A Bit About Me

• Academic credentials
  – Ph.D. and M.A. in Statistics, Yale University
  – M.S. in Ops Research, George Washington University
  – B.S. in Ops Analysis, U.S. Naval Academy

• (Some) research interests
  – Biosurveillance – how to quickly detect a bioterrorist attack
  – Survey design and analysis
  – Quality control and statistical process control
  – Military manpower and personnel issues: recruiting, retention, effects of deployment

• Contact information
  – Phone: 831-656-3048 (DSN: 756-3048)
  – E-mail: rdfricker@nps.edu
  – Web site: http://faculty.nps.edu/rdfricke

1/28/13
Outline

• Discuss advantages and disadvantages of nonparametric tests

• Describe some nonparametric tests
  – One sample data
  – Paired data
  – Multiple groups

• Illustrate application with various real-world datasets

• Show how to implement them in Excel, JMP, and R
Parametric vs. Nonparametric Hypothesis Testing

- **Parametric** hypothesis testing:
  - Statistic distributions are specified (often normal)
  - Often follows from Central Limit Theorem, but sometimes CLT assumptions don’t fit/apply
    - Perhaps sample size too small or underlying population distribution too skewed
    - For ANOVA, CLT doesn’t apply – data is assumed to follow normal distribution

- **Nonparametric** hypothesis testing:
  - Does not assume a particular probability distribution
    - Often called “distribution free”
  - Generally based on ordering or order statistics
Challenges in Parametric Hypothesis Testing

- Sometimes experiments give responses they defy exact quantification
  - Want to rank the “utility” of four weapons systems
    - Gives an ordering, but can be impossible to say things like “System A is twice as useful as B”

- Sometimes distributional assumptions of parametric tests violated
  - Want to compare two LVS maintenance programs, but data clearly non-normal
    - If the data do not fit the assumptions of the (parametric) tests, what to do?
Advantages of Nonparametric Tests

• Nonparametric tests make less stringent demands on the data
  – E.g., they require fewer assumptions
    • Usually require independent observations (or independence of paired differences)
    • Sometimes assumes continuity of the measure

• Can be more appropriate:
  – When measures are not precise
  – For ordinal data where scale is not obvious
  – When only ordering of data is available
Disadvantages of Nonparametric Tests

• They may “throw away” information
  – E.g., Sign test only uses the signs (+ or -) of the data, not the numeric values
  – If the other information is available and there is an appropriate parametric test, that test will be more powerful

• The trade-off:
  – Parametric tests are more powerful if the assumptions are met
  – Nonparametric tests provide a more general result if they are powerful enough to reject
(Some) Nonparametric Tests

<table>
<thead>
<tr>
<th>Parametric Test</th>
<th>“Equivalent” Nonparametric Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-sample t-test</td>
<td>Sign test, signed rank test</td>
</tr>
<tr>
<td>Two-sample t-test</td>
<td>Sign test, Wilcoxon rank sum test/ Mann-Whitney U test</td>
</tr>
<tr>
<td>Paired t-test</td>
<td>Sign test, signed rank test</td>
</tr>
<tr>
<td>One-way ANOVA</td>
<td>Kruskall-Wallis test</td>
</tr>
</tbody>
</table>

• Also:
  – Friedman test
  – Bootstrap-based methods
  – Spearman’s rank correlation coefficient
  – Kolmogorov-Smirnov test for comparing two distributions
  – McNemar’s test

1/28/13
Sign Test

Hypotheses and Assumptions

• Basic nonparametric test procedure to assess the plausibility of the null

• Assumption: Observations $x_1, \ldots, x_n$ are independent and identically distributed from some unknown distribution $F$

• Hypotheses:

$$ H_0 : F(x) = p_0 \quad H_a : F(x) \neq p_0 $$

• Common to test $p_0 = 0.5$; i.e., test the median
  – For symmetric distributions, equivalent to testing the mean
  – Can also test quartiles or any other percentile

