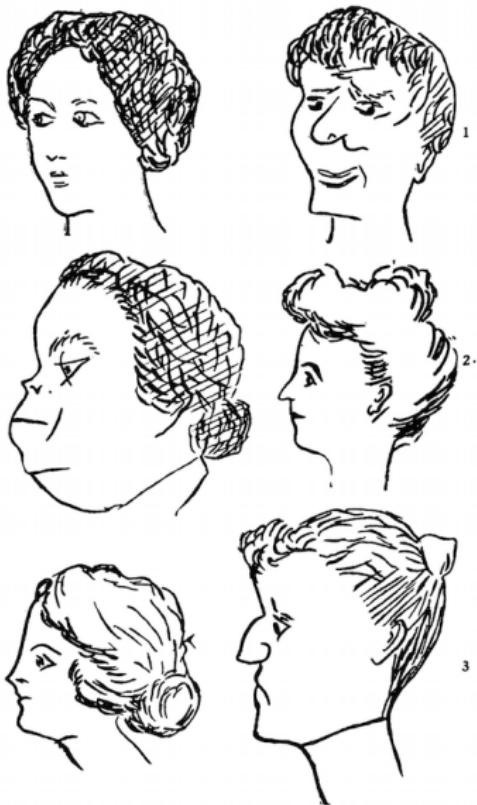


E-411-PRMA

Lecture 1

Christopher David Desjardins

17 August 2015



THE PSYCHOLOGICAL CLINIC is indebted for the loan of these cuts and those on p. 228 to the courtesy of Dr. Oliver P. Cornman, Associate Superintendent of Schools of Philadelphia, and Chairman of Committee on Backward Children Investigation. See Report of Committee, Dec. 31, 1910, appendix.



*Extreme desespoir.
Eusserste Yerzweiflung*



*Colere meslée de Crainte.
Zorn mit Furcht vermischt.*



achieve
more®

SAT®

► Topics

- ▶ Statistics, Classical Test Theory, Reliability, Validity, Item Response Theory, Generalizability Theory, Equating, and assessments/issues specific to various fields

► Assessments

- ▶ R computer assignments (30%)
- ▶ Item writing activity (5%)
- ▶ Midterm exam (25%)
- ▶ **Final exam (50%)**

The screenshot shows the RStudio interface. The left pane contains an R script named "26_June_intro.R" with the following code:

```
140 # view all data
141 precol
142
143 # If you want to see a specific observation, say #20, and you want to see responses from
144 # variables 2 through 3
145 precol[20,2:3]
146
147 # Or if you want to see variables 1, 3, and 6 for observations 19 through 21.
148 precol[19:21, c(1,3,6)]
149
150 # Don't worry R won't be so bad ;
151
152 # To get information about the data set and the variable types
153 str(precol)
154
155 # To get the variable labels
156 attr(precol, "variable.labels")
157
158 # [1] "unreliable"
159
160 # [2] "unreliable"
161
162 # [3] "unreliable"
163
164 # [4] "unreliable"
165
166 # [5] "unreliable"
167
168 # [6] "unreliable"
169
170 # [7] "unreliable"
171
172 # [8] "unreliable"
173
174 # [9] "unreliable"
175
176 # [10] "unreliable"
177
178 # [11] "unreliable"
179
180 # [12] "unreliable"
181
182 # [13] "unreliable"
183
184 # [14] "unreliable"
185
186 # [15] "unreliable"
187
188 # [16] "unreliable"
189
190 # [17] "unreliable"
191
192 # [18] "unreliable"
193
194 # [19] "unreliable"
195
196 # [20] "unreliable"
197
198 # [21] "unreliable"
```

The right pane displays the R documentation for the `separate` function. The title is "Separate one column into multiple columns." The description states: "Given either regular expression or a vector of character positions, separate() turns a single character column into multiple columns." The usage is shown as:

```
separate(data, col, into, sep = "[[:alnum:]]+", remove = TRUE,
         convert = FALSE, extra = "error", ...)
```

The arguments are listed as:

- data: A data frame.
- col: Bare column name.
- into: Names of new variables to create as character vector.
- sep: Separator between columns.

R: <https://www.r-project.org>
RStudio: <https://www.rstudio.com>

Why should I learn R?

- ▶ It's free and open-source

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses
- ▶ Helps you learn statistics better

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses
- ▶ Helps you learn statistics better
- ▶ Learn reproducible research

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses
- ▶ Helps you learn statistics better
- ▶ Learn reproducible research
- ▶ Extremely marketable skill

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses
- ▶ Helps you learn statistics better
- ▶ Learn reproducible research
- ▶ Extremely marketable skill
- ▶ High quality graphics

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses
- ▶ Helps you learn statistics better
- ▶ Learn reproducible research
- ▶ Extremely marketable skill
- ▶ High quality graphics
- ▶ Everyone is doing it

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses
- ▶ Helps you learn statistics better
- ▶ Learn reproducible research
- ▶ Extremely marketable skill
- ▶ High quality graphics
- ▶ Everyone is doing it
- ▶ Steep learning curve

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses
- ▶ Helps you learn statistics better
- ▶ Learn reproducible research
- ▶ Extremely marketable skill
- ▶ High quality graphics
- ▶ Everyone is doing it
- ▶ Steep learning curve
 - ▶ Will provide nearly all the code

Why should I learn R?

- ▶ It's free and open-source
- ▶ Statistics and psychometrics analyses
- ▶ Helps you learn statistics better
- ▶ Learn reproducible research
- ▶ Extremely marketable skill
- ▶ High quality graphics
- ▶ Everyone is doing it
- ▶ Steep learning curve
 - ▶ Will provide nearly all the code
- ▶ No SPSS in this class

Resources for R

- ▶ Icelandic resources

<http://kennslubanki.hi.is/search/efni/r>

<http://kennslubanki.hi.is/tolfraedi/myndbond/rrstudio-inngangur>

<http://kennslubanki.hi.is/tolfraedi/myndbond/rrstudio-fyrstu-skrefin>

- ▶ Please watch the last two videos before next class
- ▶ Please install R and RStudio before next class
- ▶ Next class will be an R workshop

Concepts

What is **measurement**?

Concepts

What is **measurement**?

What is a **test**?

Concepts

What is **measurement**?

What is a **test**?

What is a **scale**?

Concepts

What is **measurement**?

Assignment of numerical values based on a set of rules

What is a **test**?

What is a **scale**?

Concepts

What is **measurement**?

Assignment of numerical values based on a set of rules

What is a **test**?

An instrument used to measure

What is a **scale**?

Concepts

What is **measurement**?

Assignment of numerical values based on a set of rules

What is a **test**?

An instrument used to measure

What is a **scale**?

A set of numbers used to categorize or quantify variables ("things")

Concepts

What is **measurement**?

Assignment of numerical values based on a set of rules

What is a **test**?

An instrument used to measure

What is a **scale**?

A set of numbers used to categorize or quantify variables ("things")

