

## Parametric and nonparametric tests

A parametric statistical test is a test whose model specifies certain conditions about the parameters of the population from which the research sample was drawn. Since these conditions are not ordinarily tested, they are assumed to hold. The meaningfulness of the results of a parametric test depends on the validity of these assumptions. Parametric tests also require that the scores under analysis result from measurement in the strength of at least an interval scale.

A nonparametric statistical test is a test whose model does not specify conditions about the parameters of the population from which the sample was drawn. Certain assumptions are associated with most nonparametric statistical tests, i.e. that the observations are independent and that the variable under study has underlying continuity, but these assumptions are fewer and much weaker than those associated with parametric tests. Moreover, nonparametric tests do not require measurement so strong as that required for the parametric tests. Most nonparametric tests apply to data in an ordinal scale, and some apply also to data in a nominal scale.

## Advantages of Nonparametric tests

Probability statements obtained from most nonparametric

Statistical tests are exact probabilities (except in the case of large samples, where excellent approximations are available), regardless of the shape of the population distribution from which the random sample was drawn.

2. If sample sizes as small as  $N=6$  are used, there is NO alternative to using a nonparametric statistical test unless the nature of the population distribution is known exactly.
3. Nonparametric statistical tests are available to treat data which are inherently in ranks as well as data whose seemingly numerical scores have the strength of ranks. That is, the researcher may only be able to say of his subjects that one has more or less of the characteristic than another, without being able to say how much more or less.
4. Nonparametric methods are available to treat data which are simply classificatory, i.e. measured in nominal scale. No parametric test technique applies to such data.
5. Nonparametric statistical tests are typically much easier to learn & to apply than are parametric tests.

## Disadvantages of nonparametric Stat., etc.

1. If all the assumptions of the parametric statistical model are in fact met in the data, and if the measurement is of the required strength, then non parametric stat. tests are wasteful of data.
2. There are as yet No nonparametric methods for testing interactions in ANOVA model, unless special assumptions are made about additivity.

### I) Chi-square One-sample test

Frequently researches is undertaken in which the researcher is interested in the number of subjects, objects, or responses which fall in various categories. e.g. persons may be categorized according to whether they are 'favour of', 'indifferent to' or 'opposed to' some statement of opinion, to enable the researcher to test the hypothesis that these responses will differ in frequency.

#### Method :

In order to be able to compare an observed with an expected group of frequencies, we must of course be able to state what frequencies would be expected. The null hypothesis states the proportion of objects falling in each of the categories in the presumed population. The  $\chi^2$  technique tests whether the observed frequencies are sufficiently close to the expected ones to be likely to have occurred under  $H_0$ .

7.4

(4)

The null hypothesis may be tested by

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{k-1}$$

where  $O_i$  : observed # of cases categorized in  $i^{th}$  category  
 $E_i$  : expected # of cases categorized in  $i^{th}$  category  
 $\sum_{i=1}^k$  directs one to sum over all ( $k$ ) categories

If the agreement between the observed and expected frequencies is close, the difference  $(O_i - E_i)$  will be small & consequently  $\chi^2$  will be small.  
 Roughly speaking, the larger  $\chi^2$  is, the more likely it is that the observed frequencies did not come from the population on which null hypothesis is based.

e.g.  $\alpha = 0.05$

$H_0$ : there is no difference in expected # of winners starting from each of the post positions

1 day has 8 races, 144 is the total # of winners in 18 days of races

Post Position

	1	2	3	4	5	6	7	8	Total
expected # wins	18	18	18	18	18	18	18	18	144
obs. No. of wins	29	19	18	25	17	10	15	11	144

75

(5)

$$\chi_{\text{obs}}^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} = \frac{(29-18)^2}{18} + \frac{(19-18)^2}{18} + \frac{(18-18)^2}{18} \\ + \frac{(25-18)^2}{18} + \frac{(17-18)^2}{18} + \frac{(10-18)^2}{18} + \frac{(15-18)^2}{18} \\ + \frac{(11-18)^2}{18} = 16.3$$

$$\chi_{0.01}^2 (7) = 18.48$$

$$\chi_{\text{obs}}^2 = 16.3 < 18.48$$

Do NOT reject  $H_0$  at 0.01 level.

## II) Mc Nemar test for significance of changes

The Mc Nemar test for the significance of changes is particularly applicable to those "before and after" designs in which each person is used as his own control and in which measurement is in the strength of either a nominal or ordinal scale.