1/28/13
Sign Test: Methodology

- First, delete all observations equal to \( x \) (and decrement \( n \) accordingly)
- Calculate \( S(x) = \# x_i \leq x \) and note that if the null hypothesis is true \( S(X) \sim \text{Bin}(n, p_0) \)
- Test the appropriate alternative hypothesis:

<table>
<thead>
<tr>
<th>Alternative Hypothesis</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_a: p &gt; p_0 )</td>
<td>( \Pr(S(X) \geq S(x)) = 1 - \text{BINOMDIST}(S(x) - 1, n, p_0, 1) )</td>
</tr>
<tr>
<td>( H_a: p &lt; p_0 )</td>
<td>( \Pr(S(X) \leq S(x)) = \text{BINOMDIST}(S(x), n, p_0, 1) )</td>
</tr>
</tbody>
</table>
| \( H_a: p \neq p_0 \)  | \( p \)-value = \( \begin{cases} 
2 \times \Pr(S(X) \geq S(x)), & \text{if } S(x) > np_0 \\
2 \times \Pr(S(X) \leq S(x)), & \text{if } S(x) < np_0 
\end{cases} \) |
Large Sample Approximation

• For large $n$, the binomial can be approximated by the normal:

$$\frac{S(x)}{n} \sim N \left( F(x), \frac{F(x)(1-F(x))}{n} \right) = N \left( p_0, \frac{p_0(1-p_0)}{n} \right)$$

• Requirements: $np_0 > 5$ and $n(1-p_0) > 5$

• Then, for a two-sided test, $p$-value $\approx 2 \times \Phi \left( -\left| z \right| \right)$

with

$$z = \frac{S(x) - np_0}{\sqrt{np_0(1-p_0)}}$$
Example #2: MK-48 LVS Down Time

- Data set has 9,505 observations
  - Clearly not normally distributed
- Mean of all 9,505 obs = 63.25 days and the median is 35 days
- Think of this set as the “population”
Example #2, continued

• What if only had 30 observations
• Okay to assume mean is normally distributed?
• Nope; in this case, \( n=30 \) not sufficient for CLT →

Histogram of Means of Random Samples of Size \( n=30 \)
Example #2, continued

• Then let’s use the sign test to test whether the median down time is greater than 10 days

• Since \( np_0 = 30 \times 0.5 = 15 > 5 \), can do large sample calc:

\[
z = \frac{S(x) - np_0}{\sqrt{np_0(1 - p_0)}} = \frac{21 - 15}{\sqrt{50(0.5)(1 - 0.5)}} = \frac{6}{2.74} = 2.19
\]

So, \( p\)-value \( \approx \Phi(-2.19) = 0.014 \)
Sign Test in Other Software: R

- Just use the `binom.test()` function
- From previous example:

```r
> binom.test(x=21,n=30,p=0.5,alternative="greater")

Exact binomial test

data:  21 and 30
number of successes = 21, number of trials = 30, p-value = 0.02139
alternative hypothesis: true probability of success is greater than 0.5
95 percent confidence interval:
  0.5349273 1.0000000
sample estimates:
probability of success
  0.7
```
Sign Test for Paired Observations

• Consider $n$ pairs of observations $(x_i, y_i)$
  - $x_1, ..., x_n$ is a random sample from $F$, and
  - $y_1, ..., y_n$ is a random sample from $G$, where $G(y) = F(y - \theta)$ for some unknown value $\theta$

• Want to test the hypothesis that the distribution of the $X$s and $Y$s is the same except perhaps for the location:
  \[ H_0 : \theta = 0 \text{ vs. } H_a : \theta \neq 0 \]

• Can easily adapt the sign test to this problem
  - Idea: Define $d_i = x_i - y_i$. Then under the null hypothesis, the probability that $d_i$ is positive is 0.5
Location Shift Assumption

\[ G(y) = F(y - \theta) \]
Example #3: INSURV
Material Inspections vs. Ship Trials

- For 2008-2011 data, test whether there is a location difference in percent SAT for ship trials versus MIs

Solution:

