Nonparametric Statistics

J. Lozano

University of Goettingen

Department of Genetic Epidemiology

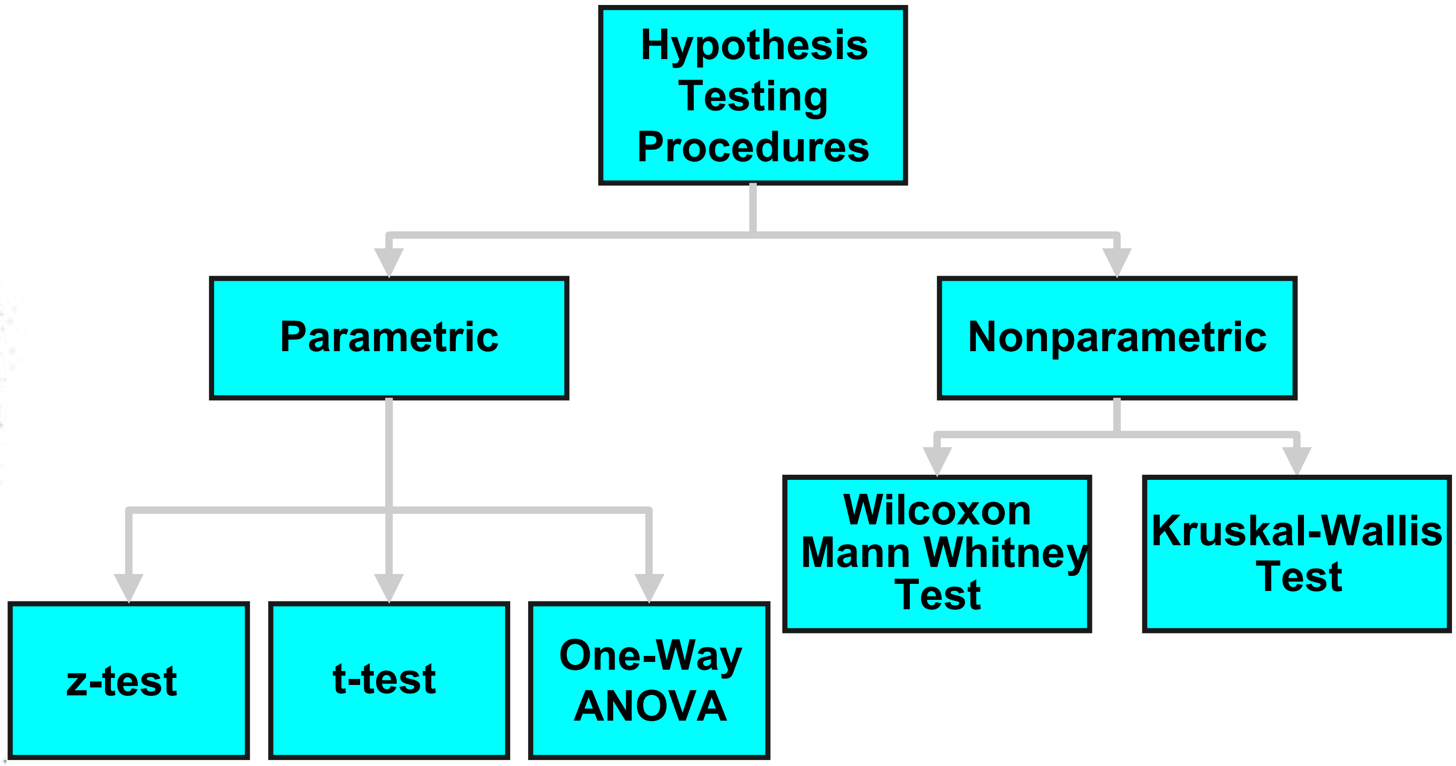
Interdisciplinary PhD Program in
Applied Statistics & Empirical Methods

Graduate Seminar in Applied Statistics

July 3, 2006


Binomial test
Friedman's ANOVA
Wilcoxon-Mann-Whitney
Mc Nemar
Median test
Sign Test
Fisher's Exact
Spearman
Wilcoxon Signed Ranks Test
Run's Test
Kruskal-Wallis
Kappa
Run's Test

Parametric vs. Nonparametric



Many More Tests Exist!

