**Assignment stats 1**

Q1. The marks awarded for an assignment set for a Year 8 class of 20 students were as

follows:

6 7 5 7 7 8 7 6 9 7 4 10 6 8 8 9 5 6 4 8

Ans.



Mean (µ)= 6+7+5+7+7+8+7+6+9+7+4+10+6+8+8+9+5+6+4+8 /20

= 137 / 20

= 6.85

In this Case,

n is even

Then,



Let’s arrange in ascending order

4 4 5 5 6 6 6 6 7 7 7 7 7 8 8 8 8 9 9 10

Median = 7+7 /2 = 14 /2 = 7

Mode is the number with the largest frequency

|  |  |
| --- | --- |
| Marks | Frequency |
| 4 | 2 |
| 5 | 2 |
| 6 | 4 |
| 7 | 5 |
| 8 | 4 |
| 9 | 2 |
| 10 | 1 |

Hence, Mode = 7

|  |  |  |
| --- | --- | --- |
| x | x- µ | (x- µ)^2 |
| 6 | -0.85 | 0.7225 |
| 7 | 0.15 | 0.0225 |
| 5 | -1.85 | 3.4225 |
| 7 | 0.15 | 0.0225 |
| 7 | 0.15 | 0.0225 |
| 8 | 1.15 | 1.3225 |
| 7 | 0.15 | 0.0225 |
| 6 | -0.85 | 0.7225 |
| 9 | 2.15 | 4.6225 |
| 7 | 0.15 | 0.0225 |
| 4 | -2.85 | 8.1225 |
| 10 | 3.15 | 9.9225 |
| 6 | -0.85 | 0.7225 |
| 8 | 1.15 | 1.3225 |
| 8 | 1.15 | 1.3225 |
| 9 | 2.15 | 4.6225 |
| 5 | -1.85 | 3.4225 |
| 6 | -0.85 | 0.7225 |
| 4 | -2.85 | 8.1225 |
| 8 | 1.15 | 1.3225 |
| Total |  | 50.55 |

Standard deviation = 2

=

= 1.59

Q2. The number of calls from motorists per day for roadside service was recorded for a

particular month:

28, 122, 217, 130, 120, 86, 80, 90, 140, 120, 70, 40, 145, 113, 90, 68, 174, 194, 170,

100, 75, 104, 97, 75,

123, 100, 75, 104, 97, 75, 123, 100, 89, 120, 109

Ans.



Mean(µ)= 28+122+217+130+120+86+80+90+140+120+70+40+145+113+90+68+174

+194+170+100+75+104+97+123+100+89+120+109 /35

= 3189/ 35

= 91.11

In this Case,

n is odd

Then,



Let’s arrange in ascending order

28, 40, 68, 70, 75, 75, 75, 75, 80, 86, 89, 90, 90, 97, 97, 100, 100, 100, 104, 104,109,

113, 120, 120, 120, 122, 123, 123, 130, 140, 145, 170, 174, 194, 217

Median = 100

Here, the number with the highest frequency is 75 which is occurring 4 times

i.e. Mode = 75

|  |  |  |
| --- | --- | --- |
| x | x- µ | (x- µ)^2 |
| 28 | -63.11 | 3982.8721 |
| 40 | -51.11 | 2612.2321 |
| 68 | -23.11 | 534.0721 |
| 70 | -21.11 | 445.6321 |
| 75 | -16.11 | 259.5321 |
| 75 | -16.11 | 259.5321 |
| 75 | -16.11 | 259.5321 |
| 75 | -16.11 | 259.5321 |
| 80 | -11.11 | 123.4321 |
| 86 | -5.11 | 26.1121 |
| 89 | -2.11 | 4.4521 |
| 90 | -1.11 | 1.2321 |
| 90 | -1.11 | 1.2321 |
| 97 | 5.89 | 34.6921 |
| 97 | 5.89 | 34.6921 |
| 100 | 8.89 | 79.0321 |
| 100 | 8.89 | 79.0321 |
| 100 | 8.89 | 79.0321 |
| 104 | 12.89 | 166.1521 |
| 104 | 12.89 | 166.1521 |
| 109 | 17.89 | 320.0521 |
| 113 | 21.89 | 479.1721 |
| 120 | 28.89 | 834.6321 |
| 120 | 28.89 | 834.6321 |
| 120 | 28.89 | 834.6321 |
| 122 | 30.89 | 954.1921 |
| 123 | 31.89 | 1016.9721 |
| 130 | 38.89 | 1512.4321 |
| 140 | 48.89 | 2390.2321 |
| 145 | 53.89 | 2904.1321 |
| 170 | 78.89 | 6223.6321 |
| 174 | 82.89 | 6870.7521 |
| 194 | 102.89 | 10586.3521 |
| 217 | 125.89 | 15848.2921 |
| Total |  | 61018.2914 |