<table>
<thead>
<tr>
<th>Year</th>
<th>MI Percent SAT ($x_i$)</th>
<th>Ship Trials Percent SAT ($y_i$)</th>
<th>$d_i = x_i - y_i$</th>
<th>Indicator ($d_i &gt; 0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>88%</td>
<td>90%</td>
<td>-2%</td>
<td>0</td>
</tr>
<tr>
<td>2009</td>
<td>81%</td>
<td>85%</td>
<td>-4%</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>78%</td>
<td>100%</td>
<td>-22%</td>
<td>0</td>
</tr>
<tr>
<td>2011</td>
<td>80%</td>
<td>87%</td>
<td>-7%</td>
<td>0</td>
</tr>
<tr>
<td>2012</td>
<td>81%</td>
<td>88%</td>
<td>-7%</td>
<td>0</td>
</tr>
</tbody>
</table>

$\sum = 0$

$p$-value = 0.03125 = BINOMDIST(E7,5,0.5,1)
Example #4: AAAV Maintenance Rates at I and III MEF

• Test whether the AAAV maintenance rates are different between the two MEFs
• Perhaps could use a parametric approach to test the mean maintenance rates:
Example #4: AAAV Maintenance Rates at I and III MEF

• Paired t-test (parametric test) result
  – Pick your preferred software output

![Paired t-test result table and graph]

```
> t.test(aaav$MEF.1, aaav$MEF.3, paired=TRUE)

Paired t-test

data:  aaav$MEF.1 and aaav$MEF.3
t = 8.7576, df = 14, p-value = 4.704e-07
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  0.001089258 0.001795836
sample estimates:
mean of the differences
  0.0014428547
```
Example #4: AAAV Maintenance Rates at I and III MEF

- Sign test confirms I MEF rates higher – without having to assume normality

<table>
<thead>
<tr>
<th>Quarter i</th>
<th>I MEF</th>
<th>III MEF</th>
<th>$d_i = x_i - y_i$</th>
<th>Indicator ($d_i &gt; 0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.00359516</td>
<td>0.00167062</td>
<td>0.0019</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.00442939</td>
<td>0.00175995</td>
<td>0.0027</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0.00384213</td>
<td>0.00125686</td>
<td>0.0026</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.00350904</td>
<td>0.00145815</td>
<td>0.0021</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0.00314295</td>
<td>0.00220157</td>
<td>0.0009</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0.0027121</td>
<td>0.00090214</td>
<td>0.0018</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0.00248196</td>
<td>0.00095776</td>
<td>0.0015</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0.00210491</td>
<td>0.00105747</td>
<td>0.0010</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0.00454197</td>
<td>0.00403405</td>
<td>0.0005</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0.00362621</td>
<td>0.00262377</td>
<td>0.0010</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0.00372042</td>
<td>0.0024293</td>
<td>0.0013</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0.00293129</td>
<td>0.00206736</td>
<td>0.0009</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0.00480068</td>
<td>0.00359857</td>
<td>0.0012</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>0.00473005</td>
<td>0.00349907</td>
<td>0.0012</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>0.00444314</td>
<td>0.00345659</td>
<td>0.0010</td>
<td>1</td>
</tr>
</tbody>
</table>

$S = 15$

$p$-value $= 6.104E-05$ $= 2 \times (1 - \text{BINOMDIST(E17-1,15,0.5,1)})$
Signed Rank Test
Hypotheses and Assumptions

• Most often used to make inferences about median of unknown distribution \( F \)

• Assumptions:
  – Observations are independent and identically distributed from \( F \)
  – If distribution is symmetric, median = mean and both can be denoted as \( \mu = F^{-1}(0.5) \)

• Hypotheses: \( H_0 : \mu = \mu_0 \quad H_a : \mu \neq \mu_0 \)

• Signed rank more sensitive than Sign test
  – It uses the magnitudes of the deviations from the median as well as the signs in its calculations
Signed Rank Test: Methodology (1)

• First, calculate deviations from hypothesized mean value: $d_i(\mu_0) = x_i - \mu_0$