Learning Objectives:

1. Differentiate nonparametric from parametric statistics

2. Discuss the advantages and disadvantages of nonparametric statistics

3. Enumerate and differentiate the different nonparametric tests

4. Apply some commonly used nonparametric test to hypothesis testing problems

Parametric Test Procedures

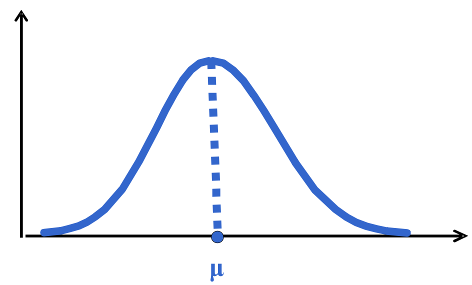
1. Involve Population Parameters

   Example: Population Mean

$$t = \frac{\bar{X} - \mu_0}{\left(S_x / \sqrt{n}\right)}$$

2. Have Stringent Assumptions

   Example: Normal Distribution



3. Require Interval Scale or Ratio Scale

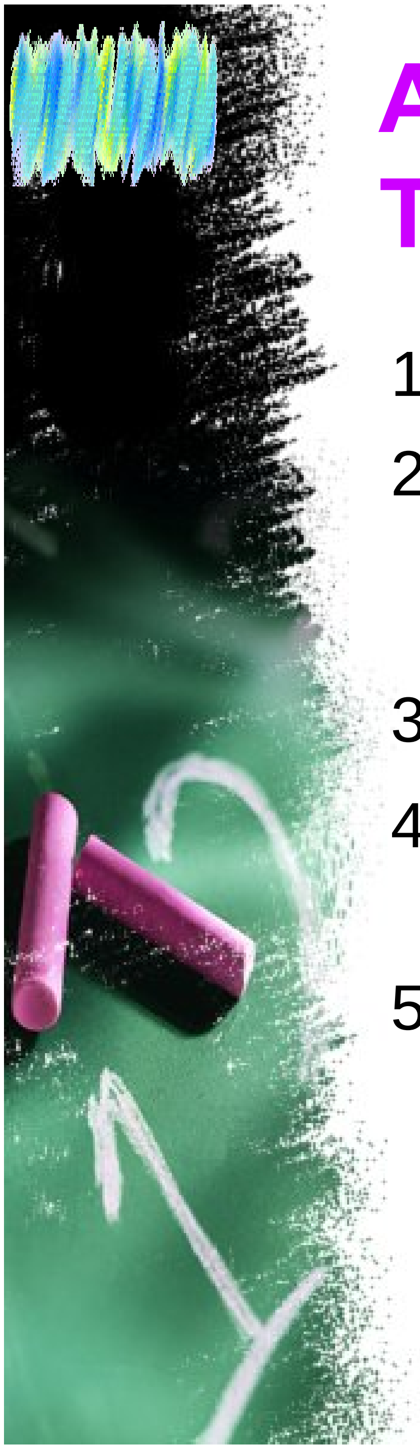
   Whole Numbers or Fractions

   Example: Height in Inches (72, 60.5, 54.7)

4. Examples: z-Test, t-Test, ANOVA

Nonparametric Test Procedures

1. Do Not Involve Population Parameters
2. No Stringent Distribution Assumptions
   “Distribution-free”
3. Data Measured on Any Scale
   Ratio or Interval
   Ordinal
      Example: Good-Better-Best
   Nominal
      Example: Male-Female
4. Example: Wilcoxon-Mann-Whitney Test