		After	
		+	-
Before	+	A	B
	-	C	D

$H_0$ : No change before & after

Notice that those cases which show changes between the first and second response appear in cells B and C. An individual is tallied in cell B if he/she changed from + to -. He/she is tallied in C if he/she changed from - to +. If no change is observed, he/she is tallied in either cell A (+ responses both before & after) or cell D (- responses both before & after).

7.6

(6)

Since  $B+C$  represents the total number of persons who changed, the expectation under the null hypothesis would be that  $\frac{1}{2}(B+C)$  cases changed in one direction and  $\frac{1}{2}(B+C)$  cases changed in the other. In other words  $\frac{1}{2}(B+C)$  is the expected frequency under  $H_0$  in both cell  $B$  and cell  $C$ .

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{\left(B - \frac{B+C}{2}\right)^2}{\frac{B+C}{2}} + \frac{\left(C - \frac{B+C}{2}\right)^2}{\frac{B+C}{2}}$$

$$= \frac{(B-C)^2}{(B+C)} \quad \text{with df} = 1 \quad \text{--- (1)}$$

That is, the sampling distribution under  $H_0$  (No change between before and after) of  $\chi^2$  is distributed as approximately as Chi-square with  $df = 1$ .

Correction for continuity. The approximation by the chi-square distribution of the sampling distribution of eqt (1) becomes an excellent one if a correction for continuity is performed.

$$\chi^2 = \frac{(|B-E| - 1)^2}{B+C} \quad \text{with df} = 1$$

e.g. (Poisit's book p207) BSE After interview

BSE Before Yes  
interview No

	Yes	No
Yes	15	0
No	5	30

$$\chi^2 = \frac{(15 - 0)^2}{5+0} = \frac{16}{5} = 3.2$$

For the Mc Nemar test, there is always 1 degree of freedom.  $\chi^2_{0.05}(1) = 3.84$ ,  $3.2 < 3.84$

We do NOT reject null hypothesis. The change in % of women who practised BSE is NOT stat. sig. at the 0.05 level.

### Wilcoxon Signed - Rank test

The Wilcoxon signed - rank test uses a random sample of matched pairs of observations. We wish to test the null hypothesis that the two population distributions are identical or the null hypothesis that the distribution of differences is centred at 0.

Procedure :

- (a) discard pairs for which the difference is 0 and rank the absolute differences in ascending order.
- (b) calculate the sum of the ranks for +ve differences and for -ve differences. The observed Wilcoxon signed - rank test statistic  $t_0$  is the smaller of these two sums.
- (c) When sample size  $\geq 15$ , the observed statistic  $Z = \frac{t_0 - \mu_T}{\sigma_T} \sim N(0, 1)$ ,  $\mu_T = \min(T_p, T_N)$

$$\text{where } \mu_T = \frac{n(n+1)}{4}, \quad \sigma_T^2 = \frac{n(n+1)(2n+1)}{24}$$

Case 1 : If the alternative hypothesis is one-sided,  
reject  $H_0$  if  $\bar{z} < -z_\alpha$ ,  $\alpha$ : level of significance

Case 2 : If the alternative hypothesis is two-sided,  
reject  $H_0$  if  $|\bar{z}| > z_{\alpha/2}$ .

Any tied values are given the same rank, which  
is the average rank of the tied values.

e.g. A soft drink bottler has produced a new drink  
using two different recipes, one of which is much  
sweeter. The bottler asks 20 individuals to taste  
both drinks and rate the drinks on a scale of 1  
to 10, where 10 means the individual likes the  
drink very much. Use the Wilcoxon signed-rank  
test to test the null hypothesis that neither  
drink is preferred over the other. Use a two-  
tailed test and a 5% level of significance.

Sol. : The sum of the ranks for the positive difference  
is  $T_p = 154$  and for negative differences,  $T_n = 17$ .  
The Wilcoxon signed-rank test statistic is  
the smaller of these two values.

$$T_0 = \min(T_p, T_n) = \min(154, 17) = 17$$

$$\mu_T = \frac{n(n+1)}{4} = \frac{18(19)}{4} = 85.5$$

$n = 18$  because 2 pairs of observations deleted due to 0 diff.