Nominal

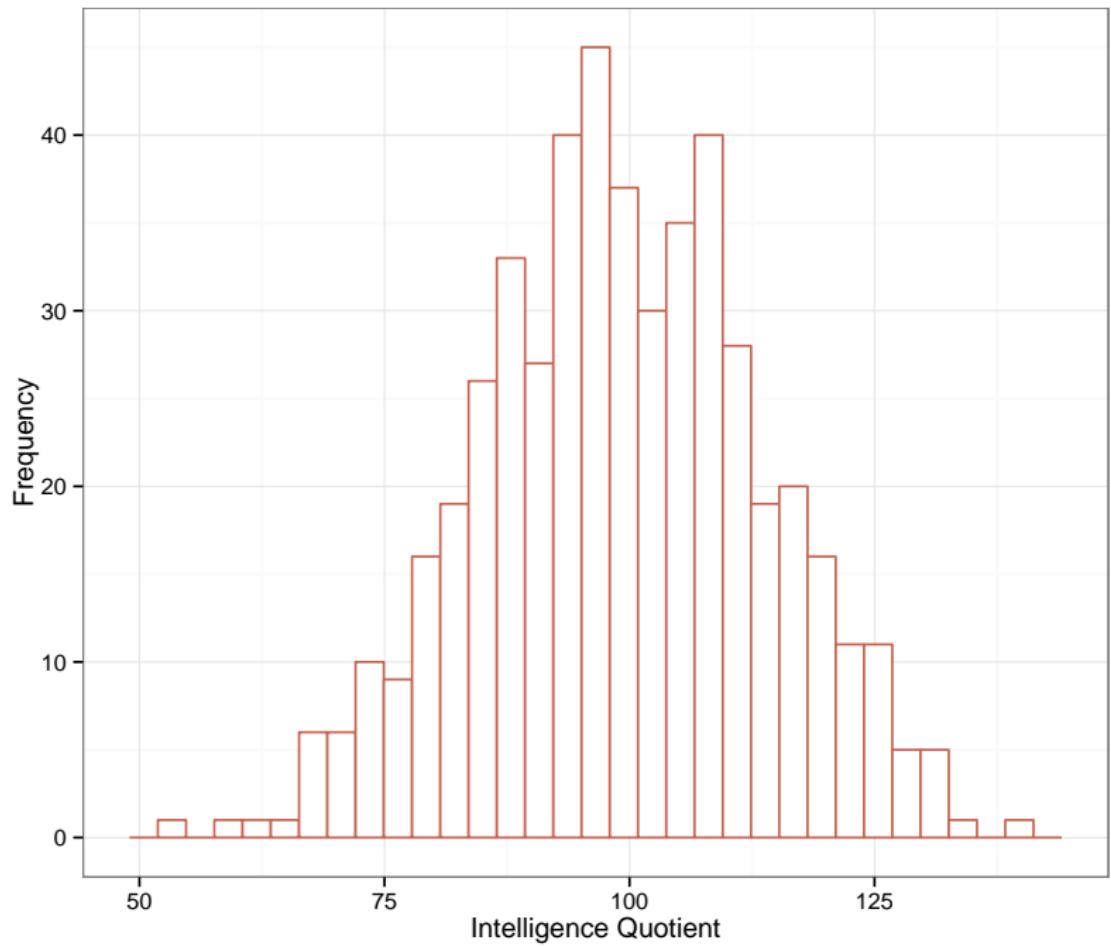
Ordinal

Ratio

Interval

What kind of scales are these?

- ▶ Temperature
- ▶ Height
- ▶ Grade Point Average
- ▶ Color
- ▶ Ethnic group
- ▶ Likert-type items
- ▶ Job satisfaction

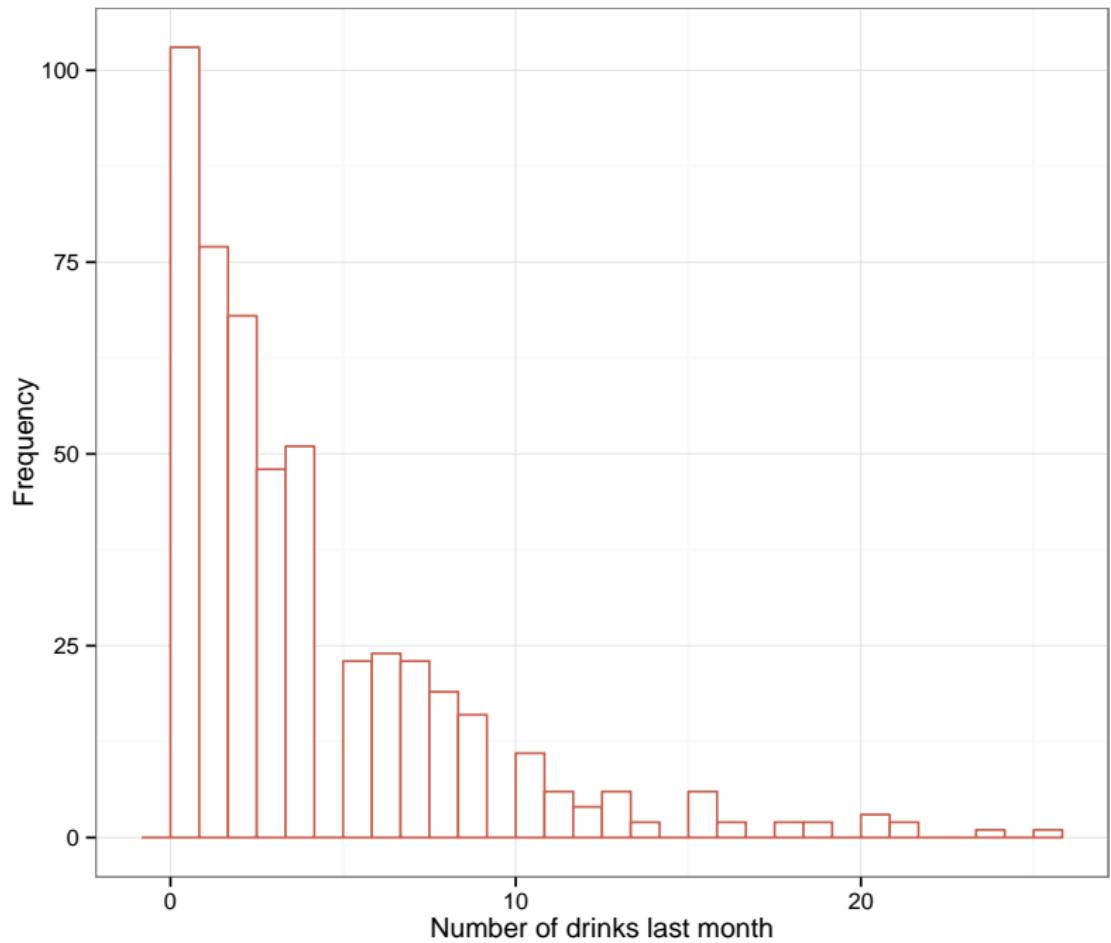


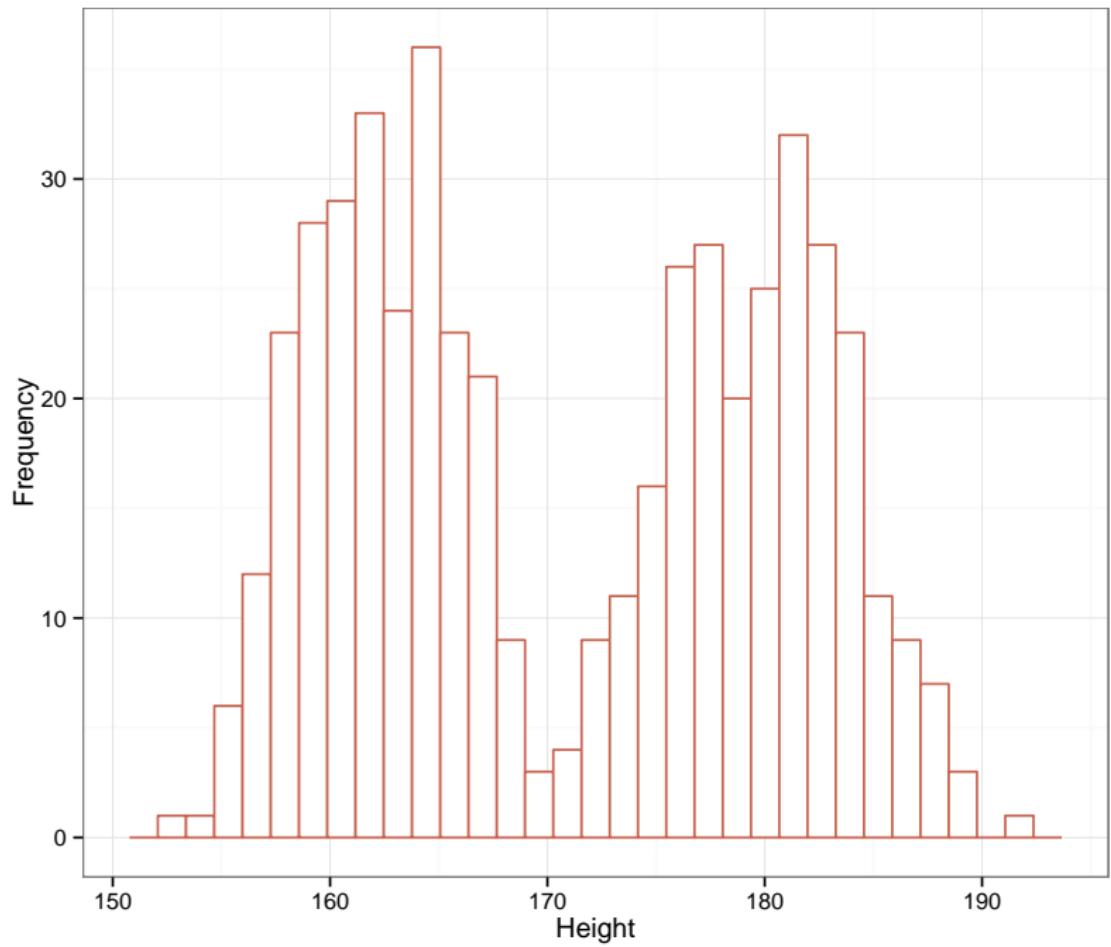
```
# Load the library
set.seed(101)
library("ggplot2")

# Set up the parameters
sample_size <- 500
mean <- 100
standard_deviation <- 15

# Generate random numbers
x <- rnorm(sample_size, mean, standard_deviation)

# Plot the data
qplot(x, fill = I("white"), color = I("#c96552")) +
  theme_bw() + xlab("Intelligence Quotient") +
  ylab("Frequency")
```





