Stats Recap

Prof. Maureen Ritchey

Approaches to analyzing data

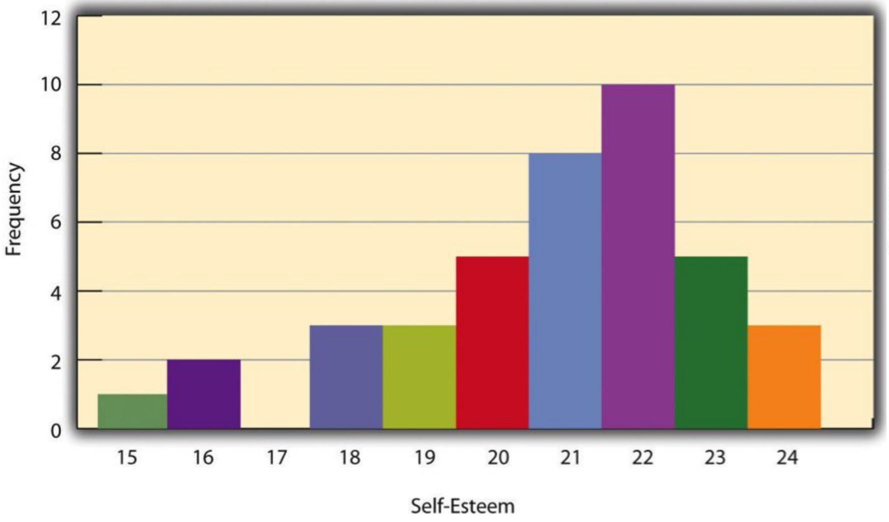
Descriptive statistics

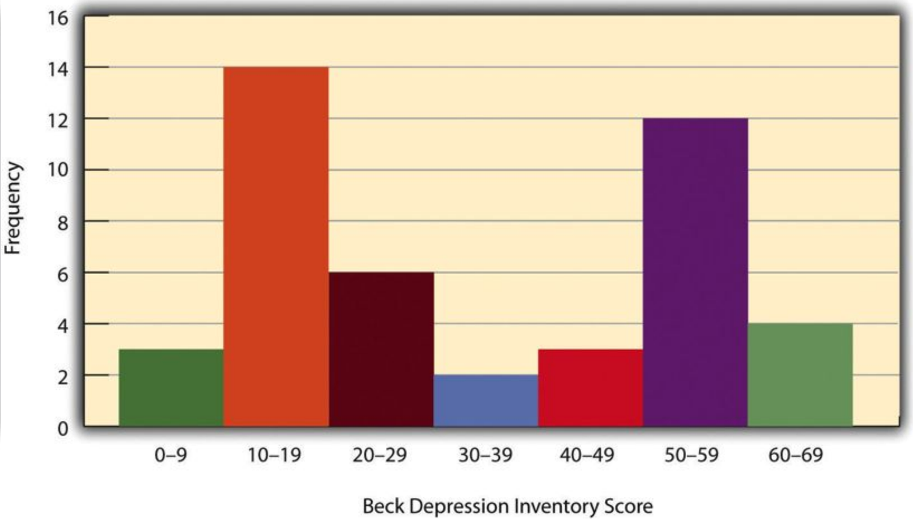
- Summarize the features of the data
- e.g., mean, standard deviation, histograms & other visualizations

Inferential statistics

- Allow you to draw conclusions about a **population** based on a **sample**
- Why use inferential statistics?