7.9

(9)

$$\sigma_T^2 = \frac{n(n+1)(2n+1)}{24} = \frac{(18)(19)(37)}{24} = 527.25$$

$$z = \frac{t_0 - \mu_T}{\sigma_T} = \frac{17 - 85.5}{\sqrt{527.25}} = -2.98$$

$$z_{\alpha/2} = 1.96,$$

Critical value of the test =  $-z_{\alpha/2} = -1.96$

Since  $-2.98 < -z_{\alpha/2} = -1.96$  reject  $H_0$  at 0.05 level.

Drink A	Drink B	Difference	Sign <sup>n</sup> A-B	Rank 1/2c	Rank 1/2d
10	6	4	+	13	
8	5	3	+	8.5	
6	2	4	+	13	
8	2	6	+	16	
-7	4	3	+	8.5	
-3	6	-1	-	2	
1	4	-3	-	8.5	
3	5	-2	-	4.5	
9	9	0	Omit		
-7	8	-1	-	2	
-4	2	2	+	4.5	
5	2	3	+	8.5	
8	1	7	+	18	
6	3	3	+	8.5	
-8	2	6	+	16	
7	6	1	+	2	
4	1	3	+	8.5	
8	2	6	+	16	
9	5	4	Omit	13	
3	3	0		T <sub>D</sub> = 154	T <sub>n</sub> = 17

7.9a

SPSS example for Chapter 7.9 data:

-> get file='d:\stat601.14\polit\17.9sav'.

-> npar tests wilcoxon=drinka with drinkb.

----- Wilcoxon Matched-Pairs Signed-Ranks Test

DRINKA  
with DRINKB

Mean Rank	Sum of Ranks	Cases	
11.00	154.0	14	- Ranks (DRINKB LT DRINKA)
4.25	17.00	4	+ Ranks (DRINKB GT DRINKA)
		2	0 Ties (DRINKB EQ DRINKA)
		--	
		20	Total

Z = -3.0003      2-Tailed P = .0027

-> t-test pairs=drinka with drinkb.

t-tests for Paired Samples

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
DRINKA	20	.256	.276	6.3000	2.386	.534
DRINKB				3.9000	2.292	.512

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
2.4000	2.854	.638	3.76	19	.001
95% CI (1.064, 3.736)					

7.10

(10)

When  $n \leq 15$ , the normal distribution does not necessarily provide a good approximation to the distribution of random variable  $T$ . For small values of  $n$ , tables of probabilities for  $T$  are available. For given values of  $n$  and  $\alpha$

$$P(T \leq T_\alpha) = \alpha. \quad \text{Reject } H_0 \text{ if } t_0 < T_\alpha.$$

e.g. Two makes of tires are tested on the rear wheel of 6 different cars. The number of miles traveled until a tire fails is recorded. Because one tire of each make is used on each car, the observations occurred in matched pairs. Use the Wilcoxon signed-rank test and a 5% level of significance to test the null hypothesis

$H_0$ : The population means are equal

vs  $H_A$ : The population means are NOT equal.

Car	Tire A (in miles)	Tire B (in miles)	A-B	Rank for A	Rank for B
1	20000	19000	-1000	1	
2	24600	23000	-1600	2	
3	32500	37000	-4500		3
4	36000	36100	-100		
5	37000	25500	-11500	4	
6	23000	39500	-16500	5	
					6

Soln:  $T_p = 12, T_n = 9, t_0 = \min(12, 9) = 9$

$T_{0.05} = 1$ .  $t_0 = 9 > 1$  Do NOT reject  $H_0$  at 0.05 level.

7.10a

A20

Appendix A • Tables

TABLE A.10 Critical values of the Wilcoxon test statistic

The following table gives critical values of  $T$  in the Wilcoxon matched-pair signed-rank test.