Central Tendency Measures

Mean

$$\bar{X} = \frac{\sum X_i}{n}$$

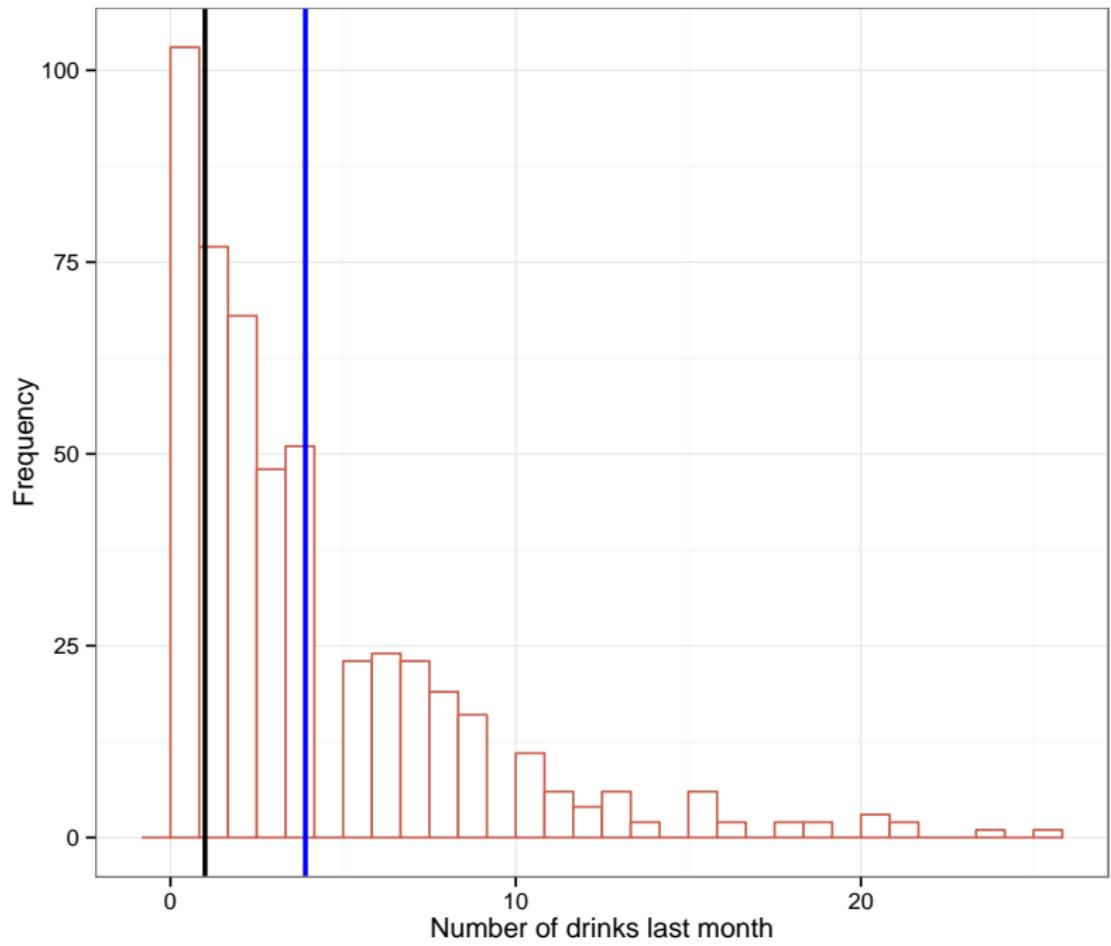
Median

$$P(X \leq m) \geq \frac{1}{2} \text{ and } P(X \geq m) \geq \frac{1}{2}$$

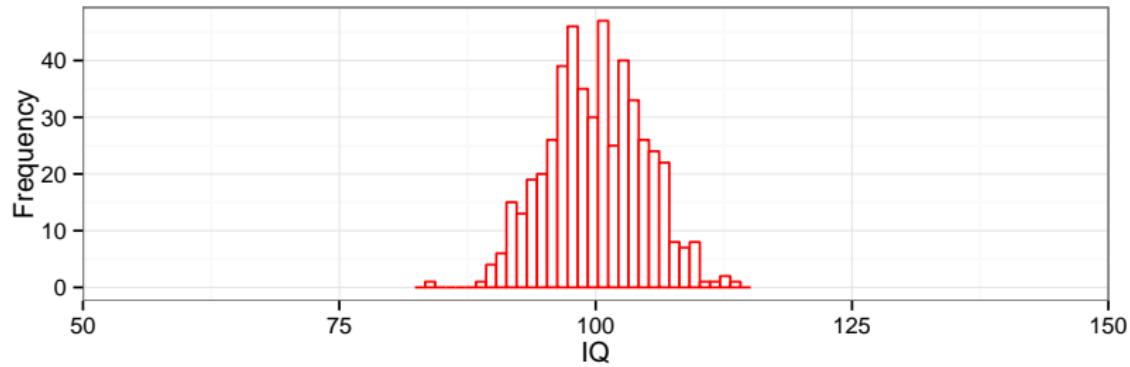
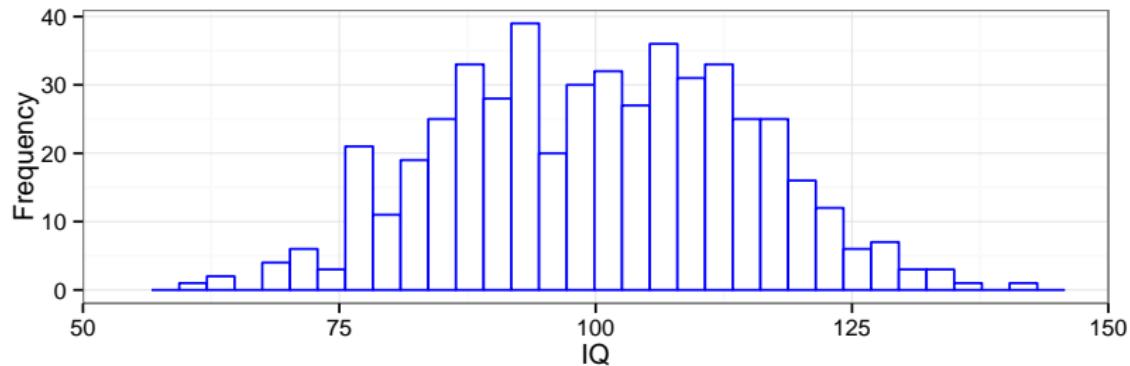
Mode

The most frequently occurring value

Which of these statistics is most robust to outliers?



