BASIC NON-PARAMETRIC STATISTICAL TOOLS*

Prepared for GCMA 2001

Gregory S. Rash
Peter M. Quesada

* Examples presented in these notes were obtained from Primer of Biostatistics by Stanton S. Glantz (McGraw Hill Text; ISBN: 0070242682)

Ordinal Data – Evaluating Two Interventions on Two Different Groups

- Mann-Whitney Rank-Sum Test
- Based on ranking of all observations without regard to group associated with each observation
- Can also be used with interval or ratio data that are not normally distributed
- Test statistic, T, is sum of all ranks for the smaller group

$$T = \sum_{i=1}^{n_S} R_i$$

- where R_i is the rank of the i^{th} observation of the smaller group and n_S is the number of observations in the smaller group
- To determine T must first rank all observations from both groups together
- Tied ranks receive average of ranks that would have been spanned (e.g. if 3 observations are tied following rank 4, then each of the tied observations would receive the average of ranks 5, 6 and 7, or (5+6+7)/2 = 6; the next observation would receive rank 8)
- Critical values of T are based on the tails of the distribution of all possible T values (assuming no ties)

Example: 2 groups with 3 observations in one and 4 observations in the other

Group 1		Group 2	
Observed Value	Overall Rank	Observed Value	Overall Rank
1000	1	1400	6
1380	5	1600	7
1200	3	1180	2
		1220	4

T (based on Group 1) = 1 + 5 + 3 = 9

- To determine probability of obtaining a particular T value consider all possible rankings for group 1 observations.

1st Observation Ranks	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2
2nd Observation Ranks	2	2	2	2	2	3	3	3	3	4	4	4	5	5	6	3	3	3
3rd Observation Ranks	3	4	5	6	7	4	5	6	7	5	6	7	6	7	7	4	5	6
Sum of Ranks	6	7	8	9	10	8	9	10	11	10	11	12	12	13	14	9	10	11

1st Observation Ranks	2	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	5
2nd Observation Ranks	3	4	4	4	5	5	6	4	4	4	5	5	6	5	5	6	6
3rd Observation Ranks	7	5	6	7	6	7	7	5	6	7	6	7	7	6	7	7	7
Sum of Ranks	12	11	12	13	13	14	15	12	13	14	14	15	16	15	16	17	18

- Thirty five possible combinations of ranks
- If all observations are actually drawn from the same population, then each combination is equally possible, and the distribution is as shown below

						X						
				X	X	X	X	X				
			X	X	X	X	X	X	X			
		X	X	X	X	X	X	X	X	X		
X	X	X	X	X	X	X	X	X	X	X	X	X
6	7	8	9	10	11	12	13	14	15	16	17	18

- If all observations were truly from a single population then there would be a 2/35 = 0.057 (5.7%) probability of obtaining one of the two extreme T values (6 or 18).
- Similarly, there would be a 4/35 = 0.114 (11.4%) probability for a T value ≤ 7 or ≥ 17.
- Note that these probabilities are discrete in nature.
- In present example $T = 9$ is associated with probability of 14/35 = 0.4 (40%), which would not be extreme enough to reject null hypothesis that all observations were drawn from the same population.
- When the larger sample contains eight or more observations, distribution of T approximates a normal distribution with mean

$$\mu_T = \frac{n_S(n_S + n_B + 1)}{2}$$

where n_B is the number of samples in the bigger group, and standard deviation

$$\theta_T = \sqrt{\frac{n_S n_B (n_S + n_B + 1)}{12}}$$

- Can then construct test statistic, z_T

$$z_T = \frac{T - \mu_T}{\theta_T}$$

which can be compared with t-distribution with infinite degrees of freedom (d.o.f.)

- This comparison is more accurate with a continuity correction where

$$z_T = \frac{|T - \mu_T| - \frac{1}{2}}{\theta_T}$$

Ordinal Data – Evaluating Three or more Interventions on Different Groups of Individuals

- Kruskal-Wallis Statistic
- Based on ranking of all observations without regard to group associated with each observation
- Test statistic, *H*, is a normalized, weighted sum of squared differences between each group's mean rank and the overall mean rank
- To determine *H* must first rank all observations without regard for groups
- Tied ranks receive average of ranks that would have been spanned
- Mean rank is determined for each group, *j*, as

$$\bar{R}_j = \frac{\sum_{i=1}^{n_j} R_{ji}}{n_j}$$

where R_{ji} is the rank of the i^{th} observation of the j^{th} group and n_j is the number of group *j* observations