• Compute the ranks of the absolute values of the deviations: $r_i(\mu_0) = \text{rank}(|d_i|)$
  – Delete any observations equal to $\mu_0$
  – If two observations are tied, give them both the average rank

• Calculate $S_+(\mu_0)$, the sum of the ranks of the positive deviations
Interpretation of the Signed Rank Test Statistic

\[ S_+ (\mu_0) = \sum \text{sum of ranks} \]

For \( S_+ (\mu_0) > \frac{n(n+1)}{4} \), suggests \( \mu > \mu_0 \)

\[ d_i(\mu_0) = x_i - \mu_0 < 0 \quad d_i(\mu_0) = x_i - \mu_0 > 0 \]

For \( S_+ (\mu_0) < \frac{n(n+1)}{4} \), suggests \( \mu < \mu_0 \)

\[ d_i(\mu_0) = x_i - \mu_0 < 0 \quad d_i(\mu_0) = x_i - \mu_0 > 0 \]
Signed Rank Test: Methodology (3)

- If $S_+(\mu_0)$ is far from $n(n+1)/4$, then reject null
  - For small sample sizes, tables/software necessary to determine “far”
  - If sample size is large, use normal approximation to judge “far”
- The test statistic
  \[
  Z_+(\mu_0) = \frac{S_+(\mu_0) - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}
  \]
  is approximately standard normally distributed
**Example #2, continued**

- **Calcs give** \( S_+(\mu_0) = 407.5 \)
  - Table: reject at significance level \( \alpha < 0.01 \)
  - Large sample calcs:
    
    \[
    Z_+(\mu_0) = \frac{S_+(\mu_0) - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}
    \]
    
    \[
    = \frac{407.5 - 30 \times 31 / 4}{\sqrt{30 \times 31 \times 61 / 24}}
    \]
    
    \[
    = 3.6
    \]
    
    So, \( p \)-value \( \approx \Phi(-3.6) = 0.0002 \)
  - Conclude median > 10
Signed Rank Variants

• Can also use signed rank test for paired data
  – Do test on differences of the pairs

| Quarter i | I MEF  | II MEF  | $d_i = x_i - y_i$ | Sorted Differences $d_i$ | Deviations $d(0) = x_i - 0$ | Absolute Deviations $|d(0)|$ | Ranks $r_i$ |
|-----------|--------|---------|-------------------|----------------------------|--------------------------------|-------------------------------|------------|
| 1         | 0.00360| 0.00167 | 0.00192           | 0.00051                    | 0.00051                        | 0.00051                        | 1          |
| 2         | 0.00443| 0.00176 | 0.00267           | 0.00086                    | 0.00086                        | 0.00086                        | 2          |
| 3         | 0.00384| 0.00126 | 0.00259           | 0.00094                    | 0.00094                        | 0.00094                        | 3          |
| 4         | 0.00351| 0.00146 | 0.00205           | 0.00099                    | 0.00099                        | 0.00099                        | 4          |
| 5         | 0.00314| 0.00220 | 0.00094           | 0.00100                    | 0.00100                        | 0.00100                        | 5          |
| 6         | 0.00271| 0.00090 | 0.00181           | 0.00105                    | 0.00105                        | 0.00105                        | 6          |
| 7         | 0.00248| 0.00096 | 0.00152           | 0.00120                    | 0.00120                        | 0.00120                        | 7          |
| 8         | 0.00210| 0.00106 | 0.00105           | 0.00123                    | 0.00123                        | 0.00123                        | 8          |
| 9         | 0.00454| 0.00403 | 0.00051           | 0.00129                    | 0.00129                        | 0.00129                        | 9          |
| 10        | 0.00363| 0.00262 | 0.00100           | 0.00152                    | 0.00152                        | 0.00152                        | 10         |
| 11        | 0.00372| 0.00243 | 0.00129           | 0.00181                    | 0.00181                        | 0.00181                        | 11         |
| 12        | 0.00293| 0.00207 | 0.00086           | 0.00192                    | 0.00192                        | 0.00192                        | 12         |
| 13        | 0.00480| 0.00360 | 0.00120           | 0.00205                    | 0.00205                        | 0.00205                        | 13         |
| 14        | 0.00473| 0.00350 | 0.00123           | 0.00259                    | 0.00259                        | 0.00259                        | 14         |
| 15        | 0.00444| 0.00346 | 0.00099           | 0.00267                    | 0.00267                        | 0.00267                        | 15         |