Variability



Measures of variability

Range

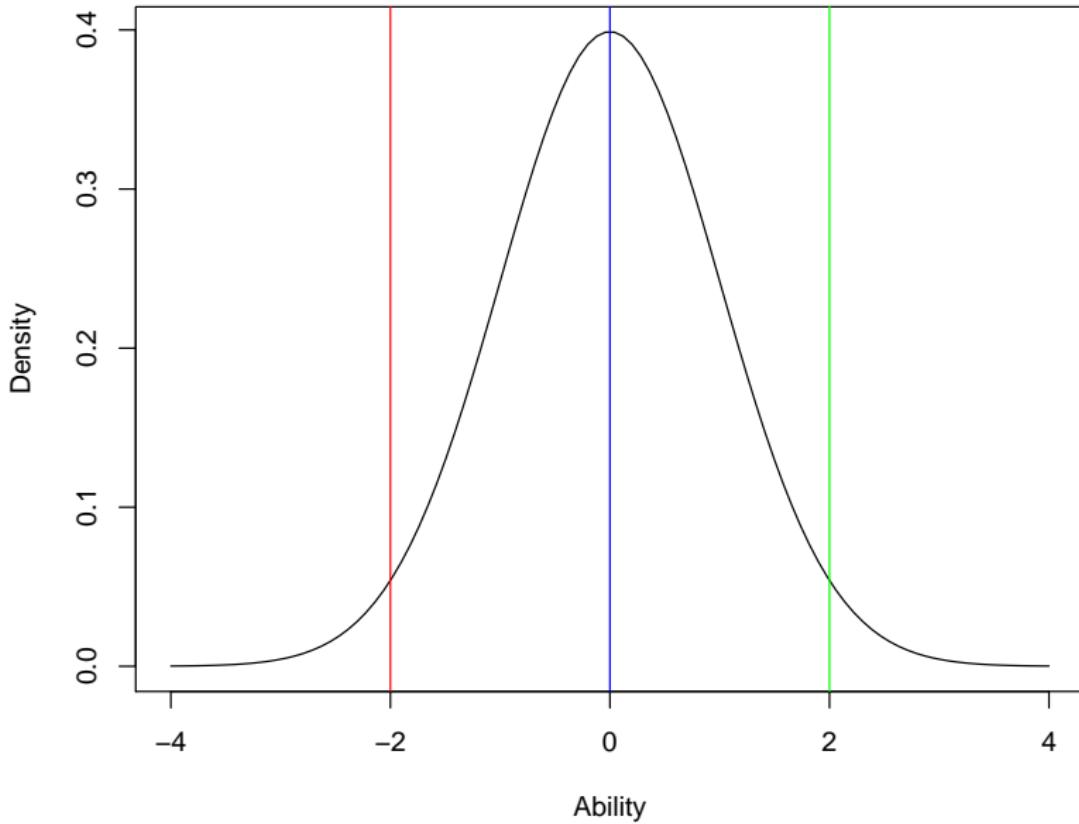
Interquartile range (Q_1, Q_2, Q_3)

Standard Deviation and Variance

$$s = \sqrt{\frac{\sum X_i - \bar{X}}{n - 1}}$$

$$s^2 = \frac{\sum X_i - \bar{X}}{n - 1}$$

Normal Distribution

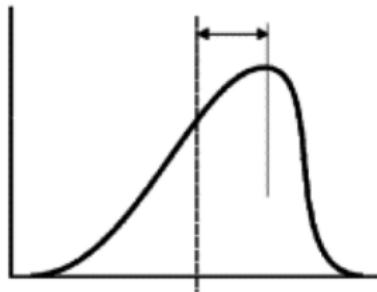


Distributions, skewness, kurtosis

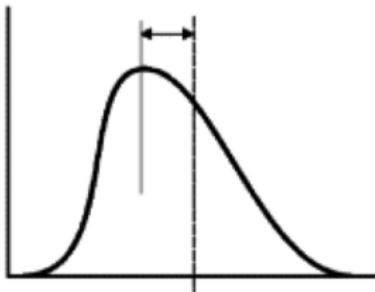
- ▶ What is a probability distribution
 - ▶ Assigns a probability, likeliness of occurrence, of a score to all possible scores
 - ▶ May be parametric or non-parametric
- ▶ What skew might you expect these outcomes to look like?
 - ▶ Reaction time in a psychological experiment
 - ▶ Number of children in a family
 - ▶ Scores on an easy test
 - ▶ Height in Iceland
- ▶ Platykurtic, mesokurtic, and leptokurtic
- ▶ Plot your data, rely less on statistics!