- Overall mean rank is

$$\bar{R}_T = \frac{\sum_{i=1}^{N} i}{N} = \frac{N+1}{2}$$

where *N* is the total number of observations ($N = \sum_{j=1}^{m} n_j$ where *m* is the number of groups)

- The weighted sum of squared differences is

$$D = \sum_{j=1}^{m} n_j \left(\bar{R}_j - \bar{R}_T\right)^2$$

- *H* is computed by dividing *D* by $N(N+1)/12$ which results in a test statistic value that does not depend on sample size

$$H = \frac{D}{N(N+1)/12} = \frac{12}{N(N+1)} \sum_{j=1}^{m} n_j \left(\bar{R}_j - \bar{R}_T\right)^2$$

- If no real difference exists between interventions then mean group ranks should be close to overall mean rank; *D* and, subsequently, *H* should be smaller values that would preclude rejection of the null hypothesis
- Critical values of *H* are based on the tails of the distribution of all possible *H* values (assuming no ties)
- If sample sizes are sufficiently large ($n_j \geq 5$ for $m = 3$; $N > 10$ when $m = 4$) then the distribution of *H* approximates the χ^2 distribution with d.o.f., $\nu = m$-1.

Example: 3 groups with different number observations

Group 1		Group 2		Group 3	
Observed Value	Overall Rank	Observed Value	Overall Rank	Observed Value	Overall Rank
2.04	1	5.30	12	10.36	25
5.16	10	7.28	19	13.28	29
6.11	15	9.98	21	11.81	28
5.82	14	6.59	16	4.54	6
5.41	13	4.59	8	11.04	26
3.51	4	5.17	11	10.08	24
3.18	2	7.25	18	14.47	31
4.57	7	3.47	3	9.43	23
4.83	9	7.60	20	13.41	30
11.34	27				
3.79	5				
9.03	22				
7.21	17				
Rank Sum	146	Rank Sum	128	Rank Sum	222
Mean Rank	11.23	Mean Rank	14.22	Mean Rank	24.67

Overall Mean Rank = (31 + 1) / 2 = 16

$$
\begin{aligned}
H &= \frac{12}{N(N+1)}\sum_{j=1}^{m} n_j\left(\overline{R}_j - \overline{R}_T\right)^2 \\
&= \frac{12}{31(31+1)}\left[13(11.23-16)^2 + 9(14.22-16)^2 + 9(24.67-16)^2\right] \\
&= 12.107 > \chi_{.01,v=2} = 9.201
\end{aligned}
$$

- Reject null hypothesis that all observations from a single population
- To determine where differences exist perform pair-wise Mann-Whitney tests with Bonferoni adjustments and continuity corrections.

Ordinal Data – Evaluating Two Interventions on the Same Group of Individuals

- Wilcoxon Signed-Rank Test
- Based on ranking of absolute differences between two observations for each individual
- Test statistic, *W*, is sum of all ranks of differences

$$W = \sum_{i=1}^{n} \frac{\Delta_i}{|\Delta_i|} R_i$$

where n is the number of individuals, Δ_i is the difference between observations for the i^{th} individual, and R_i is the rank of the absolute difference for the i^{th} individual (note: the fraction in front of the ranks will always have magnitude, 1, and will have the sign of the difference)

- If no real difference exists between individuals' observations, then the signs of the observed differences should occur by random chance; *W* would then compute to a number close to zero.
- Extreme values of *W* in either positive or negative sense, thus, lead to rejection of the null hypothesis that no difference exists between observations.

Example: 1 group with 2 observations for each individual

Individual	Observation One	Observation Two	Difference	Rank of Difference	Signed Rank of Difference
1	1600	1490	-110	5	-5
2	1850	1300	-550	6	-6
3	1300	1400	+100	4	+4
4	1500	1410	-90	3	-3
5	1400	1350	-50	2	-2
6	1010	1000	-10	1	-1
					W = -13

- For six individuals there are 65 possible combinations of signed ranks (assuming no ties).
- If no real difference exists between observations, then each combination is equally possible, and the distribution is as shown below

									X	X	X	X									
						X	X	X	X	X	X	X	X	X	X						
					X	X	X	X	X	X	X	X	X	X	X	X					
			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
-21	-19	-17	-15	-13	-11	-9	-7	-5	-3	-1	1	3	5	7	9	11	13	15	17	19	21

