Advanced Statistics Project

Contents:-

5.1 Earlier experience of Zingaro with this particular client is favorable as the stone surface was found to be of adequate hardness. However, Zingaro has reason to believe now that the unpolished stones may not be suitable for printing. Do you think Zingaro is justified in thinking so?

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

7.1 Test whether there is any difference among the dentists on the implant hardness. State the null and alternative hypotheses. Note that both types of alloys cannot be considered together. You must state the null and alternative hypotheses separately for the two types of alloys.?

7.2 Before the hypotheses may be tested, state the required assumptions. Are the assumptions fulfilled? Comment separately on both alloy types.?

7.3 Irrespective of your conclusion in 2, we will continue with the testing procedure. What do you conclude regarding whether implant hardness depends on dentists? Clearly state your conclusion. If the null hypothesis is rejected , is it possible to identify which pairs of dentists differ?

7.4 Now test whether there is any difference among the methods on the hardness of dental implant, separately for the two types of alloys. What are your conclusions? If the null hypothesis is rejected, is it possible to identify which pairs of methods differ?

7.5 Now test whether there is any difference among the temperature levels on the hardness of dental implant, separately for the two types of alloys. What are your conclusions? If the null hypothesis is rejected, is it possible to identify which levels of temperatures differ?

7.6 Consider the interaction effect of dentist and method and comment on the interaction plot, separately for the two types of alloys?

7.7 Now consider the effect of both factors, dentist, and method, separately on each alloy. What do you conclude? Is it possible to identify which dentists are different, which methods are different different , and which interaction levels are different ?

List of tables :-

List of Figures:-

Problem 1

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

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

Solution:- a=145/235=0.62
The probability that a randomly chosen player would suffer an injury 0.62

1.2 What is the probability that a player is a forward or a winger?
Solution:- b=94/235+29/235 = 0.52

The probability that a player is a forward or a winger 0.52

1.3 What is the probability that a randomly chosen player plays in a striker position and has a foot injury?
Solution: c=45/77=0.58

The probability that a randomly chosen player plays in a striker position and has a foot injury 0.58

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

Solution:- d=45/145=0.31

The probability that a randomly chosen injured player is a striker 0.31

a. What is the probability that a randomly chosen injured player is either a forward or an attacking mi dfielder?

Solution:- e=56/145+24/145=0.55

The probability that a randomly chosen injured player is either a forward or an attacking midfielder 0.55

Problem 2

An independent research organization is trying to estimate the probability that an accident at a nuclear power plant will result in radiation leakage. The types of accidents possible at the plant are, fire hazards, mechanical failure, or human error. The research organization also knows that two or more types of accidents cannot occur simultaneously.

According to the studies carried out by the organization, the probability of a radiation leak in case of a fire is 20%, the probability of a radiation leak in case of a mechanical 50%, and the probability of a radiation leak in case of a human error is 10%. The studies also showed the following;

- The probability of a radiation leak occurring simultaneously with a fire is 0.1%.
- The probability of a radiation leak occurring simultaneously with a mechanical failure is 0.15%.

- The probability of a radiation leak occurring simultaneously with a human error is 0.12%.

On the basis of the information available, answer the questions below:

2.1 What are the probabilities of a fire, a mechanical failure, and a human error respectively?

Solution:-
P(F)= 0.001/0.20= 0.005

P(M)= 0.0015/0.50= 0.003

P(H)= 0.0012/10= 0.00012

The probabilities of a fire hazards is 0.005 a mechanical failure is 0.003 and a human error is 0.00012

2.2 What is the probability of a radiation leak?

Solution:-

P(RL) =0.001+0.0015+0.0012=0.0037

The Probability of a Radiation Leak is 0.0037

2.3 Suppose there has been a radiation leak in the reactor for which the definite cause is not known. What is t he probability that it has been caused by:

- A Fire.
- A Mechanical Failure.
- A Human Error.

