Are you Teaching your Students Fairy Tales about Normality and Homoscedasticity? Introduce Robust Statistics Instead

Robust Statistics Lab
cribbie.info.yorku.ca

Rob Cribbie
Quantitative Methods
Department of Psychology
York University

QUANTITATIVE METHODS
Dept. of Psychology - York University

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Discussion Overview

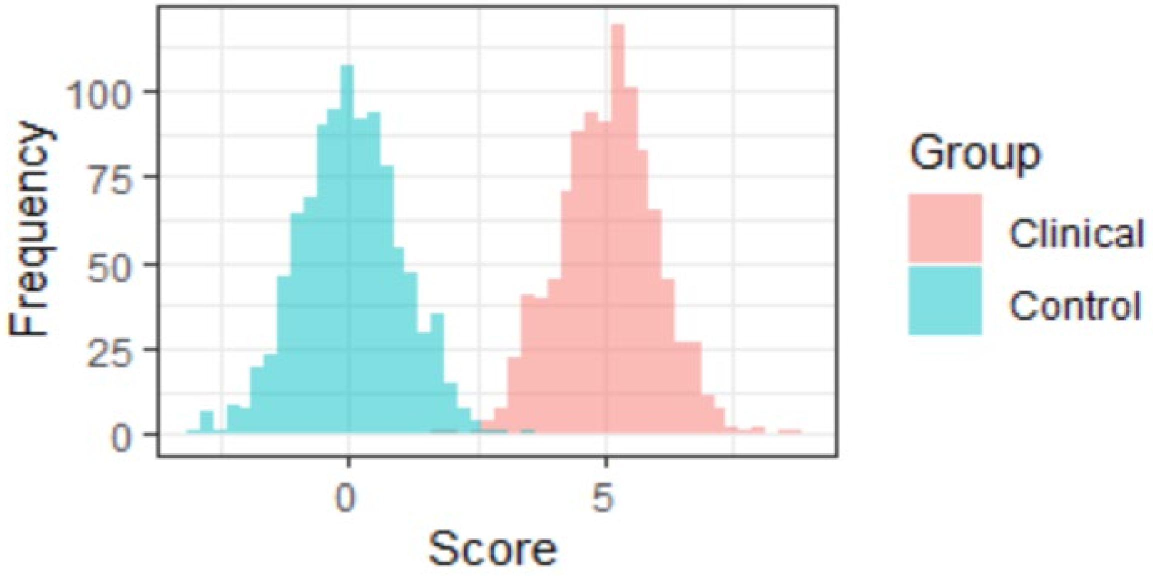
- Are Normality and Homoscedasticity Really Fairy Tales?
- Introduction to Robust Measures
- Do Robust Measures Make a Difference?
- Implementing Robust Statistics in the Classroom

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Are Normality and Homoscedasticity Really Fairy Tales?

