# **STATISTICS MAJOR ASSIGNMENT SOLUTIONS**

**1.According to a study, the daily average time spent by a user on a**

**social media website is 50 minutes. To test the claim of this study,**

**Ramesh, a researcher, takes a sample of 25 website users and finds**

**out that the mean time spent by the sample users is 60 minutes and**

**the sample standard deviation is 30 minutes.**

**Based on this information, the null and the alternative hypotheses**

**will be:**

**Ho = The average time spent by the users is 50 minutes**

**H1 = The average time spent by the users is not 50 minutes**

**Use a 5% significance level to test this hypothesis.**

**Solution:**

Sample of website users (n)= 25

Mean time spent = 60 minutes

Sample standard deviation (SD)= 30 minutes

Ho = The average time spent by the users is 50 minutes

H1 = The average time spent by the users is not 50 minutes

|  |  |
| --- | --- |
| MEAN | 60 |
| STANDARD DEVIATION | 30 |
| COUNT | 25 |
| STANDARD ERROR MEAN (SEM) | 6 |
| DEGREE OF FREEDOM (DF) | 24 |
| HYPOTHESIZED MEAN | 50 |
| T- TEST | 1.666667 |
| V-VALUE | 0.05429 |

Since the p-value is more than the significance level which is 0.20, we have to accept the null hypothesis. So the average time spent by the users is 50 minutes.

**2. Height of 7 students (in cm) is given below. What is the median? 168 170 169 160 162 164 162**.

**SOLUTION:**

The height of 7 students are

168 170 169 162 164 162

Arrange them in ascending order

=160 162 162 164 168 169 170

=median =164(middle value)

**3. Below are the observations of the marks of a student. Find the value of mode. 84 85 89 92 93 89 87 89 92**

**SOLUTION:**

84 85 89 92 93 89 87 89 92

89 appeared 3 times

So the mode is 89

4. **From the table given below, what is the mean of marks obtained by 20 students?**

**SOLUTION:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Xi** | **Fi** | **XiFi** |
|  | **3** | **1** | **3** |
|  | **4** | **2** | **8** |
|  | **5** | **2** | **10** |
|  | **6** | **4** | **24** |
|  | **7** | **5** | **35** |
|  | **8** | **3** | **24** |
|  | **9** | **2** | **18** |
|  | **10** | **1** | **10** |
| **TOTAL** | **52** | **20** | **132** |

(MEAN) =

=

=6.6

**5. For a certain type of computer, the length of time between charges of the battery is normally distributed with a mean of 50 hours and a standard deviation of 15 hours. John owns one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours.**

**Solution:**

(Mean) = 50 hours

(Standard deviation) = 15 hours

Finding the probability

P(50 < X < 70) = P(X < 70) - P(X < 50)

Converting the problem in standard form

Converting to Z, we get P(Z < (70 - 50)/15 ) - P( Z < (50 - 50)/15 )

= P( Z < 1.33) - P( Z < 0) = 0.0.9088 - 0.5 = 0.4088

**6.Find the range of the following. g = [10, 23, 12, 21, 14, 17, 16, 11, 15, 19]**

**Solution:**

Range = max(x) -min(x)

=23-10

=13

**7. It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails, and the probability for a false positive (a non-spam 2 email detected as spam) is 5%. Now if an email is detected as spam, then what is the probability that it is in fact a non-spam email?**

**Solution:**

A = event that an email is detected as spam,

B = event that an email is a spam,

BC = event that an email is not spam

We know P(B) = P(BC) =0.5

P(A/B) = 0.99

P(A/BC) =0.05

Hence by Baye’s formula we have

P(BC/A) =

=

=

=0.0480

**8. Given the following distribution of returns, determine the lower quartile: {10 25 12 21 19 17 16 11 15 19}**

**Solution:**

First arrange them in ascending order

10 11 12 15 16 17 19 19 21 25

N=10

Q1= ¼(n+1)th value

=1/4(10+1)

