Fundamentals of Statistical Inference

Measuring Uncertainty

Aaron R. Baggett, Ph.D.

Department of Psychology
University of Mary Hardin-Baylor

August 10, 2017

Outline

Outline

- Descriptive vs. inferential statistics
- The normal distribution
- Comparing groups
- Statistical/practical significance

Resources

- Slides, data, and handouts available at:

`bit.ly/umhb_dpt`

Statistics

Statistics

Stigler (1986, p. 1):

> *[Modern statistics provides]...the logic and methodology for the measurement of uncertainty and for examination of the consequences of that uncertainty in the planning and interpretation of experimentation and observation.*

Statistics

- Experimentation and observation:
    1. Measurement of uncertainty
    2. Examination of the consequences of that uncertainty

Statistics

- Two fundamental branches
  1. Descriptive statistics
     - Summarize data
     - Condense larger themes
  2. Inferential statistics
     - Infer meaning
     - Test predictions

Example

Low Birth Weight Study

- Baystate Medical Center, Springfield, MA.
- Sample of 189 births in 1986
- Risk factors in low birth weight babies

Low Birth Weight Study

Age	Weight	Race	Smoker?	Birth Weight
19	182	Black	Non-Smoker	5.56
33	155	Other	Non-Smoker	5.62
20	105	White	Smoker	5.64
21	108	White	Smoker	5.72
18	107	White	Smoker	5.73
21	124	Other	Non-Smoker	5.78

Descriptive Statistics

Descriptive Statistics

- How many babies were considered low birth weight (< 5.5 lbs.)?
- How many mothers smoked during pregnancy?
- How much did the average baby weigh?
  - By smoking status
  - By race

Descriptive Statistics

Question:

Do babies born to mothers who smoked during pregnancy weigh less than those born to mothers who did not?

Descriptive Statistics

Question:

Do babies born to mothers who smoked during pregnancy weigh less than those born to mothers who did not?

- How should we answer this question?

Descriptive Statistics

Question:

[ON AVERAGE], do babies born to mothers who smoked during pregnancy weigh less than those born to mothers who did not?

Smoking Status	n	Min.	Max.	$\bar{X}$	SD
Non-Smoker	115	2.25	11.00	6.74	1.66
Smoker	74	1.56	9.34	6.11	1.46

Descriptive Statistics

Question:

1. Based on our sample, what are we left to assume about the weights of babies *in the population* born to smoking and non-smoking mothers?

Descriptive Statistics

Question:

1. Based on our sample, what are we left to assume about the weights of babies *in the population* born to smoking and non-smoking mothers?
   - That the sample estimates represent the population parameters

Smoking Status	n	Min.	Max.	$\bar{X}$	SD
Non-Smoker	115	2.25	11.00	6.74	1.66
Smoker	74	1.56	9.34	6.11	1.46

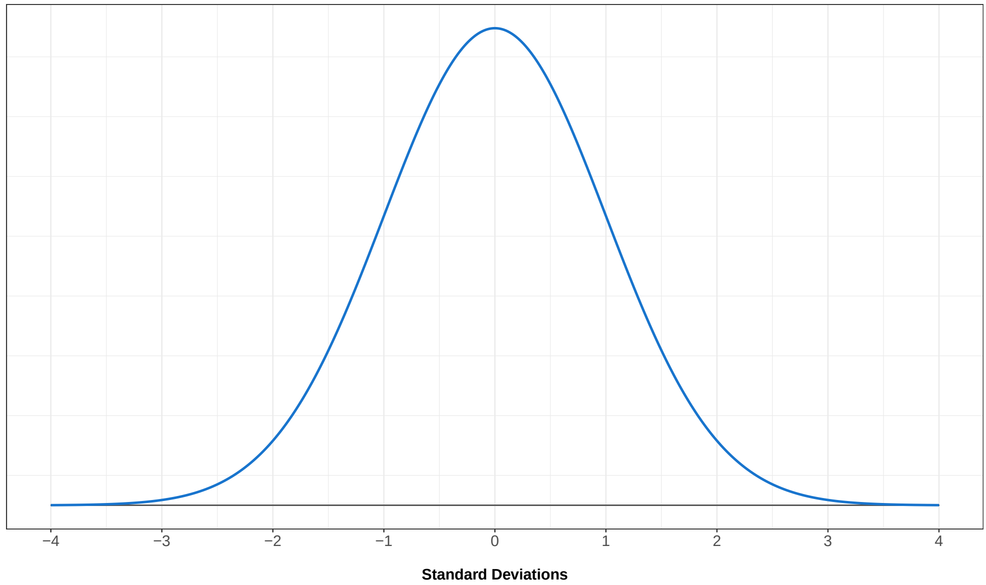
Descriptive Statistics

- In fact, we assume that the population distribution of baby weights is "normal"