 $n = 5, 6, 7, \dots, 50$ 

1-sided	2-sided	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$\alpha = .05$	$\alpha = .10$	1	2	4	6	8	11
$\alpha = .025$	$\alpha = .05$		1	2	4	6	8
$\alpha = .01$	$\alpha = .02$			0	2	3	5
$\alpha = .005$	$\alpha = .01$				0	2	3
1-sided	2-sided	$n = 11$	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$
$\alpha = .05$	$\alpha = .10$	14	17	21	26	30	36
$\alpha = .025$	$\alpha = .05$	11	14	17	21	25	30
$\alpha = .01$	$\alpha = .02$	7	10	13	16	20	24
$\alpha = .005$	$\alpha = .01$	5	7	10	13	16	19
1-sided	2-sided	$n = 17$	$n = 18$	$n = 19$	$n = 20$	$n = 21$	$n = 22$
$\alpha = .05$	$\alpha = .10$	41	47	54	60	68	75
$\alpha = .025$	$\alpha = .05$	35	40	46	52	59	66
$\alpha = .01$	$\alpha = .02$	28	33	38	43	49	56
$\alpha = .005$	$\alpha = .01$	23	28	32	37	43	49
1-sided	2-sided	$n = 23$	$n = 24$	$n = 25$	$n = 26$	$n = 27$	$n = 28$
$\alpha = .05$	$\alpha = .10$	83	92	101	110	120	130
$\alpha = .025$	$\alpha = .05$	73	81	90	98	107	117
$\alpha = .01$	$\alpha = .02$	62	69	77	85	93	102
$\alpha = .005$	$\alpha = .01$	55	61	68	76	84	92
1-sided	2-sided	$n = 29$	$n = 30$	$n = 31$	$n = 32$	$n = 33$	$n = 34$
$\alpha = .05$	$\alpha = .10$	141	152	163	175	199	201
$\alpha = .025$	$\alpha = .05$	127	137	148	159	171	183
$\alpha = .01$	$\alpha = .02$	111	120	130	141	151	162
$\alpha = .005$	$\alpha = .01$	100	109	118	128	138	149
1-sided	2-sided	$n = 35$	$n = 36$	$n = 37$	$n = 38$	$n = 39$	
$\alpha = .05$	$\alpha = .10$	214	228	242	256	271	
$\alpha = .025$	$\alpha = .05$	195	208	222	235	250	
$\alpha = .01$	$\alpha = .02$	174	186	198	211	224	
$\alpha = .005$	$\alpha = .01$	160	171	183	195	208	

Example: If  $n = 30$ , then  $P(T \geq 120) = .01$  and  $P(T \geq 109) = .005$ .

7.106

A21

## ix A • Tables

TABLE A.10 Critical values of the Wilcoxon test statistic (*continued*)*n* = 5, 6, 7, . . . , 50

1-sided	2-sided	<i>n</i> = 40	<i>n</i> = 41	<i>n</i> = 42	<i>n</i> = 43	<i>n</i> = 44	<i>n</i> = 45
$\alpha = .05$	$\alpha = .10$	287	303	319	336	353	371
$\alpha = .025$	$\alpha = .05$	264	279	295	311	327	344
$\alpha = .01$	$\alpha = .02$	238	252	267	281	297	313
$\alpha = .005$	$\alpha = .01$	221	234	248	262	277	292
1-sided	2-sided	<i>n</i> = 46	<i>n</i> = 47	<i>n</i> = 48	<i>n</i> = 49	<i>n</i> = 50	
$\alpha = .05$	$\alpha = .10$	389	408	427	446	466	
$\alpha = .025$	$\alpha = .05$	361	379	397	415	434	
$\alpha = .01$	$\alpha = .02$	329	345	362	380	398	
$\alpha = .005$	$\alpha = .01$	307	323	339	356	373	

From Wilcoxon, F. and R. A. Wilcox. "Some Rapid Approximate Statistical Procedures," 1964. Reprinted by permission of Lederle Labs, a division of the American Cyanamid Co.

7.11

(11)

## Cochran's Q-Test (Polit p209)

Cochran's Q-test can be used to test for population differences in proportions when the dependent variable is dichotomous and when there are three or more repeated observations or correlated groups. For example, suppose a sample of 10 elderly patients with constipation problems was put on a special fiber-rich diet, and bowel movement were recorded for 3 consecutive days, beginning with the day the treatment was initiated.

Patient	Day 1	Day 2	Day 3	Row Sum	Row sum $^2$
1	1	0	1	2	4
2	0	1	1	2	4
3	0	1	0	1	1
4	0	1	1	2	4
5	1	1	1	3	9
6	0	0	1	1	1
7	0	1	0	1	1
8	1	0	1	2	4
9	0	1	1	2	4
10	0	1	1	2	4

$$S_{C_1} = 3 \quad S_{C_2} = 7 \quad S_{C_3} = 8 \quad S_R = 18 \quad S_R^2 = 36$$

Codes : 0 = No bowel movement

1 = bowel movement

According to this table, a total of three patients had bowel movements on Day 1, 7 patients & 8 patients had bowel movements on Day 2 & Day 3 respectively. The research question is whether the differences

7.12

(2)

reflect true population changes or are the result of sampling fluctuations.

$$Q = \frac{k(k-1) \sum_{i=1}^k (S_{C_i} - M_R)^2}{k(S_R) - S_R^2}$$

where  $k$ : # of times of obs.

