

Bootstrap Lab #2

1 Stats Lab # 2: Robust Tests of Central Location

1.1 Initialize R

Enter the following commands in R:

```
> source(url("http://www-rcf.usc.edu/~rwilcox/Rallfun-v9"))
> load(url("http://psycserv.mcmaster.ca/bennett/rdata/sgData.Rdata"))
```

The first line loads Wilcox's functions for doing robust analyses. The second loads a data file.

1.2 Mean

Use the functions `mean`, `sd`, and `t.test` to calculate the mean, standard deviation, and 95% confidence interval for the data in `sgData`. Use `boxplot` to inspect the data visually.

- Do the mean and confidence interval reflect what you think represents a good measure of a typical score? Explain.
- Let's define an "outlier" as any score that is more than 3 standard deviations from the mean. Now, identify the outliers in `sgData`. Has this procedure worked satisfactorily? Why or why not?

1.3 Trimmed Mean

Use the commands `mean` and `tmean` to calculate the 10% and 20% trimmed means. Next, use `trimci` and `trimpb` to calculate 95% confidence intervals for the 10% and 20% trimmed means.

- How do the confidence intervals for the trimmed means compare to the confidence interval for the mean? Which do you think is a more accurate estimate of the correct confidence interval? Why?
- Do the confidence intervals for the 10% and 20% trimmed means differ? Does this make sense?
- Do the intervals calculated by `trimci` and `trimpb` differ? What might this result imply about the assumptions that underlie `trimci`?

1.4 M-estimators

- Use the commands `onestep` and `mom` to calculate the one-step M-estimator and the modified one-step M-estimator for `sgData`. How do these values compare the mean and trimmed means?
- Use `momci` and `onesampb` to calculate 95% confidence interval for $\hat{\mu}_{os}$ and $\hat{\mu}_{mom}$. How do these intervals compare to the intervals for the mean and trimmed means?
- Explain how the MOM is computed for the `sgData`.
- Are the data in `sgData` skewed? Might the presence or absence of skew influence your choice of using a trimmed mean or an M-estimator? Explain.