Descriptive statistics: Distributions

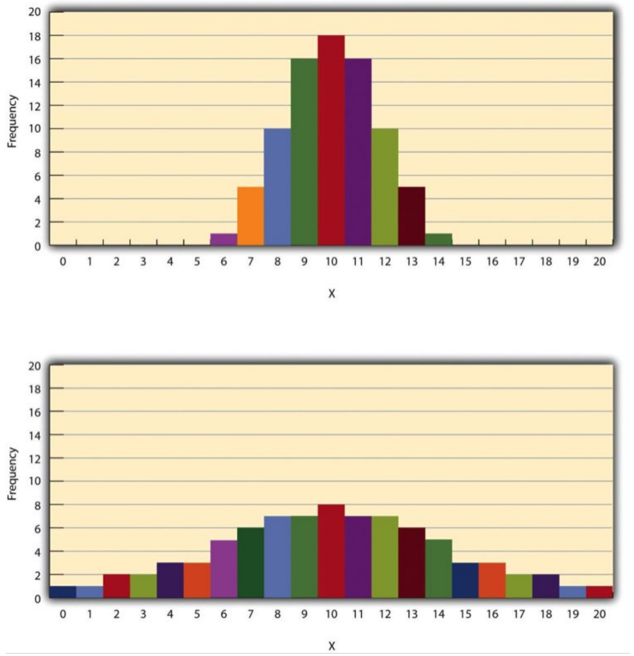




Example of histograms

Recap summary stats: *mean, median, mode*

Descriptive statistics: Distributions



Example of histograms

Recap summary stats: *range, variance*

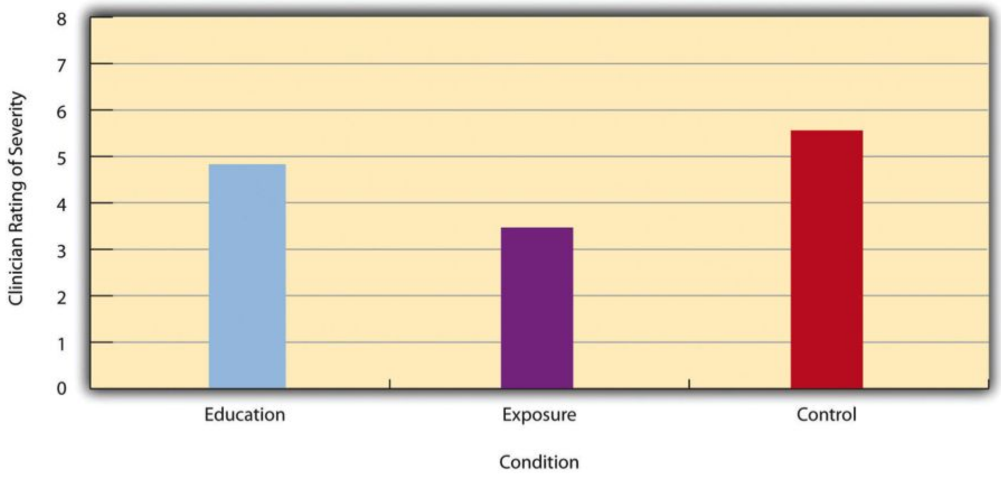
Descriptive statistics: Differences between conditions

Often expressed with bar graphs

How do we know if a difference is "big"?

Effect sizes!

