Calculate the mean, median, mode and standard deviation for the problem statements 1& 2.

**Problem Statement 1:**

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

| x | No:of Occurences | Mean | median | Mode | variance | |
|---|---|---|---|---|---|---|
| 4 | 2 | 137/20 = 6.85 | 7 | 7 | 16+16=32 | |
| 5 | 2 | | | | 25+25=50 | |
| 6 | 4 | | | | 36+36+36+36=144 | |
| 7 | 5 | | | | 49(5)=245 | |
| 8 | 4 | | | | 64(4)=256 | |
| 9 | 2 | | | | 81(2)=163 | |
| 10 | 1 | | | | 100 | |
| | | | | | 989/20=49.45 | 49.45-46.92=2.53 |

1) Arrange the given value in Ascending Order

2) **Mean** = ∑x/total no of occurrences (n) ➔ 137/20=6.85

3) **Median** = Middle value = 7

4) **Mode** ➔ based on repetition of each number; as 7 occurred more number of times == > 7

To calculate Standard deviation:

Before calculating Standard deviation, we need to calculate variance

Arithmetic mean = 7 subtract from original marks

**Variance** = σ^2 = 2.53

**Standard Deviation** =σ = √variance = 1.59

**import scipy**

**from scipy import stats**

**import numpy as np**

stats.mode(arr)

```
>>>ModeResult(mode=array([7]), count=array([5]))

np.mean(arr)
>>>6.85
np.median(arr)
>>>7.0
np.var(arr)
>>>2.5275000000000003
np.std(arr)
>>>1.5898113095584647
```

**Problem Statement 2:**

The number of calls from motorists per day for roadside service was recorded for a particular month:

Q2

| x | No:of Occurences | Mean | median | Mode | | Variance |
|---|---|---|---|---|---|---|
| 28 | 1 | 3763/35= 107.51 | 100 | 75 | 107.51^2 = 11558.4001 | 784 |
| 40 | 1 | | | | | 1600 |
| 68 | 1 | | | | | 4624 |
| 70 | 1 | | | | | 4900 |
| 75 | 4 | | | | | 5625 |
| 75 | | | | | | 5625 |
| 75 | | | | | | 5625 |
| 75 | | | | | | 5625 |
| 80 | 1 | | | | | 6400 |
| 86 | 1 | | | | | 7,396 |
| 89 | 1 | | | | | 7,921 |
| 90 | 2 | | | | | 8100 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 90 | | | | | | 8100 |
| 97 | 2 | | | | | 9409 |
| 97 | | | | | | 9409 |
| 100 | 3 | | | | | 10000 |
| 100 | | | | | | 10000 |
| 100 | | | | | | 10000 |
| 104 | 2 | | | | | 10816 |
| 104 | | | | | | 10816 |
| 109 | 1 | | | | | 11881 |
| 113 | 1 | | | | | 12769 |
| 120 | 3 | | | | | 14400 |
| 120 | | | | | | 14400 |
| 120 | | | | | | 14400 |
| 122 | 1 | | | | | 14884 |
| 123 | 2 | | | | | 15129 |
| 123 | | | | | | 15129 |
| 130 | 1 | | | | | 16900 |
| 140 | 1 | | | | | 19600 |
| 145 | 1 | | | | | 21025 |
| 170 | 1 | | | | | 28900 |
| 174 | 1 | | | | | 30276 |
| 194 | 1 | | | | | 37636 |
| 217 | 1 | | | | | 47089 |
| | | | | | | =>(457193/35)=13062.65 |

**Mean**: 107.51

**Median**: 100

**Mode**: 75

**Variance**: 13062.65 – (107.51)^2 =. 1503.514

**Standard deviation**: √1503.514 = 38.77

arr2 = [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]

**stats.mode(arr2)**

```
>>>ModeResult(mode=array([75]), count=array([4]))
```

**np.mean(arr2)**

```
>>> 107.51428571428572
```

**np.median(arr2)**

```
>>>100.0
```

**np.var(arr2)**

```
>>>1503.3355102040816
```

**np.std(arr2)**

```
>>>38.77287080168403
```

## Problem Statement 3:

The numbers 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.

| x | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| F(x) | 0.09 | 0.15 | 0.40 | 0.25 | 0.10 | 0.01 |