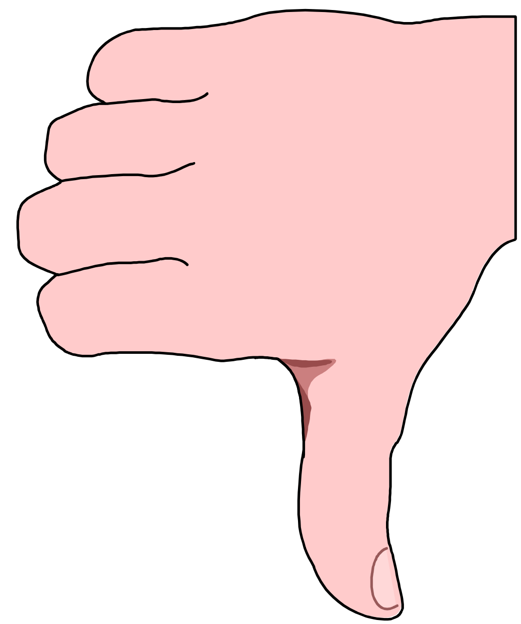


Advantages of Nonparametric Tests

1. Used With All Scales
2. Easier to Compute
   Developed Originally Before
   Wide Computer Use
3. Make Fewer Assumptions
4. Need Not Involve Population Parameters
5. Results May Be as Exact as Parametric Procedures

Disadvantages of Nonparametric Tests

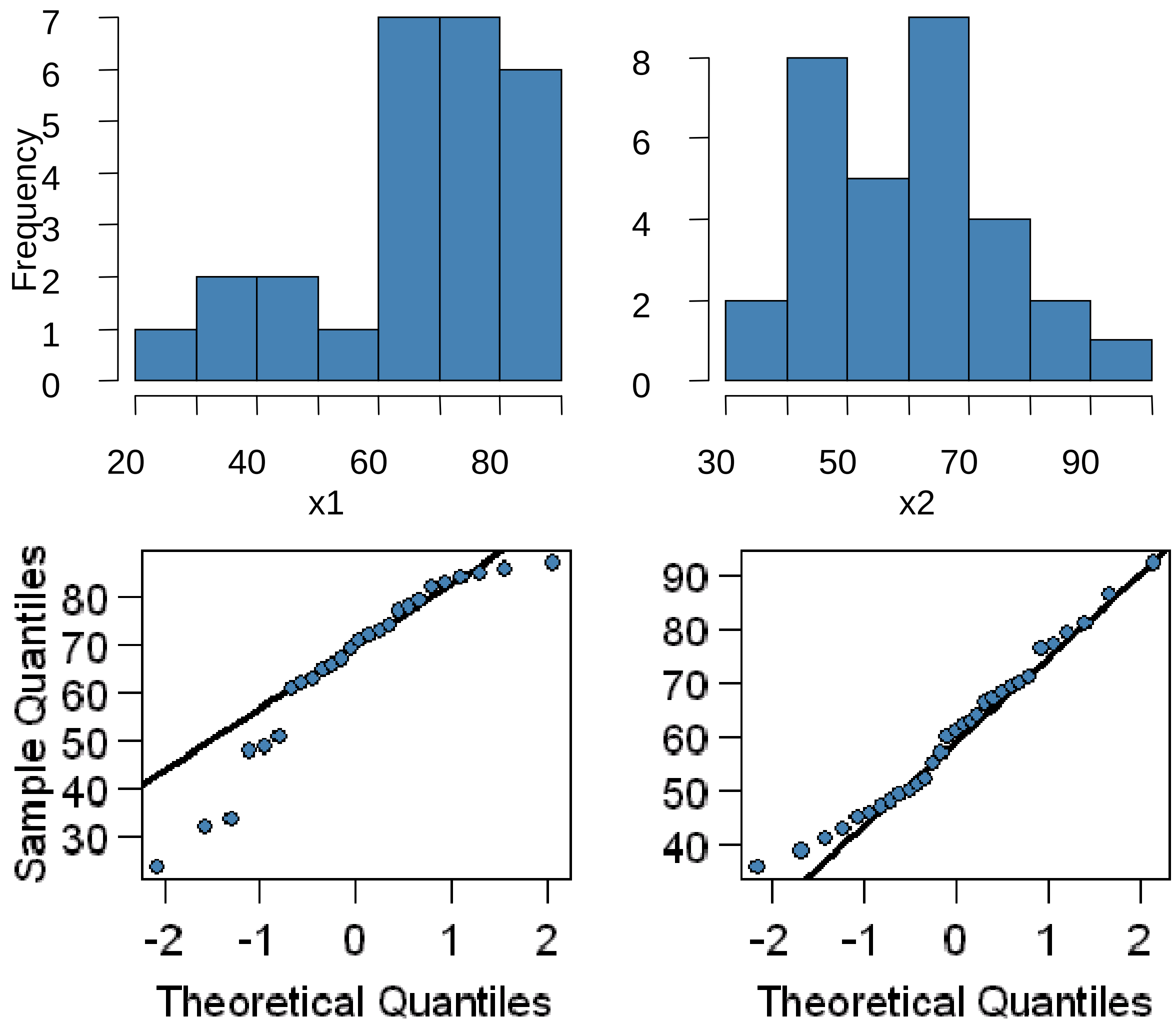
1. May Waste Information
   - If Data Permit Using Parametric Procedures
   - Example: Converting Data From Ratio to Ordinal Scale
2. Require a larger sample size than the corresponding parametric test in order to achieve the same power
3. Difficult to Compute by Hand for Large Samples
4. Stat tables are not readily available