Normality and Homoscedasticity

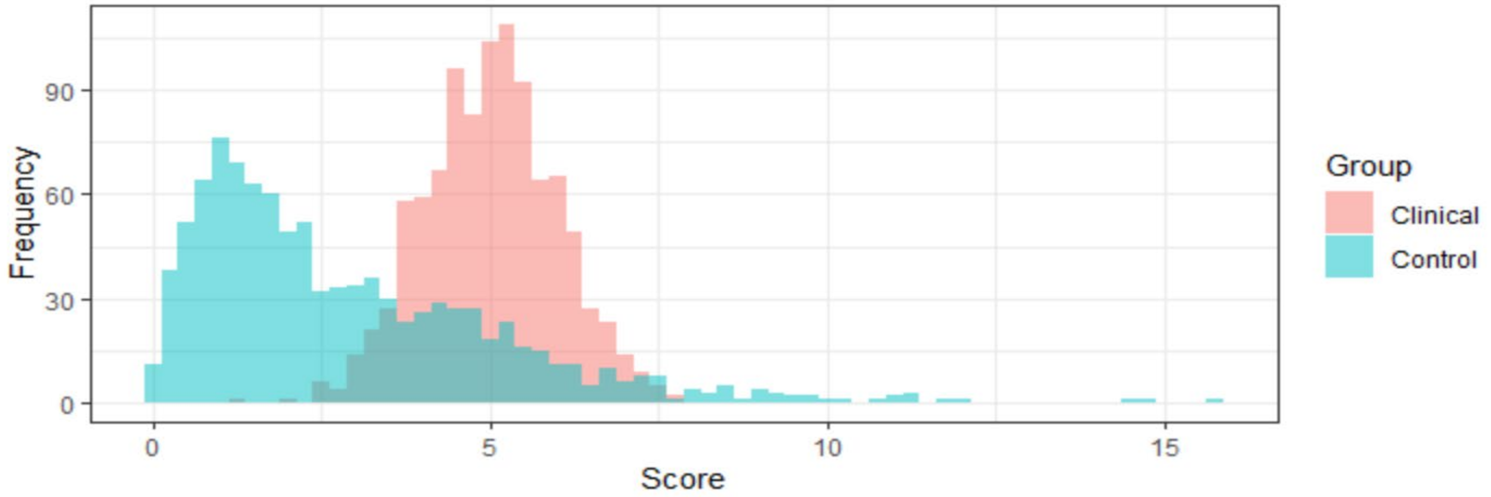


- Wouldn't it be great if all our distributions looked like this?
  - Normal distributions with equal variability
- The *Intro Stats Fallacy*

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Normality and Homoscedasticity



- However, many studies have reviewed the distributions of variables in Psychology and found normality and homoscedasticity to be rare

The Unicorn, The Normal Curve, and Other Improbable Creatures

Theodore Micceri
Department of Educational Leadership
University of South Florida

Robust Statistics Lab
cribbie.info.yorku.ca

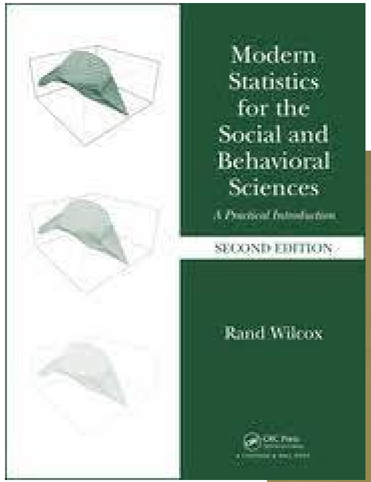
QUANTITATIVE METHODS
Dept. of Psychology - York University

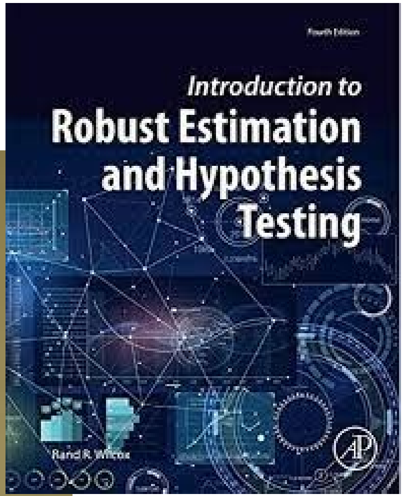
Introduction to Robust Measures

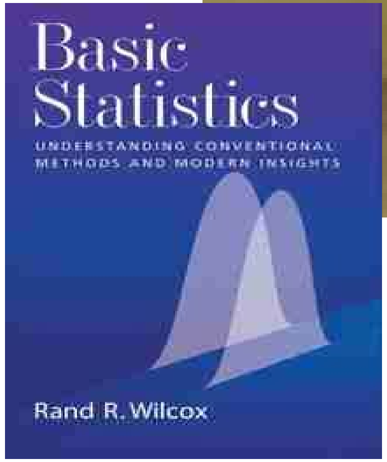
Rand Wilcox: Robust Statistics Guru!