X P

0 * 0.09 = 0

1*0.15 = 0.15

2 * 0.40 = 0.80

3 * 0.25 = 0.75

4 * 0.10 = 0.40

5 * 0.01 = 0.05

Mean = 0+0.15+0.80+0.75+0.40+0.05 ➔ 2.15

Mean = $\mu$ = $\Sigma xp = 2.15$

Variance == $Var(X) = \Sigma x^2 p - \mu^2$

| | | | | | | |
|---|---|---|---|---|---|---|
| x | 0 | 1 | 2 | 3 | 4 | 5 |
| X^2 | 0 | 1 | 4 | 9 | 16 | 25 |
| F(x) | 0.09 | 0.15 | 0.40 | 0.25 | 0.10 | 0.01 |
| X^2p | 0 | 0.15 | 1.60 | 2.25 | 1.6 | 0.25 |

$\mu$^2 = 4.6225

$\Sigma x^2 p = 5.85$

Variance == $Var(X) = \Sigma x^2 p - \mu^2$

$var(x) = 1.2275$

**Problem Statement 5:**

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.
Answer :

**Solution:** This is a **binomial** distribution because the reason is that there are only two outcomes (the LED works or not).

N = 6

P(getting a defective bulb) = 30/100 = 0.3
Q(getting a working bulb) = 1-0.3 = 0.7

Suppose X has a bionomial distribution, the probability of x success in n-Bernoulli trials,

Here in this case getting a defective bulb is success,

$$P(X=x)={}_nC_x * p^x * q^{n-x}$$

where in x=0,1,2,...,n some finite number of required successes out of some finite number of trials(n)
N = 6 trials
N = 2 defective bulbs
6C2 * (0.3)^2*(0.7)^4 ➔ 0.32 ==32%

**Hence probability of getting 2 defective bulb is 32%**

If *X* has a binomial distribution with *n* trials and probability of success *p* on each trial, then:

1. The mean of *X* is

   $\mu = np.$

   Mean = 6 * 0.32 = 1.92
2. The variance of *X* is

   $\sigma^2 = np(1-p).$ == 6 * 0.32(1-0.32) ➔ 1.3056
3. The standard deviation of *X* is

   $\sigma = \sqrt{np(1-p)}.$ ➔ √1.3056 ➔1.14

**Problem Statement – 6:**

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.

**Gaurav :** N = 8,p=0.75 q=0.25

P(X=5) = 8c5(0.75)^5(0.25)^3➔0.2072

P(X=4) = 8c4(0.75)^4(0.25)^4 ➔ 70 * 0.3154 * 0.0039 == 0.0862

P(x=6) == 8c6(0.75)^6(0.25)^2➔ 28 * 0.1779 * 0.0625 = 0.31

Chances of solving more number of correct questions is higher when comparing to that of solving less correct solutions

**Bhargav :** N=12 p = 0.45 q=0.55

P(X=5) = 12c5(0.45)^5(0.55)^7 ==>792 * 0.01845 * 0.01522 = > 0.22

P(X=4) = 12c4 (0. 45)^4 (0. 55)^8 ➔ 495 * 0.041 * 0.008373 == 0.1699

P(x=6) == 12c6(0. 45)^6 (0. 55)^6 ➔ 924 * 0.008373 * 0.0276 = 0.22

Chances of solving more number of correct questions is higher when comparing to that of solving less correct solutions

## Problem Statement 7:

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.

## Solution: This uses Poisson distribution

No: of customers coming to my shop in 1 hour ➔ 72

μ = 72

for 1 min = 1.2 approx = 1

for 4 min == approx. **5** customers

**Exactly 5 customers for 4 min**

P(X = 5) = e^-5 (5)^5/5! = (0.0067* 3125)/5! == (20.9375)/120 == 0.17 %

**B) Not more than 3 customers**

P(X<=3) = p(x=0) + p(x=1)+p(x=2)+p(x=3)

=e^-5 * 5^0 /0! + e^-5 * 5^1 /1! + e^-5 * 5^2 /2! + e^-5 * 5^3 /3!

= 0.0067 + 0.0067*5 + (0.0067 * 25) / 2 + (0.0067 * 125)/3

= 0.0067 + 0.0335 + 0.08375 + 0.2791

= 0.40305

**c) More than 3 customers ➔** 1- not more than 3 customers

= 1- 0.40305 ➔ 0.59695

## Problem Statement: 8

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 $\lambda$ affected?

How does it influence the PMF?

