Inferential Statistics

PROJECT REPORT

Gowdham S
PGPDSA.O.DEC23.A

Table of Contents

List of Figures

List of Equations

List of Tables

Data Dictionary of Problem 3

Column Name	Description	Data Type
Unpolished	Brinell's hardness index	float64
Treated and Polished	Brinell's hardness index	float64

Data Dictionary of Problem 4

Column Name	Description	Data Type	Allowed Values
Dentist	ID of the Dentist	int64	1, 2, 3, 4, 5
Method	ID of the Method	int64	1, 2, 3
Alloy	ID of the Alloy	int64	1, 2
Temp	ID of the Temperature	int64	1500, 1600, 1700
Response	Target Variable	int64	Any

Problem 1

Context

A physiotherapist with a male football team is interested in studying the relationship between foot injuries and the positions at which the players play from the data collected.

	Striker	Forward	Attacking Midfielder	Winger	Total
Players Injured	45	56	24	20	145
Players Not Injured	32	38	11	9	90
Total	77	94	35	29	235

Table 1 Players Data

Based on the above data, answer the following questions.

1.1 What is the probability that a randomly chosen player would suffer an injury?

1.2 What is the probability that a player is a forward or a winger?

1.3 What is the probability that a randomly chosen player plays in a striker position and has a foot injury?

1.4 What is the probability that a randomly chosen injured player is a striker?

1.1 What is the probability that a randomly chosen player would suffer an injury?

Solution:

We can find the probability of a randomly chosen player who would suffer an injury by using the below formula,

P = Total Number of Injured Players / Total Number of Players

Equation 1 Probability of a player being injured

- Total number of players = 235
- Number of injured players = 145
- Probability of a player being injured = 145 / 235 = 0.617

The probability that a randomly chosen player would suffer an injury is **61.7 %.**

1.2 What is the probability that a player is a forward or a winger?

Solution:

We can find the probability that a player is a forward or a winger by using the below formula,

*P(Forward or Winger) = (*Total Number of Forwards+Total Number of Wingers) / Total Number of Players

Equation 2 Probability that a player is a forward or a winger

- Total number of players = 235
- Number of forwards = 94
- Number of wingers = 29
- Probability of being a forward or a winger = (94 + 29) / 235 = 0.5234

The probability that a player is a forward or a winger is **52.34 %.**

1.3 What is the probability that a randomly chosen player plays in a striker position and has a foot injury?

Solution:

We can find the probability that a randomly chosen player plays in a striker position and has a foot injury by using the below formula,

P(Injured Striker) = Number of injured strikers / Total number of players

Equation 3 Probability of choosing a Injured Striker

- Total number of players = 235
- Number of Injured Strikers = 45
- Probability that a randomly chosen player is a striker and has injury = (45/235) = 0.1915

The probability that a randomly chosen player plays in a striker position and has injury is **19.15 %**.

1.4 What is the probability that a randomly chosen injured player is a striker?

Solution:

We can find the probability that a arandomly chosen injured player is a striker by using the below formula,

P = Number of Injured Strikers / Total Number of Injured Players

Equation 4 Probability of Injured player is a Striker

- Number of Injured strikers: 45
- Total number of injured players: 145
- Probability of injured player is a striker = (45/145) = 0.3103

The probability that a randomly chosen injured player is a striker is **31.03 %.**

Problem 2

Context

The breaking strength of gunny bags used for packaging cement is normally distributed with a mean of 5 kg per sq. centimeter and a standard deviation of 1.5 kg per sq. centimeter. The quality team of the cement company wants to know the following about the packaging material to better understand wastage or pilferage within the supply chain. Answer the questions below based on the given information (Provide an appropriate visual representation of your answers, without which marks will be deducted)

2.1 What proportion of the gunny bags have a breaking strength of less than 3.17 kg per sq cm?
2.2 What proportion of the gunny bags have a breaking strength of at least 3.6 kg per sq cm.?
2.3 What proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq cm.?
2.4 What proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq cm.?

2.1 What proportion of the gunny bags have a breaking strength of less than 3.17 kg per sq cm?

Solution:

From the Problem Statement,

- Gunny bags used for packaging cement is normally distributed
- Mean is 5 kg per sq. centimeter
- Standard Deviation is 1.5 kg per sq. centimeter

For this, we have used **cdf** method from scipy library to determine the proportion of the gunny bags have a breaking strength of less than 3.17 kg per sq cm. The cdf method takes three parameters,