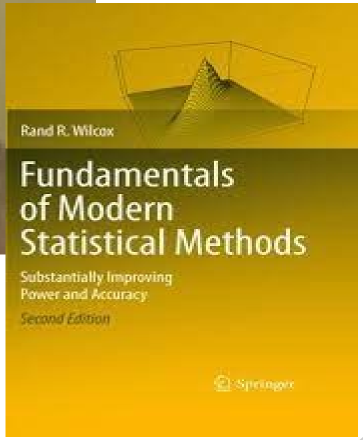
"Absolute Legend"

- Andy Field



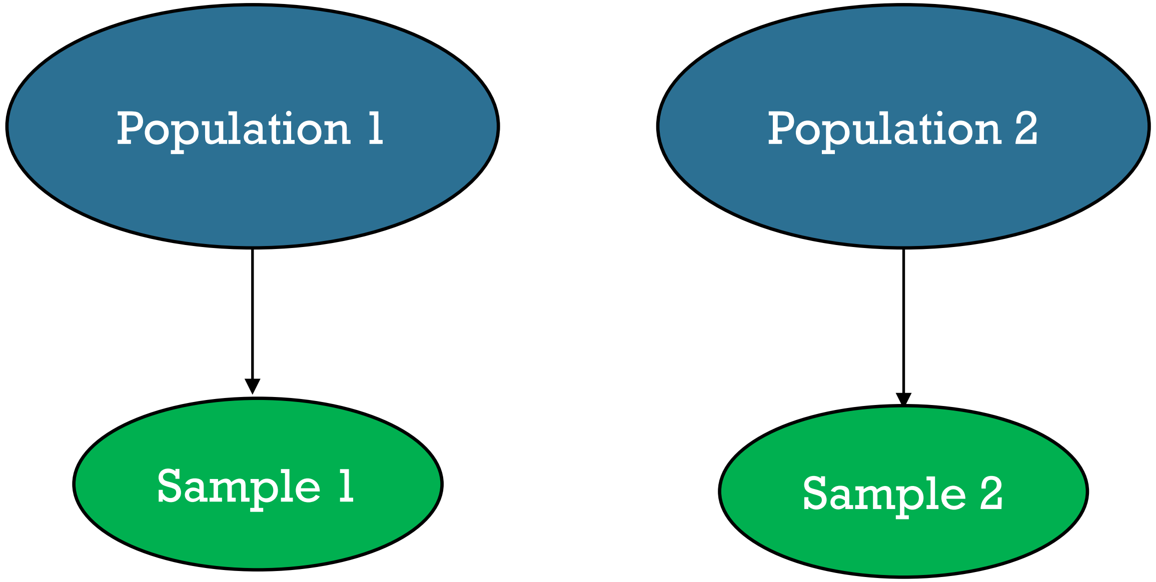






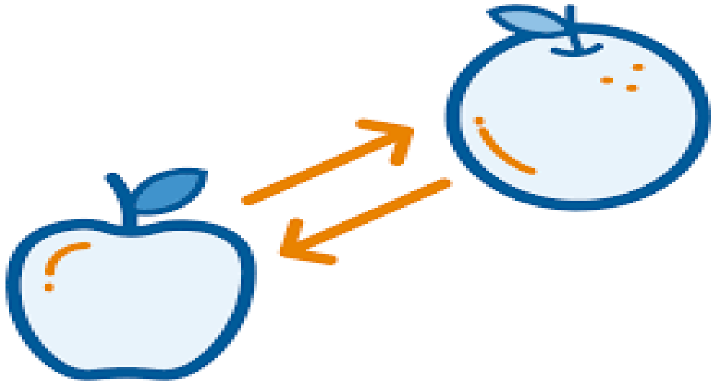
Brief Note: Design

I am going to focus exclusively on the **two independent samples design**



Brief Note: Transformations

Although **transformations** are an alternative to robust statistics, in addition to interpretation issues, they are not applicable in a wide range of situations



Introduction to Robust Statistics

- The primary goal of robust statistics is to ***limit the impact of superficial aspects of the data*** on the conclusions drawn from the data
  - For example, limit the effects of *nonnormality, outliers, unequal variances*, etc. on conclusions pertaining to the magnitude of effects
- This is often phrased in terms of analyzing the *bulk of the data*

