

Assignment 3

Due on 2015-08-09, 23:55 IST

Submitted assignment

1) Which statistical test/ technique would you use?

1 point

We are an online shopping portal. We can tell if someone who is on our website is a mac user or a PC user. Test the hypothesis that the ratio of people who buy out of those who logon (as a percentage) is higher for mac users than PC users

- ☐ Two sample t-test, unpaired, unequal variances
- ☒ Two sample proportion z- test
- ☐ One sample proportion z- test
- ☐ Single variable regression

2) What statistical test/ technique would you use?

1 point

We are an online shopping portal. We can tell if someone who is on our website is a mac user or a PC user. We want to test the hypothesis that among people who purchase something from our website mac users tend to spend more money than PC users.

- ☒ Two sample, unpaired t-test, unequal variances
- ☐ Two sample paired t-test
- ☐ Two sample z-test
- ☐ One sample Chi-square test

3) What statistical test/ technique would you use? **1 point**

We are interested in understanding the difference between siblings. Specifically, we would like to study cases where there are non-twins but only two children in the family. We intend to test the hypothesis that the IQ of the elder sibling is equal to younger one.

- ☐ Two sample unpaired t-test, unequal variances
- ☒ The paired t-test
- ☐ ANOVA
- ☐ The randomization test

4) Which statistical test/ technique would you use? **1 point**

There are 6 sales managers who are deployed to six different cities. They are randomly rotated every 3 months, but it is ensured that each manager has only one stint in each city. After six such rotations each manager has been to each city once. Their sales numbers (revenue generated in Rupees) need to be compared to determine if the managers performances are statistically different from each other.

- ☐ Regression
- ☐ Classification
- ☒ ANOVA
- ☐ Randomization test

5) Which statistical test/technique would you use? **1 point**

Two different weight loss drugs are tried on participants who come in with a certain disease. Each participant is randomly assigned one of the two drugs. The reduction in weight (as a percentage) is recorded for each participant. We would like to see if one drug is more effective than the other.

- ☒ Two sample unpaired t-test
- ☐ The paired t-test
- ☐ Two sample proportion test
- ☐ Randomization test

6) Which statistical test/technique would you use? **1 point**

We would like to test the hypothesis that the distance from a certain mall/plaza in a particular city affects the price/square foot of houses within a 10 kilometer radius of the mall.

- ☐ Two sample unpaired t-test
- ☐ Single sample t-test
- ☒ Simple regression
- ☐ Chi-Square test of independence

7) Which statistical test/technique would you use? **1 point**

We are interested in analyzing the 2011 by-elections in Tamil Nadu which fielded two popular parties/coalitions. You are a psephologist who is asked to see if there is any relationship between gender and the affiliation towards either of the two parties. However, you have no historical data and need to conduct a survey. You ask 100 randomly sampled voters and decide to do a statistical analysis to determine if there is any relationship.

- ☒ Single sample proportion z-test
- ☐ Simple regression
- ☐ Paired t-test
- ☐ Chi-squared test of independence

8) In a simple linear regression we use statistical inference to: **1 point**

- ☐ Fine tune the value of the coefficients in the linear model
- ☒ Ensure that the unexplained error captured in the Mean Squared Error (MSE) is minimized
- ☐ Test the hypothesis that each co-efficient in the linear model is statistically different from 0
- ☐ Establish the intercept and slope values in the linear fit based on some objective for fitting a line through data

9) In a simple linear regression we use optimization to: **1 point**

- ☐ Determine if the explanatory variable should be included in the model or not.
- ☐ Determine if a constant (in the form of an intercept) is needed or not.
- ☐ Test the hypothesis that each co-efficient in the linear model is statistically different from 0.

- ☒ Establish the intercept and slope values in the linear fit based on some objective for fitting a line through the data.

10) A single sample t-test was conducted to test the hypothesis that $H_0: \mu_0=13$. The data analysis resulted in the discovery of a sample mean of 14.5 and a p-value of 0.05. If we selected an $\alpha(\text{alpha})=0.1$, We would

1 point

- ☒ Reject the null hypothesis
- ☐ Accept the null hypothesis
- ☐ Fail to reject the null hypothesis
- ☐ Accept the alternate hypothesis

11) A single sample t-test was conducted to test the hypothesis that $H_0: \mu_0=13$. The data analysis resulted in the discovery of a sample mean of 14.5 and a p-value of 0.05. The 95% confidence interval around the mean would be

1 point

- ☐ 11.5 to 14.5
- ☐ 13 to 14.5
- ☐ 12.5 to 14.5
- ☒ 13 to 16

12) We run a statistical test to see if the user satisfaction of two different cell phone UIs (User Interface) are equal. We fail to reject the null hypothesis (that they are equal). We want to increase the probability of finding any difference if there is one. In principle we want to:

1 point

- ☐ Reduce Type 1 error
- ☒ Reduce Type 2 error
- ☐ Increase Type 1 error
- ☐ Reduce power