

BUS 230: Business and Economics Research and Communication

Instructor: Dr. James Murray

SPSS Handout: Statistical Significance and Univariate and Bivariate Tests

1 Univariate Tests

1.1 One-Sample T-test of a Mean

Example: Public School Spending

- Dataset: average pay for public school teachers and average public school spending per pupil for each state and the District of Columbia in 1985.
- Download dataset `eduspending.sav`.
- Conduct the following exercises:
 - Show some descriptive statistics for teacher pay and expenditure per pupil.
 - Is there statistical evidence that teachers make less than \$25,000 per year?
 - Is there statistical evidence that expenditure per pupil is more than \$3,500?

Opening the data:

1. Save `eduspending.sav` somewhere.
2. Open SPSS.
3. Click radio button `Open an existing data source`.
4. Double-click `More files...`
5. Go find and double click `eduspending.sav`.

Descriptive Statistics

1. Click `Analyze` menu, select `Descriptive Statistics`, then select `Descriptives`.
2. Click on `Pay` and click right arrow button.
3. Click on `Spending` and click right arrow button.
4. Click `Options`.
 - (a) Check any options you find interesting.
 - (b) Click `OK`
5. Click `OK`

Test Hypotheses

1. Click **Analyze** menu, select **Compare Means**, then select **One-Sample T test**.
2. Select **Pay**, then click right arrow.
3. Enter in Test Value text box 25000.
4. Output tables show descriptive statistics for pay, and hypothesis test results.

1.2 One-Sample Tests for Proportions

Example: Economic Outlook

- Data from Montana residents in 1992 concerning their outlook for the economy.
- All data is ordinal or nominal:
 - AGE = 1 under 35, 2 35-54, 3 55 and over
 - SEX = 0 male, 1 female
 - INC = yearly income: 1 under \$20K, 2 20-35\$K, 3 over \$35K
 - POL = 1 Democrat, 2 Independent, 3 Republican
 - AREA = 1 Western, 2 Northeastern, 3 Southeastern Montana
 - FIN = Financial status 1 worse, 2 same, 3 better than a year ago
 - STAT = 0, State economic outlook better, 1 not better than a year ago
- Do the majority of Montana residents feel their financial status is the same or better than one year ago?
- Do the majority of Montana residents have a more positive economic outlook than one year ago?
- Open the dataset *econoutlook.sav*.
- Parametric approach:
 1. If the variable is a zero or one, the *sample mean* is the same as the *proportion of the sample* that has a value equal to 1.
 2. Convert Financial Status (FIN) to a 0 if worse, and 1 if same or better:
 - (a) Click **Transform** menu, select **Recode into Different Variables**.
 - (b) Select FIN on left and click right arrow button.
 - (c) Click **Old and New Values** button.
 - (d) First transform FIN=1 into value 0:
 - i. On the left under **Old Value**, click radio box for **Value**. In textbox enter 1.
 - ii. On the right under **New Value**, click radio box for **Value**. In textbox enter 0.

- iii. Click **Add** button.
 - (e) Next transform FIN=2 or FIN=3 into value 1:
 - i. On the left under **Old Value**, click radio box for **Range**. In textboxes enter 2 and 3.
 - ii. On the right under **New Value**, click radio box for **Value**. In textbox enter 1.
 - iii. Click **Add** button.
 - (f) Click **Continue**
 - (g) In original Recode window, under **Output Variable**, type a name.
 - (h) Click **Change**
 - (i) Click **OK!**
- 3. Do a simple T-Test!
- Binomial approach:
 1. Independent observations that are a 0 or 1 have a binomial distribution (regardless of sample size).
 2. Click **Analyze** menu, select **Nonparametric Tests**, select **Legacy Dialogs**, then select **Binomial**.
 3. On the left, select variable to test and click right arrow button.
 4. In Test Proportion, enter 0.5.
 5. Click **OK!**

1.3 One-Sample Nonparametric Tests for Medians

Example: Attitudes Grade School Kids

- Dataset: 438 students in grades 4 through 6 were sampled from three school districts in Michigan. Students ranked from 1 (most important) to 5 (least important) how important grades, sports, being good looking, and having lots of money were to each of them.
- Open dataset `gradschools.sav`. Choose second worksheet, titled **Data**.
- Answer some of these questions:
 - Is the median importance for grades is greater than 3?
 - Is the median importance for money less than 3?
- 1. Click **Analyze** menu, select **Nonparametric Tests**, select **Legacy Dialogs**, then select **Binomial...**
- 2. Click on **Grades** (or a different variable of interest), then click on right arrow.
- 3. Click radio button for **Cut point** and enter 3 into text box.
- 4. Do you want the exact (binomial distribution) p-value or asymptotic distribution (normal distribution)?