Shapes of distributions

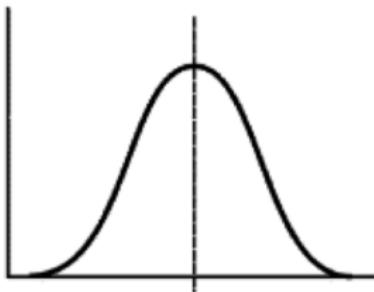
a Negatively skewed distribution



Positively skewed distribution



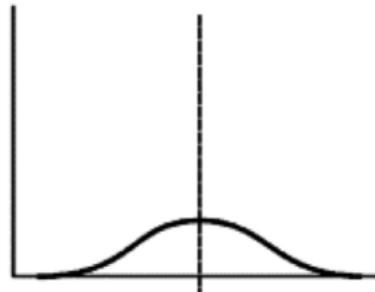
Normal distribution



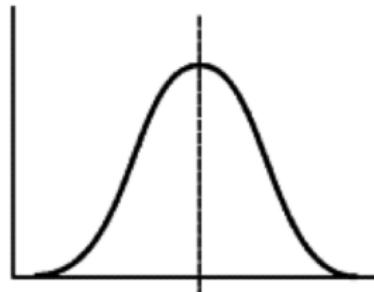
b Leptokurtic distribution



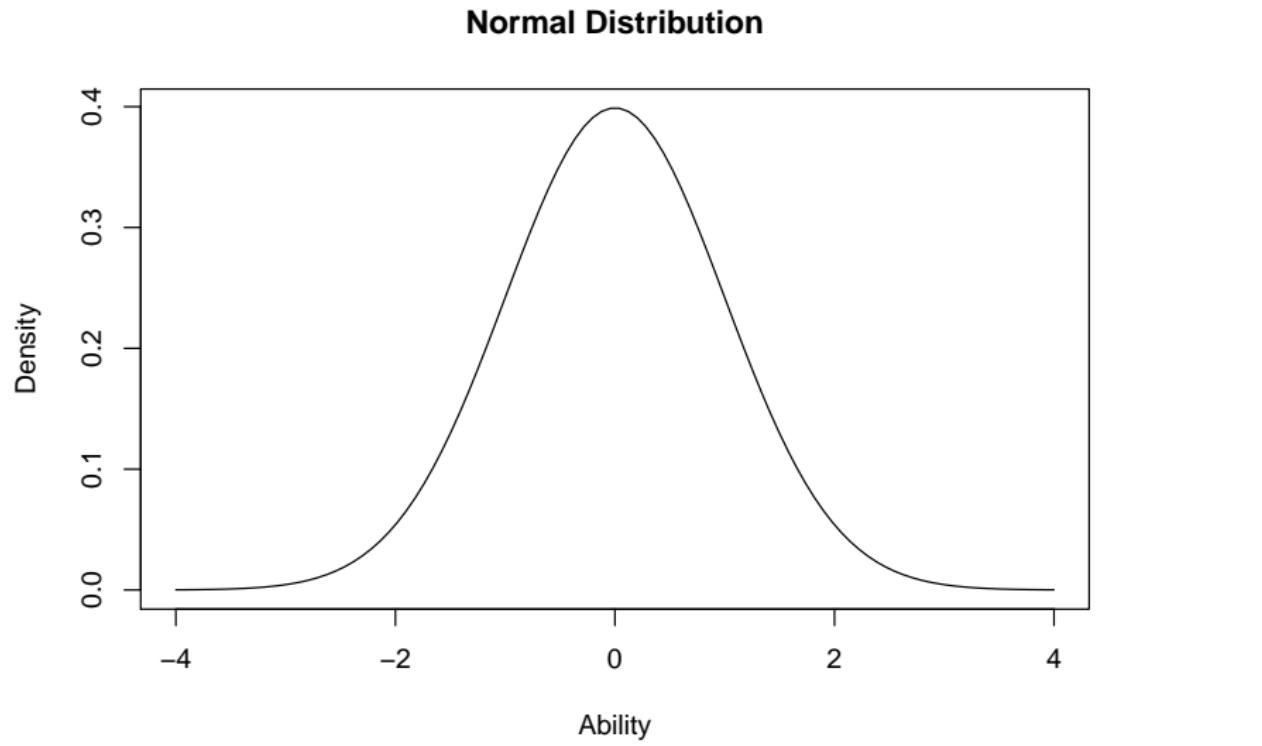
Platykurtic distribution



Normal distribution



Normal Distribution



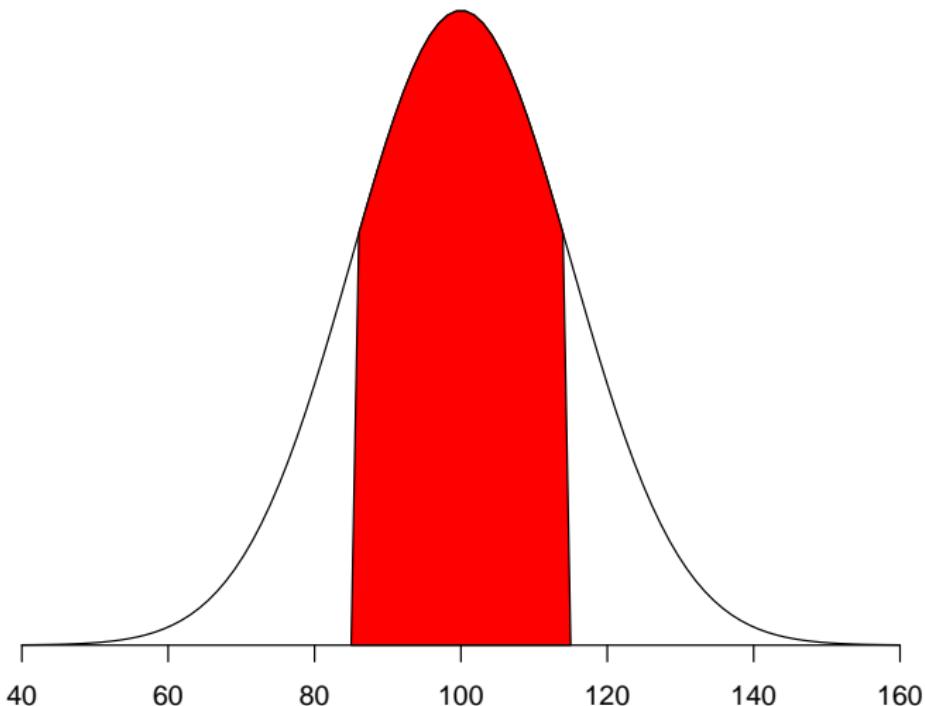
R Normal distribution applet

1. Open RStudio
2. Open normal_applet.R
3. Click the "Source" button

IQ - 1 Standard Deviation

Normal Distribution

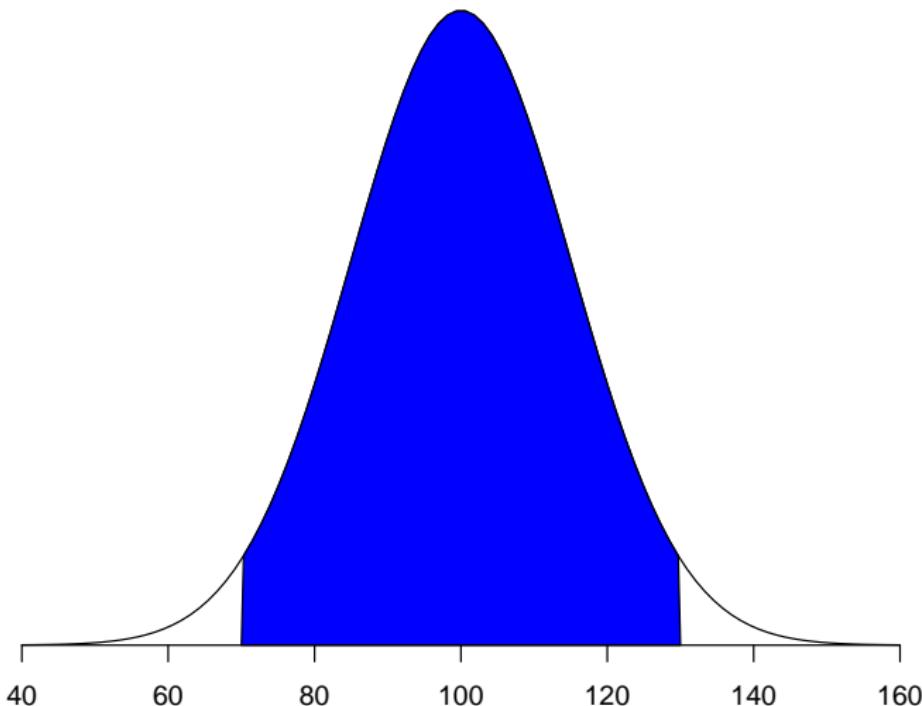
$$P(85 < \text{IQ} < 115) = 0.683$$



IQ - 2 Standard Deviation

Normal Distribution

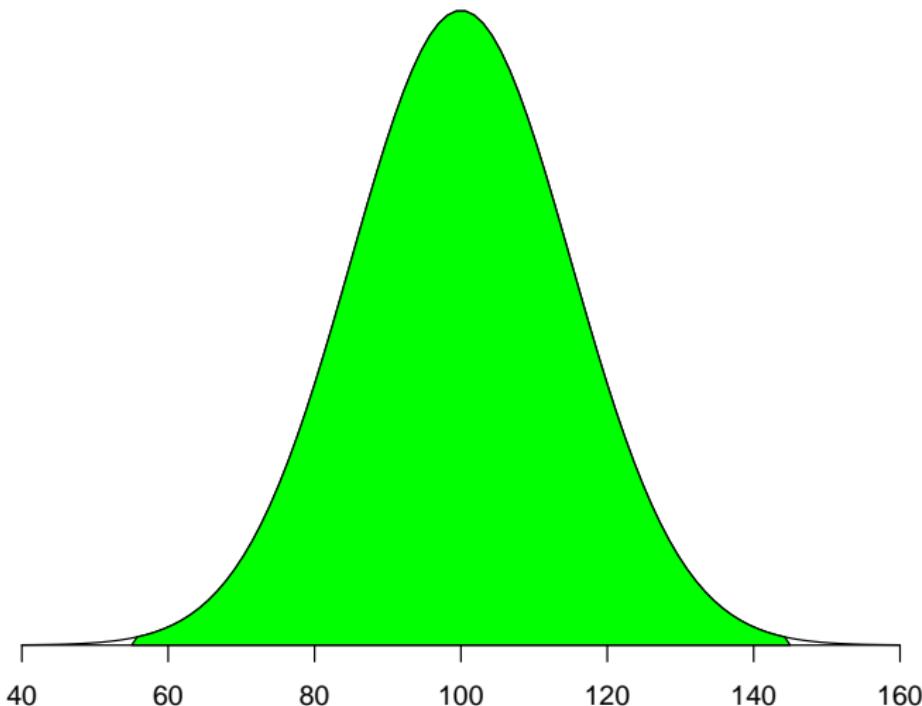
$$P(70 < \text{IQ} < 130) = 0.954$$



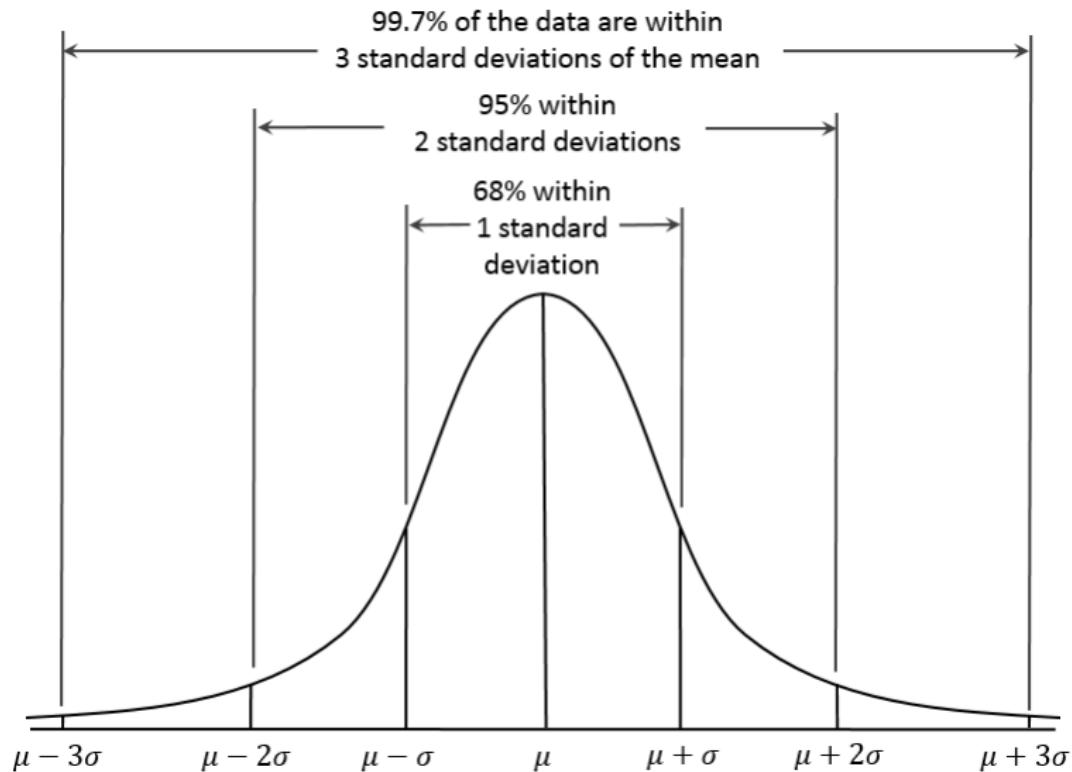
IQ - 3 Standard Deviation

Normal Distribution

$$P(55 < \text{IQ} < 145) = 0.997$$



Characteristics of the Normal distribution



Z-distribution

- ▶ The standard normal distribution is known as the z-distribution

Z-distribution

- ▶ The standard normal distribution is known as the z-distribution
- ▶ What is the mean and the standard deviation of this distribution?

Z-distribution

- ▶ The standard normal distribution is known as the z-distribution
- ▶ What is the mean and the standard deviation of this distribution?
- ▶ All normal distributions can be converted to the z-distribution

Z-distribution

- ▶ The standard normal distribution is known as the z-distribution
- ▶ What is the mean and the standard deviation of this distribution?
- ▶ All normal distributions can be converted to the z-distribution
- ▶ A raw score can be converted to a z-score.

Z-distribution

- ▶ The standard normal distribution is known as the z-distribution
- ▶ What is the mean and the standard deviation of this distribution?
- ▶ All normal distributions can be converted to the z-distribution
- ▶ A raw score can be converted to a z-score.

$$z = \frac{x-\mu}{\sigma}$$

SAT

The SAT is an aptitude test that high school students take. It is one of the criteria that is used in a college's decision to admit a student. It is composed of a math and a verbal section. Each has a mean of 500 and a standard deviation of 110 and is normally distributed.

- ▶ What are the scores on the test that corresponds to 3, 2, 1, 0, -1, -2, -3 standard deviations?

SAT

The SAT is an aptitude test that high school students take. It is one of the criteria that is used in a college's decision to admit a student. It is composed of a math and a verbal section. Each has a mean of 500 and a standard deviation of 110 and is normally distributed.

- ▶ What are the scores on the test that corresponds to 3, 2, 1, 0, -1, -2, -3 standard deviations?
- ▶ Assume 1000 people took the SAT,

SAT

The SAT is an aptitude test that high school students take. It is one of the criteria that is used in a college's decision to admit a student. It is composed of a math and a verbal section. Each has a mean of 500 and a standard deviation of 110 and is normally distributed.

- ▶ What are the scores on the test that corresponds to 3, 2, 1, 0, -1, -2, -3 standard deviations?
- ▶ Assume 1000 people took the SAT,
 - ▶ If Jon got a 700 on the math section, how many people scored above him?

SAT

The SAT is an aptitude test that high school students take. It is one of the criteria that is used in a college's decision to admit a student. It is composed of a math and a verbal section. Each has a mean of 500 and a standard deviation of 110 and is normally distributed.

- ▶ What are the scores on the test that corresponds to 3, 2, 1, 0, -1, -2, -3 standard deviations?