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

# Solution:

77 per minute== per hour ➔ 77 * 60 = 4620 == 6 errors per hour

Reduce the error per minute = 1 error per minute == 77 words with 1 error

77* 5.9 == 455 words

Probability Per 455 words

(E ^ -6 * 6^2)* 2! ➔( 0.0024 * 36)/2 ➔ 0.0432

Per 1000 words :

λ =Total words/Error

when the no. of words increases/decreases (in case of 1000 words, 255 words)?

For 1000 words ==== 4620 words – 6 error

1000 words - ?

(1000 * 6)/4620 ➔ 1.3 errors on an averge

For 255 words ==== 4620 words – 6 error

255 words - ?

(255 * 6)/4620 ➔ 0.33 errors on an average

**Problem Statement 10:**

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)

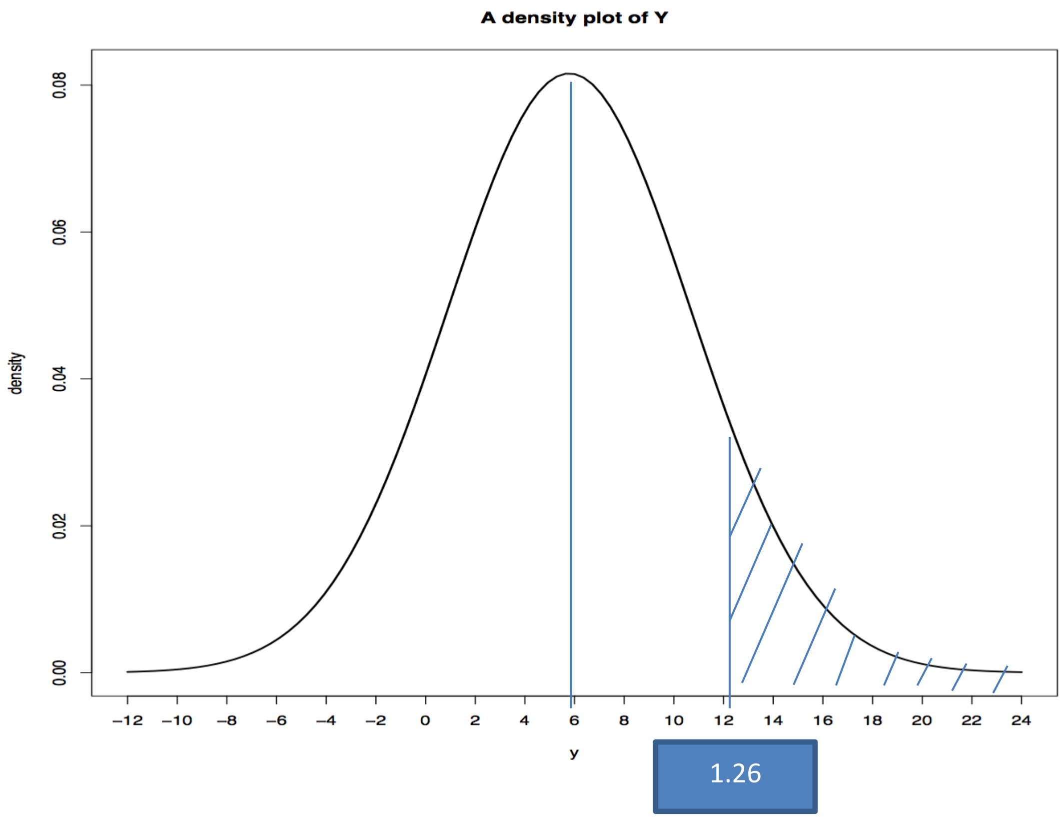
b) Find the value $z$ such that $P(Z > z) = 0.05$