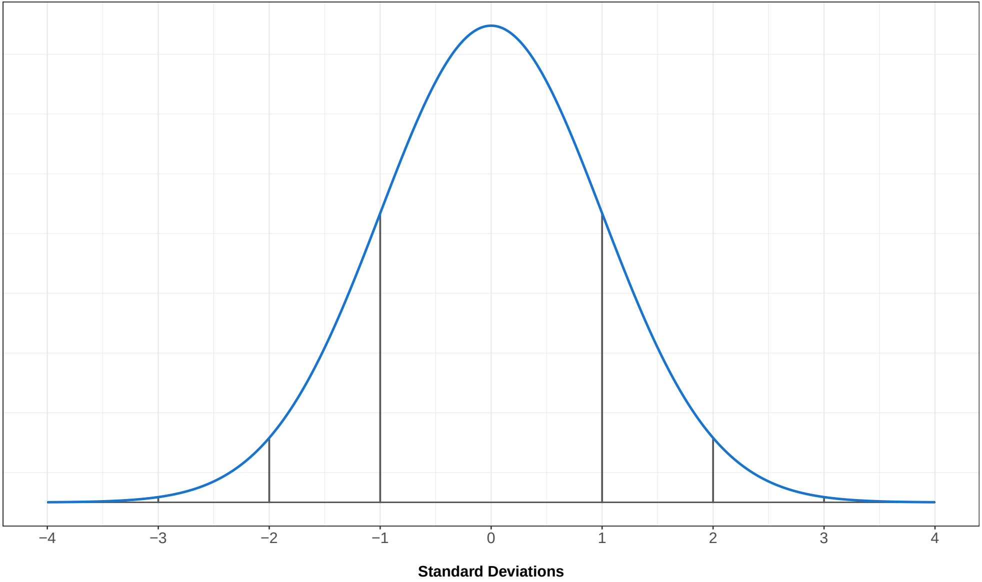
Normal Distribution

INSERT: /Users/AB/Dropbox/Talks/DPT Series/Figures/normal_curve.pdf

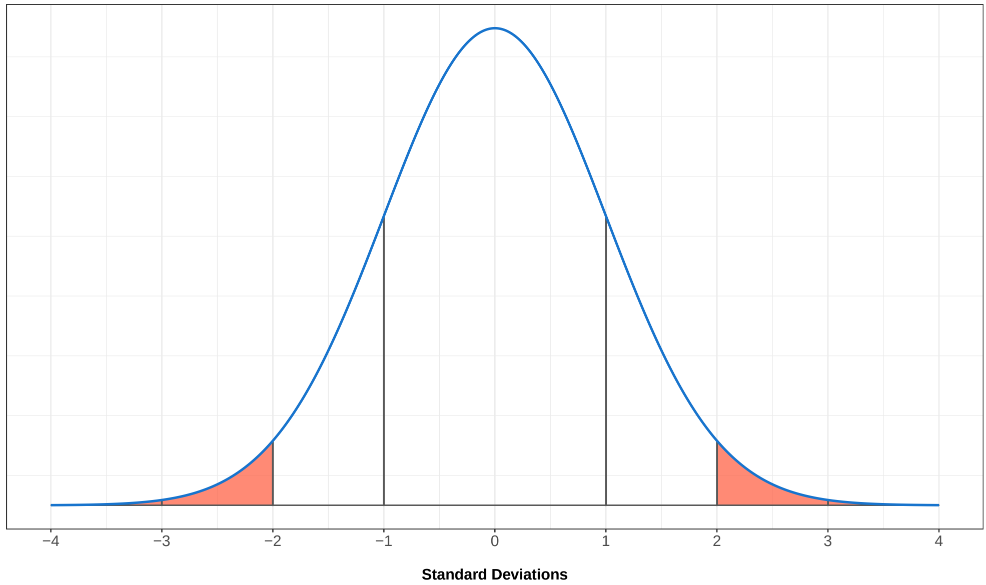
Normal Distribution



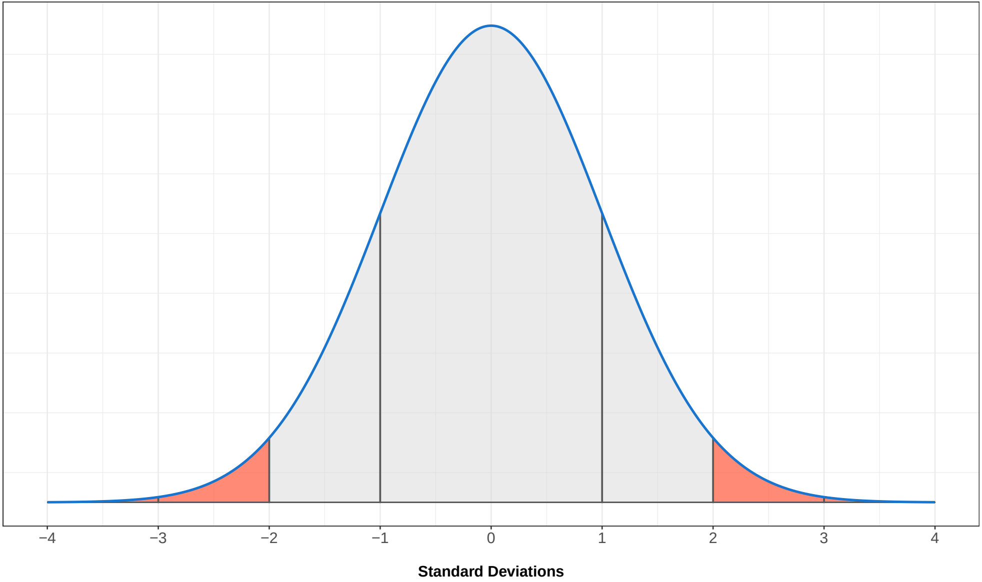
Normal Distribution



Normal Distribution



Normal Distribution



Inferential Statistics

Inferential Statistics

- More useful than descriptives
- Allow for making predictions or generalizations
- Key to hypothesis testing

Inferential Statistics

Question:

Do babies born to mothers who smoked during pregnancy weigh less than those born to mothers who did not?

Inferential Statistics

Question:

Do babies born to mothers who smoked during pregnancy weigh [STATISTICALLY SIGNIFICANTLY] less than those born to mothers who did not?

- How should we answer this question?

Inferential Statistics

- What do we mean by statistical significance?
- Observed differences which exceed “normality.”

Inferential Statistics

- We usually consider differences beyond ± 2 *SDs* from *M* to be "statistically significant"
- NOTE: Statistical significance $\neq$ practical significance