Standard deviation = 2

=

= 41.75

Q3. The number of times I go to the gym in weekdays, are given below along with its

associated probability:

x = 0, 1, 2, 3, 4, 5

f(x) = 0.09, 0.15, 0.40, 0.25, 0.10, 0.01

Calculate the mean no. of workouts in a week. Also evaluate the variance involved in

it.

Ans.

|  |  |  |
| --- | --- | --- |
| x | f(x) | x\*f(x) |
| 0 | 0.09 | 0 |
| 1 | 0.15 | 0.15 |
| 2 | 0.40 | 0.8 |
| 3 | 0.25 | 0.75 |
| 4 | 0.10 | 0.4 |
| 5 | 0.01 | 0.05 |
| Total | 1 | 2.15 |

Mean no. of workouts in a week =

= 2.15

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| x | f(x) | x-µ | (x-µ)2 | f(x)\*(x-µ)2 |
| 0 | 0.09 | -2.15 | 4.6225 | 0.416025 |
| 1 | 0.15 | -1.15 | 1.3225 | 0.198375 |
| 2 | 0.40 | 0.15 | 0.0225 | 0.009 |
| 3 | 0.25 | 0.85 | 0.7225 | 0.180625 |
| 4 | 0.10 | 1.85 | 3.4225 | 0.34225 |
| 5 | 0.01 | 2.85 | 8.1225 | 0.081225 |
| Total | 1 |  | 18.2350 | 1.2275 |

Variance = = 1.2275

Q4. Let the continuous random variable D denote the diameter of the hole drilled in an

aluminum sheet. The target diameter to be achieved is 12.5mm. Random

disturbances in the process often result in inaccuracy.

Historical data shows that the distribution of D can be modelled by the PDF (𝑑) =

20𝑒−20(𝑑−12.5), 𝑑 ≥ 12.5. If a part with diameter > 12.6 mm needs to be scrapped,

what is the proportion of those parts? What is the CDF when the diameter is of 11

mm? What is your conclusion regarding the proportion of scraps?

Ans. Target Diameter = 12.5mm

P(X>12.6) = =

= -

= 0.135

Proportion of scraps is between 12.5 and 12.6 mm

= P(12.5<X<12.6)

= 1- P(X>12.6) = 1 – 0.135 = 0.865

Cumulative Density Function(CDF) when diameter = 11mm

= 0 for x < 12.5

= = 1 - for x ≥ 12.5

CDF is given by,

=

Q5. A company manufactures LED bulbs with a faulty rate of 30%. If I randomly select 6

chosen LEDs, what is the probability of having 2 faulty LEDs in my sample?

Calculate the average value of this process. Also evaluate the standard deviation

associated with it.

Ans. Probability of a faulty LED(p) = 30 /100 = 0.3

Probability of a non- faulty LED(q) = 1 - p

1 - 0.3 = 0.7

No. of randomly chosen LEDs (n) = 6

Binomial Distribution = P(X=k) = pk(1-p)n-k

P(X=2) = (0.3)2(0.7)4 = 0.324

Probability of having 2 faulty LEDs in 6 chosen LEDs = 0.324

Mean = n\*p = 6\*0.3 = 1.8

Standard deviation = n\* p(1-p) = 8\*0.3(0.7) = 1.68

Q6. Gaurav and Barakha are both preparing for entrance exams. Gaurav attempts to

solve 8 questions per day with a correction rate of 75%, while Barakha averages

around 12 questions per day with a correction rate of 45%. What is the probability

that each of them will solve 5 questions correctly? What happens in cases of 4 and 6

correct solutions? What do you infer from it? What are the two main governing

factors affecting their ability to solve questions correctly? Give a pictorial

representation of the same to validate your answer.