$S_r (mu_r) = 120$
Signed Rank Test in Other Software: JMP

The distribution of values in each column
Select Columns | Cast Selected Columns into Roles | Action
--- | --- | ---
I MEF | Y, Columns | OK
III MEF | optional numeric | OK
d | optional | OK

Distribution

The distribution of values in each column
Select Columns | Cast Selected Columns into Roles | Action
--- | --- | ---
I MEF | Y, Columns | OK
III MEF | optional numeric | OK
d | optional | OK

Distributions

Specify Hypothesized Mean
Enter True Standard Deviation to do z-test rather than t test
If you also want a nonparametric test:
- Wilcoxon Signed Rank

Test Mean

Hypothesized Value: 0
Actual Estimate: 0.001442
DF: 14
Std Dev: 0.000842

Signed-Rank Test

Test Statistic: 60.0000
Prob > t: <0.0001*
Signed Rank Test in Other Software: R

• Use the `wilcox.test()` function for **two sample tests**

```r
> diff <- c(0.00192, 0.00267, 0.00259, 0.00205, 0.00094, 0.00181, 0.00152, 0.00105,
>           0.00051, 0.00100, 0.00129, 0.00086, 0.00120, 0.00123, 0.00099)
> wilcox.test(diff)

Wilcoxon signed rank test

data:  diff
V = 120, p-value = 6.104e-05
alternative hypothesis: true location is not equal to 0
```
Rank Sum Test
Hypotheses and Assumptions

• One- or two-sided test for the hypotheses of the means of two independent samples
  – $x_1,\ldots,x_n$ is a random sample from $F$, and
  – $y_1,\ldots,y_m$ is a random sample from $G$, where $G(y) = F(y - \theta)$ for some unknown value $\theta$

• Want to test the hypothesis that the distribution of the $X$s and $Y$s is the same except perhaps for the location:

$$H_0 : \theta = 0 \text{ vs. } H_a : \theta \neq 0$$
Idea of the Test

• If two samples come from same distribution, the data ought to be “mixed in” together
  – If not, likely that one sample of data will be systematically higher or lower than other sample

• Use ranks to quantify where data falls
  – I.e., smallest observation gets rank 1, next smallest gets rank 2, etc.

• Sum of the ranks quantifies whether whole sample is systematically different
Calculating the Test Statistic

- In JMP, assuming no ties, the test statistic is

\[ Z = \frac{S_x - E(S_x)}{\sqrt{\text{Var}(S_x)}} \]

where

\[ E(S_x) = \frac{m(m + n + 1)}{2}, \]

and

\[ \text{Var}(S_x) = \frac{mn(m + n + 1)}{12} \]

- There are other ways to calculate
  - And there are details for handling ties
  - We’ll just let JMP or R calculate for us…
Example #5

• Comparison of age distribution of US vs. non-US Iraq war casualties
  – 4,123 casualties from 3-21-03 to 10-10-07
  – 3,820 US (n) and 301 non-US (m)
Calculating in JMP

• Analyze > Fit Y by X
  – Be sure X is nominal and Y is continuous
• Under red triangle > Nonparametric > Wilcoxon Test
• Check the calcs:

  \[ E(S_x) = \frac{m(m+n+1)}{2} \]
  \[ = \frac{295(4,112)}{2} = 606,520 \]

  \[ \text{Var}(S_x) = \frac{mn(m+n+1)}{12} \]
  \[ = \frac{295 \times 3,816 \times 4,112}{12} = 385,746,720 \]

  \[ Z = \left[ \frac{S_x - E(S_x)}{\sqrt{\text{Var}(S_x)}} \right] \]
  \[ = \left[ \frac{778,733.5 - 606,520}{\sqrt{385,746,720}} \right] = 8.768314 \]

