These are the answers to the review exercises for Exam #2.
1. The mean breaking strength for each of three types of rope (named Deluxe, Econ, and Nogood) is being studied. A researcher is interested in investigating a possible relationship between breaking strength and rope type. A 0.10 significance level is chosen for all hypothesis testing. Pieces of rope are randomly selected from each type, and the breaking strengths in pounds of force are recorded as follows:

Deluxe 165 162 159 162
Econ 156 163 158
Nogood 151 154 160

Complete the following table by listing each variable involved in the analysis in the first column, indicating in the second column whether each variable should be treated as qualitative or quantitative, and indicating in the third column whether each variable is a dependent variable, independent variable, or neither:

<table>
<thead>
<tr>
<th>Variable(s)</th>
<th>Type</th>
<th>Dependent / Independent / Neither</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Rope</td>
<td>Qualitative</td>
<td>Independent Variable</td>
</tr>
<tr>
<td>Breaking Strength</td>
<td>Quantitative</td>
<td>Dependent Variable</td>
</tr>
</tbody>
</table>

Use SPSS to perform the necessary calculations for the appropriate parametric hypothesis test, and give the name of the hypothesis test selected.

one-way analysis of variance
State the null and alternative hypothesis, and indicate whether the test performed should be one-sided or two-sided.

$H_0$: the mean breaking strength is equal for all rope types
(or there is no relationship between rope type and breaking strength)

$H_1$: the mean breaking strength is different for at least one rope type
(or there is a relationship between rope type and breaking strength)

The one-way ANOVA $f$ test is always one sided

State the assumption(s) that must be satisfied for the hypothesis test.

The data consist of independent random samples selected from normal distributions all having the same variance

Circle which of the following graphical displays is most appropriate.

- pie chart
- bar chart
- stacked bar chart
- box plot(s)
- scatter plot

Write the results of the selected hypothesis test in a format that is suitable for a paper that you are submitting for publication; be sure to include relevant information about the test statistic, sample statistics, and $p$-value, and an appropriate summary of follow-up analyses, if necessary.

At the 0.05 level, Levene’s $f$ test showed no statistically significant difference in standard deviation of breaking strength among rope types Deluxe, Econ, and Nogood ($f_{2, 7} = 0.951, f_{2, 7; 0.10} = 3.26, 0.10 < p \ OR \ p = 0.431$). The results of Levene’s $f$ test suggest that the one-way ANOVA $f$ test can be used.

The one-way ANOVA $f$ test was statistically significant at the 0.05 level ($f_{2, 7} = 3.419, f_{2, 7; 0.10} = 3.26, 0.05 < p < 0.10 \ OR \ p = 0.092$). We conclude that mean breaking strength is different for at least one rope type.

From Bonferroni’s method with a 0.10 significance level, we find that the mean breaking strength for Deluxe rope (mean = 162 lbs., $n = 4$) is statistically significantly higher than for Nogood rope (mean = 155 lbs., $n = 3$). No statistically significant difference was found between in mean breaking strength between Econ rope (mean = 159 lbs., $n = 3$) and either of the other two rope types.

Indicate what alternative procedures, if any, could be used in the event that required assumptions for the selected hypothesis test were not satisfied.

Kuskal-Wallis rank sum test
2. Randomly selected customers of a car dealership are polled in order to collect data concerning area of residence and preferred color of a certain type of sports car. For each customer, an area of residence is recorded as rural, suburban, or urban, and color preference is recorded chosen from red, green, yellow, and white. The data are stored in the SPSS data file sportscar. The data are to be used to perform a hypothesis test to see if there is any evidence of an association between area of residence and preferred color of the sports car. A 0.05 significance level is chosen.

Complete the following table by listing each variable involved in the analysis in the first column, indicating in the second column whether each variable should be treated as qualitative or quantitative, and indicating in the third column whether each variable is a dependent variable, independent variable, or neither:

<table>
<thead>
<tr>
<th>Variable(s)</th>
<th>Type</th>
<th>Dependent / Independent / Neither</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area of Residence</td>
<td>Qualitative</td>
<td>Independent Variable</td>
</tr>
<tr>
<td>Car Color Preference</td>
<td>Qualitative</td>
<td>Dependent Variable</td>
</tr>
</tbody>
</table>

Use SPSS to perform the necessary calculations for the appropriate hypothesis test, state whether this hypothesis test is considered parametric or nonparametric, and give the name of the hypothesis test selected.

**chi-square test for association, which is considered a nonparametric test**
State the null and alternative hypothesis, and indicate whether the test performed should be one-sided or two-sided.

**H₀**: there is no association between area of residence and car color preference (or color preference distribution is the same for each area of residence)

**H₁**: there is an association between area of residence and car color preference (or color preference distribution is not the same for each area of residence)

The chi-square test for association is always one-sided

State the assumption(s) that must be satisfied for the hypothesis test.

Each of the expected frequencies is greater than 5.

Circle which of the following graphical displays is most appropriate.

- pie chart
- bar chart
- stacked bar chart
- box plot(s)
- scatter plot

Write the results of the selected hypothesis test in a format that is suitable for a paper that you are submitting for publication; be sure to include relevant information about the test statistic, sample statistics, and p-value, and an appropriate summary of follow-up analyses, if necessary.