Ans. In case of Gaurav,

Probability of solving questions correctly(p) = 75 /100 = 0.75

Probability of solving questions incorrectly(q) = 1-p = 1 - 0.75 = 0.25

No. of question attempts in a day(n) = 8

Binomial Distribution = P(X=k) = pk(1-p)n-k

P(X=5) = (0.75)5(0.25)3  = 0.207

P(X=4) = (0.75)4(0.25)4  = 0.086

P(X=6) = (0.75)6(0.25)2  = 0.311

In case of Barakha,

Probability of solving questions correctly(p) = 45 /100 = 0.45

Probability of solving questions incorrectly(q) = 1-p = 1 - 0.45 = 0.55

No. of question attempts in a day(n) = 12

Binomial Distribution = P(X=k) = pk(1-p)n-k

P(X=5) = (0.45)5(0.55)7  = 0.222

P(X=4) = (0.45)4(0.55)8  = 0.170

P(X=6) = (0.45)6(0.55)6  = 0.212

Probability of solving 4 questions correctly is high in case of Barakha in camparison to Gaurav and

solving 6 questions correctly is high in case of Gaurav in camparison to Barakha.

Two governing factors to solve questions correctly are:

1) No. of question attempts in a day

2) Correction Rate

Barakha

Gaurav

Question attempts in a day = 8

Question attempts in a day = 12

Correction Rate = 0.75

In Correction Rate = 0.45

Correction Rate = 0.55

In Correction Rate = 0.25

P(X=5) = 0.222

P(X=4) = 0.170

P(X=6) = 0.212

0.0

P(X=5) = 0.207

P(X=4) = 0.086

P(X=6) = 0.311

Q7. Customers arrive at a rate of 72 per hour to my shop. What is the probability of 𝑘

customers arriving in 4 minutes? a) 5 customers, b) not more than 3 customers, c)

more than 3 customers. Give a pictorial representation of the same to validate your

answer.

Ans. Average no. of customers arriving per hour in a shop = 72

i.e. average no. of customers arriving in 4 min in a shop(n) = 72 \* 4/60 = 4.8

Poisson Distribution = P(X=k) = λk e-λ

k!

a) P(X=5) = 4.85 e-4.8 = 0.175

5!

b) P(X<=3) = P(X=1) + P(X=2) + P(X=3)

= 4.81 e-4.8 + 4.82 e-4.8 + 4.83 e-4.8

1! 2! 3!

= 0.03950278583 + 0.094806686 + 0.1516906976

= 0.28600016943 = ~0.286

c) P(X>3) = 1 - P(X<=3) = 1 - 0.286 = 0.714

Average no. of customers arriving per hour in a shop = 72

Average no. of customers arriving in 4 min in a shop = 4.8

P(X=5) = 0.175

P(X<=3) = 0.286

P(X=3) = 0.714

Q8. I work as a data analyst in Aeon Learning Pvt. Ltd. After analyzing data, I make

reports, where I have the efficiency of entering 77 words per minute with 6 errors per

hour. What is the probability that I will commit 2 errors in a 455-word financial report?

What happens when the no. of words increases/decreases (in case of 1000 words,

255 words)?

How is the 𝜆 affected?

How does it influence the PMF?

Give a pictorial representation of the same to validate your answer.

Ans. No. of words to be efficiently entered per minute = 77

No. of errors committed per hour = 6

No. of errors committed per minute = 6/60 = 0.1

We have to find out the probability that 2 errors will be committed in a 455-word financial report

Average no. of errors committed in a 455-word financial report(λ) =

455\*0.1 /77 = 0.59

Poisson Distribution = P(X=k) = λk e-λ

k!

P(X=2) = 0.592 e-0.59 = 0.096

2!

If there are 1000-word financial report-

Average no. of errors committed in a 1000-word financial report(λ) =

1000\*0.1 /77 = 1.29

P(X=2) = 1.292 e-1.29 = 0.229

2!

λ has increased once we have increased the no. of pages

PMF also increased

If there are 255-word financial report-

Average no. of errors committed in a 255-word financial report(λ) =

255\*0.1 /77 = 0.33

P(X=2) = 0.332 e-0.33 = 0.039

2!