• Conclusion: Location different
Calculating in R

• Use the `wilcox.test()` function with option `paired=FALSE`

```r
> wilcox.test(iraq$Age~iraq$US_ind,paired=FALSE)

Wilcoxon rank sum test with continuity correction

data:  iraq$Age by iraq$US_ind
W = 735073.5, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0
```
Kruskal-Wallis Test
Hypotheses and Assumptions

• Generalization of the rank sum test for $k > 2$ samples
  – $n_T$ observations divided into $k$ groups
  – Nonparametric alternative to one-way ANOVA

• Null hypothesis is that all $k$ samples are from the same population

• Assumption: error terms are identically distributed
  – Need not be a normal distribution
Kruskal-Wallis Test: Methodology

• Idea: combine $k$ samples into one large sample, order them, and rank observations from 1 to $n_T$

• Denote rank of observation $x_{ij}$ as $r_{ij}$, $\bar{r}_1$, $\ldots$, $\bar{r}_k$

$1 \leq i \leq k, 1 \leq j \leq n_i$

• Further denote the average ranks as where

$$\bar{r}_i = \frac{r_{i1} + \cdots + r_{in_i}}{n_i}$$
Calculate the Test Statistic

The test statistic is

\[ H = \frac{12}{n_T(n_T + 1)} \sum_{i=1}^{k} n_i \bar{r}_i^2 - 3(n_T + 1) \]

Calculate \( p \)-value as \( \Pr(X > H) \) where \( X \) has a chi-square distribution with \( k-1 \) df
Example #6: Helo Back Pain Survey

- Survey of 683 US Navy helo pilots and co-pilots about back pain in the cockpit
  - Question 29: “How often do you have back and/or neck pain during or immediately after a flight?” (1=Rarely, 5=Very Often)
  - Compare by question 11: “In what type of helicopter do you have the most flight hours?” (H-46, H-53, TH-57, H60)
  - 553 respondents answered both questions:

```r
> table(helo)

     Q29
Q11  H-46 H-60 TH-57 H-53
   1  1  23   3   1
   2  1  1   16  15  1
   3  6  88  144  13
   4  3  131  16  16  8
   5  2  71   7   3
```

1/28/13
Kruskall-Wallis Test in JMP

• Analyze > Fit Y by X > red triangle > Nonparametric > Wilcoxon Test
  – Same as Wilcoxon rank sum, just need more than two levels in $X$

```
<table>
<thead>
<tr>
<th>Level</th>
<th>Count</th>
<th>Score Sum</th>
<th>Score</th>
<th>Score Mean</th>
<th>(Mean-Mean0)/Std0</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-46</td>
<td>13</td>
<td>3632.00</td>
<td>3601.00</td>
<td>279.385</td>
<td>0.055</td>
</tr>
<tr>
<td>H-53</td>
<td>26</td>
<td>7662.50</td>
<td>7202.00</td>
<td>294.712</td>
<td>0.599</td>
</tr>
<tr>
<td>H-60</td>
<td>457</td>
<td>127212</td>
<td>126589</td>
<td>278.363</td>
<td>0.453</td>
</tr>
<tr>
<td>TH-57</td>
<td>57</td>
<td>14674.5</td>
<td>15789.0</td>
<td>257.447</td>
<td>-1.010</td>
</tr>
</tbody>
</table>
```

1-way Test, ChiSquare Approximation

<table>
<thead>
<tr>
<th>ChiSquare</th>
<th>DF</th>
<th>Prob&gt;ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2968</td>
<td>3</td>
<td>0.7299</td>
</tr>
</tbody>
</table>