Robust Statistics Lab
cribbie.info.yorku.ca

Quantitative Methods
Dept. of Psychology - York University

Central Tendency: Traditional Measure

- Mean (M)
  - Arithmetic average
  - $M = \frac{\sum X}{N}$

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Central Tendency: Robust Measures

- Median (Mdn)
  - Middle score of a sorted numeric variable
- Trimmed Mean (Mt)
  - Mean of a variable after removing extreme observations
    - 20% Trimmed Mean = mean after removing the most extreme 20% of observations from each tail

~~2 5 7~~ 12 14 15 15 15 17 21 22 25 34 ~~42 58 72~~

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Variability: Traditional Measure

- Standard Deviation (SD)
  - Square root of the average squared deviation from the mean
  - $SD = \sqrt{\frac{\sum(X - \bar{X})^2}{N-1}}$

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Variability: Robust Measures

- Median Absolute Deviation (MAD)
  - Median of the absolute deviations from the median
- Winsorized Standard Deviation (WSD)
  - SD after replacing the extreme observations with less extreme observations
    - 20% WSD = SD after replacing the lowest and highest 20% of cases from each tail with the lowest and highest nonextreme value

~~2 5 7~~ 12 14 15 15 15 17 21 22 25 34 ~~42 58 72~~

12 12 12 12 14 15 15 15 17 21 22 25 34 34 34 34

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Effect Sizes: Traditional Measure

- $$d = \frac{M_1 - M_2}{\sqrt{\frac{(n_1-1)SD_1^2 + (n_1-1)SD_1^2}{n_1+n_2-2}}}$$

Ordinary Cohen's *d*: Not Robust

Effect Sizes: Robust Measures

- $d_W = \frac{M_1 - M_2}{\sqrt{\frac{SD_1^2 + SD_2^2}{2}}}$

Welch Cohen's *d*: Robust to Variance Heterogeneity

- $d_r = .642 \frac{M_{t1} - M_{t2}}{\sqrt{\frac{WSD_1^2 + WSD_2^2}{2}}}$

Robust Cohen's *d*: Robust to Heteroscedasticity and Nonnormality

The multiplier .642 puts d_r on the same scale as a measure not incorporating WSD, with 20% trimming

Test Statistics: Traditional Measure

- $$t = \frac{M_1 - M_2}{\sqrt{\frac{(n_1-1)SD_1^2+(n_1-1)SD_1^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Ordinary *t* Test: Not Robust

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Test Statistics: Robust Measures

- $$t_W = \frac{M_1 - M_2}{\sqrt{\frac{SD_1^2}{n_1} + \frac{SD_2^2}{n_2}}}$$

Welch t Test: Robust to Heteroscedasticity

- $$t_Y = \frac{M_{t1} - M_{t2}}{\sqrt{\frac{(n_1 - 1)WSD_1^2}{h_1(h_1 - 1)} + \frac{(n_2 - 1)WSD_2^2}{h_2(h_2 - 1)}}}$$

h = sample size after trimming

Yuen t Test: Robust to Heteroscedasticity and Nonnormality

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Do Robust Measures Make a Difference?

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Descriptive Statistics

- Effects of Extreme Values on the Mean and Standard Deviation
  - X1 = {2 2 4 5 7 7 8 9 9 14}
  - X2 = {2 2 4 5 7 7 8 9 9 38}