- ▶ Assume 1000 people took the SAT,
 - ▶ If Jon got a 700 on the math section, how many people scored above him?
 - ▶ If 300 people scored below Anna on the verbal section, what was Anna's score?

SAT

The SAT is an aptitude test that high school students take. It is one of the criteria that is used in a college's decision to admit a student. It is composed of a math and a verbal section. Each has a mean of 500 and a standard deviation of 110 and is normally distributed.

- ▶ What are the scores on the test that corresponds to 3, 2, 1, 0, -1, -2, -3 standard deviations?
- ▶ Assume 1000 people took the SAT,
 - ▶ If Jon got a 700 on the math section, how many people scored above him?
 - ▶ If 300 people scored below Anna on the verbal section, what was Anna's score?
 - ▶ How many people got scores between 390 and 610?

SAT

The SAT is an aptitude test that high school students take. It is one of the criteria that is used in a college's decision to admit a student. It is composed of a math and a verbal section. Each has a mean of 500 and a standard deviation of 110 and is normally distributed.

- ▶ What are the scores on the test that corresponds to 3, 2, 1, 0, -1, -2, -3 standard deviations?
- ▶ Assume 1000 people took the SAT,
 - ▶ If Jon got a 700 on the math section, how many people scored above him?
 - ▶ If 300 people scored below Anna on the verbal section, what was Anna's score?
 - ▶ How many people got scores between 390 and 610?
 - ▶ If Sigga got a 350 on the math section, how many people scored below her?

SAT

The SAT is an aptitude test that high school students take. It is one of the criteria that is used in a college's decision to admit a student. It is composed of a math and a verbal section. Each has a mean of 500 and a standard deviation of 110 and is normally distributed.

- ▶ What are the scores on the test that corresponds to 3, 2, 1, 0, -1, -2, -3 standard deviations?
- ▶ Assume 1000 people took the SAT,
 - ▶ If Jon got a 700 on the math section, how many people scored above him?
 - ▶ If 300 people scored below Anna on the verbal section, what was Anna's score?
 - ▶ How many people got scores between 390 and 610?
 - ▶ If Sigga got a 350 on the math section, how many people scored below her?
 - ▶ If Einar was in the 98% percentile in math, what was Einar's score?

Other standard scores

- ▶ T scores have a mean of 50 and a standard deviation of 10.

Other standard scores

- ▶ T scores have a mean of 50 and a standard deviation of 10.
 - ▶ What would T scores of 30 and 70 be as z-scores?

Other standard scores

- ▶ T scores have a mean of 50 and a standard deviation of 10.
 - ▶ What would T scores of 30 and 70 be as z-scores?
- ▶ stanine, range from 1 to 9, are centered at 5 with a standard deviation of 2. Each stanine, corresponds to 1/2 a standard deviation and the 5th stanine is at the mean.

Other standard scores

- ▶ T scores have a mean of 50 and a standard deviation of 10.
 - ▶ What would T scores of 30 and 70 be as z-scores?
- ▶ stanine, range from 1 to 9, are centered at 5 with a standard deviation of 2. Each stanine, corresponds to 1/2 a standard deviation and the 5th stanine is at the mean.
 - ▶ If you were in the 3rd stanine, what would your z-score be?

Other standard scores

- ▶ T scores have a mean of 50 and a standard deviation of 10.
 - ▶ What would T scores of 30 and 70 be as z-scores?
- ▶ stanine, range from 1 to 9, are centered at 5 with a standard deviation of 2. Each stanine, corresponds to 1/2 a standard deviation and the 5th stanine is at the mean.
 - ▶ If you were in the 3rd stanine, what would your z-score be?
 - ▶ How many people would be below you?