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

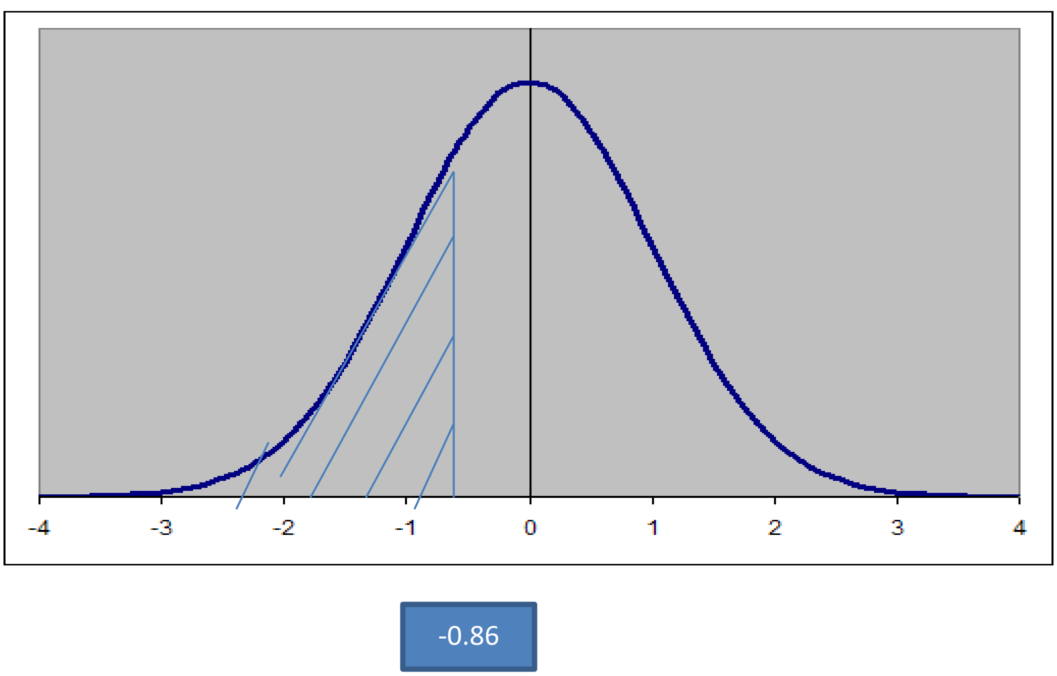
P(Z > 1.26) = 1- P(Z < 1.26)

= 1 – 0.89617

= 0.10383



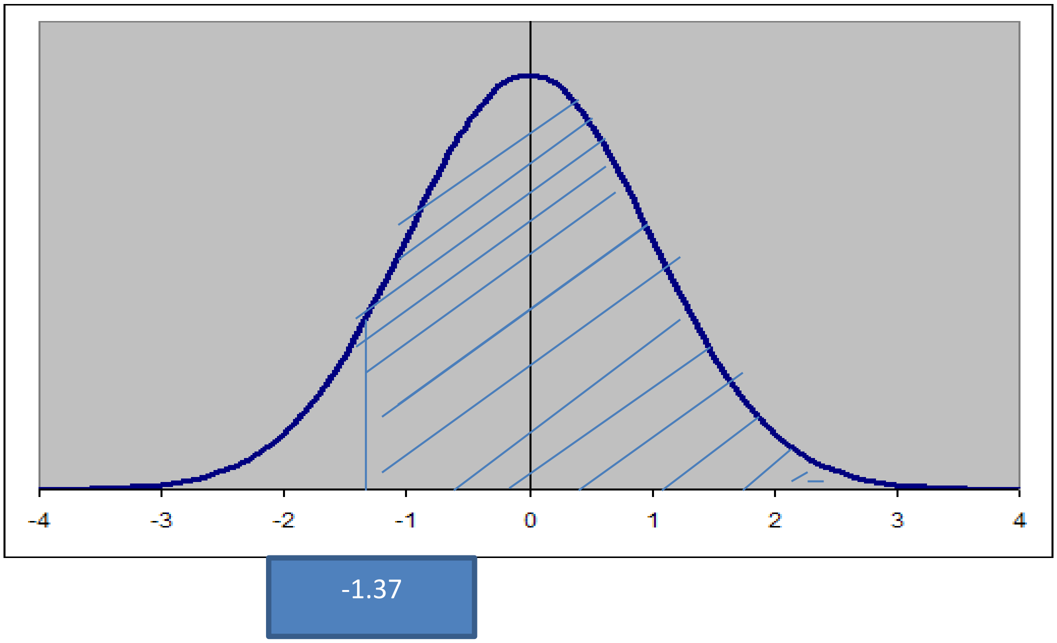

P(Z < −0.86) = 0.19489



P(Z > −1.37) = 1 - P(Z < −1.37)

