|  |  |
| --- | --- |
| A picture of a winding road and trees  Advance Statistics Project  **(2022-2023)**  McCombs School of Business | **Course Name**  **Post Graduate Program in**  **Data Science and Business Analytics**  **Batch Id**  **(PGP-DSBA-June22C)**  ***Submitted by***  ***Jayant Singh*** |

TABLE OF CONTENTS

|  |  |  |
| --- | --- | --- |
| **S.No** | **Topics** | **Page No.** |
| 1 | **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 | 2 |
| 2 | **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. | 3 |
| 3 | **Problem 3** The breaking strength of gunny bags used for packaging cement is normally distributed with a mean of 5 kg per sq. centimetre and a standard deviation of 1.5 kg per sq. centimetre. 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; | 4 - 5 |
| 4 | **Problem 4** Grades of the final examination in a training course are found to be normally distributed, with a mean of 77 and a standard deviation of 8.5. Based on the given information answer the questions below. | 6 - 7 |
| 5 | **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 | 8 |
| 6 | **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%) | 9 |
| 7 | **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. | 10 - 11 |

1

**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?***

**Ans: 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?***

**Ans: Probability that a player is a forward or a winger: 52.34%**

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

**Ans: Probability that a randomly chosen player plays in a striker position**

**and has a foot injury: 19.15%**

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

**Ans: probability that a randomly chosen injured player is a striker: 31.03%**

***# - 1.5 What is the probability that a randomly chosen injured player is either a forward or an attacking midfielder?***

**Ans: probability that a randomly chosen injured player is either a forward or an attacking midfielder : 55.17 %**

2

**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%.



**Let’s Define the events:**

F =Fire

M = Mechanical Failure

H = Human Error

R = Radiation Leak

N = No Accident

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

1. Ans : The probability that it has been caused by a Fire = 0.005
2. Ans: The probability that it has been caused by a mechanical failure = 0.003
3. Ans: the probability that it has been caused by a human error=0.012

**2.2 What is the probability of a radiation leak?**

Ans : The probability of a radiation leak = 0.0037

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

**1 Fire.**

**2 Mechanical Failure.**

**3 Human Error.**

**Answer**

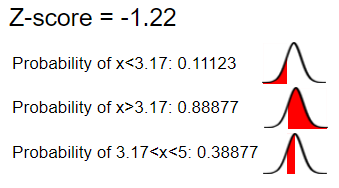
1. The probability that it has been caused by a **Fire** = 0.2702702702702703
2. The probability that it has been caused by a **Mechanical failure** = 0.4054054054054054
3. The probability that it has been caused by a **Human error** = 0.32432432432432434

3

**Problem 3**

The breaking strength of gunny bags used for packaging cement is normally distributed with a mean of 5 kg per sq. centimetre 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?



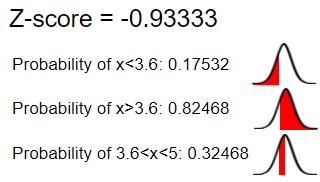


|  |
| --- |
|  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Z score= | xi-μ | 3.17-5 | -1.22 | =0.88877 |
| σ | 1.5 |

**The gunny bags will have 88% breaking strength less then 3.17 kg per sq cm**

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

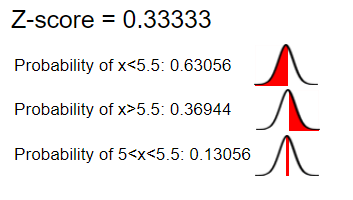


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Z score= | xi-μ | 3.6-5 | 0.9333 | =0.32468 |
| σ | 1.5 |

**The gunny bags will have 32% breaking strength less then 3.6 kg per sq cm**

4

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

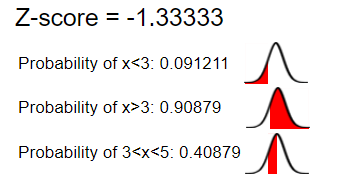


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Z score= | xi-μ | 5.5-5 | 0.333 | =0.13056 |
| σ | 1.5 |

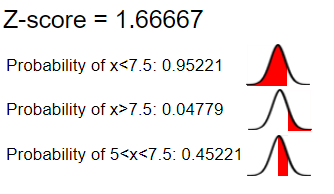
1

**The gunny bags will have 13% breaking strength between 5 & 5.5kg per sq**

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



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Z score= | xi-μ | 3-5 | -1.3333 | =0.091211 |
| σ | 1.5 |