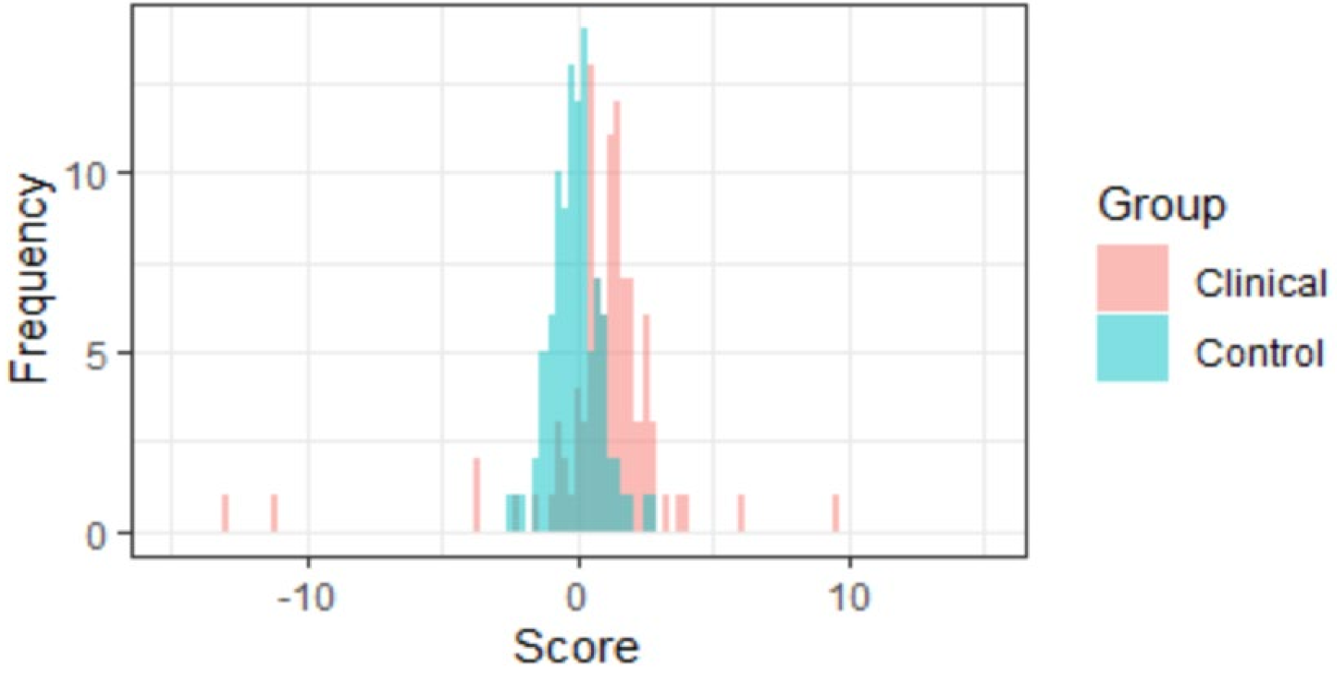
	Mean	Mdn	M_t	SD	MAD	WSD
X1	6.7	7	6.4	3.7	3.0	4.7
X2	9.1	7	6.4	10.5	3.0	4.7

- The *breakdown point* of the M and SD is 1!

Robust Statistics Lab
cribbie.info.yorku.ca
QUANTITATIVE METHODS
Dept. of Psychology - York University

Effect Sizes

Effects of Nonnormality/ Heteroscedasticity on Cohen's *d*



- One group has some extreme scores
  - Imagine pre-post changes over an intervention for a control group and a clinical group

Effect Sizes

- **Effects of Extreme Scores on Cohen's *d***

- $$d = \frac{M_1 - M_2}{\sqrt{\frac{(n_1-1)SD_1^2+(n_1-1)SD_1^2}{n_1+n_2-2}}} = -.53$$

- $$d_r = .642\,\frac{M_{t1} - M_{t2}}{\sqrt{\frac{WSD_1^2+WSD_1^2}{2}}} = -1.40$$

Test Statistics

- Traditional *t* Test
  - Nonnormality and/or heteroscedasticity can have deleterious effects on the **Type I error rates** and **power** of traditional test statistics (e.g., *t*, *F*)
  - If NHST-based results are a primary outcome, robust test statistics are definitely recommended
    - E.g., Yuen test

Test Statistics

- *p*-Values for the Clinical vs Control Data
  - Ordinary *t* Test
    - $p = .051$
  - Welch *t* Test
    - $p = .009$
  - Yuen *t* Test
    - $p < .001$

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

What can we do in the classroom?