1. Target value
2. Mean
3. Standard Deviation

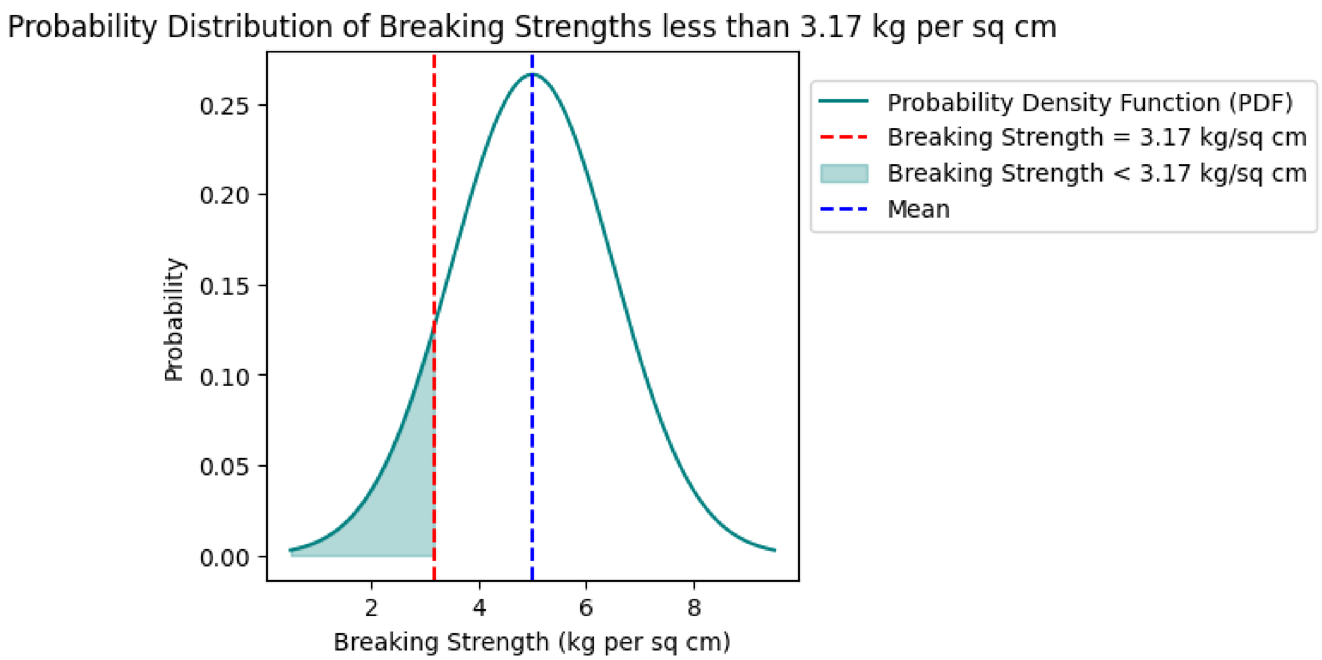

Figure 1 Probability Distribution of Breaking Strengths less than 3.17 kg per sq cm

On applying the parameter values in the cdf method, we got 0.1112. So the proportion of the gunny bags have a breaking strength of less than 3.17 kg per sq cm is **11.12 %**

2.2 What proportion of the gunny bags have a breaking strength of at least 3.6 kg per sq cm?

Solution:

From the Problem Statement,

- Gunny bags used for packaging cement is normally distributed
- Mean is 5 kg per sq. centimeter
- Standard Deviation is 1.5 kg per sq. centimeter

For this, we have used **cdf** method from scipy library to determine the proportion of the gunny bags have a breaking strength of at least 3.6 kg per sq cm. The cdf method takes three parameters,

1. Target value
2. Mean
3. Standard Deviation

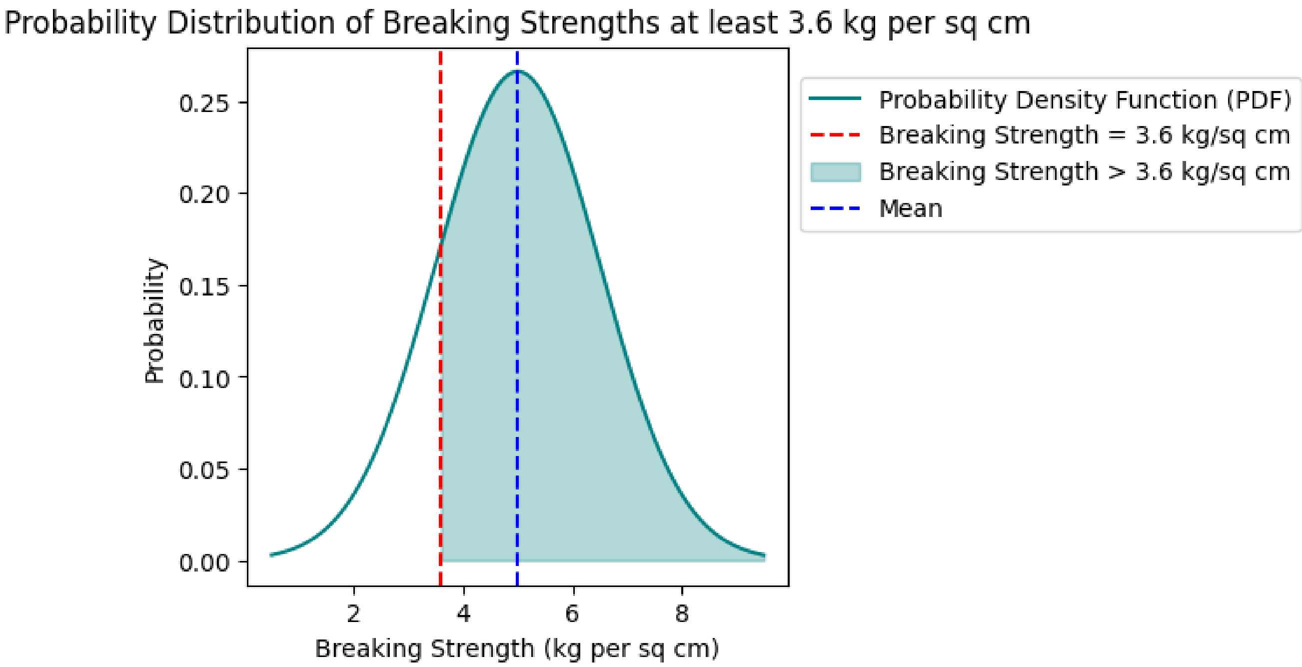

Figure 2 Probability Distribution of Breaking Strengths atleast 3.17 kg per sq cm

As the cdf method gives the value to the left of Target variable, we have to use 1 – cdf to find the proportion of the gunny bags have a breaking strength of at least 3.6 kg per sq cm.

