0.0351$

$\Rightarrow 1 - 0.0351 = \boxed{0.9648}$

$\Rightarrow \boxed{96.48\%}$