Solution:-

P(FH) =0.001/0.0037

MF=0.0015/0.0037

HR=0.0012/0.0037
*The Probability that it has been caused by a Fire is 0.27

*The Probability that it has been caused by a Mechanical Failure is 0.40

*The Probability that it has been caused by a Human Error is 0.32

Problem 3:

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)**

3.1 What proportion of the gunny bags have a breaking strength less than 3.17 kg per sq cm?
Solution:-
z=(3.17-5)/1.5= -1.22
Calculating the P value using the formula: - stats.norm.cdf(-1.22)=0.11

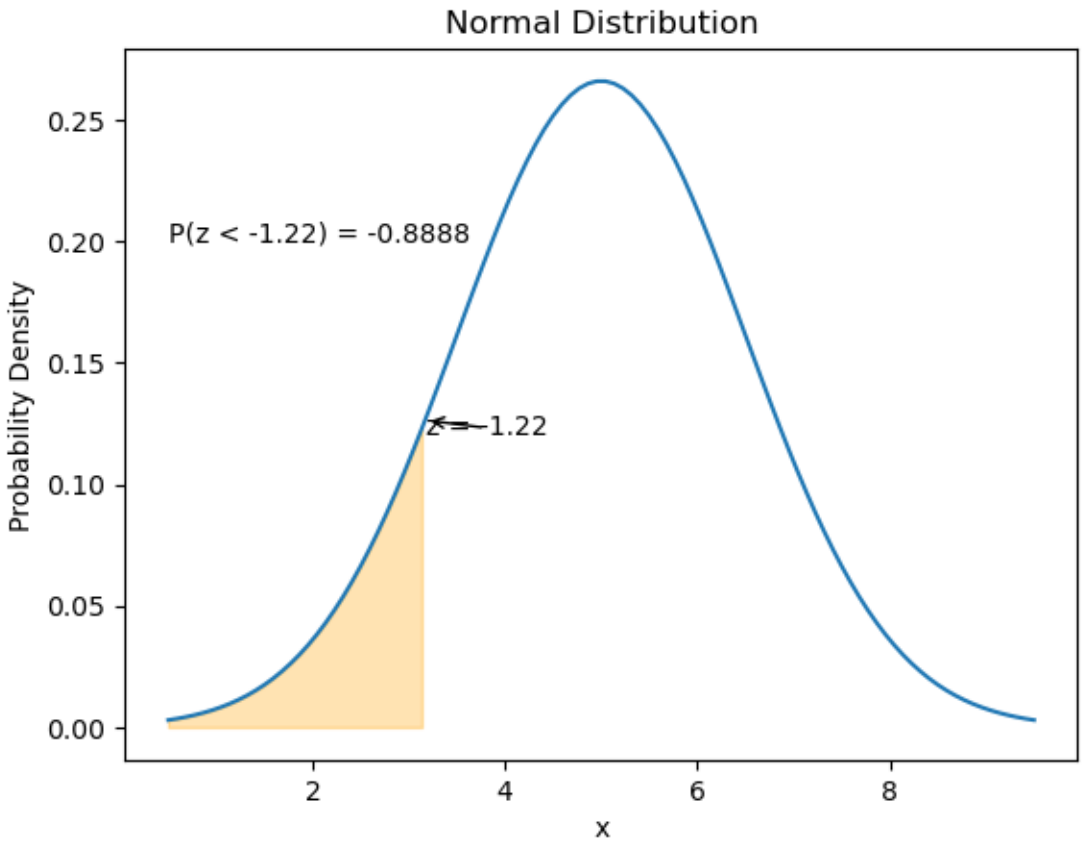


Figure:-1

1

The gunny bags have a breaking strength less than 3.17 kg per sq cm is 11.1%

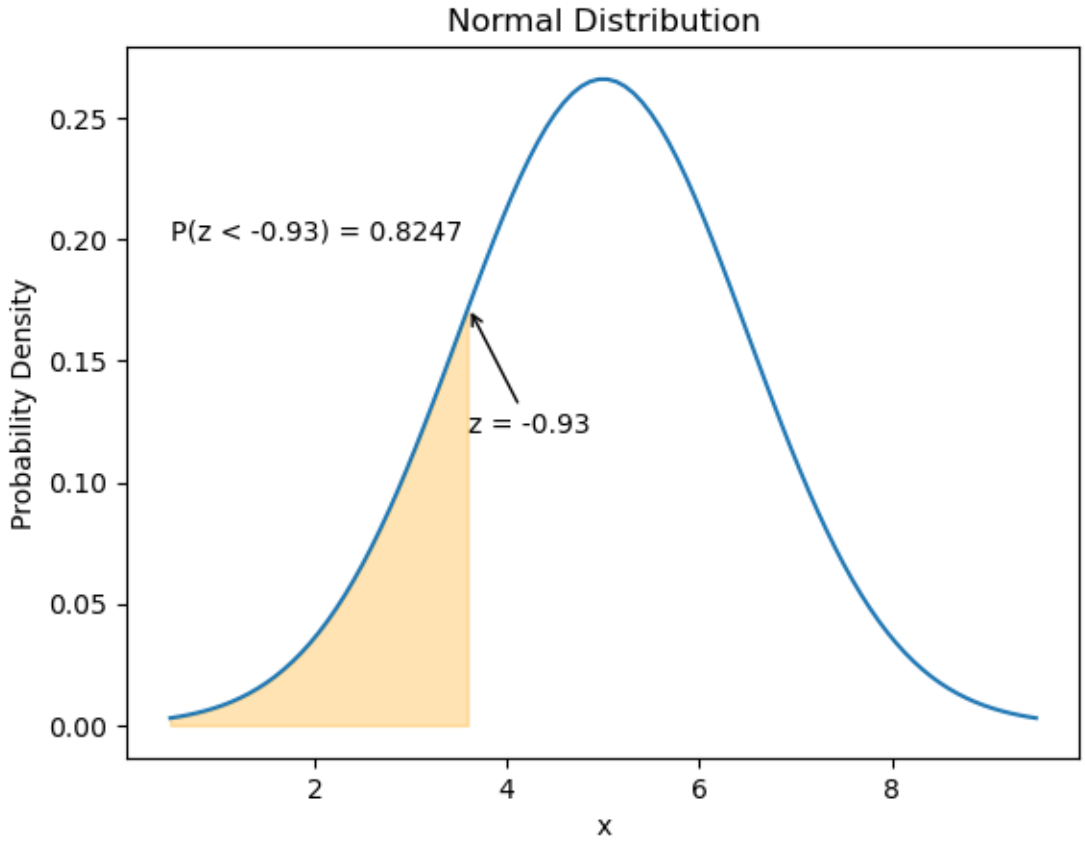
3.2 What proportion of the gunny bags have a breaking strength at least 3.6 kg per sq cm.?