Summary Table of Statistical Tests

Level of Measurement	Sample Characteristics					Correlation
	1 Sample	2 Sample		K Sample (i.e., >2)		
		Independent	Dependent	Independent	Dependent	
Categorical or Nominal	**Bi-nomial** **X^2**	**X^2** **Fisher Exact**	**Mc Nemar's**	**X^2**	**Cochran's Q**	**Kappa Agreement Test**
Rank or Ordinal	**Run's test** **Kolmo-gorov Smirnov**	**Median Test** **Wilcoxon-Mann-Whitney**	**Sign Test** **Wilcoxon Signed Ranks**	**Kruskal-Wallis**	**Friendman's ANOVA**	**Spearman's rho** **Kendall Rank**
Parametric (Interval & Ratio)	z-test or t-test	t- test between groups	Paired t-test	1 way ANOVA between groups	1 way ANOVA (within or repeated measure)	Pearson's r
		Factorial ANOVA				

Application of Commonly Used Nonparametric Statistics



=> nonparametric statistics…

Wilcoxon-Mann-Whitney Test

Wilcoxon-Mann-Whitney Test

- Also known as Wilcoxon-test, Wilcoxon rank sum test, U-test, Mann-Whitney-U-test
- Tests Two Independent Population
  - compare medians
- Corresponds to t-Test for 2 Independent Means
- Assumptions
  - Independent, Random Samples
  - Populations Are Continuous
- Can Use Normal Approximation If $n_i \geq 20$

Wilcoxon-Mann-Whitney Test Procedure

1. Assign Ranks, r_i, to the $n_1 + n_2$ Sample Observations
   If unequal sample sizes, let n_1 refer to smaller-sized sample
   Smallest Value = 1
   Average Ties

2. Sum the Ranks, R_i, for Each Sample

3. Test Statistic
   - Null hypothesis: both samples come from the same underlying distribution

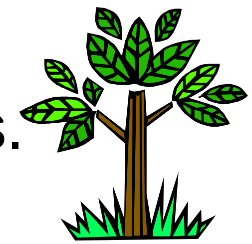
$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \qquad \text{where } \mathrm{R}_1 \text{ is the sum of ranks in sample1}$$

Compute also ($n_1 n_2$ –U). If this expression is larger than the above U, then this becomes the final U statistic.

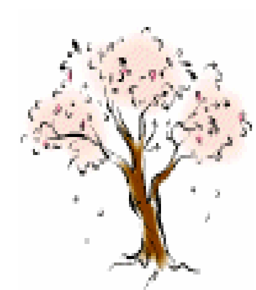
** At N>=20, U begins to approximate t, so the test stat changes to a t-value.

Wilcoxon-Mann-Whitney Test Example

- You're an agriculturist. You want to determine if there was a difference in the biomass of male and female Juniper trees.



- Randomly select 6 trees of each gender from the field. Dry them to constant moisture and weigh in kg.



- Male trees Data: 71, 73, 78, 75, 72, 74
- Female trees Data: 80, 73, 83, 84, 82, 79

Wilcoxon-Mann-Whitney Test Solution

- H_0:
- H_a:
- α =
- n_1 = n_2 =
- Critical Value:

Test Statistic:

Decision:

Conclusion:

Wilcoxon-Mann-Whitney Test Computation Table

Male Trees		Female Trees	
Weight	Rank	Weight	Rank
71		80	
73		73	
78		83	
75		84	
72		82	
74		79	

Rank Sum

Wilcoxon-Mann-Whitney Test Solution

Median weights =

- H_a: Medians not =
- α =0.05
- n_1 = 6 n_2 = 6
- Critical Value:

U table = 31

* Depends on the U table you are using

Test Statistic:

U = (6)(6) + (6)(7)/2 – 24.5 = 32.5

n1n2 - U = (6)(6) – 32.5 = 3.5

***!!! final U_{Calc} = 32.5**

Decision:

Reject H_0 at α = .05 since $U_{calc} > U_{table}$

Conclusion:

There is evidence that the medians are not equal.

