Q 1) Calculate probability from the given dataset for the below cases

Data \_set: Cars.csv

Calculate the probability of MPG of Cars for the below cases.

MPG <- Cars$MPG

* 1. P(MPG>38)
  2. P(MPG<40)
  3. P (20<MPG<50)

Ans) > mean(Cars$MPG)

[1] 34.42208

P(MPG>38):

> sd(Cars$MPG)

[1] 9.131445

> pnorm(38,34.42,9.13)

[1] 0.652513

P(MPG>38)=1-P(MPG<38)(PS: Z-table gives you only less than probabilities)

>1 - 0.65

[1] 0.35

P(MPG<40):

pnorm(40,34.42,9.13)

[1] 0.7294571

P (20<MPG<50):

> pnorm(50,34.42,9.13)-pnorm(20,34.42,9.13)

[1] 0.8989178

Q 2) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution

Dataset: Cars.csv

Ans) Follows Normal distribution as indicated by qq-plot.



1. Check Whether the Adipose Tissue (AT) and Waist Circumference(Waist) from wc-at data set follows Normal Distribution

Dataset: wc-at.csv

Ans) waist follows Normal Distribution from the below QQ-plot

> qqnorm(wc\_at$Waist)

> qqline(wc\_at$Waist)



Adipose Tissue follows normal distribution

qqnorm(wc\_at$AT)

qqline(wc\_at$AT) 

Q 3) Calculate the Z scores of 90% confidence interval,94% confidence interval, 60% confidence interval

Ans) 90% 🡪> qnorm(0.95)

[1] 1.644854

94% 🡪> qnorm(0.97)

[1] 1.880794

60% 🡪> qnorm(0.8)

[1] 0.8416212

Q 4) Calculate the t scores of 95% confidence interval, 96% confidence interval, 99% confidence interval for sample size of 25

Ans) 95% 🡪> qt(0.975,24)

[1] 2.063899

96% 🡪> qt(0.98,24)

[1] 2.171545

99% 🡪 qt(0.995,24)

[1] 2.79694

Q 5**)** A Government company claims that an average light bulb lasts 270 days. A researcher randomly selects 18 bulbs for testing. The sampled bulbs last an average of 260 days, with a standard deviation of 90 days. If the CEO's claim were true, what is the probability that 18 randomly selected bulbs would have an average life of no more than 260 days?

Hint:

Rcode → pt (tscore, df)

df → degrees of freedom

Ans) µ=270, =260, SD=90, n=18, df=n-1=18-1= 17

tscore= = = -10/21.23= -0.47

> pt(-0.47,17)

[1] 0.3221639

Required probability = 0.32=32%

1. The time required for servicing transmissions is normally distributed with *μ* = 45 minutes and *σ* = 8 minutes. The service manager plans to have work begin on the transmission of a customer’s car 10 minutes after the car is dropped off and the customer is told that the car will be ready within 1 hour from drop-off. What is the probability that the service manager cannot meet his commitment?
2. 0.3875
3. 0.2676
4. 0.5
5. 0.6987

Ans) B

let time taken for service transmission= T

T is normally distributed with *μ* = 45 minutes and standard deviation *σ* = 8 minutes.

Time delay= 10 minutes

Time available to finish the work= 60-10=50 minutes.

Therefore from the equation Z=(T-µ)/ *σ*

P(T≤50)=p(Z≤(50-45)/8)=p(Z≤0.625)= 0.7324(using z table)

Therefore p(T>50)=1-p(≤50)= 1-0.7324= 0.2676

(Or)

Using R-function : [1-pnorm(50,45,8)]=

|  |
| --- |
| > 1-pnorm(50,45,8)  [1] 0.2659855 |
|  |
| |  | | --- | |  | |

1. The current age (in years) of 400 clerical employees at an insurance claims processing center is normally distributed with mean *μ* = 38 and Standard deviation *σ* =6. For each statement below, please specify True/False. If false, briefly explain why.
2. More employees at the processing center are older than 44 than between 38 and 44.

Ans) False.

• 68% of the data falls within one standard deviation of the mean (µ+*σ)*.

Here µ=38, *σ* =6

Then, µ+*σ= 38+*6=44

1. A training program for employees under the age of 30 at the center would be expected to attract about 36 employees.

Ans) True

Z=(X-µ)/ *σ*

P(X≤30)=p(Z≤(30-38)/6)=p(Z≤-1.33)= 0.0918(using z table)

Expected count=0.0918\*400= 36.72

1. If *X1*~ *N*(μ, σ2) and *X*2 ~ *N*(μ, σ2) are *iid*normal random variables, then what is the difference between 2 *X*1 and *X*1 + *X*2? Discuss both their distributions and parameters.

Ans) 2 is simply a larger scale version of the random variable *X1.* If is normally distributed then 2X1 is also normally distributed.

*X*1 and *X*2 are normal distributed, the associated sums and random samples are exactly (and not just approximately) normal, with the appropriate parameters.

1. Let X ~ N(100, 202). Find two values, *a* and *b*, symmetric about the mean, such that the probability of the random variable taking a value between them is 0.99.
2. 90.5, 105.9
3. 80.2, 119.8
4. 22, 78
5. 48.5, 151.5
6. 90.1, 109.9

Ans) D

Here we need range of 99% data which lies between 3rd standard deviation of the mean.

Here µ=100, *σ* =20

From empirical rule, µ±3*σ= 100±3\*20=>(100-60, 100+60)=>(40,160).*

1. Consider a company that has two different divisions. The annual profits from the two divisions are independent and have distributions Profit1 ~ N(5, 32) and Profit2 ~ N(7, 42) respectively. Both the profits are in $ Million. Answer the following questions about the total profit of the company in Rupees. Assume that $1 = Rs. 45
2. Specify a Rupee range (centered on the mean) such that it contains 95% probability for the annual profit of the company.
3. Specify the 5th percentile of profit (in Rupees) for the company
4. Which of the two divisions has a larger probability of making a loss in a given year?

Ans) let, X is the sum of two random variables having normal distribution.

E[X]= E[45\*(profit 1+profit 2)]= 45\*(5+7)=540 million rupees

SD[X]= SD[profit 1 +profit 2]=> 45\*()

= 45\*= 225 million rupees.

**Therefore, X~ N(540,)**

1. From the empirical rule, Approximately 95% of the data falls within two standard deviation of the mean.

μ ± 2σ = 540±2\*225=> (540-450, 540+450)=> **(90,990)**

**B)**



From the above normal distribution we can say that to find 5th percentile from the left side we can use the formula,

μ - 1.5σ => 540-(1.5\*225) =>202.5 million rupees.

c) this question concerns the original profit distributions.

For division1= Z score for a profit of zero: Z=(X-µ)/ *σ =>*  (0-5)/3 => -1.66=0.0485

(or)

> pnorm(0,5,3)

[1] 0.04779035

For division2= Z score for a profit of zero: Z=(X-µ)/ *σ*  =(0-7)/4 => -1.75= .0401

> pnorm(0,7,4)

[1] 0.04005916

Division2 has a higher probability of making a loss.