- For this distribution there is a 4/65 = 0.0625 (6.25%) chance of obtaining a value of *W* at or beyond 19 (or –19) if no real difference exists.
- For present example *W* = -13 is not extreme enough to reject null hypothesis.
- As with other parametric methods, p-values for the Wilcoxon Signed-Rank Test are discrete in nature.
- For large number of individuals, however, distribution of *W* values approximate a normal distribution with mean

$$\mu_W = 0$$

and standard deviation

$$\theta_W = \sqrt{\frac{n(n+1)(2n+1)}{6}}$$

- From which test statistic, z_w can be computed as

$$z_W = \frac{W - \mu_W}{\theta_W} = \frac{W}{\sqrt{n(n+1)(2n+1)/6}}$$

which can be compared with t-distribution with infinite degrees of freedom (d.o.f.)

which with a continuity correction becomes

$$z_W = \frac{|W| - \frac{1}{2}}{\sqrt{n(n+1)(2n+1)/6}}$$

- To handle tied ranks, must first identify type of tie.
- Ties in which difference is zero result in individual being dropped from sample entirely.
- Ties in which difference is non-zero are handled as before.

Ordinal Data – Evaluating Three or More Interventions on the Same Group of Individuals

- Friedman Statistic
- Based on rankings of each individual's observations associated with each intervention
- Initially, S is determined as the sum of squared differences between each intervention's observed rank sum, $R_{Tj} = \sum_{i=1}^{n} R_{ji}$, and the expected rank sums, $n(k+1)/2$ where n is the number of individuals and k is the number of interventions.

$$S = \sum_{j=1}^{k} \left(R_{Tj} - n(k+1)/2\right)^2$$

- Friedman statistic is then formed by dividing S by $nk(k+1)/12$ to obtain a statistic whose distribution approximates a χ^2 distribution with $\nu = k$-1.

$$\chi_r^2 = \frac{12\sum_{j=1}^{k} \left(R_{Tj} - n(k+1)/2\right)^2}{nk(k+1)}$$

- If no real difference exists between individuals' observations, then observed rank sums should be close to expected rank sums; thus squared differences should be small, and S & χ_r^2 should be close to zero.
- If $n < 9$ for $k = 3$ or $n < 4$ for $k = 4$, distribution of Friedman statistic does not approximate χ^2 distribution; must use actual distribution of Friedman statistic to determine discrete critical values (see Table below)

Example: 1 large group with 6 observations for each individual

	1st Observation		2nd Observation		3rd Observation		4th Observation		5th Observation		6th Observation	
Individual	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
1	193	4	217	6	191	3	149	2	202	5	127	1
2	206	5	214	6	203	4	169	2	189	3	130	1
3	188	4	197	6	181	3	145	2	192	5	128	1
4	375	3	412	6	400	5	306	2	387	4	230	1
5	204	5	199	4	211	6	170	2	196	3	132	1
6	287	3	310	5	304	4	243	2	312	6	198	1
7	221	5	215	4	213	3	158	2	232	6	135	1
8	216	5	223	6	207	3	155	2	209	4	124	1
9	195	4	208	6	186	3	144	2	200	5	129	1
10	231	6	224	4	227	5	172	2	218	3	125	1
R_T		44		53		39		20		49		10

$$n(k+1)/2 = 10(6+1)/2 = 35$$

$$\chi_r^2 = \frac{12\sum_{j=1}^{k} \left(R_{Tj} - n(k+1)/2\right)^2}{nk(k+1)} = \frac{12\left[(44-35)^2 + (53-35)^2 + (39-35)^2 + (20-35)^2 + (49-35)^2 + (10-35)^2\right]}{(10)(6)(6+1)}$$

$$\chi_r^2 = 38.63 > \chi_{.001,\nu=5}^2 = 20.515$$

- Reject null hypothesis that no difference exists between interventions

Example: 1 small group with 3 observations for each individual

	1st Observation		2nd Observation		3rd Observation	
Individual	Value	Rank	Value	Rank	Value	Rank
1	22.2	3	5.4	1	10.6	2
2	17.0	3	6.3	2	6.2	1
3	14.1	3	8.5	1	9.3	2
4	17.0	3	10.7	1	12.3	2
R_T		12		5		7

$$n(k+1)/2 = 4(3+1)/2 = 8$$

$$\chi_r^2 = \frac{12\sum_{j=1}^{k}\left(R_{Tj} - n(k+1)/2\right)^2}{nk(k+1)} = \frac{12\left[(12-8)^2 + (5-8)^2 + (7-8)^2\right]}{(4)(3)(3+1)} = 6.5$$