Wilcoxon-Mann-Whitney Test Solution (SPSS Output)

NPar Tests

Mann-Whitney Test

Ranks

	Gender of Tree	N	Mean Rank	Sum of Ranks
Weight in kg	Male	6	4.08	24.50
	Female	6	8.92	53.50
	Total	12		

Test Statistics[b]

	Weight in kg
Mann-Whitney U	3.500
Wilcoxon W	24.500
Z	-2.326
Asymp. Sig. (2-tailed)	.020
Exact Sig. [2*(1-tailed Sig.)]	.015[a]

a. Not corrected for ties.

b. Grouping Variable: Gender of Tree

$p < \alpha$; Reject Ho

Sign Test

Sign Test

1. Tests One Population Median, η (eta)
2. Corresponds to t-Test for 1 Mean
3. Assumes Population Is Continuous
4. Small Sample Test Statistic: # Sample Values Above (or Below) Median
5. Can Use Normal Approximation If $n \geq 10$

Sign Test Example

- You're a marketing analyst for Chefs-R-Us. You've asked 8 people to rate a new ravioli on a 5-point Likert scale

  1 = terrible to
  5 = excellent

The ratings are:

2 4 1 2 1 1 2 1

At the **.05** level, is there evidence that the **median** rating is **at least 3**?

Sign Test Solution

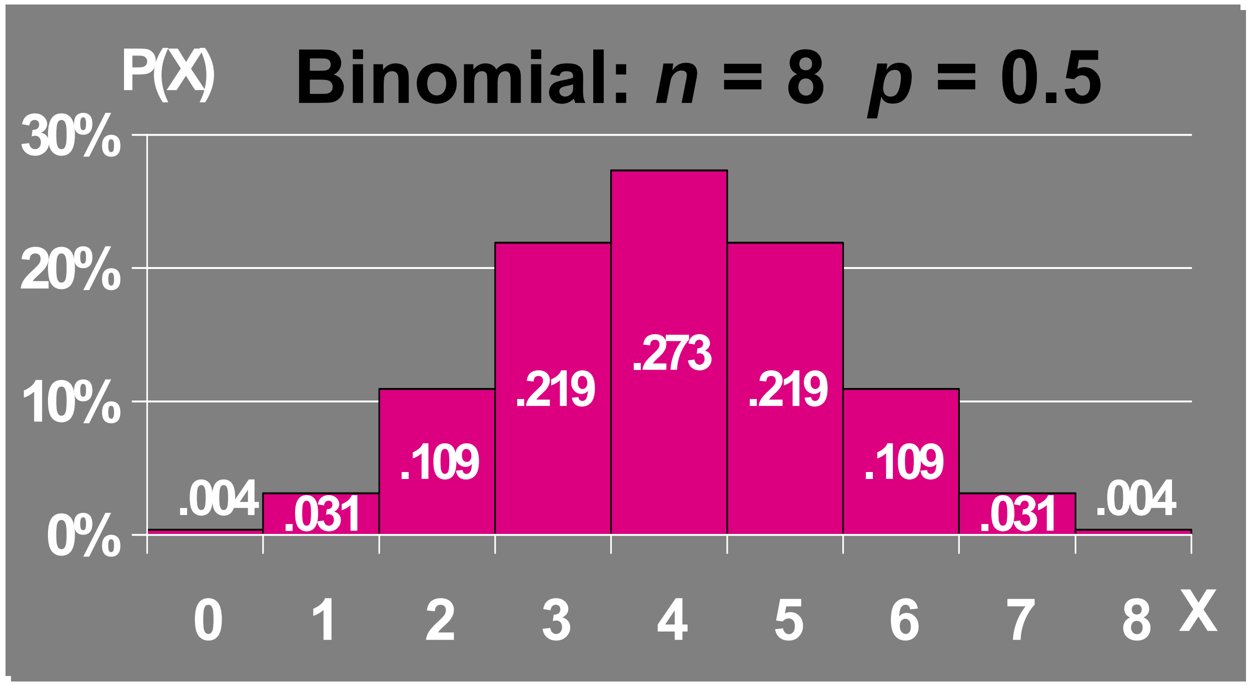
- **H_0:**
- **H_a:**
- α **=**
- **Test Statistic:**

