





## Chapter 7: Nonparametric Methods

- The Kruskal-Wallis Test (Section 7.2) and Rank Analysis of Covariance (Section 7.6) have been covered elsewhere

### 7.3 Friedman's Chi-Square Test


- Friedman's test (1937) is nonparametric method for analyzing randomized complete block design
- Experimental units are divided into blocks such that units within block are relatively homogeneous. Treatments assigned at random to blocks such that each treatment appears only once in each block

- 
- Standard parametric ANOVA for analyzing randomized complete block design requires assumption of normally distributed experimental errors
  - Friedman's test depends only on ranks of observations within each block, sometimes called two-way analysis of variance by ranks

- 
- For small samples: can use StatXact package (Mehta and Patel 1991)
  - For larger samples (as # of blocks increases): distribution of Friedman statistic approaches chi-square distribution with  $(s - 1)$  degrees of freedom ( $s$  is number of treatments), which is Mantel-Haenszel mean score statistic with rank scores and one subject per treatment group in each block

- Example: Data from experiment to compare performance of five electrode types. All five types were applied to arms of 16 subjects so that each subject is a block.

Subject	Electrode Type				
	1	2	3	4	5
1	500	400	98	200	250
2	660	600	600	75	310
3	250	370	220	250	220
4	72	140	240	33	54
5	135	300	450	430	70
⋮	⋮	⋮	⋮	⋮	⋮
12	160	200	300	300	220
13	250	400	50	50	92
14	170	310	230	20	150
15	66	1000	1050	280	220
16	107	48	26	45	51

- 
- To use Mantel-Haenszel mean score statistic to perform Friedman's test:

```
proc freq;  
    tables subject*type*resist / noprint cmh2  
        scores=rank;  
run;
```


- Results show there is little difference among the five types of electrodes (row mean score chi-square = 5.452, 4 df,  $p=0.244$ )




## **7.4 Aligned Ranks Test for Randomized Complete Blocks**

- When number of blocks or treatments is small, Friedman's test has relatively low power
- Alternative to Friedman's is to use aligned ranks to make blocks more comparable:
  1. Subtract from each observation within a block the average or median of observations
  2. Resulting differences are aligned observations
  3. Rank complete set (not within each block) to get aligned ranks



- 
- Aligned rank test introduced by Hodges and Lehmann (1962), and also by Koch and Sen (1968)
  - Exact distribution is cumbersome to compute. However, null distribution of test statistic is approximately chi-square with  $(s - 1)$  df, where  $s$  is number of treatments (or block size when there is one observation per treatment in each block)

- 
- Example: Consider data in section 7.3 from experiment to compare performance of five electrode types. Compute aligned rank statistic using the following statements:

```
proc standard mean = 0;  
    by subject;  
    var resist;  
proc rank;  
    var resist;  
proc freq;  
    tables subject*type*resist / noprint cmh2;  
run;
```




SUMMARY STATISTICS FOR TYPE BY RESIST  
CONTROLLING FOR SUBJECT

Cochran-Mantel-Haenszel Statistics (Based on Table Scores)

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	4.978	0.026
2	Row Mean Scores Differ	4	13.600	0.009

Total Sample Size = 80

- 
- Results show that there is a significant difference among the five electrode types (row mean score chi-square = 13.6, 4 df,  $p = 0.009$ )
  - In section 7.3, Friedman's test was not significant, this example illustrates potentially greater power of aligned ranks test



## **7.5 Analyzing Incomplete Data**

- When responses are missing, Mantel-Haenszel statistics can still be used to analyze the data.
- The methodology can accommodate a variable number of observations per patient (under the assumption that missing values are missing completely at random and that the test statistic is specified with either table scores or ranks).
- Example: Consider a response variable  $V_{\max}$ , measured at up to four pH levels (6.5, 6.9, 7.4, 7.9) for each of 25 patients. Interest is in assessing whether average  $V_{\max}$  differs among the pH levels. Data are found on Pages 182-183.

GLM procedure can be used to calculate weighted mean Vmax between the pH groups:

```
proc glm;  
    class subject ph;  
    model vmax = subject ph;  
    lsmeans ph;  
    estimate 'direction' ph -3 -1 1 3 /divisor=3;  
run;
```

Least Squares Means

ph	vmax LSMEAN
6.5	193.695986
6.9	230.458679
7.4	255.840000
7.9	265.998419

Dependent Variable: vmax

Parameter	Estimate	Standard Error	t Value	Pr >  t
direction	80.7628733	4.85912530	16.62	<.0001

## SAS code to request Mantel-Haenszel Statistics:

```
proc freq;  
  tables subject*ph*vmax / noprint cmh2;  
  tables subject*ph*vmax / noprint cmh2 scores=modridit;  
run;
```

There is evidence of differing average Vmax by pH (row mean score) and of significant linear trend (correlation statistic):

### Table Scores

	DF	Value	Prob
Nonzero Correlation	1	27.3891	<.0001
Row Mean Scores Differ	3	27.7431	<.0001

### Modified Ridit Scores

Nonzero Correlation	1	35.3818	<.0001
Row Mean Scores Differ	3	34.7945	<.0001