X

NPTEL

Courses » **Introduction to Data Analytics**

Announcements   **Course**   Ask a Question   Progress

Unit 4 - Week 3 - Inferential Statistics

Course outline

How to access the portal

Week 1 - Course Overview and Descriptive Statistics

Week 2 - Probability Distributions & Inferential Statistics

Week 3 - Inferential Statistics

○ Two Sample tests

○ Type 1 and Type 2 Errors

○ Confidence Intervals

○ ANOVA and Test of Independence

○ Short Introduction to Regression

○ Quiz : Assignment 3

○ Feedback for week 3

○ Assignment 3: Solution

Week 4 - Machine Learning

Week 5 - Supervised Learning (Regression and Classification Techniques) - I

Week 6 : Supervised Learning (Regression and Classification Techniques)-II

Week 7 - Association Rule Mining and Big Data

Week 8 - Clustering Analysis and Prescriptive Analytics

Course Summary+ Insight into the Final Exam

Assignment 3

The due date for submitting this assignment has passed.   **Due on 2017-08-16, 23:55 IST**
As per our records you have not submitted this assignment.

1) Based on the data generated from two different approaches to producing tooth paste, you need to decide   *1 point*
which one to select. To do this task, you will choose techniques from

- ○ descriptive statistics
- ○ inferential statistics
- ○ predictive analytics

No, the answer is incorrect.
Score: 0

Accepted Answers:
inferential statistics

2) In which among the following two sample tests, can the number of data points in the two samples differ?   *1 point*
(Note: More than one options can be correct)

- ☐ two sample z-test
- ☐ two sample t-test
- ☐ paired t-test
- ☐ F-test

No, the answer is incorrect.
Score: 0

Accepted Answers:
two sample z-test
two sample t-test
F-test

3) A study was conducted to test the effect of a special training program over employees. Each employee was   *1 point*
given a test twice, both before and after completing the training program. Let ΔX denotes the difference between the first
and second test scores of each employee. It means, if mean of ΔX is zero, the training program has no effect on the
average. The data of 20 employees has been recorded and provided here in the table.

Employee - ID	Pre-training score	Post-training score	ΔX
1	18	22	+4
2	21	25	+4
3	16	17	+1
4	22	24	+2
5	19	16	-3
6	24	29	+5
7	17	20	+3
8	21	23	+2
9	23	19	-4
10	18	20	+2
11	14	15	+1
12	16	15	-1
13	16	18	+2
14	19	26	+7
15	18	18	0
16	20	24	+4
17	12	18	+6
18	22	25	+3
19	15	19	+4
20	17	16	-1

Your task is to test the hypothesis of no effect against the hypothesis of positive effect. What will be the null and alternative hypothesis?

○ $H_0 : \mu = 0; H_1 : \mu \neq 0$

○ $H_0 : \mu = 0; H_1 : \mu < 0$

○ $H_0 : \mu \geq 0; H_1 : \mu < 0$

○ $H_0 : \mu = 0; H_1 : \mu > 0$

No, the answer is incorrect.
Score: 0

Accepted Answers:
$H_0 : \mu = 0; H_1 : \mu > 0$

4) In the previous question, the value of the test statistic is: *1 point*

○ 2.05
○ 2.837
○ 3.231
○ 0.634

No, the answer is incorrect.
Score: 0

Accepted Answers:
3.231

5) Using the test statistic calculated in previous question and appropriate degree of freedom, make a decision   *1 point* with $\alpha = 0.05$. The decision is:

(Hint: use the following
z-table: www.stat.ufl.edu/ athienit/Tables/Ztable.pdf or
t-table: www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf or
f-table: www.stat.purdue.edu/ jtroisi/STAT350Spring2015/tables/FTable.pdf as required).

○ reject H_0 and conclude that effect is positive

○ failed to reject H_0H_0 and conclude that there was no effect

○ cannot make any inference as the data is incomplete

No, the answer is incorrect.
Score: 0

Accepted Answers:
reject H_0 and conclude that effect is positive

6) If the decision you made in the previous question is incorrect, what type of error has been made?   *1 point*

○ type I error
○ type II error
○ type IV error
○ both (a) and (b)

No, the answer is incorrect.
Score: 0

Accepted Answers:
type I error

7) Suppose that a sample (shown in table) of n = 5 was selected from the price of properties sold by The New   *1 point*
Okhla Industrial Development Authority in the National Capital Region of India.

Year	Price of 5 samples (Rs./acre)				
2000	30000	34000	36000	38000	4000
2001	30000	35000	37000	38000	4000
2002	40000	41000	43000	44000	5000

Your task is to conduct ANOVA over this data to check whether you get evidence that prices over the land were not same for the three years considered. The F-statistic for the given data is:

○ 8.96
○ 4.312
○ 6.834
○ 1.337
○ none of the above

No, the answer is incorrect.
Score: 0

Accepted Answers:
6.834

8) By using the data given in previous question, can we reject the null hypothesis (null hypothesis is :   *1 point*
$\mu_2 000 = \mu_2 001 = \mu_2 002$) at 0.01 level? What about 0.05 level? (Hint: use the following
z-table: www.stat.ufl.edu/ athienit/Tables/Ztable.pdf or
t-table: www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf or
f-table: www.stat.purdue.edu/ jtroisi/STAT350Spring2015/tables/FTable.pdf as required).

○ reject the null hypothesis at 0.01 level; reject the null hypothesis at 0.05 level
○ reject the null hypothesis at 0.01 level; failed reject the null hypothesis at 0.05 level
○ failed to reject the null hypothesis at 0.01 level; reject the null hypothesis at 0.05 level
○ failed to reject the null hypothesis at 0.01 level; failed to reject the null hypothesis at 0.05 level

No, the answer is incorrect.
Score: 0

Accepted Answers:
failed to reject the null hypothesis at 0.01 level; reject the null hypothesis at 0.05 level

9) What is the purpose of a multiple regression?   *1 point*

○ To predict scores on a dependent variable from scores on multiple independent variables.
○ To predict scores on an independent variable from scores on multiple dependent variables
○ To assess whether there is a significant difference between repeated measures

○ To predict scores on a dependent variable from scores on a single independent variable

○ To predict scores on an independent variable from scores on a single dependent variable

○ To assess whether there is a significant difference between independent groups

No, the answer is incorrect.
Score: 0

Accepted Answers:
To predict scores on a dependent variable from scores on multiple independent variables.

10) For a chi-square test, a 4 × 5 contingency table will have how many degrees of freedom?   *1 p*

○ 12

○ 8

○ 9

○ 6

No, the answer is incorrect.
Score: 0

Accepted Answers:
12

Previous Page   End