P-Value:

Decision:

Conclusion:

Sign Test Uses P-Value to Make Decision



P-Value is the probability of getting an observation at Least as extreme as we got.

If 7 of 8 Observations 'Favor' H_a, Then P-Value = $P(x \geq 7)$ = .031 + .004 = .035.
If α = .05, Then Reject H_0 Since P-Value $\leq \alpha$.

Sign Test Solution

- H_0: $\eta = 3$
- H_a: $\eta < 3$
- $\alpha = .05$
- Test Statistic:

S = 7

(Ratings 1 & 2 are less than $\eta = 3$: 2 4 1 2 1 1 2 1)

P-Value:

***P*(*x* ≥ 7) = .031 + .004 = .035**

(Binomial Table, *n* = 8, *p* = 0.50)

$$P(X) = \frac{n!}{X!(n-X)!}\pi^x(1-\pi)^{n-X}$$

Decision:

Reject H_0 at $\alpha = .05$

Conclusion:

There is evidence that the median is less than 3

Sign Test (R Output)

R version 2.3.1 (BSDA and eV10 package must be loaded)

```
> x <- c(2,4,1,2,1,1,2,1)
> sign.test(x,md=3, alternative="less")

        One-sample Sign-Test

data:  x
s = 1, p-value = 0.03516
alternative hypothesis: true median is less than 3
95 percent confidence interval:
 -Inf    2
sample estimates:
median of x
        1.5

$Confidence.Intervals
                     Conf.Level  L.E.pt U.E.pt
Lower Achieved CI     0.8555       -Inf      2
Interpolated CI       0.9500       -Inf      2
Upper Achieved CI     0.9648       -Inf      2

Warning message:
multi-argument returns are deprecated in: return(rval, Confidence.Intervals)
```

Mc Nemar Change Test

McNemar Change Test (MCT)

- Typical examples:
  - Testing the shift in the proportion of abnormal responses from before and after treatment in the same group of patients
  - Comparing two ocular treatments when both are given to each patient, one in each eye

Data Layout: MCT

		Condition 2		
		No. of 'Responders'	No. of 'Non-Responders'	TOTAL
Condition 1	No. of 'Responders'	A	B	A + B
	No. of 'Non-Responders'	C	D	C + D
	TOTAL	A + C	B + D	N = A +B +C +D

McNemar Change Test (MCT)

- The hypothesis of interest is the equality of response proportions, p_1 and p_2, under conditions 1 and 2, respectively.
- The test statistic is based on the difference in the discordant cell frequencies (B, C).

Test Summary: MCT

- Null hypothesis: $H_0 : p_1 = p_2$
- Alternative hypothesis: $H_a : p_1 \neq p_2$
- Test Statistic:

$$\chi^2 = \frac{(B - C)^2}{B + C}$$

- Decision Rule: Reject H_0 if $\chi^2 > \chi_1^2(\alpha)$
  - where $\chi_1^2(\alpha)$ is the critical value from the chi-square table with significance level α and 1 degree of freedom

Example: MCT

- Bilirubin Abnormalities Following Drug Treatment

86 patients were treated with an experimental drug for 3 months.

Pre-post study clinical laboratory results showed abnormally high total bilirubin values.

Is there evidence of a change in the pre- to post-treatment rates of abnormalities?

Solution: MCT

Let p_1 and p_2 represent the proportions of patients with abnormally high bilirubin values (‘Y’) before and after treatment, respectively.