On applying the parameter values in the cdf method, we got 0.8247. So the proportion of the gunny bags have a breaking strength of at least 3.6 kg per sq cm is **82.47 %**

2.3 What proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq cm?

Solution:

From the Problem Statement,

- Gunny bags used for packaging cement is normally distributed
- Mean is 5 kg per sq. centimeter
- Standard Deviation is 1.5 kg per sq. centimeter

For this, we have used **cdf** method from scipy library to determine the proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq cm. The cdf method takes three parameters,

1. Target value
2. Mean
3. Standard Deviation

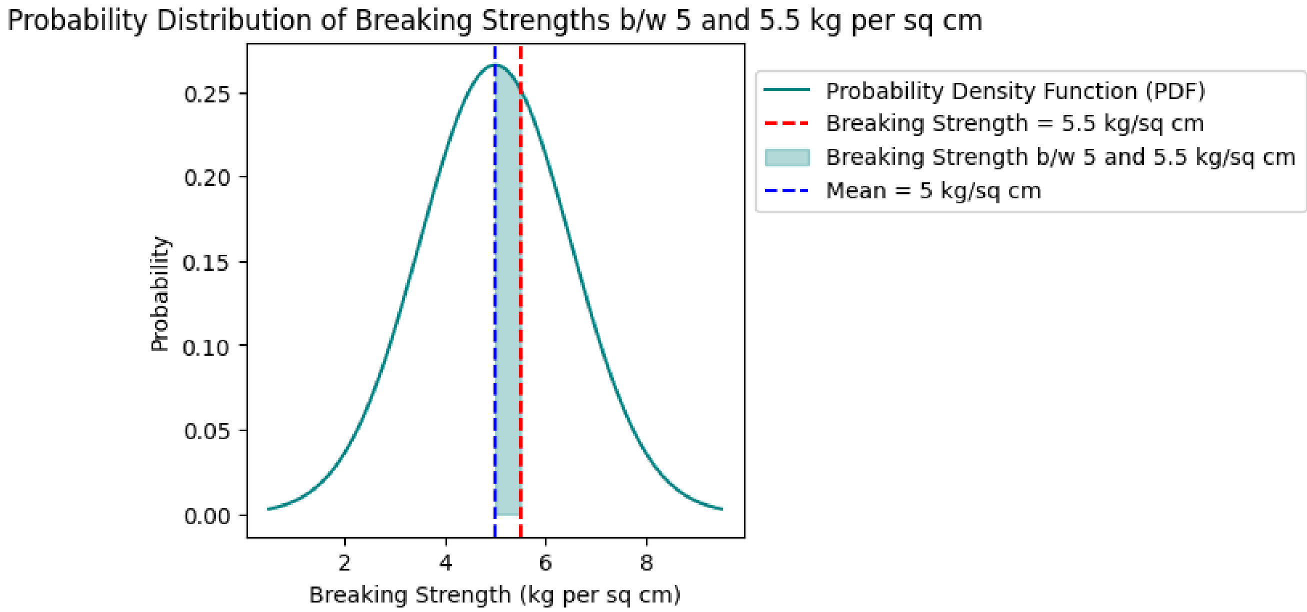

Figure 3 Probability Distribution of Breaking Strength between 5 and 5.5 kg per sq cm

As we have to calculate the proportion between two target variables, we have to find the cdf values for those two variables and subtract cdf value of 5 from 5.5.

On applying the parameter values in the cdf method and subtracting them, we got 0.1306. So the proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq cm is **13.06%**

2.4 What proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq cm?

Solution:

From the Problem Statement,

- Gunny bags used for packaging cement is normally distributed
- Mean is 5 kg per sq. centimeter
- Standard Deviation is 1.5 kg per sq. centimeter

For this, we have used **cdf** method from scipy library to determine the proportion of the gunny bags have a breaking strength not between 3 and 7.5 kg per sq cm. The cdf method takes three parameters,

1. Target value
2. Mean
3. Standard Deviation

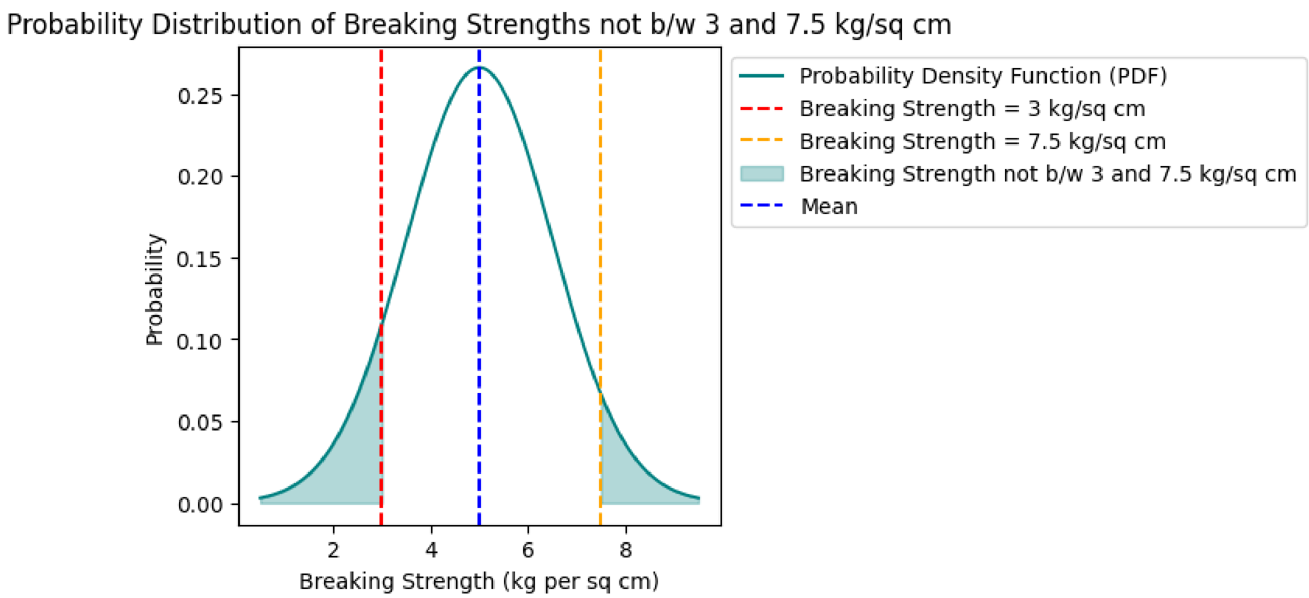

Figure 4 Probability Distribution of Breaking Strength not between 3 and 7.5 kg per sq cm

As we have to calculate the proportion not between two target variables, we have to find the cdf values for those two variables and subtract the cdf value of target variable 7.5 from 1 and finally adding the two cdf values.