- (a) Exact: click on **Exact...**
 - (b) Click **Exact** radio button.
 - (c) Click **Continue**.
5. Click **OK**
- Output table shows exact p-value and normal distribution p-value for a two-tailed test.
 - What is your conclusion?

2 Bivariate Statistics

2.1 Independent Samples T-Test for Differences in Means

Example

- Dataset: average pay for public school teachers and average public school spending per pupil for each state and the District of Columbia in 1985.
 - Test the following hypotheses:
 - Does spending per pupil differ in the North (region 1) and the South (region 2)?
 - Does teacher salary differ in the North and the West (region 3)?
 - Do you see any weaknesses in our statistical analysis?
1. Open `eduspending.sav` in SPSS.
 2. Click **Analyze** menu, select **Compare Means**, then select **Independent-Samples T test**.
 3. Select **Pay** or **Spending**, depending on which you are currently interested in.
 4. Click the right arrow that is just to the left of **Test Variables**.
 5. Select **Area** and click on right arrow to the left of **Grouping Variable**.
 6. You need to tell SPSS what your grouping variable means and what groups you are interested in:
 - (a) Click on **Define Groups**
 - (b) Click radio button **Use specified values**.
 - (c) Enter in the appropriate numbers for Group 1 and Group 2 (i.e. if you want the North to be group 1, type a 1 in Group 1 text box, and if you want the West to be group 2, type a 3 in the Group 2 text box).
 - (d) Click **Continue**.
 7. Click **OK!**

- The first output table shows some descriptive statistics for each group.
- The next output table shows:
 - Statistical evidence about whether the variances are different.
 - Statistical evidence about whether the means are different.
 - Descriptive statistics about the difference in the means.
 - Confidence intervals for the difference in the means.

2.2 Mann-Whitney Test: Independent Samples Nonparametric Test for Difference in Medians

1. Use the same Education Spending dataset as above.
 2. Click **Analyze** menu, select **Nonparametric Tests**, select **Legacy Dialogs**, then select **2 Independent Samples...**
 3. Move **Pay** (or whatever you are interested in) into **Test Variable List**
 4. Move **Area** into **Grouping Variable**
 5. Define groups as before.
 6. You can get exact p-values if absolutely necessary (takes more time).
 7. Click **OK**.
- The value of the Mann-Whitney test statistic can be huge and this does not have much interpretation (it's equal the smaller of the two sums of ranks).
 - The Significance is the p-value.
 - What is your conclusion?

2.3 Paired Samples T-test for a Difference in Means

- Dataset: Job Satisfaction measures for Israeli public workers.
 - Includes information on a number of measures of job satisfaction including, (1) overall, (2) colleagues, (3) supervision, (4) salary, and (5) promotion.
 - Includes background information like ethnicity, years experience, and marital status.
 - Includes workplace information like location and workload.
- Open dataset **JobSatisfaction.sav**.
- Answer some of these questions:
 1. Are employees' mean level of satisfaction different for colleagues versus supervision?

2. Are employees' mean level of satisfaction different for salary versus overall?

1. Click **Analyze** menu, select **Compare Means**, then select **Paired Samples T Test**.
 2. Move **Colleague** and **Supervision** into the **Paired Variables** list.
 3. Click **OK**.
- The first table of output shows the sample means for **Colleague** and **Supervision**.
 - The second table shows the correlation between the two variables - ignore this.
 - The third table shows:
 - The average difference between the two variables.
 - The standard deviation of the difference.
 - A confidence interval for the difference.
 - A t-statistic (similar to a z-statistic) for the null hypothesis that the difference=0.
 - A p-value for the null hypothesis that difference=0, versus the alternative hypothesis that the difference $\neq 0$.

2.4 Wilcoxon Signed Rank Test for a Difference in Medians

- Dataset: sample **JobSatisfaction.sav**.
 - Answer one of these questions:
 - Are employees' median levels of satisfaction different for colleagues versus supervision?
 - Are employees' median level of satisfaction different for salary versus overall?
1. Click **Analyze** menu, select **Nonparametric Tests**, select **Legacy Dialogs**, select **2 Related Samples**.
 2. Move **Colleague** and **Supervision** into the **Test Pairs** list.
 3. Make sure **Wilcoxon** is checked.
 4. Click **OK**.
- Null hypothesis: Medians are the same.
 - Alternative hypothesis: Medians are different.
 - P-value is given in the last row of the last table.