		POST-Treatment		TOTAL
		‘N’	‘Y’	
PRE-Treatment	‘N’	60	14	74
	‘Y’	6	6	12
	TOTAL	66	20	86

‘Y’ = T.Bilirubin above upper limit of normal range

Test Summary: MCT

- Null hypothesis: $H_0 : p_1 = p_2$
- Alternative hypothesis: $H_a : p_1 \neq p_2$
- Test Statistic: $\chi^2 = \frac{(B-C)^2}{B+C} = \frac{(14-6)^2}{14+6} = \frac{64}{20} = 3.20$
- Decision: Do not reject H_0 since

$$\chi^2 < \chi_1^2(\alpha_{0.05}) = 3.841$$

- Conclusion:

There is no sufficient evidence to conclude at 0.05 level of significance that a shift in abnormality rates occurs with treatment.

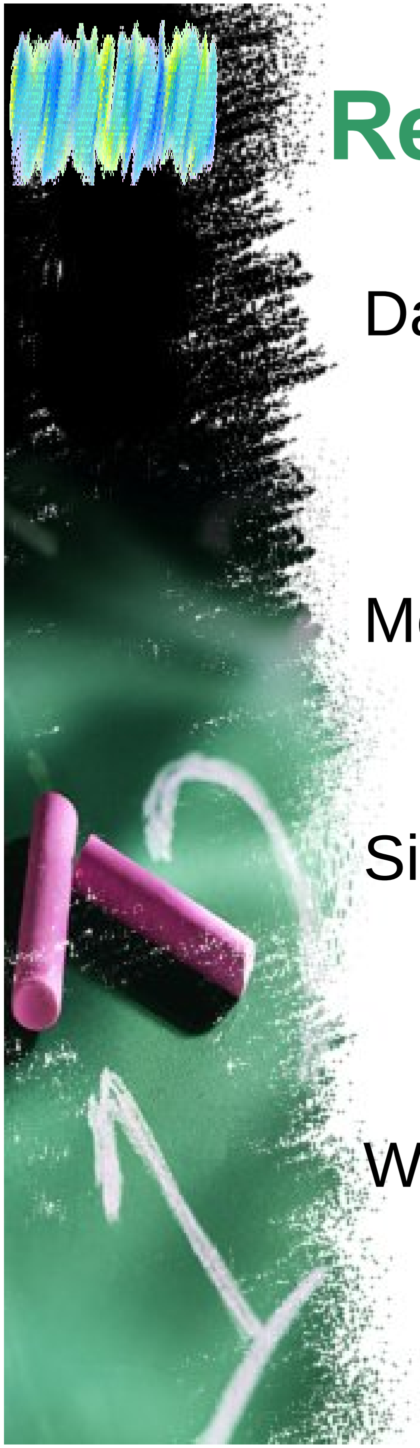
SAS Computer Output for MCT

```
                   McNemar's Test
Example: Bilirubin Abnormalities Following Drug Treatment
                  TABLE OF PRE BY PST

           PRE(PRE)     PST(PST)

           Frequency,
           Percent  ,
           Row Pct  ,
           Col Pct  ,N       ,Y       ,  Total
           ffffffffˆfffffffˆfffffffˆ
           N        ,     60 ,     14 ,     74
                    ,  69.77 ,  16.28 ,  86.05
                    ,  81.08 ,  18.92 ,
                    ,  90.91 ,  70.00 ,
           ffffffffˆfffffffˆfffffffˆ
           Y        ,      6 ,      6 ,     12
                    ,   6.98 ,   6.98 ,  13.95
                    ,  50.00 ,  50.00 ,
                    ,   9.09 ,  30.00 ,
           ffffffffˆfffffffˆfffffffˆ
           Total          66       20       86
                       76.74    23.26   100.00

           STATISTICS FOR TABLE OF PRE BY PST
                       McNemar's Test
                      --------------
  Statistic = 3.200        DF = 1        Prob = 0.074
```



References

Dawson-Saunders, B. and Trapp, R.G. Basic and Clinical Biostatistics. Connecticut: Prentice Hall International, Inc. (1990).

McGill, J.J. Chap. 14: Nonparametric Statistics Presentation. (May 2006).

Siegel S. and Castellan N.J. Nonparametric Statistics for the Behavioral Sciences (2nd edition). New York: McGraw Hill (1988).

Walker G.A. Common Statistical Methods for Clinical Research with SAS Examples, SAS Institute, Inc, Cary NC (1997).