On applying the parameter values in the cdf method, subtracting cdf of 7.5 from 1 and finally adding cdf og 3 and cdf of 7.5, we got 0.139. So the proportion of the gunny bags have a breaking strength not between 3 and 7.5 kg per sq cm is **13.9%**

Problem 3

Context

Zingaro stone printing is a company that specializes in printing images or patterns on polished or unpolished stones. However, for the optimum level of printing of the image, the stone surface has to have a Brinell's hardness index of at least 150. Recently, Zingaro has received a batch of polished and unpolished stones from its clients. Use the data provided to answer the following (assuming a 5% significance level);

3.1 Zingaro has reason to believe that the unpolished stones may not be suitable for printing. Do you think Zingaro is justified in thinking so?
3.2 Is the mean hardness of the polished and unpolished stones the same?

3.1 Zingaro has reason to believe that the unpolished stones may not be suitable for printing. Do you think Zingaro is justified in thinking so?

Solution:

Step 1: Define the null and alternative hypothesis

- Ho: Unpolished Stone has adequate hardness >= 150
- Ha: Unpolished does not have adequate hardness <150

Step 2: Decide the significance level

- From the problem statement, we are given Level of Significance (α) = 0.05.

Step 3: Select the appropriate test

- From the dataset we have,
  - There are no null values
  - Sample size is 75 & it is greater than 30 i.e., n > 30
  - We are not given Population Mean and Standard Deviation
- Hence we are proceeding with one sample t-test.

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

Figure 5 Equation for One Sampled T Test

Where,

- $\bar{X}$ is the sample mean.
- μ is the population mean.
- S is the sample standard deviation.
- n is the sample size.

Step 4: Calculate the test statistics and p-value

- We used **ttest_1samp** method from scipy library to calculate the t value and p value. We can aslo manually determine the t value using the Figure 5 formula
- We passed the Unpolished sample data and mean hardness value inside ttest_1samp method
- t value = -4.165 , p-value for 1 tailed = 0.0000417

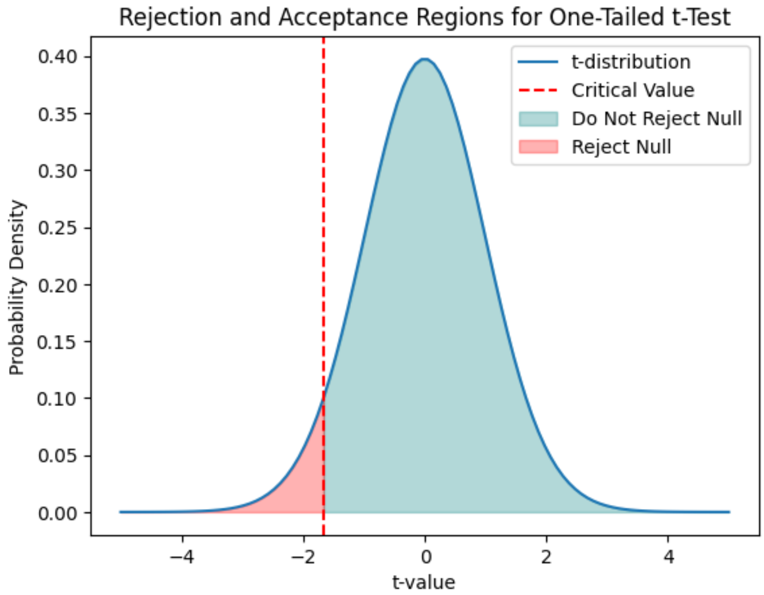

Figure 6 Rejection and Acceptance Regions for One-Tailed t-Test

Step 5: Decide to Reject or Accept the null hypothesis

- As the p-value is less than Alpha 0.05 and the t value lies in the rejection region, we reject the null hypothesis and conclude that Unpolished stone is not suitablefor printing and has Brinell's hardness index of less than 150. Hence Zingaro's thinking is justified.

3.2 Is the mean hardness of the polished and unpolished stones the same?

Solution:

Step 1: Define the null and alternative hypothesis

- Ho: Unpolished Stone mean hardness = Polished Stone mean hardness
- Ha: Unpolished Stone mean hardness ≠ Polished Stone mean hardness

Step 2: Decide the significance level

- From the problem statement, we are given Level of Significance (α) = 0.05.

Step 3: Select the appropriate test

- From the dataset we have,
  - There are no null values
  - Sample size is 75 & it is greater than 30 i.e., n > 30
  - We are not given Population Mean and Standard Deviation

- Hence we are proceeding with Independent two sample t test as the directionality is not known.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}\right)}}$$

Figure 7 Equation for Independent Two Sampled T Test

Where,

- $\bar{x}_1$ and $\bar{x}_2$ are the sample means of the two groups.
- s1 and s2 are the sample standard deviations of the two groups.
- n1 and n2 are the sample sizes of the two groups.