- From table below χ_r^2 matches value with $p = .042$ for $n = 4$ and $k = 3$ $\Rightarrow$ Reject null hypothesis

k = 3 interventions			k = 4 interventions		
n	χ_r^2	p	n	χ_r^2	p
3	6.00	.028	2	6.00	.042
4	6.50	.042	3	7.00	.054
	8.00	.005		8.20	.017
5	5.20	.093	4	7.50	.054
	6.40	.039		9.30	.011
	8.40	.008	5	7.80	.049
6	5.33	.072		9.93	.009
	6.33	.052	6	7.60	.043
	9.00	.008		10.20	.010
7	6.00	.051	7	7.63	.051
	8.86	.008		10.37	.009
8	6.25	.047	8	7.65	.049
	9.00	.010		10.35	.010
9	6.22	.048			
	8.67	.010			
10	6.20	.046			
	8.60	.012			
11	6.54	.043			
	8.91	.011			
12	6.17	.050			
	8.67	.011			
13	6.00	.050			
	8.67	.012			
14	6.14	.049			
	9.00	.010			
15	6.40	.047			
	8.93	.010			

Ordinal Data – Evaluating Association Between Two Variables

- Spearman Rank Correlation Coefficient
- Based on association between rankings of each variable
- Initially, must rank each variable in either ascending or descending order
- Spearman Rank Correlation Coefficient, r_S is then essentially determined as the Pearson product-moment correlation between the ranks, rather than the actual values of the variables.
- Alternatively, r_S can be computed using the equation

$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n}$$

where d_i is the difference between variable ranks for the i^{th} individual and n is the number of individuals.

- If no real association exists between variables, then the sum of squared differences will tend toward larger values, and r_S will tend toward zero.
- As r_S approaches 1, it becomes less likely that r_S value was obtained by random chance for two variables with no association between them.
- Critical values for r_S are identified from Spearman Rank Correlation Coefficient table depending on acceptable p-value (i.e. chance of falsely concluding that an association exists) and number of individuals (samples).
- If $n > 50$, however, can compute a t-value as

$$t = \frac{r_S}{\sqrt{\left(1 - r_S^2\right) / (n-2)}}$$

which can be evaluated for significance based on $v = n - 2$.

Example:

	Variable 1		Variable 2		
Individual	Value	Rank	Value	Rank	Rank Diff.
1	31	1	7.7	2	-1
2	32	2	8.3	3	-1
3	33	3	7.6	1	2
4	34	4	9.1	4	0
5	35	5.5	9.6	5	0.5
6	35	5.5	9.9	6	-0.5
7	40	7	11.8	7	0
8	41	8	12.2	8	0
9	42	9	14.8	9	0
10	46	10	15.0	10	0

$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n} = 1 - \frac{6\left[(-1)^2 + (-1)^2 + 2^2 + 0^2 + 0.5^2 + (-0.5)^2 + 0^2 + 0^2 + 0^2 + 0^2\right]}{10^3 - 10}$$

$$r_S = 0.96 > r_{S_{p=.001,n=10}} = 0.903$$

- Reject null hypothesis that no association exists between variable 1 and variable 2