λ has decreased once we have decreased the no. of pages

PMF also decreased

λ and PMF has direct relationship as one increases the other also increases and vice-versa

Average no. of errors committed in a 455-word financial report(λ) = 0.59

PMF = 0.096

Average no. of errors committed in a 1000-word financial report(λ) = 1.29

PMF = 0.229

Average no. of errors committed in a 255-word financial report(λ) = 0.33

PMF = 0.039

Q9. Please compute the following:

a) P(Z > 1.26), P(Z < -0.86), P(Z > -1.37), P(-1.25 < Z < 0.37), P(Z ≤ -4.6)

b) Find the value 𝑧 such that 𝑃(𝑍 > 𝑧) = 0.05

c) Find the value of 𝑧 such that 𝑃(-𝑧 < 𝑍 < 𝑧) = 0.99

Ans. a) P(Z > 1.26) = 1 - P(Z <= 1.26) = 1 - 0.8962 = 0.1038 = 10.38%

P(Z < -0.86) = 0.1949 = 19.49%

P(Z > -1.37) = 1 - P(Z <= -1.37) = 1 - 0.0853 = 0.9147 = 91.47%

P(-1.25 < Z < 0.37) = P(Z <= 0.37) - P(Z <= -1.25) = 0.6443 - 0.1056 = 0.5387 = 53.87%

P(Z <= -4.6) = 0 = 0%

b) P(Z > z) = 0.05

Value of z = -1.64

c) P(-z < Z < z) = 0.99

P(-z < Z < z) = P(Z <= z) – P(Z <= -z)

Value of z = 3

Q10. The current flow in a copper wire follow a normal distribution with a mean of 10 𝑚A

and a variance of 4 (𝑚𝐴)2.

What is the probability that a current measurement will exceed 13 𝑚𝐴? What is the

probability that a current measurement is between 9 and 11mA? Determine the

current measurement which has a probability of 0.98.

Ans. Mean(µ) = 10

Variance = 4

Standard deviation(σ) = = 2

P(X > 13)

z score of (X > 13) = X - µ /σ = 13 – 10 /2 = 1.5

P(Z > 1.5) = 0.9332 = 93.32%

P(9 < X < 11) = P(X <= 11) - P(X <=9)

P(X <= 11)

z score of (X <= 11) = X - µ /σ = 11 – 10 /2 = 0.5

P(Z < 0.5) = 0.6915 = 69.15%

P(X <= 9)

z score of (X <= 9) = X - µ /σ = 9 – 10 /2 = -0.5

P(Z > 0.5) = 0.3085 = 30.85%

P(9 <= X <= 11) = 69.15 – 30.85 = 38.3%

P(Z < z) = 0.98

Value of current measurement(z) = 2.06

Q11. The shaft in a piston has its diameter normally distributed with a mean of 0.2508 inch

and a standard deviation of 0.0005 inch. The specifications of the shaft are 0.2500 ∓

0.0015 inch. What proportion of shafts are in sync with the specifications? If the

process is centered so that the mean is equal to the target value of 0.2500, what

proportion of shafts conform to the new specifications? What is your conclusion from

this experiment?

Ans. Mean(µ) = 0.2508

Standard deviation (σ) = 0.0005

P(0.2485 <= X <= 0.2515)

= P(X <= 0.2515) - P(X <= 0.2485)

P(X <= 0.2515)

z score of (X <= 0.2515) = X - µ /σ = 0.2515 – 0.2508 /0.0005 = 1.4

P(Z < 1.4) = 0.9192 = 91.92%

P(X <= 0.2485)

z score of (X <= 0.2485) = X - µ /σ = 0.2485 – 0.2508 /0.0005 = -4.6

P(Z < -4.6) = 0 = 0%

P(0.2485 <= X <= 0.2515) = 0.9192 – 0 = 0.9192 = 91.92%

i.e. 91.92% of the shafts confirm to the specifications

If Mean(µ) = 0.2500

P(X <= 0.2515)

z score of (X <= 0.2515) = X - µ /σ = 0.2515 – 0.2500 /0.0005 = 3

P(Z < 1.4) = 1 = 100%

P(X <= 0.2485)

z score of (X <= 0.2485) = X - µ /σ = 0.2485 – 0.2500 /0.0005 = -3

P(Z < -4.6) = 0 = 0%

P(0.2485 <= X <= 0.2515) = 1 – 0 = 1 = 100%

i.e. 100% of the shafts confirm to the specifications in this case