Other standard scores

- ▶ T scores have a mean of 50 and a standard deviation of 10.
 - ▶ What would T scores of 30 and 70 be as z-scores?
- ▶ stanine, range from 1 to 9, are centered at 5 with a standard deviation of 2. Each stanine, corresponds to 1/2 a standard deviation and the 5th stanine is at the mean.
 - ▶ If you were in the 3rd stanine, what would your z-score be?
 - ▶ How many people would be below you?
 - ▶ What percent of the people are between the 3rd and the 6th stanines?

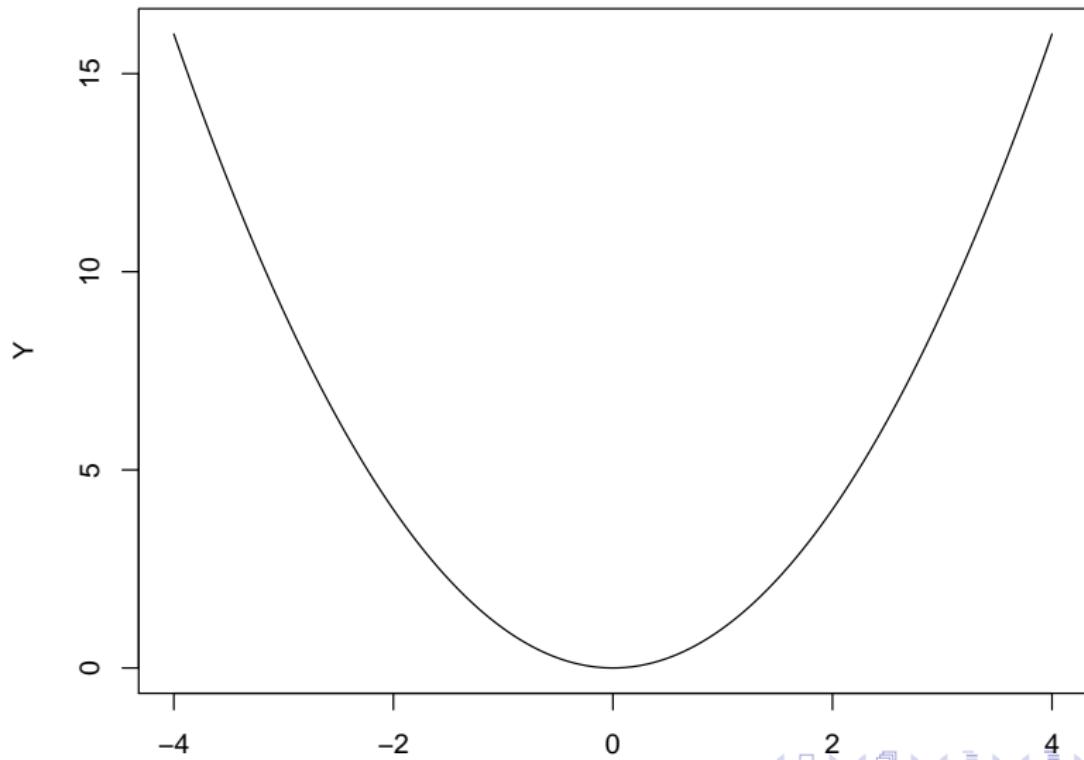
Other standard scores

- ▶ T scores have a mean of 50 and a standard deviation of 10.
 - ▶ What would T scores of 30 and 70 be as z-scores?
- ▶ stanine, range from 1 to 9, are centered at 5 with a standard deviation of 2. Each stanine, corresponds to 1/2 a standard deviation and the 5th stanine is at the mean.
 - ▶ If you were in the 3rd stanine, what would your z-score be?
 - ▶ How many people would be below you?
 - ▶ What percent of the people are between the 3rd and the 6th stanines?
- ▶ Various linear and non-linear transformations are done to create scores and scores may be normalized.

What is a correlation?

- ▶ Is it an association?
- ▶ Does it imply causation?
- ▶ Is a correlation necessary for causation?
- ▶ Does it need linearity?
- ▶ Is it affected by variability?
- ▶ Is it affected by outliers?
- ▶ Is it related to the simple linear regression?

What is the Pearson correlation coefficient?



Pearson correlation coefficient

$$\frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2} \sqrt{\sum((Y - \bar{Y})^2)}}$$

Calculating Pearson correlation coefficient

X	Y
5	6
3	0
1	0
Mean	3 2

```
x <- c(5, 3, 1)
y <- c(6, 0, 0)
cor(x, y)
```

R correlation applet

1. Open RStudio
2. Open correlation_applet.R
3. Click the "Source" button

Spearman's rho

- ▶ Non-parametric measure of association
- ▶ Appropriate when at least one of your variables is ordinal variables
- ▶ Don't use Pearson's correlation with ordinal variables!

Simple Linear Regression

- ▶ If you are interested in predicting height given someone's weight, what would you do?

Simple Linear Regression

- ▶ If you are interested in predicting height given someone's weight, what would you do?
- ▶ We could consider a regression model.

Simple Linear Regression

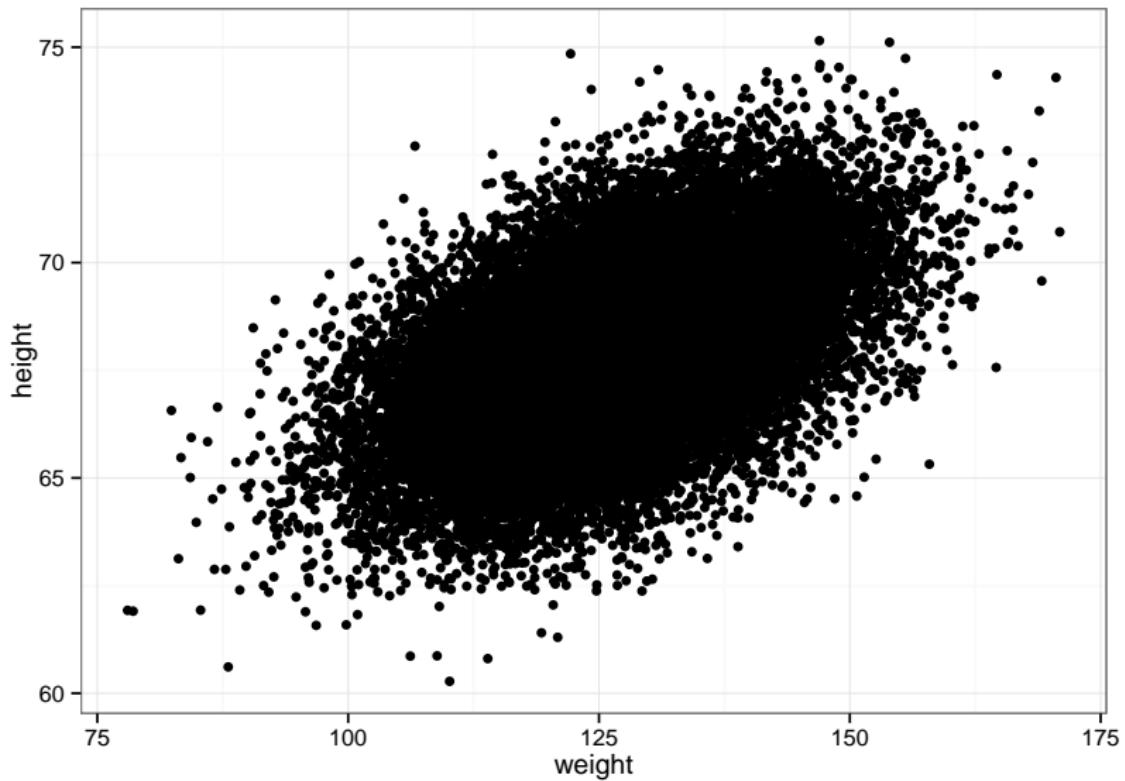
- ▶ If you are interested in predicting height given someone's weight, what would you do?
- ▶ We could consider a regression model.
- ▶ $Y_i = \beta_0 + \beta_1 * X_i$

Simple Linear Regression

- ▶ If you are interested in predicting height given someone's weight, what would you do?
- ▶ We could consider a regression model.
- ▶ $Y_i = \beta_0 + \beta_1 * X_i$
- ▶ How could we assess if this is appropriate?