Step 4: Calculate the test statistics and p-value

- We have used **ttest_ind** method from scipy library to calculate the t value and p value. We can aslo manually determine the t value using the Figure 7 formula
- We passed the Unpolished & Polished sample data inside ttest_ind method
- t value = -3.242, p-value for 2 tailed = 0.001465

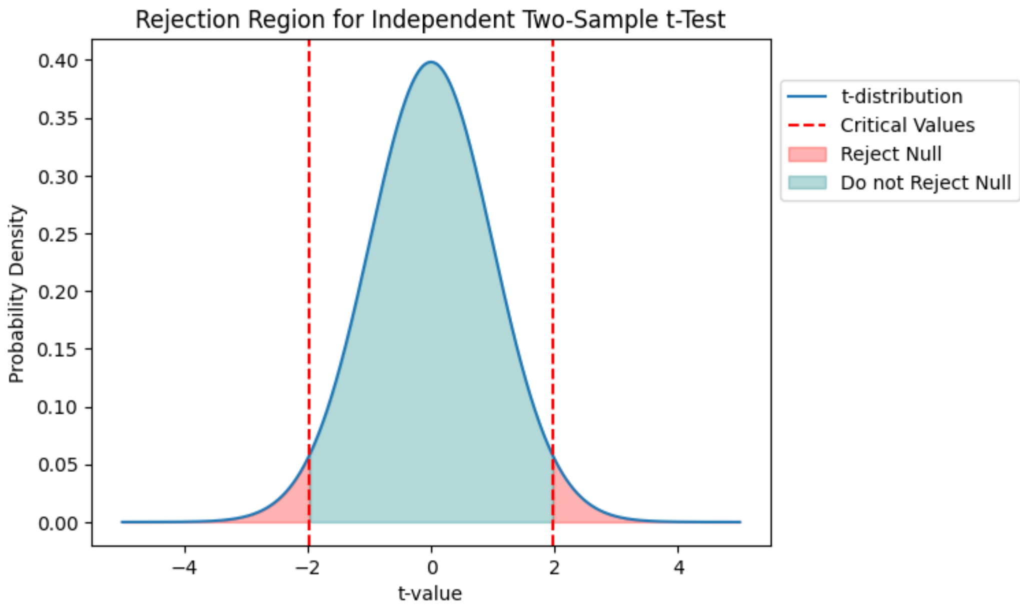

Figure 8 Rejection and Acceptance Regions for Two-Tailed t-Test

Step 5: Decide to Reject or Accept the null hypothesis

- As the p-value is less than Alpha 0.05 and the t value lies in the rejection region, we reject the null hypothesis and conclude that mean hardness of Unpolished and Polished stones are not same.

Problem 4

Context

Dental implant data: The hardness of metal implants in dental cavities depends on multiple factors, such as the method of implant, the temperature at which the metal is treated, the alloy used as well as the dentists who may favor one method above another and may work better in his/her favorite method. The response is the variable of interest.

4.1 How does the hardness of implants vary depending on dentists?
4.2 How does the hardness of implants vary depending on methods?
4.3 What is the interaction effect between the dentist and method on the hardness of dental implants for each type of alloy?
4.4 How does the hardness of implants vary depending on dentists and methods together?

Solution:

Before proceeding with solving the actual problem, we have done the initial analysis of the data. The results of the analysis are,

- The dataset has 90 rows and 5 columns
- All the columns are of type int64
- Response column is the target variable and other columns are categorical
- Mean value of Response column is 741.77
- The Response column follows fairly normal distribution with left skewness.

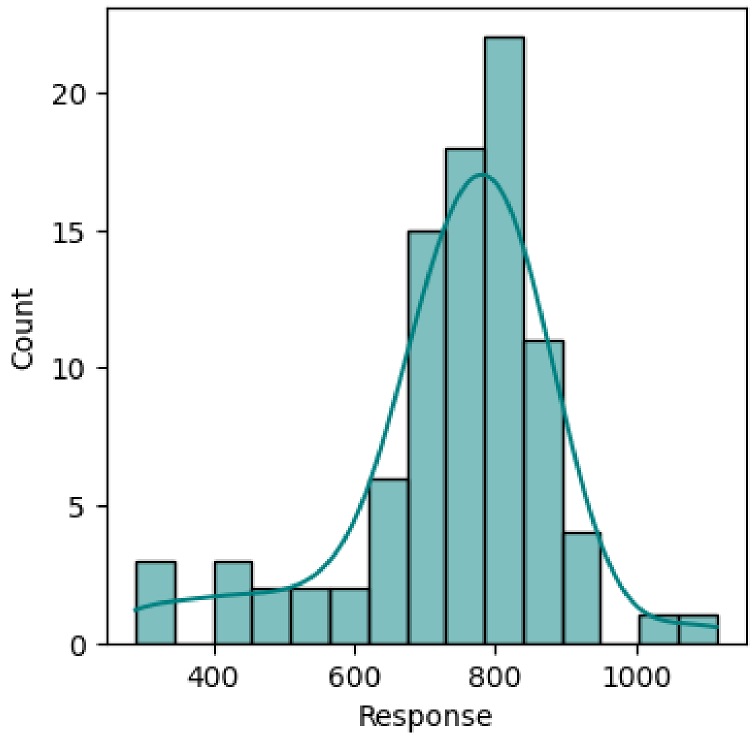

Figure 9 Distribution of Response Column

- Once the above analysis is done, we then proceeded with changing the datatype of categorical columns i.e., Dentist, Method, Alloy & Temp from int64 to categorical using the pd.Categorical() method.
- We then separated the dataset into two based on the type of alloy.

Assumptions of ANOVA

- Independent Sample - Sample should be selected randomly(Equally likely events).
- Normal Distribution - Distribution of each group should be normal
- Homogenous Group - Variance between the group should be the same.

We can test the Normality using Shapiro-Wilk's test and Homogenity using Levene's test

Shapiro-Wilk test for Alloy 1

- Null Hypothesis (Alloy 1): The population from which the sample was drawn follows a normal distribution.
- Alternative Hypothesis (Alloy 1): The population from which the sample was drawn does not follow a normal distribution.

We utilized the Shapiro method from the stats module of the SciPy library to calculate the t and p-values. For Alloy 1, we obtained t and p-values of 0.830 and 0.0000119, respectively. Since the p-value is below the significance level of 0.05, we have sufficient evidence to reject the null hypothesis. Consequently, we conclude that the population from which the sample was drawn for Alloy 1 does not adhere to a normal distribution.

