

## SPSS Lab 5 – More Plots and T -Tests

### Plots and Charts

This week we will analyze plots and charts. Use the same file “birthweights” as in lab 4. As before answer questions in the body of the email.

1. Create a new field “MAGE” with categorical data that describes the age of the mother. Categorical data has a limited number of distinct values or categories (for example, gender or religion). Categorical variables can be string (alphanumeric) or numeric variables that use numeric codes to represent categories (for example, 0 = male and 1 = female).
  - a. The new field “MAGE” should contain 4 different values depending on the mother’s age:
    - i. 13-20 Years,
    - ii. 21-30,
    - iii. 31-35,
    - iv. 36-55,We will use a new method to create this field, but what method do you know that you can use to create this new field? Describe it!
  - b. Let’s go to Transform\Visual Binning. Select mother’s age as the variable to bin. Continue and click on the Mother’s age variable. Enter “MAGE” as the “binned variable” name. Click on Make cutpoints. Enter 20 for the “First Cutpoint location” and then enter 4 for “Number of cutpoints”. Hit apply. Now edit the grid that has values and empty labels. For example enter “<20” as the label for value 20. The other interval ranges are incorrect. Fix them and label them. What about the missing numbers?
  - c. Let’s do a binned field GAGE for gestational gage with 4 intervals. Here you have missing values. You can either assign a value to them using Transform\Compute Variable or filter them out. What would be the “If” condition that you would use in either case?
  - d. We would like to make a bar chart of the *birthweight by mothers age*. Let’s try doing it like this: Go to Graphs\Legacy Dialogs\Bar... Click on Clustered. Well, what do we select now? If you add MAGE to the X-axis and define clusters by birthweight you get an ugly bar chart. Try it. Why?
  - e. What we need is a bar chart where in the x axis we only have a few clustered groups that we want to compare!
    - i. the birthweight need to be grouped! Do it using the visual binning method for variable BWEIGHT?
    - ii. The Y-axis should show the counts.

How do we do that? Write it here as you will need the steps for your lab assignment (you can go to Graphs\Legacy Dialogs\Bar... Click on Clustered. Click on the button help and see if that helps.):

---

---

---

---

---

2. Learn to use SPSS help file and tutorial files. Hit F1 from anywhere in SPSS. The Help file Window will open
  - a. You already checked out the help file. Now, let's look at the Tutorial. Click Help\Tutorial and the Tutorial window will open. Click on "Creating and Editing charts." Click on "Chart Creation Basics." You are now watching the tutorial. Learn about creating bar charts with the Chart builder. You can go to the next page by clicking the arrow at the bottom right corner. If you click the magnifying glass you will get a list of related tutorial topics. Create a bar chart using the "Chart builder"!
  - b. Now look at the tutorial list and pick "Other examples"|"pie charts" and do a pie chart for Variable= GAGE. Describe it!

### 1. One-sample T-test

We will continue to use the birthweight data set.

- a. One sample t-tests can be used to determine if the mean of a sample is different from a particular value. In this example, we will determine if the mean gestational age is different from 39.5.
- b. Write the null and alternative hypotheses in the body of the email.
- c. Run Analyze/Compare Means/One sample t-test and add value 39.5 as the test value and Gestational Age as the test variable. What conclusions can you obtain from the results of this test?
- d. Run c for the mean (test value in SPSS) 39.25. What conclusions can you obtain from the results of this test?
- e. Now let's say we only want to know if the mean gestational age is SMALLER than 39.25. Write the null and alternative hypotheses in the body of the email. Do you need a 1-sided or a 2-sided test now? SPSS does not run 1-tailed hypothesis.
- f. To run a 1-sided test for  $H_0: \mu = \mu_0$   $H_1: \mu < \mu_0$  do the following: Check your results from 1e). The second column of the output gives us the t-test value:  $(38.26 - 39.25) / (11.74 / \text{square root of } 444) = -1.782$  [if you do the calculation, the values will not match exactly because of round-off error). The fourth column tells us the *two-tailed* significance (the 2-tailed p value.) But we didn't want a two-tailed test; our hypothesis is one tailed and there is no option to specify a one-tailed test. Because this is a one-tailed test, look in a table of critical t values to determine the

critical t. The critical t with 443 degrees of freedom,  $\alpha = .05$  and one-tailed is 1.648.

- g. Determine if we can reject the null hypothesis or not. The decision rule is: if the one-tailed critical t value is more than the observed absolute value of t, then we can reject  $H_0$  (page 231 in Biostatistics book). In this example, the critical t is -1.648 (from Excel) > the observed  $t = -1.782$ . What is the decision and what does it mean with regards to our hypothesis about gestational age? How does this decision compare with the 2-tailed test results? How do you explain this?

### Lab Assignment:

Open the file testgrades.sav with the scores of last year's biostatistics tests. Follow the guidelines from the demo above to run the analysis described below, ***answer all questions from above(!) that are relevant.***

1. Do a bar chart showing the scores distribution over the three tests. The binned graded variable should contain 1="A" for 27-30, 2="B" for 24-26, 3="C" for 21-23, 4="D" for 19-20, 5="F" for the rest. Discuss the chart.
2. Run a t-test to find out if the mean of last year's first test is significantly LOWER than the mean of this year's first test, which is 22.5 out of 30 points. Also answer this question: Is the mean significantly DIFFERENT than the mean of this year first test? What mean score should the students have obtained this year to claim that their test mean score is significantly different to last year's test?