Collin Real (yhi267)

DA 6823

Kilger

Module 1: Part #1 (50 points)

**The Power of Statistics + the Levels of Measurement + the Different Classes of Variables and Determining Appropriate Statistical Technique + Basic Descriptive Measures**

**General Instructions:** In your own words, answer each of the following questions - don’t copy (e.g. cut and paste) some definition out of a book word for word. This is not a group project – you are expected to complete this module on your own. You may refer to text books, online or other sources but not your fellow classmates. If you don’t understand the question, feel free to ask the instructor in class, in office hours or in an email.

1. Provide a short definition for dependent variable. (3 points)

The response variable being measured, modeled or predicted. The dependent variable’s value depends on independent variables and may change based on these variables.

1. Provide a short definition for independent variable. (3 points)

The explanatory variable that is being manipulated. The independent variable’s value stands alone and does not change based on other variables.

1. Provide a short definition for control variable. (3 points)

A variable held constant and unchanged in an analysis to assess the relationship between the independent and dependent variables. Control variables do not influence the outcomes and are not the primary concern in an analysis.

1. Be able to describe the simple criteria for each of the four levels of measurement:
   1. Nominal (2 points)

Categorized data only

* 1. Ordinal (2 points)

Categorized and ranked data

* 1. Interval (2 points)

Categorized, ranked, and evenly spaced data

* 1. Ratio (2 points)

Categorized, ranked, equal intervals, and there is a natural zero point

1. Provide an example of a variable for each of the four measurement levels below.
   1. Nominal (2 points)

Gender

* 1. Ordinal (2 points)

Language ability (beginner, intermediate, advanced)

* 1. Interval (be careful – be sure it is interval and not ratio!) (2 points)

Temperature (Fahrenheit or Celsius)

* 1. Ratio (2 points)

Age

1. Name at least two criteria from the IDRE chart that are used in determining which statistical technique can be used in a situation. (3 points)

The number of dependent variables, the nature of independent variables, the nature of dependent variables, and assumptions of the test.

1. Briefly explain the difference between descriptive and inferential statistics. (4 points)

Descriptive statistics describe the features of populations/samples, factually organize and draw conclusions based on known data, and present results visually via graphs, tables, or charts. Inferential statistics make generalizations/predictions about larger populations based on a sample of data, draw conclusions that go beyond this sample data, and present these conclusions in probabilities.

1. Almost every statistical technique you will come across has some sort of assumptions – even non-parametric statistics.
   1. Name one benefit of that assumptions of a test provides you (2 points)

One benefit of the normality assumption for t-tests is that it identifies useful properties. These properties include the data being symmetrical around the mean, the mean, median, and mode are equal, and the standard deviation is meaningful.

* 1. Name one cost that assumptions of a test carry (2 points)

If assumptions of a test are not checked and end up being violated, the results of the analysis may not be accurate. For example, violating the normality assumption can lead to biased estimates and p-values.

1. What happens if you violate the assumptions of a statistical test? Do the statistical police come and arrest you? (4 points)

The statistical police do not arrest you (yet), but the validity of the results of the study are threatened. Less faith is put in the findings of the study due to the increased likelihood of the results being erroneous.

1. Using the IDRE chart, suggest the appropriate statistical test for each of the following business cases
   1. As a maker of colored contact lenses, you think that there may be relationship between the color of the contact lenses purchased and the gender of the purchaser. (2 points)

Chi-square test

* 1. As an auctioneer of fine art, you think that there may be a different between the price paid for a piece of art between men and women. (2 points)

Two-sample t-test

* 1. You want to better understand how different versions and price mixes of your product – the Vegematic – have on the number of product sold. You hypothesize that color of product, price, region of the country (North, South, East, West), gender of purchaser, household income of purchaser have an effect on the number of pieces sold. You may also want to make some predictions about how many products would be sold under various levels of these variables. (2 points)

Multiple regression analysis

* 1. As publisher of the popular magazine Rabbit Times, you think that there may be a relationship between the number of pages in the magazine and the number of copies of that issue sold. How do you find out the direction and how strong this relationship might be? (2 points)

Correlation analysis

* 1. You are the maker of FelineHair – a hair growing drug for hairless cats. You want to test your drug against three other drugs to see which one grows the most hairs on the cats in the experiment. You also want to see if there are other differences in the effectiveness depending upon the gender of the cat and what color coat the cat has. You end up with a drug (4) x cat gender (2) by cat coat color (black, white, brown) experimental design. What analysis technique would you use for this experiment? (2 points)

Factorial logistic regression