Shapiro-Wilk test for Alloy 2

- Null Hypothesis (Alloy 2): The population from which the sample was drawn follows a normal distribution.
- Alternative Hypothesis (Alloy 2): The population from which the sample was drawn does not follow a normal distribution.

We employed the Shapiro method from the stats module of the SciPy library to compute the t and p-values. For Alloy 2, we obtained t and p-values of 0.887 and 0.00040, respectively. Since the p-value is below the significance level of 0.05, we possess adequate evidence to reject the null hypothesis. Therefore, we infer that the population from which the sample was drawn for Alloy 2 does not conform to a normal distribution.

Levene's test

- Null Hypothesis: The population variances of the groups from which the samples were drawn are equal.
- Alternative Hypothesis: The population variances of the groups from which the samples were drawn are not equal.

We applied the Levene method from the stats module of the SciPy library to calculate the t and p-values. The obtained values were t = 1.41 and p = 0.2366. Given that the p-value exceeds the significance level of 0.05, we do not have sufficient evidence to reject the null hypothesis. Therefore, we conclude that the population variances of the groups from which the samples were drawn are equal.

4.1 How does the hardness of implants vary depending on dentists?

Solution:

Step 1: Define the null and alternative hypothesis

- Null Hypothesis: There is no difference in means among the dentists in terms of implant hardness
- Alternate Hypothesis : There is a difference in means among the dentists in terms of implant hardness

Step 2: Decide the significance level

- We are choosing the Level of Significance as 0.05 as there is no specific value given

Step 3: Performing the Test

We have used the **ols** method from Stasmodels library to perform the One-Way ANOVA test. We are choosing One-Way ANOVA as we have 1 factor (Dentist) and 1 continuous variable (Response)

Alloy 1

• Null hypothesis: There is no difference in means among the dentists in terms of implant hardness for Alloy 1.

• Alternative hypothesis: There is a difference in means among the dentists in terms of implant hardness for Alloy 1.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4	106683.7	26670.92	1.977112	0.116567
Residual	40	539593.6	13489.84	NaN	NaN

Table 2 ANOVA for Alloy 1 (Dentist)

On performing the test, we got p value 0.115 which is greater than the level of significance. So we fail to reject the null hypothesis and consider there is no difference in means among the dentists in terms of implant hardness for Alloy 1.

Alloy 2

• Null hypothesis: There is no difference in means among the dentists in terms of implant hardness for Alloy 2.

• Alternative hypothesis: There is a difference in means among the dentists in terms of implant hardness for Alloy 2.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4	56797.91	14199.48	0.524835	0.718031
Residual	40	10,82,205	27055.12	NaN	NaN

Table 3 ANOVA for Alloy 2 (Dentist)

On performing the test, we got p value 0.715 which is greater than the level of significance. So we fail to reject the null hypothesis and consider there is no difference in means among the dentists in terms of implant hardness for Alloy 2.

4.2 How does the hardness of implants vary depending on methods?

Solution:

Step 1: Define the null and alternative hypothesis

- Null Hypothesis: There is no difference in means among the methods in terms of implant hardness
- Alternate Hypothesis : There is a difference in means among the methods in terms of implant hardness

Step 2: Decide the significance level

- We are choosing the Level of Significance as 0.05 as there is no specific value given

Step 3: Performing the Test

We have used the **ols** method from Stasmodels library to perform the One-Way ANOVA test. We are choosing One-Way ANOVA as we have 1 factor (Method) and 1 continuous variable (Response)

Alloy 1

• Null hypothesis: There is no difference in means among the methods in terms of implant hardness for Alloy 1.
• Alternative hypothesis: There is a difference in means among the methods in terms of implant hardness for Alloy 1.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	2	148472.2	74236.09	6.263327	0.004163
Residual	42	4,97,805	11852.5	NaN	NaN

Table 4 ANOVA for Alloy 1 (Method)

On performing the test, we got p value 0.004 which is lesser than the level of significance. So we reject the null hypothesis and consider there is a difference in means among the methods in terms of implant hardness for Alloy 1.

Alloy 2

• Null hypothesis: There is no difference in means among the methods in terms of implant hardness for Alloy 2.

• Alternative hypothesis: There is a difference in means among the methods in terms of implant hardness for Alloy 2.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	2	499640.4	249820.2	16.4108	0.000005
Residual	42	6,39,362	15222.91	NaN	NaN

Table 5 ANOVA for Alloy 2 (Method)

On performing the test, we got p value 0.000005 which is greater than the level of significance. So we fail to reject the null hypothesis and consider there is no difference in means among the dentists in terms of implant hardness for Alloy 2.

Since the p-values for both type of alloys are lesser than the level of significance, we have to perform Tukey HSD test.

Tukey HSD test

We used the tukeyhsd() method for comparing the means of multiple groups to determine which groups are significantly different from each other.

```
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
========================================================
group1 group2  meandiff p-adj    lower    upper   reject
--------------------------------------------------------
     1      2   -6.1333  0.987  -102.714  90.4473  False
     1      3    -124.8 0.0085 -221.3807 -28.2193   True
     2      3 -118.6667 0.0128 -215.2473  -22.086   True
--------------------------------------------------------
```

Figure 10 Tukey HSD for Alloy 1

From the Figure 10 we can see,

- Group 1 vs. Group 2: The mean difference is not statistically significant (p = 0.987), suggesting no difference in means.
- Group 1 vs. Group 3: A significant mean difference exists (p = 0.0085), rejecting the null hypothesis.
- Group 2 vs. Group 3: A significant mean difference exists (p = 0.0128), rejecting the null hypothesis.

```
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
========================================================
group1 group2 meandiff p-adj   lower     upper   reject
--------------------------------------------------------
     1      2     27.0 0.8212  -82.4546  136.4546  False
     1      3   -208.8 0.0001 -318.2546  -99.3454   True
     2      3   -235.8    0.0 -345.2546 -126.3454   True
--------------------------------------------------------
```

Figure 11 Tukey HSD for Alloy 2

From the Figure 11 we can see,

- Group 1 vs. Group 2: The mean difference is 27.0, with an adjusted p-value of 0.8212, indicating no significant difference in means between the two groups.
- Group 1 vs. Group 3: The mean difference is -208.8, with an adjusted p-value of 0.0001, suggesting a significant difference in means between the two groups.
- Group 2 vs. Group 3: The mean difference is -235.8, with an adjusted p-value of 0.0, indicating a significant difference in means between the two groups.

