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 ;
   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 NPAR1WAY Procedure

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

<table>
<thead>
<tr>
<th>group</th>
<th>N</th>
<th>Sum of Scores</th>
<th>Under H0</th>
<th>Under H0</th>
<th>Std Dev</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>23</td>
<td>1094.00</td>
<td>828.0</td>
<td>81.355202</td>
<td>47.565217</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>15</td>
<td>491.50</td>
<td>540.0</td>
<td>70.964383</td>
<td>32.766667</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>33</td>
<td>970.50</td>
<td>1188.0</td>
<td>86.706081</td>
<td>29.409091</td>
<td></td>
</tr>
</tbody>
</table>

   Average scores were used for ties.

   Kruskal-Wallis Test

   Chi-Square   10.9635
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.

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

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

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