$S_C$ : sum of each column

$S_R$ : sum of all summed rows

$$M_R = \frac{S_R}{k}$$

$$M_R = \frac{18}{3} = 6$$

$$Q = \frac{3(3-1) [(3-6)^2 + (7-6)^2 + (8-6)^2]}{3(18) - 36}$$

$$= 4.67$$

When sample size  $\geq 10$ ,  $Q \sim \chi^2_{(k-1)}$

$$\chi^2_{0.05}(2) = 5.99$$

$Q < 5.99$  Do Not reject  $H_0$  at 0.05 level.

The Friedman Test (Polit p211)

Like Cochran's Q, the Friedman Test is used when there are three or more correlated groups or matched sets of observations for the same subjects. However, the Friedman test is used when the dependent variable

get file='d:\stat601.14\polit\p210.sav'.  
npar tests cochran=day1 day2 day3.

7.12 a

## NPar Tests:p210.spo

### Cochran Test

#### Frequencies

	Value	
	0	1
day1	7	3
day2	3	7
day3	2	8

#### Test Statistics

N	10
Cochran's Q	4.667 <sup>a</sup>
df	2
Asymp. Sig.	.097

a. 1 is treated as a success.

7/13

(13)

is measured on ordinal scale.

The Friedman test involves ranking the scores for each subject across the different conditions.

g. H<sub>0</sub>: Nurses' scores on aggressiveness of nursing care are unrelated to the type of patient's illness. Score (rank)

Nurse	AIDS	Cancer	Alzheimer
1	17 (2)	18 (3)	15 (1)
2	15 (2)	20 (3)	11 (1)
3	14 (3)	12 (1)	13 (2)
4	11 (1)	19 (3)	18 (2)
5	18 (2)	20 (3)	17 (1)
6	16 (3)	14 (1)	15 (2)
7	12 (1)	14 (3)	13 (2)
8	9 (1)	13 (3)	12 (2)
9	16 (2)	17 (3)	15 (1)
	<u>R<sub>1</sub></u> 17	<u>R<sub>2</sub></u> 23	<u>R<sub>3</sub></u> 14

$$\chi^2_{\text{calc}} = \left[ \frac{12 (\sum R^2)}{N k (k+1)} \right] - 3 N (k+1) \sim \chi^2(k-1)$$

k : # conditions

R : sum of the ranks for each k condition

$\sum$  : sum of squared sum of ranks ( $R^2$ ) for the k conditions

N : # of subjects

$$\chi^2_{\text{calc}} = \frac{12 (17^2 + 23^2 + 14^2)}{9 (3) (4)} - (3)(9)(4) = 4.67$$

$$\chi^2_{0.05}(2) = 5.99, \quad \chi^2_{\text{calc}} = 4.67 < 5.99 \quad \text{Do NOT reject H}_0$$

```
get file='d:\stat601.14\polit\p212.sav'.  
npar tests friedman=aids cancer alzh.
```

7.14

## NPar Tests:p212.spo

### Friedman Test

Ranks

	Mean Rank
aids	1.89
cancer	2.56
alzh	1.56

$= \frac{17}{9}$   
 $= \frac{23}{9}$   
 $= \frac{14}{9}$

Test Statistics<sup>a</sup>

N	9
Chi-Square	4.667
df	2
Asymp. Sig.	.097

a. Friedman Test

non parametric approach

parametric approach

### GLM

```
aids cancer alzh  
/WSFACTOR = disease 3 Polynomial  
/METHOD = SSTYPE(3)  
/EMMEANS = TABLES(disease)  
/PRINT = DESCRIPTIVE  
/CRITERIA = ALPHA(.05)  
/WSDESIGN = disease .
```

## General Linear Model

### Within-Subjects Factors

Measure: MEASURE\_1

disease	Dependent Variable
1	aids
2	cancer
3	alzh

### Descriptive Statistics

	Mean	Std. Deviation	N
aids	14.2222	2.99073	9
cancer	16.3333	3.12250	9
alzh	14.3333	2.29129	9

7.14a

**Multivariate Tests<sup>b</sup>**

Effect		Value	F	Hypothesis df	Error df	Sig.
disease	Pillai's Trace	.402	2.349 <sup>a</sup>	2.000	7.000	.166
	Wilks' Lambda	.598	2.349 <sup>a</sup>	2.000	7.000	.166
	Hotelling's Trace	.671	2.349 <sup>a</sup>	2.000	7.000	.166
	Roy's Largest Root	.671	2.349 <sup>a</sup>	2.000	7.000	.166

a. Exact statistic

b.