Cohen's d = (M1 - M2) / **SD**

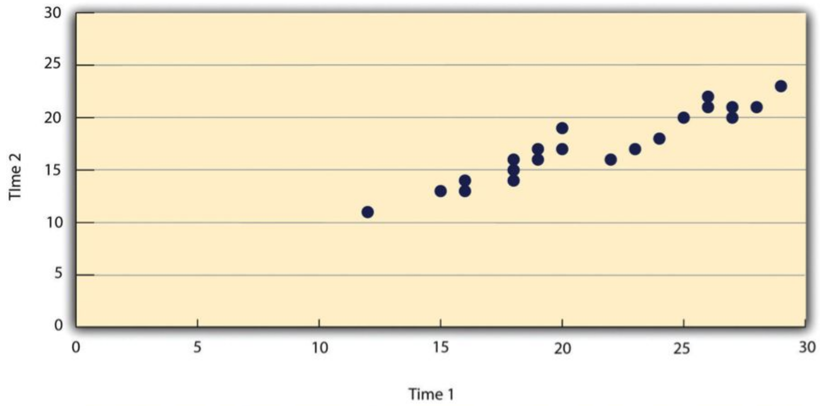


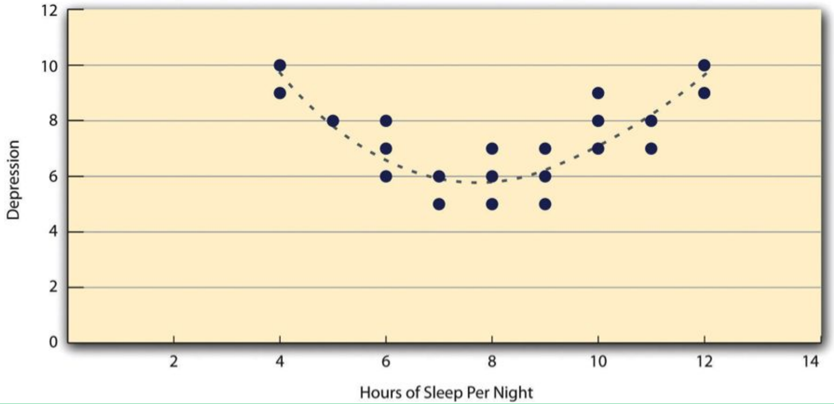
Descriptive statistics: Relationships between variables

Often expressed with scatterplots

Can be **linear** (fit by straight line) or **nonlinear** (fit by curved line)

How do we know if a relationship is "strong"?





Descriptive statistics: Relationships between variables

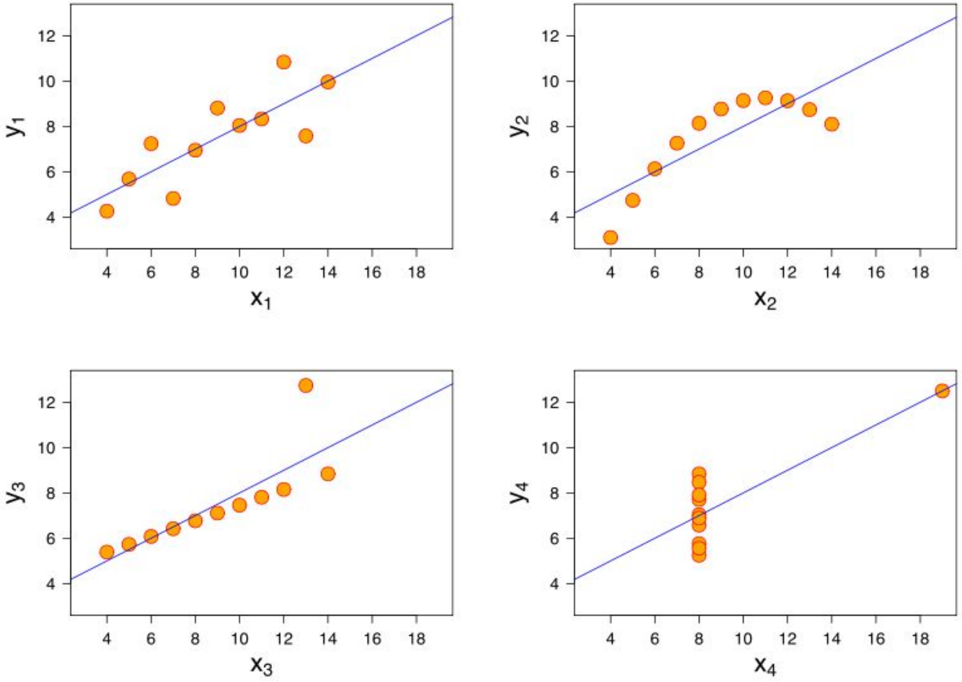
Anscombe's Quartet

These data have the same…

Mean

Variance

And Pearson's r value!



When to use descriptive statistics

- Often as a “first pass” set of analyses
- They can reveal if you have missing data, or there are any outliers in the data
- They give you a sense of the “big picture” of your data.

However… they come up short when you want to draw real conclusions.

