Bootstrap Lab #3

1 Stats Lab # 3: Robust Comparisons Among J Independent Groups

1.1 Initialize R

Enter the following commands in R:

```
> source(url("http://www-rcf.usc.edu/~rwilcox/Rallfun-v9_2"))
> load(url("http://psycserv.mcmaster.ca/bennett/rdata/anovaLab.Rdata"))
```

The first line loads Wilcox's functions for doing robust analyses. The second loads four data vectors: `y1`, `y2`, `y3`, `y4`. For this lab, we will treat `y1`···`y4` as four independent sets of data.

1.2 Mean

- Use `t.test` to compare the four group means (i.e., test the null hypothesis that each pairwise difference is zero). Does it make a difference if you assume that the group variance are equal or unequal?

  Answer: I'm not going to do *all* pairwise tests. Instead, I'll just show you how to do several of them:

```
> # assuming equal variances:
> t.test(y1,y2,var.equal=T)

        Two Sample t-test

data:  y1 and y2
t = -4.0963, df = 58, p-value = 0.0001319
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.722001 -0.934959
sample estimates:
mean of x mean of y
0.5047281 2.3332079

> t.test(y1,y3,var.equal=T)

        Two Sample t-test

data:  y1 and y3
t = 3.0807, df = 58, p-value = 0.003156
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.9850268 4.6399788
sample estimates:
 mean of x  mean of y
 0.5047281 -2.3077747
```

```
> t.test(y1,y4,var.equal=T)

        Two Sample t-test

data:  y1 and y4
t = 0.4599, df = 58, p-value = 0.6473
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.8516407  1.3597370
sample estimates:
mean of x mean of y
0.5047281 0.2506799

> # assuming unequal variances:
> t.test(y1,y2,var.equal=F)

        Welch Two Sample t-test

data:  y1 and y2
t = -4.0963, df = 44.654, p-value = 0.0001742
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.7277219 -0.9292377
sample estimates:
mean of x mean of y
0.5047281 2.3332079

> t.test(y1,y3,var.equal=F)

        Welch Two Sample t-test

data:  y1 and y3
t = 3.0807, df = 41.511, p-value = 0.003658
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.9694442 4.6555614
sample estimates:
 mean of x  mean of y
 0.5047281 -2.3077747

> t.test(y1,y4,var.equal=F)

        Welch Two Sample t-test

data:  y1 and y4
t = 0.4599, df = 57.994, p-value = 0.6473
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.851643  1.359739
sample estimates:
mean of x mean of y
0.5047281 0.2506799
```

Answer: Assuming the variances are equal changes the values of t and p for each comparison, but the bottom line – i.e., accepting or rejecting the null hypothesis of no difference between groups – does not differ.

- Use `boxplot` to examine each set of data. Do you think the normality and equal variance assumptions are valid? Explain.

  Answer: The following code was used to create Figure 1.

  ```
  > boxplot(y1,y2,y3,y4)
  ```

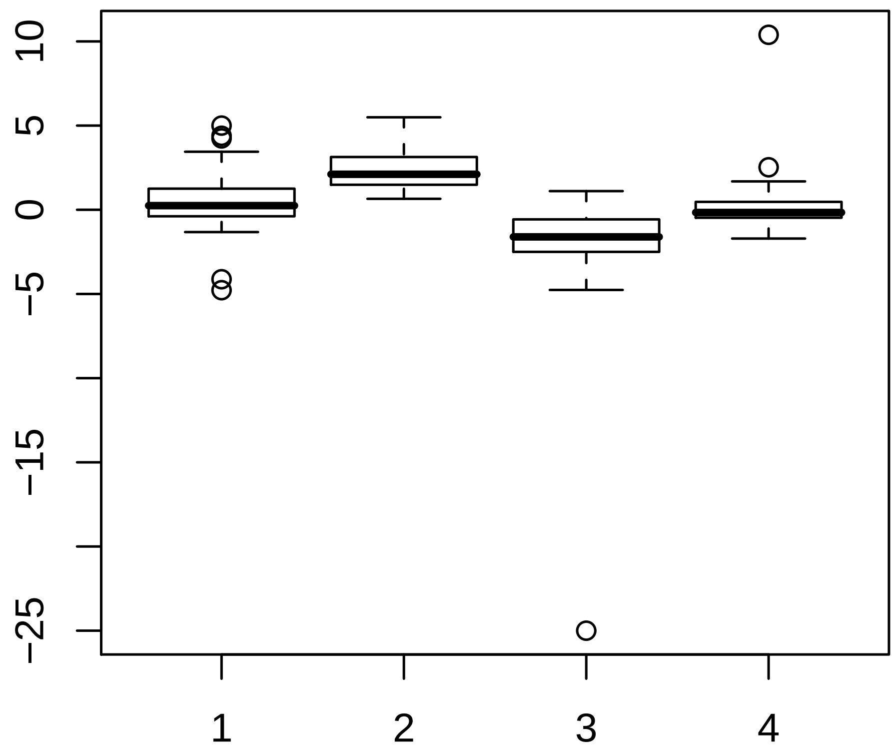

Figure 1: Boxplot of `y1`, `y2`, `y3`, `y4`.

The width of the interquartile range is approximately equal across groups, but there seem to be more outliers than I would expect based on the normality assumption at least in group 1 and, maybe, 4. Also, the outlier in group 3 may make it difficult to detect differences between the mean of that group and the other three groups.

