# Basic Statistics (Module – 4 (Part – 1))

**Name: Shahina Athar**

**Btach No: 10122020**

Q1) 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) = **40.74%**
2. P(MPG<40) = **75.30%**
3. P (20<MPG<50) = **2.46%**

Q2) Check whether the data follows normal distribution

1. Check whether the MPG of Cars follows Normal Distribution Dataset: Cars.csv

**Ans:** Yes MPG of Cars follows Normal Distribution

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 Circumference (Waist) from wc-at data set not follows Normal Distribution.

Adipose Tissue (AT) from wc-at data set follows Normal Distribution.

Q3) Calculate the Z scores of 90% confidence interval,94% confidence interval, 60% confidence interval

**Ans:** Z(90%) = Probability = (0.05-0.95) => z = +-1.645

Z(94%) = Probability = (0.03-0.97) => z = +- 1.88

Z(60%) = Probability = (0.2-0.80) => z = +- 0.841

Q4) Calculate the t scores of 95% confidence interval, 96% confidence interval, 99% confidence interval for sample size of 25

**Ans:** t(95%) = Probability = (0.025-0.975) => z = +-2.063

t(96%) = Probability = (0.02-0.98) => z = +-2.171

t(99%) = Probability = (0.005-0.995) => z = +-2.796

Q5**)** 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:** Given: μ = 270; n = 18;  x̄ = 260; s = 90

First of all we have to calculate t value:

Formula:

t = x̄- μ / s/ √ n

t = 260-270/90/(18)^(1/2)

t = -0.471

With this t- value we find Probability using t-table:

32% of probability that 18 randomly selected bulbs would have an average life of no more than 260 days.

Q6) The time required for servicing transmissions is normally distributed with  = 45 minutes and  = 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?

A. 0.3875

B. 0.2676 **- Ans**

C. 0.5

D. 0.6987

Ans: given: μ = 45; σ =8; x = 10

The work begin after 10 min, so the average time increase from 45min to 55min.

for normal distribution :-

z = (X-μ)/б

= (60-55)/8

= 0.625

As we want to find the probability of service manager cannot meet his commitment, So we should write below command.

1-pnorm(0.625)=0.2659



Q7) The current age (in years) of 400 clerical employees at an insurance claims processing center is normally distributed with mean  = 38 and Standard deviation

 =6. For each statement below, please specify True/False. If false, briefly explain why.

1. More employees at the processing center are older than 44 than between 38 and 44.
2. A training program for employees under the age of 30 at the center would be expected to attract about 36 employees.

**Ans:**

Mean = 38

SD = 6

Z score = (Value - Mean)/SD

Z score for 44  = (44 - 38)/6  = 1  =>  84.13 %  (using z-table)

=> People above 44 age = 100 - 84.13 =  15.87% ,

=>15.87\*400/100  ≈  63  out of 400

Z score for 38  = (38 - 38)/6 = 0 => 50%

Hence People between 38 & 44  age = 84.13 - 50 = 34.13 %

=>34.13\*400/100 ≈  137 out of 400

Hence More employees at the processing center are older than 44 than between 38 and 44. is F**ALSE**

Z score for 30  = (30 - 38)/6 =  -1.33  =  9.15  %

=>9.15\*400/100   ≈ 36 out of 400

Hence A training program for employees under the age of 30 at the center would be expected to attract about 36 employees - **TRUE**

Q8) If X1 ~ N(μ, σ2) and X2 ~ N(μ, σ2) are iid normal random variables, then what is the

difference between 2 X1 and X1 + X2? Discuss both their distributions and parameters.

**Ans:** As we know that if X ∼ N(µ1, σ1^2 ), and Y ∼ N(µ2, σ2^2 ) are two independent random variables then X + Y ∼ N(µ1 + µ2, σ1^2 + σ2^2 ) ,

and X − Y ∼ N(µ1 − µ2, σ1^2 + σ2^2 ).Similarly if Z = aX + bY , where X and Y are as defined above, i.e Z is linear combination of X and Y ,

then Z ∼ N(aµ1 + bµ2, a^2σ1^2 + b^2σ2^2 ).

Therefore in this question

|  |
| --- |
| 2X1~ N(2 u,4 σ^2) and |
| X1+X2 ~ N(µ + µ, σ^2 + σ^2 ) ~ N(2 u, 2σ^2 ) |
| 2X1-(X1+X2) = N( 4µ,6 σ^2) |

Q9) Let X ~ N(100, 20^2) its (100, 20 square).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.

A.

90.5, 105.9

B. 80.2, 119.8 C.

22, 78

D. 48.5, 151.5 **- Ans**

E. 90.1, 109.9

**Ans:** μ = 100, б = 20

Since we need to find out the values of a and b, which are symmetric about the mean, such that the probability of random variable taking a value between them is 0.99, we have to work out in reverse order. The Probability of getting value between a and b should be 0.99. So the Probability of going wrong, or the Probability outside the a and b area is 0.01 (ie. 1-0.99).

The Probability towards left from a = -0.005 (ie. 0.01/2).

The Probability towards right from b = +0.005 (ie. 0.01/2).

So since we have the probabilities of a and b, we need to calculate X, the random variable at a and b which has got these probabilities.

By finding the Standard Normal Variable Z (Z Value), we can calculate the X values.

Z=(X- μ) / σ

For Probability 0.005 the Z Value is -2.57 (from Z Table).

Z \* σ + μ = X

Z(-0.005)\*20+100 = -(-2.57)\*20+100 = 151.4

Z(+0.005)\*20+100 = (-2.57)\*20+100 = 48.6

Q10) Consider a company that has two different divisions. The annual profits from the two divisions are independent and have distributions Profit1 ~ N(5, 3^2) and Profit2 ~ N(7, 4^2) 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

1. Specify a Rupee range (centered on the mean) such that it contains 95% probability for the annual profit of the company.
2. Specify the 5th percentile of profit (in Rupees) for the company
3. Which of the two divisions has a larger probability of making a loss in a given year?

**Hints:**

1. Business Problem
   1. Objective
   2. Constraints (if any)
2. For each assignment the solution should be submitted in the below format
3. Research and Perform all possible steps for obtaining solution
4. For Basic Statistics explanation of the solutions should be documented in black and white along with the codes.

One must follow these guidelines as well:

* 1. Be thorough with the concepts of Probability, Central Limit Theorem and Perform the calculation stepwise
  2. For True/False Questions, explanation is must.
  3. R & Python code for Univariate Analysis (histogram, box plot, bar plots etc.) for data distribution to be attached

1. All the codes (executable programs) should execute without errors