Design: Intercept

Within Subjects Design: disease

**Mauchly's Test of Sphericity<sup>b</sup>**

Measure: MEASURE\_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
disease	.992	.059	2	.971

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

**Mauchly's Test of Sphericity<sup>b</sup>**

Measure: MEASURE\_1

Within Subjects Effect	Epsilon <sup>a</sup>		
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
disease	.992	1.000	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b.

Design: Intercept

Within Subjects Design: disease

**Tests of Within-Subjects Effects**

Measure: MEASURE\_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
disease	Sphericity Assumed	25.407	2	12.704	2.564	.108
	Greenhouse-Geisser	25.407	1.983	12.810	2.564	.109
	Huynh-Feldt	25.407	2.000	12.704	2.564	.108
	Lower-bound	25.407	1.000	25.407	2.564	.148
Error(disease)	Sphericity Assumed	79.259	16	4.954		
	Greenhouse-Geisser	79.259	15.867	4.995		
	Huynh-Feldt	79.259	16.000	4.954		
	Lower-bound	79.259	8.000	9.907		

7.15

Tests of Within-Subjects Contrasts

Measure: MEASURE\_1

Source	disease	Type III Sum of Squares	df	Mean Square	F	Sig.
disease	Linear	.056	1	.056	.011	.920
	Quadratic	25.352	1	25.352	5.363	.049
Error(disease)	Linear	41.444	8	5.181		
	Quadratic	37.815	8	4.727		

Tests of Between-Subjects Effects

Measure: MEASURE\_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	6045.037	1	6045.037	430.649	.000
Error	112.296	8	14.037		

Estimated Marginal Means

disease

Measure: MEASURE\_1

disease	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	14.222	.997	11.923	16.521
2	16.333	1.041	13.933	18.733
3	14.333	.764	12.572	16.095

7.16

(14)

### Mann-Whitney U-test

The Mann-Whitney U-test is a popular nonparametric analog of the independent groups t-test. This statistic tests the null hypothesis that two populations distributions are identical against the alternative hypothesis that the distribution are NOT identical.

Suppose we had 2 groups of burn patient (Group A and Group B) who obtained the following scores on a scale measuring positive body image:

Group A: 14, 19, 11, 22, 17

Group B: 10, 16, 15, 18, 13

Score	Group	Rank (A)	Rank (B)
10	B		1
11	A	2	
13	B		3
14	A	4	
15	B		5
16	B		6
17	A	7	
18	B		8
19	A	9	
22	A	10	
		<u><math>R_A = 32</math></u>	<u><math>R_B = 23</math></u>

$$U_A = n_A n_B + \frac{n_A(n_A+1)}{2} - R_A$$

$$U_B = n_A n_B + \frac{n_B(n_B+1)}{2} - R_B$$

7.17

where  $n_A$ : # of obs. of group A

$n_B$ : # of obs. of group B

$R_A$ : sum of ranks for group A

$R_B$ : sum of ranks for group B

$$U = \min (U_A, U_B)$$

$$U_A = 8, \quad U_B = 17$$

$$U = \min (8, 17) = 8$$

Look up critical values ( $\alpha=0.05$ ) of  $U$  statistics  
table B-5 on page 476

if  $U < U_{0.05}(n_1, n_2)$  Reject  $H_0$

$$U_{0.05}(5, 5) = 2$$

$U = 8 > 2$  Do NOT reject  $H_0$  at 0.05 level.

Note that table B-5 is appropriate only when sample size of both groups is 20 or less.