- Concatenate all of the four groups into a single variable, Y. Next, create a grouping variable, or factor, named `group`. (See the lecture notes for an example of how to create a factor). Finally, use the function `oneway.test` to evaluate the null hypothesis of no difference among group means. How does `oneway.test` differ from a traditional ANOVA?

```
> Y <- c(y1,y2,y3,y4)
> n1 <- length(y1)
> n2 <- length(y2)
> n3 <- length(y3)
> n4 <- length(y4)
> group <- as.factor(c(rep(1,n1),rep(2,n2),rep(3,n3),rep(4,n4)) )
> oneway.test(Y~group)

        One-way analysis of means (not assuming equal variances)

data:  Y and group
F = 17.4049, num df = 3.000, denom df = 59.699, p-value = 3.099e-08
```

Answer: The effect of group is significant, $F(3, 59.7) = 17.4$, $p < .001$, and therefore the null hypothesis of no difference among group means is rejected. The difference between `oneway.test` and traditional ANOVA is that the former does not assume equal variances among groups.

1.3 Trimmed Mean

- Use the commands `yuen` and `yuenbt` to evaluate the null hypothesis of no difference among 20% trimmed means in groups `y1` and `y3`. How do the analyses performed by these commands differ from each other? How do they compare to the t test?

```
> t.test(y1,y3)

        Welch Two Sample t-test

data:  y1 and y3
t = 3.0807, df = 41.511, p-value = 0.003658
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.9694442 4.6555614
sample estimates:
 mean of x  mean of y
 0.5047281 -2.3077747

> yuen(y1,y3)

$ci
[1] 1.201500 2.625527

$p.value
[1] 4.505372e-06

$dif
[1] 1.913514

$se
[1] 0.3501262
```

```
$teststat
[1] 5.465212

$crit
[1] 2.033591

$df
[1] 33.39947

> yuenbt(y1,y3)

[1] "NOTE: p-value computed only when side=T"
[1] "Taking bootstrap samples. Please wait."
$ci
[1] 1.289213 2.612600

$test.stat
[1] 5.529134

$p.value
[1] NA
```

Answer: All three tests lead to rejection of the null hypothesis of no difference between groups. In the case of the t test, we reject the hypothesis of no difference between group means. In the case of the test on trimmed means, we reject the null hypothesis of no differences between 20% trimmed means. Note, however, that the confidence interval of the between-group difference is smaller in the case of trimmed means. Why? because the trimming removes the outliers and therefore reduces the variability of the estimated group difference. The results of the two versions of the yuen test are similar, which suggests that the distributional assumption made by the function `yuen` – that the difference between trimmed mean is distributed as t – is reasonable in this case.

- Combine the four groups of data into a single *list* named `Y.list`. (See the lecture notes for an example of creating a list in R). Next, use `t1way` and `t1waybt` to evaluate the different groups. Is this one-way test better, worse, or the same as doing a series of pairwise tests? Explain.

```
> Y.list <- list()
> Y.list[[1]] <- y1
> Y.list[[2]] <- y2
> Y.list[[3]] <- y3
> Y.list[[4]] <- y4
> t1way(Y.list)

$TEST
[1] 39.36734

$nu1
[1] 3

$nu2
[1] 36.52873
```

```
$siglevel
[1] 1.553713e-11

> t1waybt(Y.list)

[1] "Taking bootstrap samples. Please wait."
[1] "Working on group  1"
[1] "Working on group  2"
[1] "Working on group  3"
[1] "Working on group  4"
$test
[1] 39.36734

$p.value
[1] 0
```

Answer: As was the case when we used `oneway.test`, reject the null the null hypothesis of no difference among group 20% trimmed means. Is this omnibus test better than a series of pairwise comparisons? The answer is, "it depends". If you want to evaluate planned hypotheses about specific pairwise differences, then you probably better off testing those hypotheses with t tests (particularly if the number of such tests is not too large). On the other hand, if you do not have a prior hypotheses about pairwise differences, then the omnibus test is better because i) the estimate of population error is more reliable when you use all of the groups; and ii) the omnibus test looks for *any* kind of difference, not just pairwise differences. Note, however, that finding a significant difference among groups does not tell you *how* the groups differ.

1.4 M-estimators

- Use the command `b1way` to evaluate the null hypothesis that the MOM does not vary across groups.

```
> b1way(Y.list,est=mom)

[1] "Taking bootstrap samples. Please wait."
[1] "Working on group  1"
[1] "Working on group  2"
[1] "Working on group  3"
[1] "Working on group  4"
$teststat
[1] 1.640362

$p.value
[1] 0
```

- Use the command `b1way` to evaluate the null hypothesis that the one-step M-estimator does not vary across groups.

```
> b1way(Y.list,est=onestep)

[1] "Taking bootstrap samples. Please wait."
[1] "Working on group  1"
[1] "Working on group  2"
[1] "Working on group  3"
```

```
[1] "Working on group  4"
$teststat
[1] 1.851392

$p.value
[1] 0
```

- What advantages and disadvantages does an M-estimator have relative to a trimmed mean?

Answer: The key advantage of an M-estimator is that the scheme for detecting outliers (and "trimming") is customized for each group. Inspection of Figure 1 suggests that the distribution of extreme scores differs across groups, and therefore using an M-estimator may provide some advantages in this case. One key *disadvantage* of using an M-estimator is that it sometimes is less stable when used in conjunction with the bootstrap and when sample size is small. Recall that M-estimators use robust measures of spread (e.g., the MAD or MADN) to define outliers. When sample size is small (e.g., $n \leq 10$), it is possible for a bootstrapped sample to consist of n identical scores. In such cases, outliers are undefined and M-estimators cannot be computed. Note that this problem does not affect the computation of trimmed means.