1993 Growth Survey of 25,000 Hong Kongese children

source: http://wiki.stat.ucla.edu/socr/index.php/SOCR_Data_Dinov_020108_HeightsWeights



Model Summary

Parameter	Estimate	SE	t-value	p-value
β_0	57.57	0.11	506.01	< .001
β_1	0.08	0.001	91.98	< .001

How does this relate to correlation?

Slope and the correlation

- ▶ There is a relationship between the estimated slope and the correlation between two variables in a simple linear regression.

Slope and the correlation

- ▶ There is a relationship between the estimated slope and the correlation between two variables in a simple linear regression.
- ▶ $r = \beta_1 \frac{sd_x}{sd_y}$

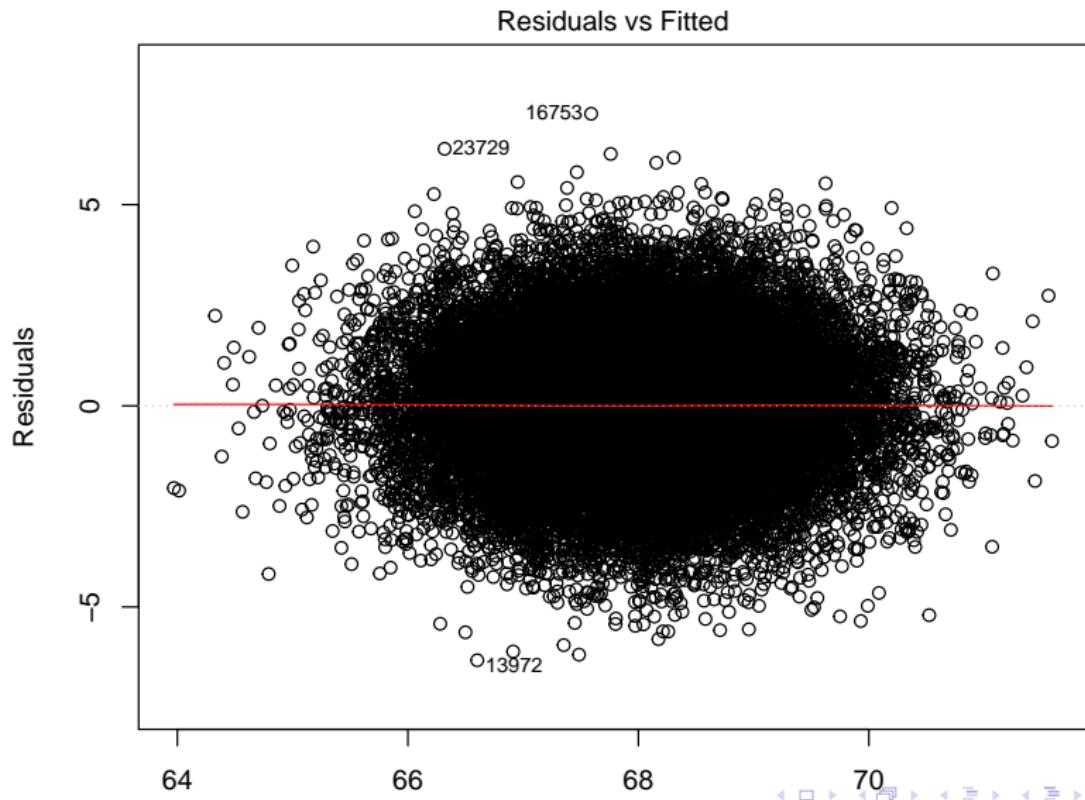
Slope and the correlation

- ▶ There is a relationship between the estimated slope and the correlation between two variables in a simple linear regression.
- ▶ $r = \beta_1 \frac{sd_x}{sd_y}$
- ▶ If $\beta_1 = 0.08$, the standard deviation of weight and height are 11.6608976 and 1.9016788, respectively, what is r?

Slope and the correlation

- ▶ There is a relationship between the estimated slope and the correlation between two variables in a simple linear regression.
- ▶ $r = \beta_1 \frac{sd_x}{sd_y}$
- ▶ If $\beta_1 = 0.08$, the standard deviation of weight and height are 11.6608976 and 1.9016788, respectively, what is r?
- ▶ 0.5028585

Always look at the residuals



Brief history of testing

- ▶ 2200 BCE, Chinese believed to use testing for determining who would get governmental jobs
- ▶ Greek and Romans categorized individuals based on personality type ("blood" or "phlegm")
- ▶ Francis Galton's classification based on "natural gift" (i.e. eugenics)
 - ▶ Contributed to development of questionnaires, rating scales, and self-report inventories
- ▶ Wilhelm Wundt's laboratory and his focus on "standardization"
 - ▶ James Cattell's mental tests
 - ▶ Charles Spearman - reliability and factor analysis

Testing in the 20th century

- ▶ 1905, Binet and Simon publish a test measuring intelligence in mental retarded school children in Paris
- ▶ 1939, Wechsler publishes a test to measure intelligence in adults (would become WAIS)
- ▶ Group intelligence test administered by the US military during WWI and WWII
- ▶ WWI personality tests used to screen recruits

Necessary test assumptions

- ▶ Psychological traits and states exist

Necessary test assumptions

- ▶ Psychological traits and states exist
- ▶ Psychological traits and states can be measured

Necessary test assumptions

- ▶ Psychological traits and states exist
- ▶ Psychological traits and states can be measured
- ▶ Behavior on tests predicts non-test behavior

Necessary test assumptions

- ▶ Psychological traits and states exist
- ▶ Psychological traits and states can be measured
- ▶ Behavior on tests predicts non-test behavior
- ▶ Measurement error is part of the process

Necessary test assumptions

- ▶ Psychological traits and states exist
- ▶ Psychological traits and states can be measured
- ▶ Behavior on tests predicts non-test behavior
- ▶ Measurement error is part of the process
- ▶ Test can be fair

Necessary test assumptions

- ▶ Psychological traits and states exist
- ▶ Psychological traits and states can be measured
- ▶ Behavior on tests predicts non-test behavior
- ▶ Measurement error is part of the process
- ▶ Test can be fair
- ▶ Test can benefit society

Norm-Referenced

- ▶ Individuals scores are relative only to some reference group

Norm-Referenced

- ▶ Individuals scores are relative only to some reference group
- ▶ This group should represent the entire pool of test takers for the tested construct

Norm-Referenced

- ▶ Individuals scores are relative only to some reference group
- ▶ This group should represent the entire pool of test takers for the tested construct
- ▶ Collectively, this group is known as a **normative sample** and data from them make up the **norms**

Norm-Referenced

- ▶ Individuals scores are relative only to some reference group
- ▶ This group should represent the entire pool of test takers for the tested construct
- ▶ Collectively, this group is known as a **normative sample** and data from them make up the **norms**
- ▶ Understanding the normative sample is very important, why?

Different Norms

- ▶ Developmental Norms
 - ▶ Age Norms
 - ▶ A 6 year old performs at the level of a 10 year old
 - ▶ This is on this material only though!
 - ▶ Grade Norms
 - ▶ School year typically 10 months in the US (and Iceland?)
 - ▶ A 4th grader is performing at the level of a 5th grader in third month
 - ▶ This is on this material only though!
- ▶ National Norms, national representative
 - ▶ Anchor norms enable two tests to be compared
 - ▶ In USA, students could take SAT or ACT for admission to college

Fixed Reference and Criterion-Related

- ▶ Fixed reference group scores are used as the basis for calculation of future administrations of the test
- ▶ SAT does this through using anchor items and equating
- ▶ Criterion-referenced, evaluate a score with reference to a set criteria or standard NOT other test takers
- ▶ **What is the fairest way to score grades in a class room?**

Sampling

- ▶ Simple random sample
- ▶ Stratified random sample
- ▶ Cluster random sample
- ▶ Purposive sample
- ▶ Convenience sample