= 1 – 0.08534

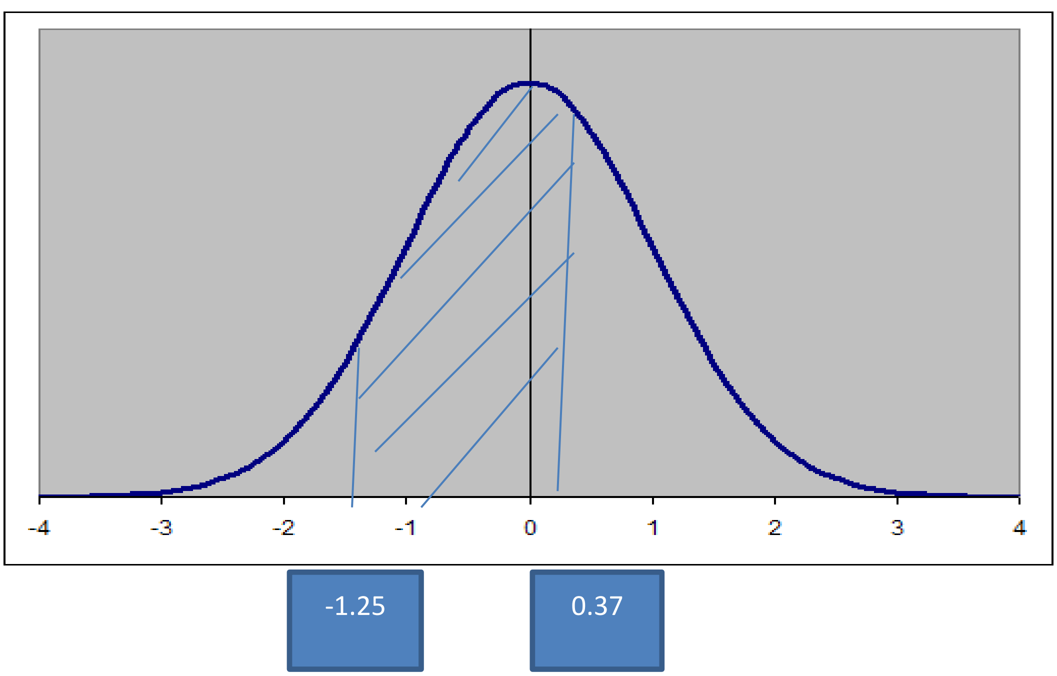
= 0.91466




P(−1.25 < Z < 0.37) = p(z<0.37)-p(-1.25)

= 0.64431-0.10565

= 0.53866



**B )** Find the value $z$ such that $P(Z > z) = 0.05$

## This is One sided case

The z score of 0.05 = (1- 0.05) = P(Z =0.95) = 1.64

**C )** Find the value of $z$ such that $P(-z < Z < z) = 0.99$

# This is Two sided case

1-0.99 = 0.01

= 0.01/2 ➔ 0.005

From normal distribution table, we found that P(Z<z)=0.995

Thus z 0.005 = 2.54

**Problem Statement 11:**

The current flow in a copper wire follow a normal distribution with a mean of 10 $m$A 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.

Let X denote the current in mA μ = 10 mA
σ = 2 so the probability is given by

P(X>13) = P((X – 10)/2 > (13 -10)/2) = P(Z > 1.5) = 1 – P(Z<= 1.5)
= 1 – 0.9332
= 0.0668

P( 9 <X < 11) = P ((9-10)/2 - (X – 10)/2 - (11 – 10)/2) = P(-0.5 < Z < 05)
= P(Z < 0.5) – P(Z<= -0.5)
= 0.6915 – 0.3085
= 0.383

b) We want to find the value of x such that P(X = x) = 0.98
P ((X – 10)/2 = (x – 10)/2) = P(Z = (x – 10)/2) = 0.98
P (Z = 2.05) = 0.98
(X-10)/2 = 2.05 ➔ 14.1 mA

## Problem Statement 12

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 is 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 conforms to the new specifications? What is your conclusion from this experiment?

## Solution:

$Z = (x - \mu)/\sigma$
$= (x - 0.2508) / 0.0005$
Target = 0.2500

Initially
$P((0.2485 – 0.2508)/0.0005 < Z < (0.2515 – 0.2508)/0.0005) = P(-4.6 < Z < 1.4) =$
$P(Z<1.4) – P(Z<-4.6) = 0.91924$
Lower Limit => 0.2500 – 0.0015 = 0.2485
Upper limit => 0.2500 + 0.0015 = 0.2515

$P(0.2485 < Z < 0.2515) = P((0.2485 – 0.2500)/0.0005 < Z < (0.2515 – 0.2500)/0.0005)$

$\Rightarrow P(-3 < Z < 3) = 0.9973$

It would be increased from 92% from 99.73%