Solution:-

z= (3.6-5)/1.5= -0.93

Calculating the P value using the formula: - 1-stats.norm.cdf (-0.93) = 0.82

Figure:-2



The gunny bags have a breaking strength at least 3.6 kg per sq cm is 82%

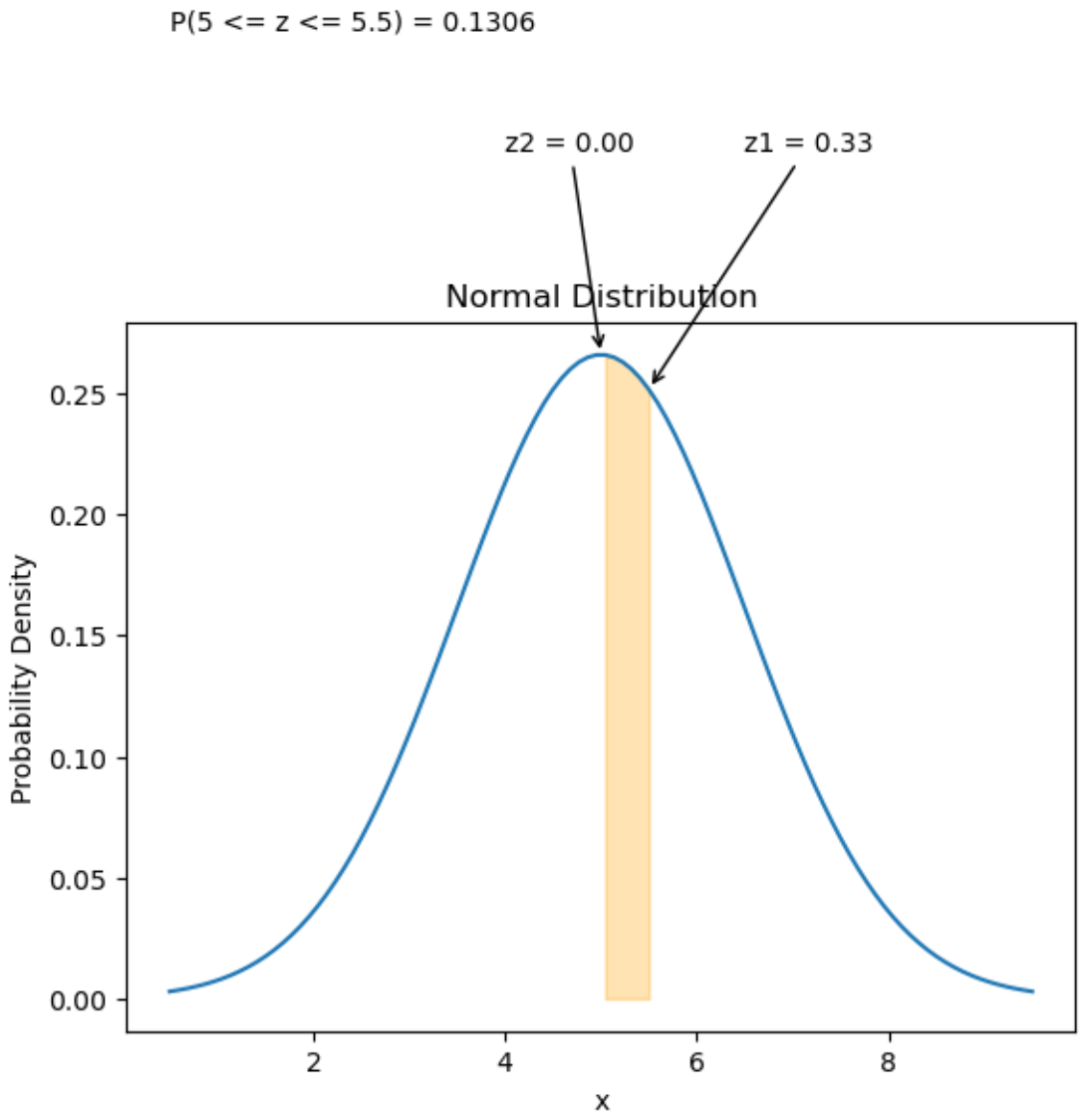
3.3 What proportion of the gunny bags have a breaking strength between 5 and 5.5 kg per sq cm.?
Solution:-

z1= (5.5-5)/1.5= 0.33

z2= (5-5)/1.5=0

Calculating the P value using the formula= stats.norm.cdf (0.33)-stats.norm.cdf(0)=0.13