We can conclude that, Method 3 has variation when compared with the other two in both type of alloys.

4.3 What is the interaction effect between the dentist and method on the hardness of dental implants for each type of alloy?

Solution:

We plotted a interaction plot between Dentist vs Response with Method as hue parameter using pointplot for both type of alloys.

Alloy 1

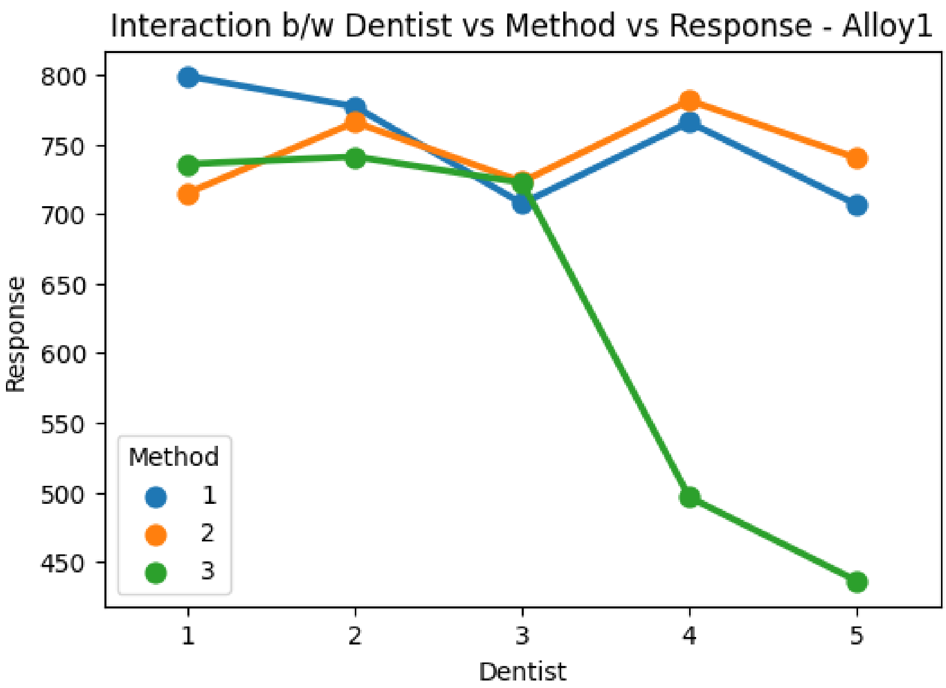

Figure 12 Interaction b/w Dentist vs Method vs Response - Alloy1

Observations

- The response pattern for Method 2 is relatively stable compared to the other methods, with no drastic increases or decreases.
- We can clearly see drastic decrease in response for Method 3 in the case of Dentist 4 and 5
- Dentist 1 shows a consistent increase in response from Method 1 to Method 3.
- Dentist 2 has the highest response for Method 1, slightly decreasing for Methods 2 and 3.
- Dentist 3's response peaks at Method 2, with Method 3 showing a significant drop.
- Dentist 5 shows a drastic decrease in response when moving from Method 1 to Method 3.
- The overall highest response is observed with Dentist 1 using Method 2.
- The overall lowest response is observed with Dentist 5 using Method 3.

Alloy 2

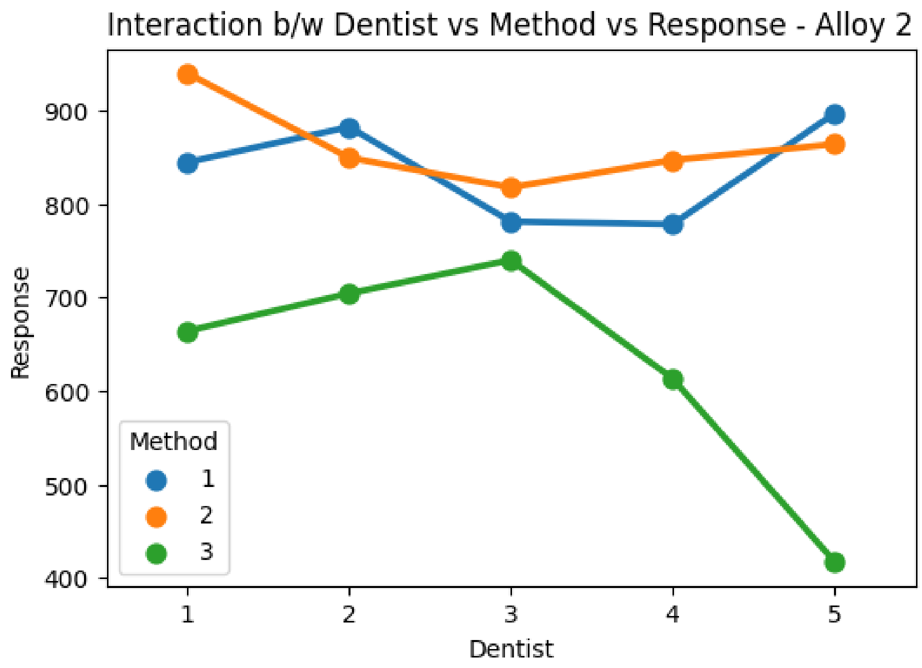

Figure 13 Interaction b/w Dentist vs Method vs Response - Alloy2

Observations

- The response pattern for Method 2 is relatively stable compared to the other methods, with no drastic increases or decreases.
- We can clearly see drastic decrease in response for Method 3 in the case of Dentist 4 and 5
- Dentist 1 shows a consistent increase in response from Method 1 to Method 2.
- Dentist 2 has the highest response for Method 1, slightly decreasing for Method 2.
- Dentist 3's response peaks at Method 2, with Method 3 showing a significant drop.
- Dentist 5 shows a drastic decrease in response when moving from Method 1 to Method 3.
- The overall highest response is observed with Dentist 1 using Method 2.
- The overall lowest response is observed with Dentist 5 using Method 3.