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Z score= | xi-μ | 7.5-5 | 1.66667 | =0.95221 |
| σ | 1.5 |

0.95221 – 0.091211 = 0.860099 = 1-860099 = 0.139001 = 13%

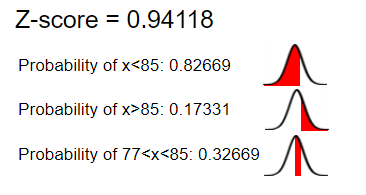
**The gunny bags will have 13% breaking strength NOT between 3 & 7.5kg per sq cm**

5

**Problem 4**

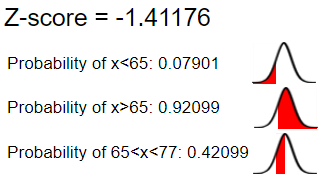
Grades of the final examination in a training course are found to be normally distributed, with a mean of 77 and a standard deviation of 8.5. Based on the given information answer the questions below.

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

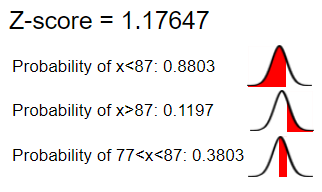


Ans: The Probability of 94% students will get below 85 on this exam

4.2 What is the probability that a randomly selected student score between 65 and 87?



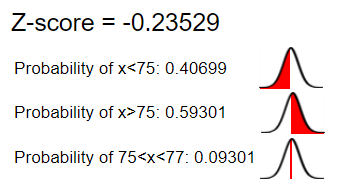
6



Answer:

80.12% is the probability that a randomly selected student score between 65 and 87.

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



Ans : 59% marks is the passing cut-off so that 75% of the students clear the exam.

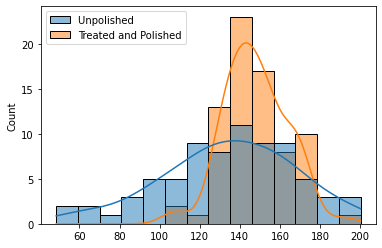
7

**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 favourable 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?

Ans: True

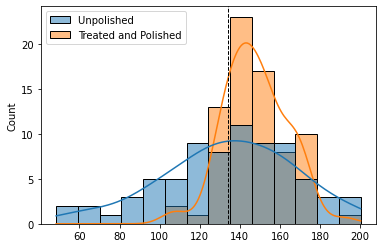
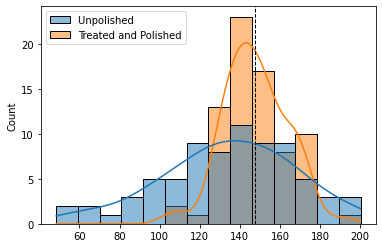


***As the stone having value is greater then 130 are polished & Treated stones in count with respect to Unpolished hence we can reject the null hypothesis.***

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

Ans: No, as we can see below that the mean has significant difference

|  |  |
| --- | --- |
| **Unpolished** | **Treated and Polished** |
| 134.11053 | 147.788117 |

This dotted line in the given figures represents MEANS of Unpolished & Treated and Polished hence we can reject the null hypothesis.

8

**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.

 Ans: tstat -19.323 P-value for one tail test: 1.1460209626255983e-35

Since our Alpha value is greater then our P value we can say that we have to reject NULL hypothesis and except alternate hypothesis hence there is a significant increase in the strength of the person after the program hence we will reject the null hypothesis as programme is successful

9

**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.?**  **Ans :**  **p value:0.000000**  **Reject null hypothesis** |  |
| **7.2 Before the hypotheses may be tested, state the required assumptions. Are the assumptions fulfilled? Comment separately on both alloy types.?**  **Ans:**  H0 : μ1=μ2=μ3 or The mean of the samples is the same.  H1 : At least one of them is different. |  |
| **7.3 Irrespective of your conclusion in 7.2, 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?**  **Ans**  p value:0.000000  Reject null hypothesis  **From the above result we can conclude that there is a significant change due to change in Dentist of metal** |  |
| **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?**  **Ans:**  p value:0.000000  Reject null hypothesis  No  **10** |  |
| **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?**  **Ans:**  p value:0.000000  Reject null hypothesis  **No** |  |
| **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, and which interaction levels are different?**  **Ans: I can not attend as there is no possibility of taking out any result** |  |
|  |  |

11