Inferential statistics

They tell you whether you can generalize your effect to the population, because it's unlikely to have occurred by chance. That means: they help you answer your research question!

Null hypothesis testing (NHST):

- Imagine you find that there is a relationship between your IV and DV. How do you know it's "real" versus just sampling error?
- NHST lets you test your hypothesis (that your IV has a real effect on the DV) against the null hypothesis (that there is no real effect, it's only error)

Step 1: Assume the null hypothesis.
Step 2: How likely are you to see the sample (observed) relationship if the null hypothesis was true? (p-value)
Step 3: If it's unlikely - **reject** the null hypothesis*. If it's not, **retain** the null.

*We never "prove" the alternative hypothesis, only "reject" the null


I CAN'T BELIEVE SCHOOLS ARE STILL TEACHING KIDS ABOUT THE NULL HYPOTHESIS.
I REMEMBER READING A BIG STUDY THAT CONCLUSIVELY DISPROVED IT *YEARS* AGO.

Inferential statistics: Testing relationships

Whether or not a relationship is significant is affected not only by how “big” or “strong” the effect is, but also by the sample size (N).

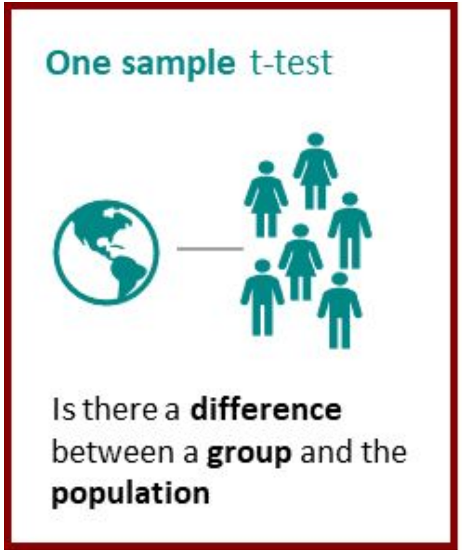
In other words, an effect size might be significant with a big sample, but not with a small sample.

Inferential statistics: Testing differences

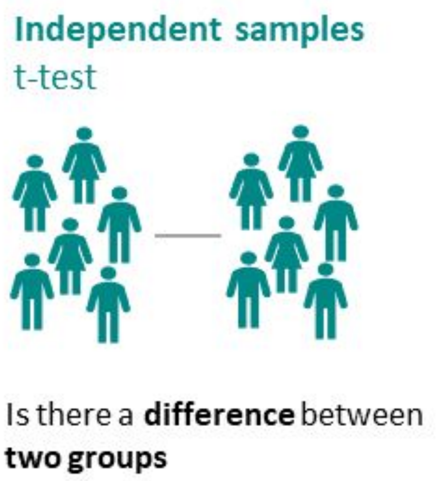
Test the difference between **two means** -> t-test!

- **One-sample t-test:** test against a known value (often zero)
- **Two-sample t-test:** test between two groups (*between-subjects*)
- **Paired samples t-test:** test between two conditions (*within-subjects*)

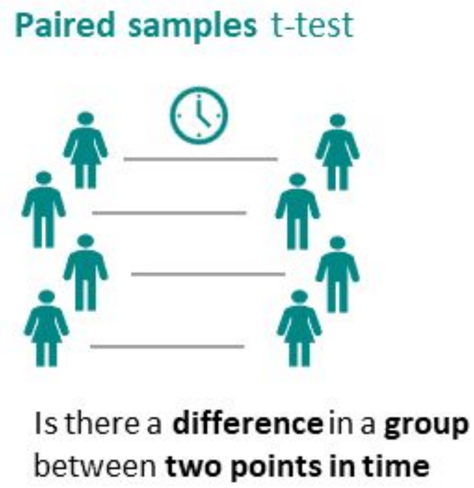
Inferential statistics: Testing differences



Or a known value



Aka two-sample t-test



Or between two conditions

Inferential statistics: Testing differences

Test the difference between **two means** -> t-test!