4.4 How does the hardness of implants vary depending on dentists and methods together?

Solution:

Step 1: Define the null and alternative hypothesis

- Null Hypothesis : There is no difference in mean hardness of implants depending on dentists and methods together
- Alternate Hypothesis : There is a difference in mean hardness of implants depending on dentists and methods together

Step 2: Decide the significance level

- We are choosing the Level of Significance as 0.05 as there is no specific value given

Step 3: Performing the Test

We have used the **ols** method from Stasmodels library to perform the Two-Way ANOVA test. We are choosing One-Way ANOVA as we have 2 factors (Dentist & Method) and 1 continuous variable (Response)

Alloy 1

• Null hypothesis: There is no difference in mean hardness of implants depending on dentists and methods together for Alloy 1.
• Alternative hypothesis: There is difference in mean hardness of implants depending on dentists and methods together for Alloy 1.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	106683.688889	26670.922222	2.591255	0.051875
C(Method)	2.0	148472.177778	74236.088889	7.212522	0.002211
Residual	38.0	391121.377778	10292.667836	NaN	NaN

Figure 14 Two-Way ANOVA for Alloy 1

Upon conducting the test, we obtained a p-value of 0.0022 for the Method factor, which is below the significance level of 0.05. Thus, we have ample evidence to reject the null hypothesis, indicating that at least one pair of Method means differs for Alloy 1, and it significantly affects the Response. Conversely, the p-value for the Dentist factor is 0.051, surpassing the 0.05 threshold. Consequently, we cannot reject the null hypothesis, suggesting no variance in the mean hardness of implants based on dentists.

From Figure 12 we can see that there is some sort of interaction between the two treatments. So, we will introduce a new term while performing the Two Way ANOVA.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	106683.688889	26670.922222	3.899638	0.011484
C(Method)	2.0	148472.177778	74236.088889	10.854287	0.000284
C(Dentist):C(Method)	8.0	185941.377778	23242.672222	3.398383	0.006793
Residual	30.0	205180.000000	6839.333333	NaN	NaN

Figure 15 Two-Way ANOVA for Alloy 1 (Interaction Variable)

Due to the inclusion of the interaction effect term, we can see a change in the p-value of the first two treatments as compared to the Two-Way ANOVA without the interaction effect terms.

- The p-value for the Dentist factor is 0.011484, which is below the significance level of 0.05. Thus, there is a significant effect of Dentist on the Response variable.
- The p-value for the Method factor is 0.000284, which is below the significance level of 0.05. Therefore, there is a significant effect of Method on the Response variable.
- The p-value for the interaction between Dentist and Method (C(Dentist):C(Method)) is 0.006793, below the significance level of 0.05. This indicates that there is a significant interaction effect between Dentist and Method on the Response variable.
- We can see that the p-value of the interaction effect term of Dentist and Method suggests that the Null Hypothesis is rejected in this case.

Alloy 2

• Null hypothesis: There is no difference in mean hardness of implants depending on dentists and methods together for Alloy 2.
• Alternative hypothesis: There is difference in mean hardness of implants depending on dentists and methods together for Alloy 2.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	56797.911111	14199.477778	0.926215	0.458933
C(Method)	2.0	499640.400000	249820.200000	16.295479	0.000008
Residual	38.0	582564.488889	15330.644444	NaN	NaN

Figure 16 Two-Way ANOVA for Alloy 2

After testing, we found a p-value of 0.000008 for the Method factor, which is below the significance level of 0.05. This suggests that at least one pair of Method means differs for Alloy 2, and it does significantly affect the Response. Conversely, the p-value for the Dentist factor is 0.458, below the 0.05 threshold. Therefore, we reject the null hypothesis, indicating variance in the mean hardness of implants based on dentists.

From Figure 13 we can see that there is some sort of interaction between the two treatments. So, we will introduce a new term while performing the Two Way ANOVA.

	df	sum_sq	mean_sq	F	PR(>F)
C(Dentist)	4.0	56797.911111	14199.477778	1.106152	0.371833
C(Method)	2.0	499640.400000	249820.200000	19.461218	0.000004
C(Dentist):C(Method)	8.0	197459.822222	24682.477778	1.922787	0.093234
Residual	30.0	385104.666667	12836.822222	NaN	NaN

Figure 17 Two-Way ANOVA for Alloy 2 (Interaction Variable)

Due to the inclusion of the interaction effect term, we can see a change in the p-value of the first two treatments as compared to the Two-Way ANOVA without the interaction effect terms.

- The p-value for the Dentist factor is 0.371833, which is greater than the significance level of 0.05. Thus, there is no significant effect of Dentist on the Response variable.
- The p-value for the Method factor is 0.000004, which is much less than 0.05. Therefore, there is a significant effect of Method on the Response variable.
- The p-value for the interaction between Dentist and Method (C(Dentist):C(Method)) is 0.093234, which is greater than 0.05. This suggests that there is no significant interaction effect between Dentist and Method on the Response variable.
- We can see that the p-value of the interaction effect term of Dentist and Method suggests that we fail to reject Null Hypothesis in this case.