=11/4

=2.75

2nd term + 0.75(3rd term – 2nd term)

=11+0.75(12-11)

=11.75(ans)

**9. For a Binomial distribution, the number of trials(n) is 25, and the probability of success is 0.3. What’s the variability of the distribution?**

**Solution:**

n=25

Mean=np=25\*0.3=6 (where p is the probability of success)

q=1-p (where q is the probability of failure)

1-0.3=0.7

Variance =n\*p\*q=25\*0.3\*0.7=5.25

Standard deviation= square root of 5.25= 2.29129

**10. Amy has two bags. Bag-I has 7 red and 2 blue balls and Bag-II has 5 red and 9 blue balls. Amy draws a ball at random and it turns out to be red. Determine the probability that the ball was from the Bag-I using the Bayes theorem**

**SOLUTION:**

Let X and Y be the events that the ball is from the bag I and bag II, respectively.

A to be the event of drawing a red ball.

probability of choosing a bag for drawing a ball is 1/2  
P(X) = P(Y) = 1/2

P(drawing a red ball from the bag I) = P(A|X) = 7/11

P(drawing a red ball from bag II) = P(A|Y) = 5/14

Using bayes theorem,

P(X/A) =

=

= =0.64

Hence, the probability that the ball is drawn is from bag I is 0.64

**11.Find the mean, mode and median of g = [10, 23, 12, 21, 14, 17, 16, 11, 15, 19, 12]**

**SOLUTION:**

First arrange the data in ascending order,

10 11 12 12 14 15 16 17 19 21 27

Mean =

=170/11 =15.4545

Median =15(middle value)

Mode =12 (occurred 2 times)

**12. The mean height of a random sample of 100 individuals from a population is 160. The Standard deviation of the sample is 10. Would it be reasonable to suppose that the mean height of the population is 165?**

**Solution:**

Given that

Sample (n)= 100

Mean= 160

Standard deviation(=10

Hypothesised mean (=165

H0=mean height of the population is 165

H1= mean height of the population is more than 165

Test statistic is given by,

Since n>30

Z test is performed

Z=

Z=-5

Level of significance,

For Z is 1.96

If calculated value is less than equals to tabulated value then accept H0

-5<1.96

-5 does not lie in the accepted region

So, null hypothesis is rejected

That means mean height of the population is more than 165

**13. In a study, physicians were asked what the odds of breast cancer would be in a woman who was initially thought to have a 1% risk of cancer but who ended up with a positive mammogram result (a mammogram accurately classifies about 80% of cancerous tumors and 90% of benign tumors.) 95 out of a hundred physicians estimated the probability of cancer to be about 75%. Do you agree?**

**Solution:**

+ = mammogram result is positive,

B = tumor is benign,

M = tumor is malignant.

Note that BC = M

It is given

P(M) = .01

so P(B) = 1 − P(M) = .99

it is also given the conditional probabilities

P(+ / M) = .80

P(− /B) = .90

where the event − is the complement of +, thus P(+ /B) = .10

P(M /+) =

=

=

=0.074766

=0.075

So the chance would be 7.5%. A far cry from a common estimate of 75%

**14. Suppose we have 3 cards identical in form except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side is colored black. The 3 cards are mixed up in a hat, and 1 3 card is randomly selected and put down on the ground. If the upper side of the chosen card is colored red, what is the probability that the other side is colored black?**

**SOLUTION:**

Given the following events

RR = All side red

BB = All side black

RB = Red and Black

R= Upturned side of the chosen card is red.

Based on the information given in the equation we have

P(R/RB)= ½

P(RB) = 1/3

P(BB) =1/3

P(RR) = 1/3

P(R/RR) =1

P(R/BB) =0

Using the results we have to find p(RB/R)

So, we can say

P(RB/R) =

=

=by substituting the values we get

P(RB/R)=

=

=

=3/9 =1/3 =0.333

The required probability is 0.333