

22S:30/105, Statistical Methods and Computing

Instructor: Cowles
Lab 8, April 22, 2009
Nonparametric methods

1 Downloading datasets

Download the following data files from the course web page:

autistic.dat

2

(example taken from Daniel *Biostatistics: A Foundation for Analysis in the Health Sciences*

Research by Singh et al. (1999) as reported in the journal *Clinical Immunology and Immunopathology* is concerned with immune abnormalities in autistic children. As part of their research, they took measurements on the serum concentration of an antigen in three samples of children, autistic children, normal children, and mentally-handicapped children (non-Down's-syndrome). All children were 10 years old or younger.

This dataset contains two variables:

```
concentration of the antigen (in units per milliliter of serum)
group,      coded A for autistic
             N for normal
             M for mentally handicapped
```

1. Read the dataset into SAS:

```
options linesize = 75 ;

data autistic ;
input conc group $ ;
datalines ;
<data>
;
run ;
```

2. Use SAS to check the assumptions of one-way ANOVA.

```
proc sort data = autistic ;
```

```
by group ;
run ;

proc univariate plot data = autistic ;
var conc ;
by group ;
run ;

proc means data = autistic ;
var conc ;
by group ;
run ;
```

- (a) Do the distributions of the sample data appear to be roughly normal?
 - (b) Is the largest sample standard deviation no more than twice as large as the smallest sample standard deviation?
3. Use a nonparametric test to compare the centers of the three population distributions. The Kruskal-Wallis test is appropriate here. It is an extension of the Wilcoxon Rank Sum test to comparisons of three or more populations.

```
proc npar1way data = autistic wilcoxon;
class group ;
var conc ;
run ;
```

The NPARIWAY Procedure

Wilcoxon Scores (Rank Sums) for Variable conc
Classified by Variable group

group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
A	23	1094.00	828.0	81.355202	47.565217
M	15	491.50	540.0	70.964383	32.766667
N	33	970.50	1188.0	86.706081	29.409091

Average scores were used for ties.

Kruskal-Wallis Test

Chi-Square 10.9635

```
DF          2
Pr > Chi-Square 0.0042
```

4. Since the overall Chi-square test has a p-value of 0.0042, we can reject the null hypothesis that the centers of all three populations are equal. We can then proceed to do pairwise Wilcoxon tests to see where the differences lie.

```
proc npar1way data = autistic wilcoxon;
class group ;
var conc ;
where group in ("A","N") ;
run ;
```

```
proc npar1way data = autistic wilcoxon;
class group ;
var conc ;
where group in ("A","M") ;
run ;
```

```
proc npar1way data = autistic wilcoxon;
class group ;
var conc ;
where group in ("N","M") ;
run ;
```