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Undergraduate Introductory Statistics Class

- In an undergraduate *Introduction to Statistics* course, we might introduce them to:

Effect of nonnormality/outliers on measures of central tendency

Effect of nonnormality/outliers and heteroscedasticity on measures of effect size and NHST-based tests

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Undergraduate Introductory Statistics Class

- In an undergraduate *Introduction to Statistics* course, we might introduce them to:

Robust measures of central tendency and variability

Robust effect size measures

Graduate Statistics Courses

In a graduate *Statistics* course, we should be introducing students to the full complement of robust statistics

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

What can we do in the Classroom?

- One way to allow students to visualize the advantages of robust statistics is via *Shiny* apps
  - Here is a tiny example demonstrating the effects of nonnormality and variance heterogeneity on traditional and robust effect sizes (Cohen's *d*)
- https://cribbie.shinyapps.io/Robust_Effect_Size/

Robust Statistics Lab
cribbie.info.yorku.ca

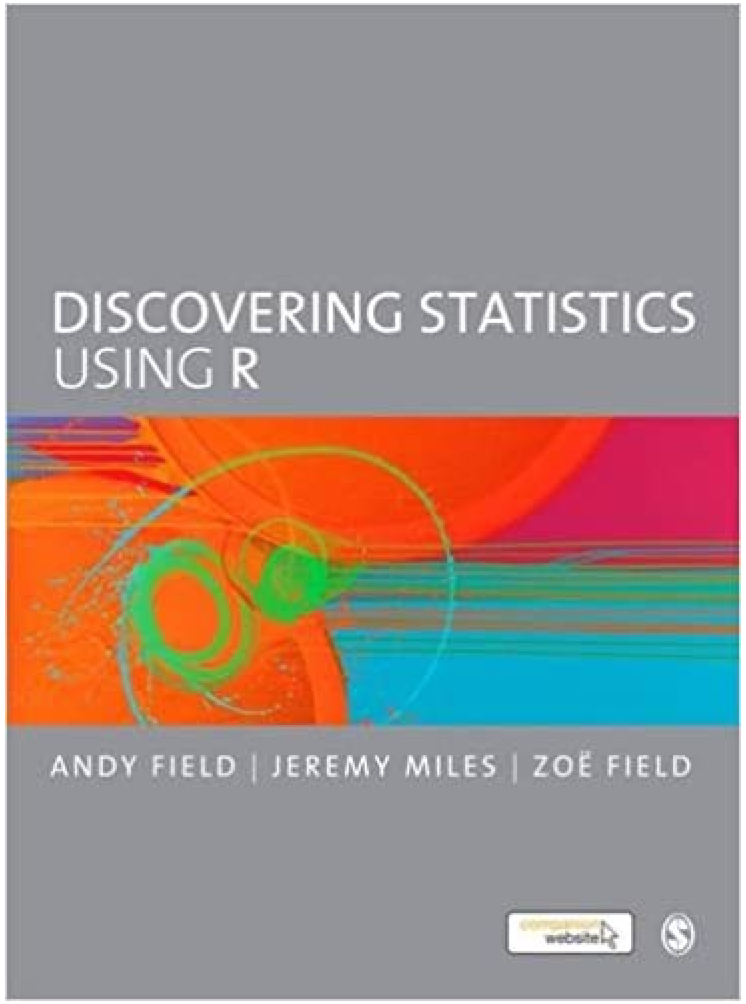
QUANTITATIVE METHODS
Dept. of Psychology - York University

What if I use Statistical Software in my Course?

- Limitless robust statistics are available

- Trimmed Mean
  - mean(x, tr=.1)
- Median Absolution Deviation
  - mad(x)
- Winsorized Variance
  - WRS2::winvar(x, tr=.1)
- Robust Cohen's *d*
  - WRS2::akp.effect(model)
- Welch *t* Test
  - t.test(dv ~ iv)
- Yuen *t* Test
  - WRS2::yuen(dv ~ iv, tr=.1)

Textbook with Excellent Coverage of Robust Statistics!



9.5.2.7. Robust methods to compare independent means ②

Robust Statistics Lab
cribbie.info.yorku.ca

QUANTITATIVE METHODS
Dept. of Psychology - York University

Rob Cribbie
cribbie@yorku.ca

Thank You!