$p$-value – cannot reject null
# Parametric vs. Nonparametric

## Wilcoxon / Kruskal-Wallis Tests (Rank Sums)

<table>
<thead>
<tr>
<th>Level</th>
<th>Count</th>
<th>Score Sum</th>
<th>Score</th>
<th>Score Mean</th>
<th>(Mean-Mean0)/Std0</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-46</td>
<td>13</td>
<td>3632.00</td>
<td>3601.00</td>
<td>279.385</td>
<td>0.055</td>
</tr>
<tr>
<td>H-53</td>
<td>26</td>
<td>7662.50</td>
<td>7202.00</td>
<td>294.712</td>
<td>0.599</td>
</tr>
<tr>
<td>H-60</td>
<td>457</td>
<td>127212</td>
<td>126589</td>
<td>278.363</td>
<td>0.453</td>
</tr>
<tr>
<td>TH-57</td>
<td>57</td>
<td>14674.5</td>
<td>15789.0</td>
<td>257.447</td>
<td>-1.010</td>
</tr>
</tbody>
</table>

## 1-way Test, ChiSquare Approximation

<table>
<thead>
<tr>
<th>ChiSquare</th>
<th>DF</th>
<th>Prob&gt;ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2968</td>
<td>3</td>
<td>0.7299</td>
</tr>
</tbody>
</table>

## Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q11</td>
<td>3</td>
<td>1.53836</td>
<td>0.51279</td>
<td>0.4282</td>
<td>0.7329</td>
</tr>
<tr>
<td>Error</td>
<td>549</td>
<td>657.41643</td>
<td>1.19748</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Total</td>
<td>552</td>
<td>658.95479</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Means for One-way Anova

<table>
<thead>
<tr>
<th>Level</th>
<th>Number</th>
<th>Mean</th>
<th>Std Error</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-46</td>
<td>13</td>
<td>3.30769</td>
<td>0.30350</td>
<td>2.7115</td>
<td>3.9039</td>
</tr>
<tr>
<td>H-53</td>
<td>26</td>
<td>3.42308</td>
<td>0.21461</td>
<td>3.0015</td>
<td>3.8446</td>
</tr>
<tr>
<td>H-60</td>
<td>457</td>
<td>3.30416</td>
<td>0.05119</td>
<td>3.2036</td>
<td>3.4047</td>
</tr>
<tr>
<td>TH-57</td>
<td>57</td>
<td>3.15789</td>
<td>0.14494</td>
<td>2.8732</td>
<td>3.4426</td>
</tr>
</tbody>
</table>
Calculating in R

• Use the `kruskal.test()` function

```r
> kruskal.test(helo$Q29~helo$Q11)

Kruskal-Wallis rank sum test

data:  helo$Q29 by helo$Q11
Kruskal-Wallis chi-squared = 1.2968, df = 3, p-value = 0.7299
```

• No statistical difference by type of helo

<table>
<thead>
<tr>
<th></th>
<th>Q11</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>H-46</td>
<td>Row %</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.69</td>
<td>7.69</td>
<td>46.15</td>
<td>23.08</td>
</tr>
<tr>
<td>H-53</td>
<td></td>
<td>1</td>
<td>1</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.85</td>
<td>3.85</td>
<td>50.00</td>
<td>30.77</td>
</tr>
<tr>
<td>H-60</td>
<td></td>
<td>23</td>
<td>88</td>
<td>144</td>
<td>131</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.03</td>
<td>19.26</td>
<td>31.51</td>
<td>28.67</td>
</tr>
<tr>
<td>TH-57</td>
<td></td>
<td>3</td>
<td>15</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.26</td>
<td>26.32</td>
<td>28.07</td>
<td>28.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td>105</td>
<td>179</td>
<td>158</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>5</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Row %</td>
<td>2</td>
<td>3</td>
<td>15.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>15.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>26</td>
<td>11.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>457</td>
<td>15.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>57</td>
<td>12.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>553</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Good References

• My OA3102 course notes: http://faculty.nps.edu/rdfricke/OA3102.htm


What We Covered

• Discussed advantages and disadvantages of nonparametric tests
• Described some nonparametric tests
  – One sample data
  – Paired data
  – Multiple groups
• Illustrated application with various real-world datasets
• Showed how to implement them in Excel, JMP, and R