Figure:-3

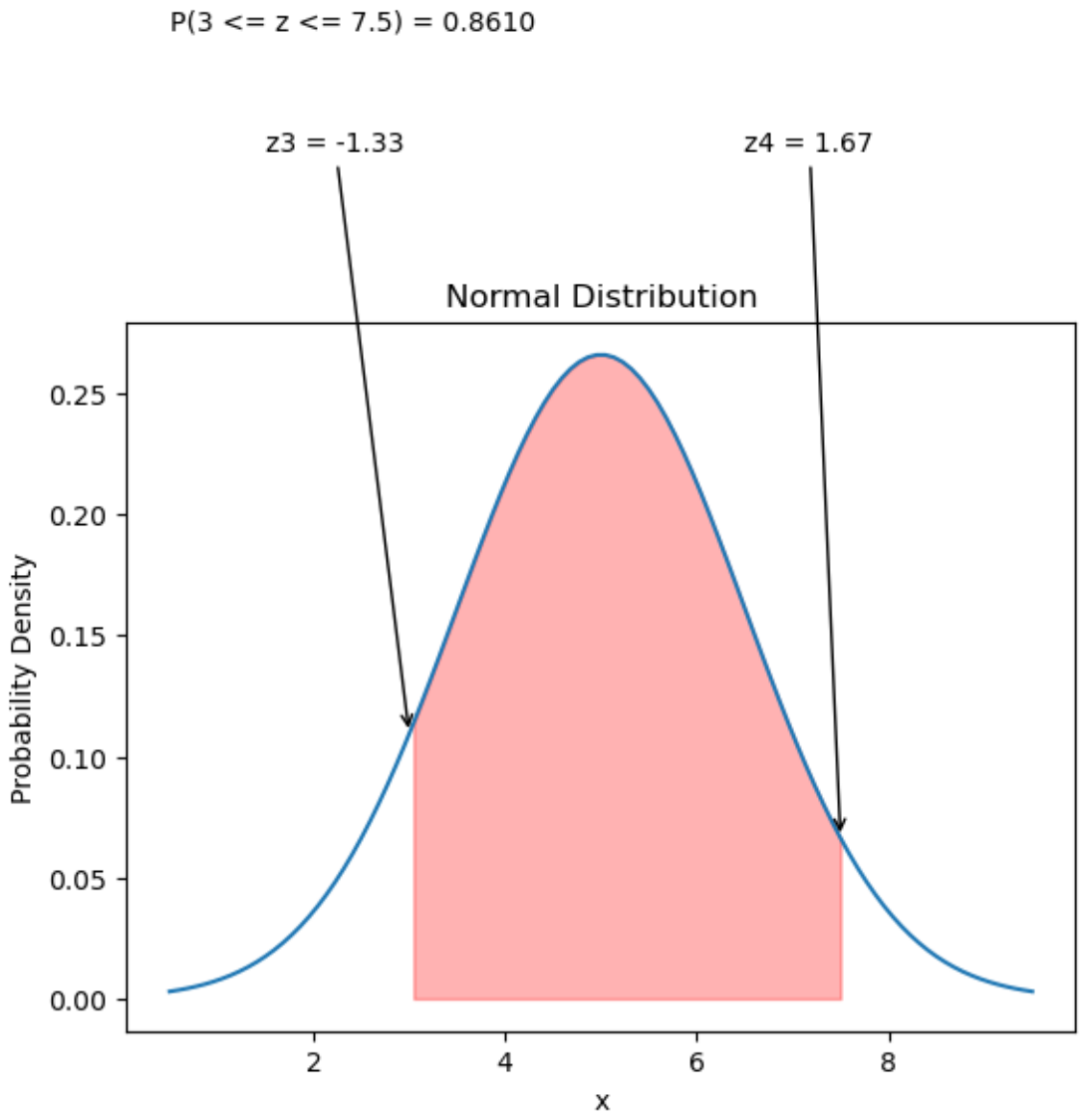


The gunny bags have a breaking strength between 5 and 5.5 kg per sq cm is 13%

3.4 What proportion of the gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq cm.?
Solution:-

z3= (3-5)/1.5= -1.33
z4= (7.5-5)/1.5= 1.66
Calculating the P value using the formula= stats.norm.cdf(1.66)-stats.norm.cdf(-1.33)=0.86

Figure:-4



The gunny bags have a breaking strength NOT between 3 and 7.5 kg per sq cm is 86%

Problem 4:

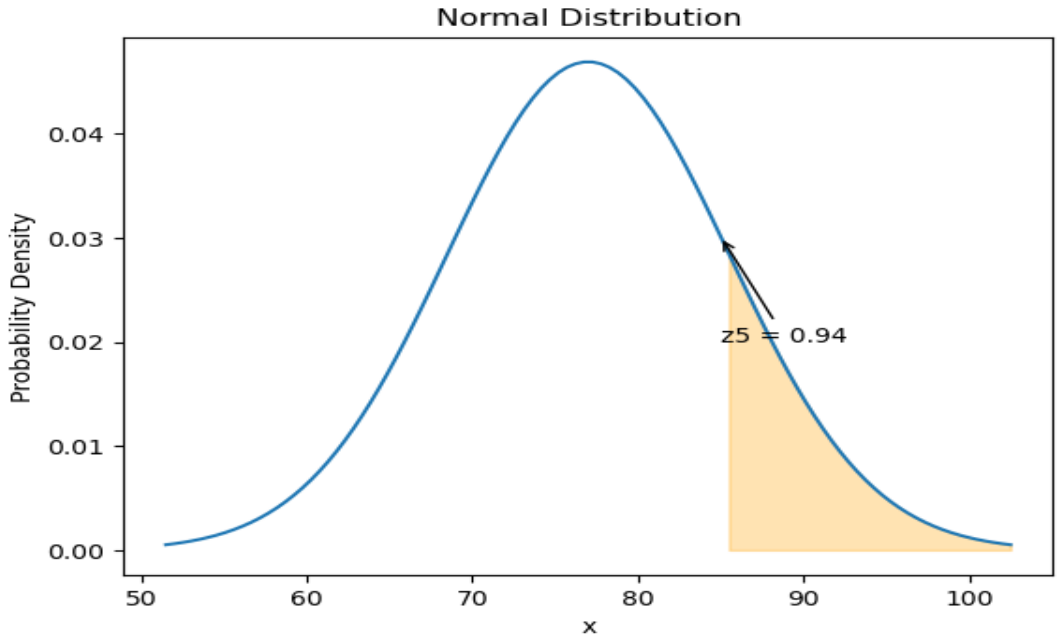
Grades of the final examination in a training course are found to be normally distributed, with a mean of 77 an d a standard deviation of 8.5. Based on the given information answer the questions below.

