1. The data set at http://ftp.uni-bayreuth.de/math/statlib/datasets/federalistpapers.txt gives data from an analysis of a series of documents. The first column gives document number, the second gives the name of a text file, the third gives a group to which the text is assigned, the fourth represents a measure of the use of first person in the text, the fifth presents a measure of inner thinking, the sixth presents a measure of positivity, and the seventh presents a measure of negativity. There are other columns that you can ignore. Test the null hypothesis that negativity is equally distributed across the groups using a Kruskal-Wallis test. (The version on line, above, has odd line breaks. A fixed version can be found at stat.rutgers.edu/home/kolassa/960-555/federalistpapers.txt).

The following R commands will do this calculation:

```r
fed<-as.data.frame(scan("federalistpapers.txt",  
    what=list(nn=0,fn="",grp=0,firstp=0,inner=0,pos=0,neg=0),  
    skip=7,flush=T))
kruskal.test(neg~as.factor(grp),data=fed)
```

The *p*-value is 0.03431. \textit{Reject the null hypothesis of identical dispersion of inner thinking between groups.}

2. The data set at http://ftp.uni-bayreuth.de/math/statlib/datasets/federalistpapers.txt gives data from an analysis of a series of documents. The first column gives document number, the second gives the name of a text file, the third gives a group to which the text is assigned, the fourth represents a measure of the use of first person in the text, the fifth presents a measure of inner thinking, the sixth presents a measure of positivity, and the seventh presents a measure of negativity. There are other columns that you can ignore. Test at $\alpha = .05$ the pairwise comparisons for negativity between groups using the Bonferroni adjustment, and repeat for Tukey’s HSD. (The version on line, above, has odd line breaks. A fixed version can be found at stat.rutgers.edu/home/kolassa/960-555/federalistpapers.txt).

The following R commands will do this calculation:

```r
fed<-as.data.frame(scan("federalistpapers.txt",  
    what=list(nn=0,fn="",grp=0,firstp=0,inner=0,pos=0,neg=0),  
    skip=7,flush=T))
pairwise.wilcox.test(fed$neg,fed$grp,method="bonferroni")
library("MultNonParam")
tukey.kruskal.test(fed$neg,fed$grp)
```

\textit{Both adjustments reject only the 1 vs. 3 comparison.}
3. The data set at http://ftp.uni-bayreuth.de/math/statlib/datasets/Plasma_Retinol gives data relating various quantities, including smoking status (1 never, 2 former, 3 current) in column 3 and beta plasma in column 13. Perform a nonparametric test to investigate an ordered effect of smoking status on beta plasma.

The following R commands will do this calculation:

\begin{verbatim}
smoke<-as.data.frame(scan("Plasma_Retinol",what=list(AGE=0, SEX=0, SMOKSTAT=0, QUETELET=0, VITUSE=0, CALORIES=0, FAT=0, FIBER=0, ALCOHOL=0, CHOLESTEROL=0, BETADIET=0, RETDIET=0, BETAPLASMA=0, RETPLASMA=0),skip=30,nmax=315))
library("clinfun")
jonckheere.test(smoke$BETAPLASMA,smoke$SMOKSTAT)
\end{verbatim}

The two-sided \( p \)-value is 0.001546.

4. The data set at http://ftp.uni-bayreuth.de/math/statlib/datasets/schizo gives data from an experiment using measurements used to detect schizophrenia, on non-schizophrenic patients. Various tests are given. Pick those data points with CS in the second column and compare the first and second gain rations, in the third and fourth columns, by taking their difference. Test the null hypothesis of zero median difference using the Wilcoxon signed rank test, for the non-schizophrenic patients.

The following commands perform these calculations in R.

\begin{verbatim}
gl<-rep(0,11)
names(gl)<-paste("g",1:11,sep=")
chizo<-as.data.frame(scan("schizo",skip=6,nmax=177,
   what=c(list(subject =0,type=""),gl)))
smaller<-chizo[chizo[[2]]=="CS",c("g1","g2")]
realsmall<-smaller[(smaller[,1]>0)&(smaller[,2]>0),]
realsmall<-realsmall[,2]-realsmall[,1]
wilcox.test(realsmall)
\end{verbatim}

The \( p \)-value is 0.0218. Do not reject the null hypothesis that the distribution of the first and second gain ratios has median 0.