- **One-sample t-test:** test against a known value (often zero)
- **Two-sample t-test:** test between two groups (*between-subjects*)
- **Paired samples t-test:** test between two conditions (*within-subjects*)

When you run a t-test, you get:

- A t-statistic (math + math = t-stat)
  - This can be directional, i.e. A>B is different from B>A
- A p-value (likelihood of that t-stat occurring by chance)

Inferential statistics: Testing differences

```
# One-sample t-test
t.test(my_data$variable, mu = 10) # compare to a mean of 10


# Independent t-test
t.test(my_data$dependent_variable ~ my_data$grouping_variable)


# Paired t-test
t.test(my_data$before, my_data$after, paired = TRUE)
```

Inferential statistics: Testing effects across multiple conditions

Test the difference across means for **multiple** levels of an IV (or multiple IVs) -> ANOVA!

- IVs can be repeated-measures or between-subjects, or mixed
- **One-way ANOVA:** one IV, multiple levels (conditions) - main effect only
  - E.g., effect of emotional valence on memory accuracy

	Emotionally negative	Emotionally neutral
All participants	accuracy Main effect	accuracy of valence

Inferential statistics: Testing effects across multiple conditions

Test the difference across means for **multiple** levels of an IV (or multiple IVs) -> ANOVA!

- IVs can be repeated-measures or between-subjects, or mixed
- **Factorial ANOVA:** more than one IV, can look at main effects & interactions
    - E.g., effect of emotional valence AND age on memory accuracy

	Emotionally negative	Emotionally neutral
Young adult (< 35)		
Middle age (35-65)		
Older age (> 65)		

Main effect of valence

Main effect of age

Interaction

Inferential statistics: Testing effects across multiple conditions

Test the difference across means for **multiple** levels of an IV (or multiple IVs) -> ANOVA!

- A F-statistic (math + math = F-stat) -> FOR EVERY M.E. & Intx!!
    - This is not directional - the stat alone doesn't tell you whether it's A>B>C or A<B<C or A>B<C
- A p-value (likelihood of that F-stat occurring by chance) -> FOR EVERY M.E. & Intx!!

After you run an ANOVA, you must test the direction of your effects (e.g., using "post-hoc" t-tests such as Tukey's HSD)

	Emotionally negative	Emotionally neutral
Young adult (< 35)		
Middle age (35-65)		
Older age (> 65)		

Main effect of valence

Main effect of age

Interaction

Inferential statistics: Testing effects across multiple conditions

```
# Installation and Loading:

install.packages("ez")
library(ez)


# One-way ANOVA (between-subjects):

anova_result <- ezANOVA(data = my_data, dv = dependent_variable, wid =
subject_id, between = grouping_variable)
print(anova_result)
```

Inferential statistics: Testing effects across multiple conditions

```
# Two-way ANOVA (between-subjects):

anova_result2 <- ezANOVA(data = my_data, dv = dependent_variable, wid =
subject_id, between = c(grouping_variable1, grouping_variable2))
print(anova_result2)
```

```
# Within-subjects ANOVA (repeated measures):

within_anova <- ezANOVA(data = my_data, dv = dependent_variable, wid =
subject_id, within = within_factor)
print(within_anova)
```

```
# Mixed-design ANOVA:

mixed_anova <- ezANOVA(data = my_data, dv = dependent_variable, wid =
subject_id, between = between_factor, within = within_factor)
print(mixed_anova)
```

Criticisms of NHST

- Using a p-value of .05 is a convention but it’s also totally arbitrary - perhaps we shouldn’t draw such harsh dividing lines between “significant” and “non-significant”
- NHST lets you reject the null hypothesis… but it doesn’t really tell you how much evidence is *in favor* of your alternative hypothesis
- Alternatives:
  - Combine p-values with effect sizes
  - Focus on confidence intervals
  - Bayesian statistics (testing the likelihood of your hypothesis)