4.1 What is the probability that a randomly chosen student gets a grade below 85 on this exam?
Solution:-

z5= (85-77)/8.5=0.94

Calculating the P value using the formula= 1-stats.norm.cdf (0.94)=0.17

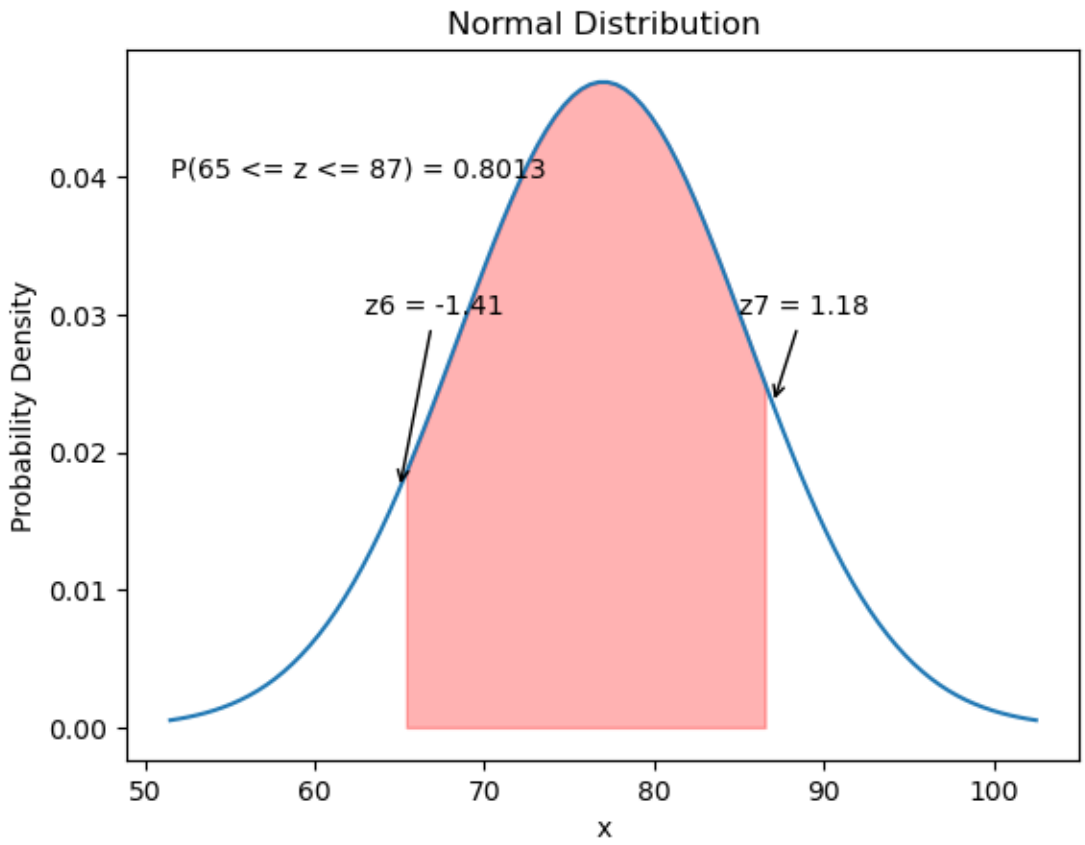
Figure:-5



Hence, a randomly chosen student gets a grade below 85 on this exam is 17%

4.2 What is the probability that a randomly selected student scores between 65 and 87?
Solution :-
z6=(65-77)/8.5= -1.41
z7=(87-77)/8.5= 1.17
Calculating the P value using the Formula: - stats.norm.cdf(1.17)-stats.norm.cdf(-1.41)=0.80

Figure:-6

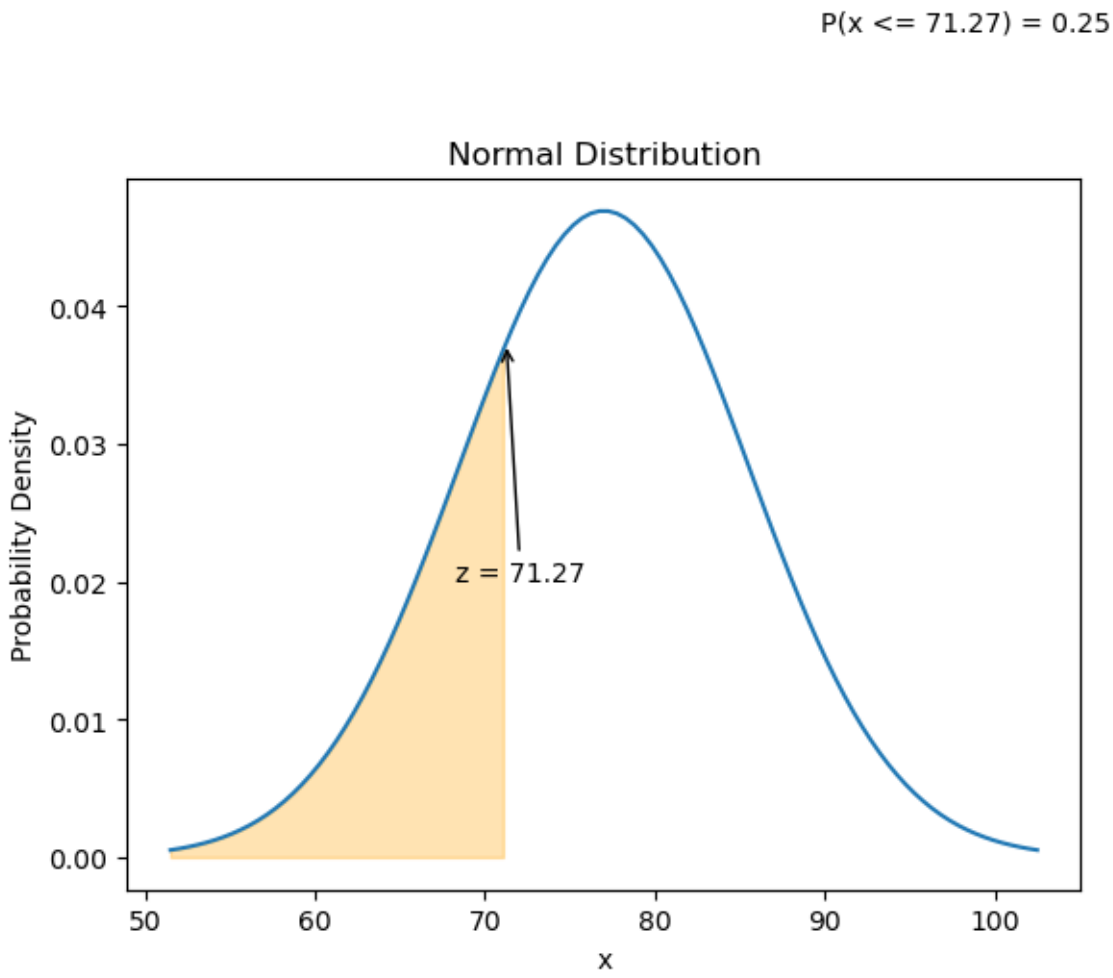


Hence, a randomly selected student scores between 65 and 87 is 80%

4.3 What should be the passing cut-off so that 75% of the students clear the exam?
Solution :-

Calculating the P value using the formula :- stats.norm.ppf(0.25,loc=77,scale=8.5)= 71.26

Figure:-7



The passing cut-off so that 75% of the students clear the exam is 71.26

Problem 5:

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);

5.1 Earlier experience of Zingaro with this particular client is favorable as the stone surface was found to be of adequate hardness. However, Zingaro has reason to believe now that the unpolished stones may not be suitab le for printing. Do you think Zingaro is justified in thinking so?