The chi-square test for association was not statistically significant at the 0.05 level ($\chi^2_6 = 8.430, \chi^2_{6; 0.05} = 12.592, p = 0.208$). We conclude that there is no between area of residence and preferred color of the sports car. Since the null hypothesis of no association is not rejected, no further analysis is required.

Indicate what alternative procedures, if any, could be used in the event that required assumptions for the selected hypothesis test were not satisfied.

If all expected frequencies are not greater than 5, categories could be combined.
3. Randomly selected customers of a car dealership are polled in order to collect data concerning area of residence and preferred color of a certain type of sports car. For each customer, an area of residence is recorded as rural, suburban, or urban, and color preference is recorded chosen from red, green, yellow, and white. The data are stored in the SPSS data file sportscar. The data are to be used to perform a hypothesis test to see if there is any evidence that the proportion preferring each color is the same among rural residents. A 0.05 significance level is chosen.

Complete the following table by listing each variable involved in the analysis in the first column, indicating in the second column whether each variable should be treated as qualitative or quantitative, and indicating in the third column whether each variable is a dependent variable, independent variable, or neither:

<table>
<thead>
<tr>
<th>Variable(s)</th>
<th>Type</th>
<th>Dependent / Independent / Neither</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car Color Preference</td>
<td>Qualitative</td>
<td>Dependent Variable</td>
</tr>
</tbody>
</table>

Use SPSS to perform the necessary calculations for the appropriate hypothesis test, state whether this hypothesis test is considered parametric or nonparametric, and give the name of the hypothesis test selected.

chi-square goodness-of-fit test, which is considered a nonparametric test
State the null and alternative hypothesis, and indicate whether the test performed should be one-sided or two-sided.

Hₐ: the proportion preferring each color is the same among rural residents

H₁: the proportion preferring each color is not the same among rural residents

The chi-square goodness-of-fit test is always one sided

State the assumption(s) that must be satisfied for the hypothesis test.

Each of the expected frequencies is greater than 5.

Circle which of the following graphical displays is most appropriate.

pie chart bar chart stacked bar chart box plot(s) scatter plot

Write the results of the selected hypothesis test in a format that is suitable for a paper that you are submitting for publication; be sure to include relevant information about the test statistic, sample statistics, and p-value, and an appropriate summary of follow-up analyses, if necessary.

The chi-square goodness-of-fit test was not statistically significant at the 0.05 level ($\chi^2 = 0.783, \chi^2_{0.05, 3} = 7.815, p = 0.854$). We conclude that the proportion preferring each color is the same among rural residents. Since the null hypothesis is not rejected, no further analysis is required.

Indicate what alternative procedures, if any, could be used in the event that required assumptions for the selected hypothesis test were not satisfied.

If all expected frequencies are not greater than 5, categories could be combined.
4. For each of the following, indicate the parametric and nonparametric statistical analysis which could be appropriate:

   (a) A standard treatment for an ailment is being compared with two new treatments labeled "New A," and "New B." A 0.05 significance level is chosen for a hypothesis test to see if there is any evidence of a difference in the proportion of complete cures. The data observed on a group of subjects treated as a simple random sample have been organized into a table with two rows (labeled “cure complete” and “cure not complete”) and three columns (labeled “standard treatment”, “new A treatment”, and “new B treatment”).

   **Pearson chi-square test about independence (or concerning association) in a contingency table**

   (b) Knowledge about good nutrition habits is to be compared among three different high schools: Central high school, Grandview high school, and Sandler high school. A 30-item questionnaire is administered to randomly selected students at each school, and the number of correct responses for each student is recorded. A 0.10 significance level is chosen for a hypothesis test to see if there is any evidence of a difference in mean score among the high schools.

   **one-way analysis of variance (ANOVA)**
   **Kruskal-Wallis test**

   (c) A 0.05 significance level is chosen for a hypothesis test to see if the mean length of ball bearings produced by the Nuketown factory is larger than for those produced by the Highville factory. A simple random sample of ball bearings produced by each of the two factories is recorded (in inches).

   **two-sample (independent samples) t test**
   **Mann-Whitney rank sum test**
   **(NOTE: the one-way ANOVA and Kruskal-Wallis test could be used)**
4. - continued

(d) A hypothesis test is to be performed to see if there is any evidence that the four political party categories Republican, Democrat, Independent, and Other, are not equally likely among voters in a state.

**Pearson chi-square goodness-of-fit test**

(e) A particular company has chosen a 0.05 significance level to see if there is any evidence that the mean number of sales per salesperson tends to be higher in Bombsville than Slumsville. Several salespersons were each assigned to both cities for one week. The number of sales for each salesperson in each city was recorded.

**paired \( t \) test**  
**Wilcoxon signed rank test**

(f) The customers who buy apples at a supermarket have a choice of red apples or green apples. A hypothesis test is to be performed to see if there is any evidence that the different types of apples are not equally preferred.

**one-sample \( z \) test about a proportion**  
**OR the Pearson chi-square goodness-of-fit test**