	Probability of Greater Value P								
n	0.5	0.2	0.1	0.05	0.02	0.01	0.005	0.002	0.001
4	0.600	1.000	1.000						
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.892	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.217	0.406	0.503	0.587	0.678	0.727	0.769	0.818	0.846
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.791	0.824
14	0.200	0.367	0.164	0.538	0.626	0.679	0.723	0.771	0.802
15	0.189	0.354	0.446	0.521	0.604	0.654	0.700	0.75	0.779
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.729	0.762
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.662	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580
31	0.126	0.236	0.301	0.356	0.418	0.459	0.496	0.541	0.571
32	0.124	0.232	0.293	0.350	0.402	0.452	0.489	0.533	0.563
33	0.121	0.229	0.291	0.345	0.405	0.4446	0.482	0.525	0.554
34	0.120	0.225	0.287	0.340	0.399	0.439	0.475	0.517	0.547
35	0.118	0.222	0.283	0.335	0.394	0.433	0.468	0.510	0.539
36	0.116	0.219	0.279	0.330	0.388	0.427	0.462	0.504	0.533
37	0.114	0.216	0.275	0.325	0.383	0.421	0.456	0.497	0.526
38	0.113	0.212	0.271	0.210	0.378	0.415	0.450	0.491	0.519
39	0.111	0.210	0.267	0.317	0.373	0.410	0.444	0.485	0.513
40	0.110	0.207	0.264	0.313	0.368	0.405	0.439	0.479	0.507
41	0.108	0.204	0.231	0.309	0.364	0.400	0.433	0.473	0.501
42	0.107	0.202	0.257	0.305	0.359	0.395	0.428	0.468	0.495
43	0.105	0.199	0.254	0.301	0.355	0.391	0.423	0.463	0.490
44	0.104	0.197	0.251	0.298	0.351	0.386	0.419	0.458	0.484
45	0.103	0.194	0.248	0.294	0.347	0.382	0.414	0.453	0.479
46	0.102	0.192	0.246	0.291	0.343	0.378	0.410	0.448	0.474
47	0.101	0.190	0.243	0.288	0.340	0.374	0.405	0.443	0.469
48	0.100	0.188	0.240	0.285	0.336	0.370	0.401	0.439	0.465
49	0.098	0.186	0.238	0.282	0.333	0.366	0.397	0.434	0.460
50	0.097	0.184	0.235	0.279	0.329	0.363	0.393	0.43	0.456

<u>Nominal Data – Evaluating Two or More Interventions on Different Groups</u>

- Chi-square Analysis of Contingency
- Based on contingency tables containing cells with numbers of individuals matching row and column specifications
- Two types of contingency tables
  - Observed (actual)
  - Expected
- Chi-Square test statistic, χ^2 , is a sum of normalized squared differences between corresponding cells of observed and expected tables

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

where i is the row index, j is the column index, O_{ij} is the number of observations in cell ij, E_{ij} is the expected number of observations in cell ij, n is the number of rows, and m is the number of columns.

- The expected number of observations for a given cell is determined from the row, column and overall observation totals from the observed table as

$$E_{ij} = \frac{R_i C_j}{T}$$

where R_i is the total number of observations in row i, C_j is the total number observations in column j, and T is the total number of observations in the entire table.

- χ^2 gets larger as observed table deviates more from expected table
- If no real difference exists between cell or row conditions, then larger χ^2 values are less likely to occur due to random chance.
- χ^2 values associated with random chance probabilities less than a critical value (p_{crit}) cause rejection of the null hypothesis.
- χ^2 probabilities obtained from table based on d.o.f. (ν)

$$\nu = (n-1)(m-1)$$

- When $\nu = 1$ (i.e. for 2 X 2 contingency table), should apply Yates correction such that

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\left(\left|O_{ij} - E_{ij}\right| - \frac{1}{2}\right)^2}{E_{ij}}$$

Example: 2 outcomes, 3 classifications (groups, interventions)

Observed Table

	Outcome 1	Outcome 2	Row Totals
Classification 1	14	40	54
Classification 2	9	14	23
Classification 3	46	42	88
Column Totals	69	96	165

Expected Table

	Outcome 1	Outcome 2	Row Totals
Classification 1	(54)(69)/165 = 22.58	(54)(96)/165 = 31.42	54
Classification 2	(23)(69)/165 = 9.62	(23)(96)/165 = 13.38	23
Classification 3	(88)(69)/165 = 36.8	(88)(96)/165 = 51.2	88
Column Totals	69	96	165

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$\chi^2 = \frac{(14-22.58)^2}{22.58} + \frac{(40-31.42)^2}{31.42} + \frac{(9-9.62)^2}{9.62} + \frac{(14-13.38)^2}{13.38} + \frac{(46-36.8)^2}{36.8} + \frac{(42-51.2)^2}{51.2}$$

$$\chi^2 = 9.625 > \chi^2_{.05}(\nu = 2) = 5.991$$

- Reject null hypothesis and conclude that there is a difference in outcomes between the classifications
- Note that results do not yet indicate where the differences are; only that they exist
- Can subdivide contingency table to perform pair-wise comparisons

Example continued:

Observed Table

	Outcome 1	Outcome 2	Row Totals
Classification 2	9	14	23
Classification 3	46	42	88
Column Totals	55	56	111

Expected Table

	Outcome 1	Outcome 2	Row Totals
Classification 2	(23)(55)/111 = 11.40	(23)(56)/111 = 11.60	23
Classification 3	(88)(55)/111 = 43.60	(88)(56)/111 = 44.40	88
Column Totals	55	56	111

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\left(\left|O_{ij} - E_{ij}\right| - \frac{1}{2}\right)^2}{E_{ij}}$$

$$\chi^2 = \frac{\left(\left|9-11.40\right| - \frac{1}{2}\right)^2}{11.40} + \frac{\left(\left|14-11.60\right| - \frac{1}{2}\right)^2}{11.60} + \frac{\left(\left|46-43.60\right| - \frac{1}{2}\right)^2}{43.60} + \frac{\left(\left|42-44.40\right| - \frac{1}{2}\right)^2}{44.40}$$

$$\chi^2 = 0.79 < \chi^2_{.05}(\nu = 1) = 3.841$$

- Cannot reject null hypothesis, so classifications 2 and 3 are deemed to be a single classification
- Classifications 2 and 3 are combined to form a new classification (4) which can then be compared with classification 1

Observed Table

	Outcome 1	Outcome 2	Row Totals
Classification 1	14	40	54
Classification 4	55	56	111
Column Totals	69	96	165

Expected Table

	Outcome 1	Outcome 2	Row Totals
Classification 1	(54)(69)/165 = 22.58	(54)(96)/165 = 31.42	54
Classification 4	(111)(55)/165 = 46.42	(88)(56)/165 = 64.58	111
Column Totals	69	96	165

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\left(\left|O_{ij} - E_{ij}\right| - \frac{1}{2}\right)^2}{E_{ij}}$$

$$\chi^2 = \frac{\left(\left|14 - 22.58\right| - \frac{1}{2}\right)^2}{22.58} + \frac{\left(\left|40 - 31.42\right| - \frac{1}{2}\right)^2}{31.42} + \frac{\left(\left|55 - 46.42\right| - \frac{1}{2}\right)^2}{46.42} + \frac{\left(\left|56 - 64.58\right| - \frac{1}{2}\right)^2}{64.58}$$

$$\chi^2 = 7.390 > \chi^2_{.01}(\nu = 1) = 6.635$$

- Reject null hypothesis, classification 1 differs significantly from the combination of classifications 2 & 3

Nominal Data – Evaluating Two Interventions on the Same Group of Individuals

- McNemar's Test for Changes
- Based on cells in 2 X 2 contingency table that represent individuals with different outcomes for each intervention (cell that represent similar outcomes for each intervention are ignored)
- Chi-Square test statistic, χ^2 , is sum of normalized squared differences between corresponding observed and expected table cells that are not ignored (with $\nu = 1$)

$$\chi^2 = \sum \frac{\left(|O - E| - \frac{1}{2} \right)^2}{E}$$

- Expected value for the remaining cells is computed as the average of the remaining cells

$$E = \frac{\sum O}{2}$$

- If no real difference exists between interventions, then larger χ^2 values are less likely to occur due to random chance.
- χ^2 values associated with random chance probabilities less than a critical value (p_{crit}) cause rejection of the null hypothesis.

Example: 2 outcomes, 3 classifications (groups, interventions)

Observed Table

	Outcome 1	Outcome 2
Outcome 1	~~81~~	48
Outcome 2	23	~~21~~

Expected Table

	Outcome 1	Outcome 2
Outcome 1		(48+23)/2 = 35.5
Outcome 2	(48+23)/2 = 35.5	

Counts of individuals with outcome 1 for both interventions or outcome 2 for both interventions are ignored; χ^2 calculation based on remaining cells.

$$\chi^2 = \sum \frac{\left(|O_{ij} - E_{ij}| - \frac{1}{2} \right)^2}{E_{ij}}$$

$$\chi^2 = \frac{\left(|48 - 35.5| - \frac{1}{2} \right)^2}{35.5} + \frac{\left(|23 - 35.5| - \frac{1}{2} \right)^2}{35.5}$$

$$\chi^2 = 8.113 > \chi^2_{.01}(\nu = 1) = 6.635$$

- Reject null hypothesis and conclude that there is a difference in outcomes between the classifications