When  $n$  for either group is  $> 20$  the value of  $U$  approaches a normal distribution.

$$Z = \frac{U - n_A n_B / 2}{\sqrt{\frac{n_A n_B}{12} (n_A + n_B + 1)}}$$

7.17a

TABLE B-5. CRITICAL VALUES OF THE  $U$  STATISTIC FOR  $\alpha = .05$  (TWO-TAILED TEST)

$n_1 \rightarrow$	1	2	3	4	5	6	7	8	9	10	11	(12)	13	14	15	16	17	18	19	20
$n_2 \downarrow$	— <sup>a</sup>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	
1	— <sup>a</sup>	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	
2	—	—	—	—	—	—	—	0	0	0	0	1	1	1	1	1	2	2	2	
3	—	—	—	—	0	1	1	2	2	3	3	3	4	4	5	5	6	6	7	
4	—	—	—	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	
5	—	—	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	
6	—	—	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	
7	—	—	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	
8	—	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	
9	—	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	
10	—	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	
11	—	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	
12	—	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	
13	—	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	
14	—	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	
15	—	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	
16	—	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	
17	—	2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	
18	—	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	
19	—	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	
20	—	2	8	13	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	

Note: To be statistically significant, the calculated  $U$  must be *equal to or less than* the tabled value.<sup>a</sup>A dash indicates that no decision is possible for the specified  $n_s$ .

sig. if  $U_{\text{calc}} < \text{the tabled value}$

This is from Polit's Page 203 (Mann-Whitney U-Test).

-> get file='d:\stat601.14\polit\p203.sav'.

-> npar test m-w=score by gp(1,2).

- - - - - Mann-Whitney U - Wilcoxon Rank Sum W Test

SCORE  
by GP

Mean Rank	Sum of Ranks	Cases	
6.40	32.00	5 GP	= 1.00
4.60	23.00	5 GP	= 2.00
	--		
	10 Total		

$$Z = \frac{8 - \frac{5 \times 5}{2}}{\sqrt{\frac{25(11)}{12}}} = \frac{-4.5}{\sqrt{22.9167}} = \frac{-4.5}{4.787}$$

$$= -0.9400$$

U	W	Exact** 2*(One-Tailed P)	Z	2-Tailed P
8.0	23.0	.4206	-.9400	.3472

\*\*This exact p-value is not corrected for ties.

-> t-test groups=gp(1,2)/variables=score.

t-tests for Independent Samples of GP

Variable	Number of Cases	Mean	SD	SE of Mean
<hr/>				
SCORE				
GP 1	5	16.6000	4.278	1.913
GP 2	5	14.4000	3.050	1.364

Mean Difference = 2.2000

Levene's Test for Equality of Variances: F= .621 P= .453

Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
					t-test for Equality of Means
Equal	.94	8	.376	2.349	(-3.218, 7.618)
Unequal	.94	7.23	.379	2.349	(-3.320, 7.720)

7.19

(16)

## The Kruskal-Wallis Test (Polit p205)

The Kruskal-Wallis test is the nonparametric counterpart of the simple one-way ANOVA. It is used to analyze the relationship between a dependent variable that is ordinal in nature and a categorical independent variable that has three or more levels.

The K-W procedure tests the null hypothesis that the population distributions for the three (or more) independent groups are identical against that there are differences in the distributions. This test should be used only if there are five or more cases per group.

Suppose that we compared the life satisfaction of patients in three nursing homes, using a six-item scale, and obtained the following scores:

Home A : 6, 12, 18, 14, 17

Home B : 15, 19, 16, 20, 10

Home C : 30, 27, 24, 25, 22

Score : 6, 10, 12, 14, 15, 16, 17, 18, 19, 20, 22, 24, 25, 27, 30

Group : A B A A B B A A B B C C C C

Rank : 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

$$R_A = 1+3+4+7+8 = 23 \rightarrow \text{mean rank} = 4.6$$

$$R_B = 2+5+6+9+10 = 32 \rightarrow \text{mean rank} = 6.6$$

$$R_C = 11+12+13+14+15 = 65 \rightarrow \text{mean rank} = 13.0$$

7.20

(17)

Kruskal and Wallis proposed the following formula for the test statistic, the H statistic:

$$H = \left[ \frac{12}{N(N+1)} \right] \left[ \sum_{i=1}^k \frac{R_i^2}{n_i} \right] - 3(N+1)$$

where N: total sample size

$R_i$ : summed ranks for group i

$n_i$ : # of observations for group i

$$H_{\text{calc}} = \left[ \frac{12}{15(16)} \right] \left[ \frac{23^2}{5} + \frac{32^2}{5} + \frac{65^2}{5} \right] - 3 \times 16$$

$$= 0.05 (105.8 + 204.8 + 845.0) - 48 = 9.78$$

The H statistic has a sampling distribution that approximates a Chi-square distribution with  $k-1$  degrees of freedom, where k is the number of groups.

$$H_{\text{calc}} = 9.78 > \chi^2_{0.05}(2) = 5.99$$

Reject the null hypothesis that the distribution of life satisfaction scores in the three nursing homes is identical.

get file='d:\stat601.14\polit\p205.sav'.  
 get file='d:\stat601.14\polit\p205.sav'.  
 npar tests k-w=satisfy by home(1,3).