Solution :-

Importing the dataset and getting the first the first 5 values of the dataset as :-
Table :-1

	Unpolished	Treated and Polished
0	164.481713	133.209393
1	154.307045	138.482771
2	129.861048	159.665201
3	159.096184	145.663528
4	135.256748	136.789227

Alpha=0.05, n=75
H0= The stone surface was found to be of adequate hardness
Ha=The unpolished stones not be suitable for printing
H0>=150
Ha<150
Calculating the P value using the formula :- stats.ttest_1samp(mydata,75)=0.0003
Since p value is less than significance level we fail to reject the null hypothesis

5.2) Is the mean hardness of the polished and unpolished stones the same?
Solution :-

Unpolished 134.110527
Treated and Polished 147.788117
The mean hardness of the polished stones is 134.11 and unpolished stones is 147.78. Thus they are not same

Problem 6:

Aquarius health club, one of the largest and most popular cross-fit gyms in the country has been advertising a rigorous program for body conditioning. The program is considered successful if the candidate is able to do more than 5 push-ups, as compared to when he/she enrolled in the program. Using the sample data provided can you conclude whether the program is successful? (Consider the level of Significance as 5%)

Note that this is a problem of the paired-t-test. Since the claim is that the training will make a difference of more than 5, the null and alternative hypotheses must be formed accordingly.

Solution –

Importing the dataset and getting the first the first 5 values of the dataset as :-

Table :-2

	Sr no.	Before	After
0	1	39	44
1	2	25	25
2	3	39	39
3	4	6	13
4	5	40	44

Alpha= 0.05, n=101
H0<=5
Ha>5

Since, it is two paired sample we can calculate P value using T test and we get the value 1.35

Problem 7:

Dental implant data: The hardness of metal implant 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 on the dentists who may favour one method above another and may work better in his/her favourite method. The response is the variable of interest.

7.1 Test whether there is any difference among the dentists on the implant hardness. State the null and alternative hypotheses. Note that both types of alloys cannot be considered together. You must state the null and alternative hypotheses separately for the two types of alloys.

Solution :-

Importing the dataset and getting the first the first 5 values of the dataset as :-

Table :-3

	Dentist	Method	Alloy	Temp	Response	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	Unnamed: 10	Unnamed: 11	Unnamed: 12	Unnamed: 13
0	1.0	1.0	1.0	1500.0	813.0	NaN	NaN	Anova: Two-Factor Without Replication	NaN	NaN	NaN	NaN	NaN	NaN
1	1.0	1.0	1.0	1600.0	792.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	1.0	1.0	1.0	1700.0	792.0	NaN	NaN	SUMMARY	Count	Sum	Average	Variance	NaN	NaN
3	1.0	1.0	2.0	1500.0	907.0	NaN	NaN	1	4	2315	578.75	523721.583333	NaN	NaN
4	1.0	1.0	2.0	1600.0	792.0	NaN	NaN	1	4	2394	598.5	584819	NaN	NaN

For Alloy 1 the hypothesis will be :-

H0-There is no difference among the dentists in implant hardness for Alloy 1

Ha-There is a difference among the dentists in implant hardness for Alloy 1.

For Alloy 2 the hypothesis will be :-

H0-There is no difference among the dentists in implant hardness for Alloy 2.

Ha-There is a difference among the dentists in implant hardness for Alloy 2.

7.2 Before the hypotheses may be tested, state the required assumptions. Are the assumptions fulfilled? Comment separately on both alloy types.?

Solution :-

The assumptions for conducting a statistical test would typically include :
Independence : The observations within each group are independent