7.21

## NPar Tests:p205.spo

### Kruskal-Wallis Test

Ranks

home	N	Mean Rank
satisfy	1.00 A	5
	2.00 B	5
	3.00 C	5
	Total	15

### Test Statistics<sup>a,b</sup>

	satisfy
Chi-Square	9.780
df	2
Asymp. Sig.	.008

a. Kruskal Wallis Test

b. Grouping Variable: home

*Nonparametric approach*

### ONEWAY

satisfy BY home  
 /STATISTICS DESCRIPTIVES HOMOGENEITY  
 /MISSING ANALYSIS  
 /POSTHOC = SNK TUKEY SCHEFFE LSD ALPHA(.05).

*parametric approach*.

### Oneway

#### Descriptives

satisfy

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
					Lower Bound	Upper Bound
1.00 A	5	13.4000	4.77493	2.13542	7.4711	19.3289
2.00 B	5	16.0000	3.93700	1.76068	11.1116	20.8884
3.00 C	5	25.6000	3.04959	1.36382	21.8134	29.3866
Total	15	18.3333	6.56470	1.69500	14.6979	21.9687

## Descriptives

satisfy

	Minimum	Maximum
1.00 A	6.00	18.00
2.00 B	10.00	20.00
3.00 C	22.00	30.00
Total	6.00	30.00

## Test of Homogeneity of Variances

satisfy

Levene Statistic	df1	df2	Sig.
.351	2	12	.711

## ANOVA

satisfy

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	412.933	2	206.467	13.013	.001
Within Groups	190.400	12	15.867		
Total	603.333	14			

## Post Hoc Tests

## Multiple Comparisons

Dependent Variable: satisfy

	(I) home	(J) home	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1.00 A	2.00 B	-2.60000	2.51926	.572	-9.3210	4.1210
		3.00 C	-12.20000*	2.51926	.001	-18.9210	-5.4790
	2.00 B	1.00 A	2.60000	2.51926	.572	-4.1210	9.3210
		3.00 C	-9.60000*	2.51926	.006	-16.3210	-2.8790
	3.00 C	1.00 A	12.20000*	2.51926	.001	5.4790	18.9210
		2.00 B	9.60000*	2.51926	.006	2.8790	16.3210
Scheffe	1.00 A	2.00 B	-2.60000	2.51926	.600	-9.6226	4.4226
		3.00 C	-12.20000*	2.51926	.002	-19.2226	-5.1774
	2.00 B	1.00 A	2.60000	2.51926	.600	-4.4226	9.6226
		3.00 C	-9.60000*	2.51926	.009	-16.6226	-2.5774
	3.00 C	1.00 A	12.20000*	2.51926	.002	5.1774	19.2226
		2.00 B	9.60000*	2.51926	.009	2.5774	16.6226
LSD	1.00 A	2.00 B	-2.60000	2.51926	.322	-8.0890	2.8890
		3.00 C	-12.20000*	2.51926	.000	-17.6890	-6.7110
	2.00 B	1.00 A	2.60000	2.51926	.322	-2.8890	8.0890
		3.00 C	-9.60000*	2.51926	.002	-15.0890	-4.1110
	3.00 C	1.00 A	12.20000*	2.51926	.000	6.7110	17.6890
		2.00 B	9.60000*	2.51926	.002	4.1110	15.0890

\*. The mean difference is significant at the .05 level.

## Homogeneous Subsets

satisfy

	home	N	Subset for alpha = .05	
			1	2
Student-Newman-Keuls <sup>a</sup>	1.00 A	5	13.4000	
	2.00 B		16.0000	
	3.00 C			25.6000
	Sig.		.322	1.000
Tukey HSD <sup>a</sup>	1.00 A	5	13.4000	
	2.00 B		16.0000	
	3.00 C			25.6000
	Sig.		.572	1.000
Scheffe <sup>a</sup>	1.00 A	5	13.4000	
	2.00 B		16.0000	
	3.00 C			25.6000
	Sig.		.600	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.