Low Birth Weight Study

Question:

- Do babies born to mothers who smoked during pregnancy weigh less than those born to mothers who did not?

Low Birth Weight Study

Hypotheses:

- H_0: There is no mean difference in the birth weight of babies born to mothers who did and did not smoke during pregnancy ($\mu_N - \mu_S = 0$)

- H_1: There is some difference in the birth weight of babies born to mothers who did and did not smoke during pregnancy ($\mu_N - \mu_S \neq 0$)

Low Birth Weight Study

- Let's test our hypothesis using an independent-samples t-test
  - IV: Mothers' smoking status (smoker, non-smoker)
  - DV: Baby birth weight

$$t = \frac{\overline{X}_{non-smokers} - \overline{X}_{smokers}}{\sqrt{\frac{s^2_{non-smokers}}{N_{non-smokers}} + \frac{s^2_{smokers}}{N_{smokers}}}}$$

Results

Table 1: Results of independent-samples *t*-test

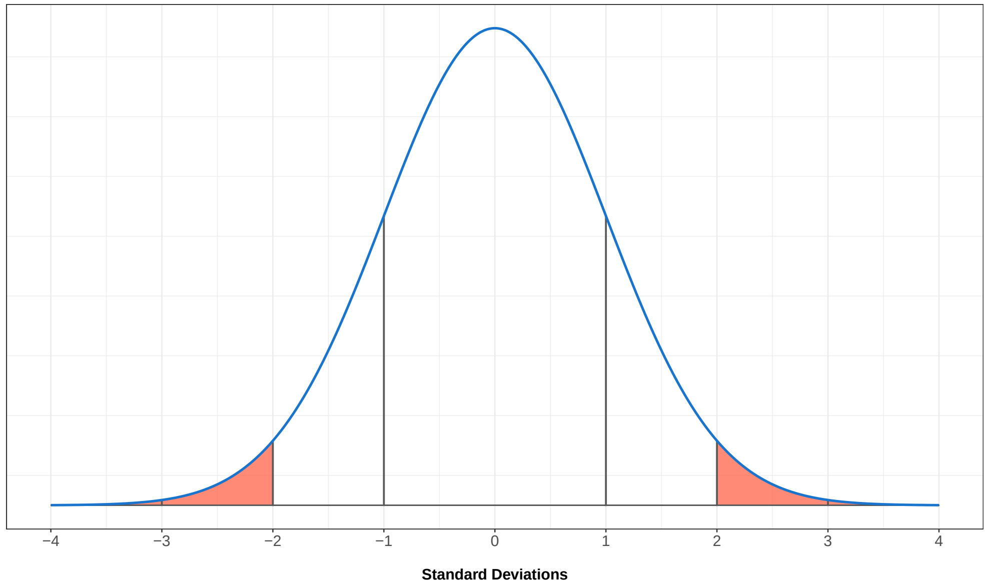
	Non-Smokers			Smokers					
	n	M	SD	n	M	SD	$t(187)$	p	ω^2
Baby birth weight	115	6.74	1.66	47	6.11	1.46	2.63	0.009	0.008

Note: M = Mean; SD = Standard deviation

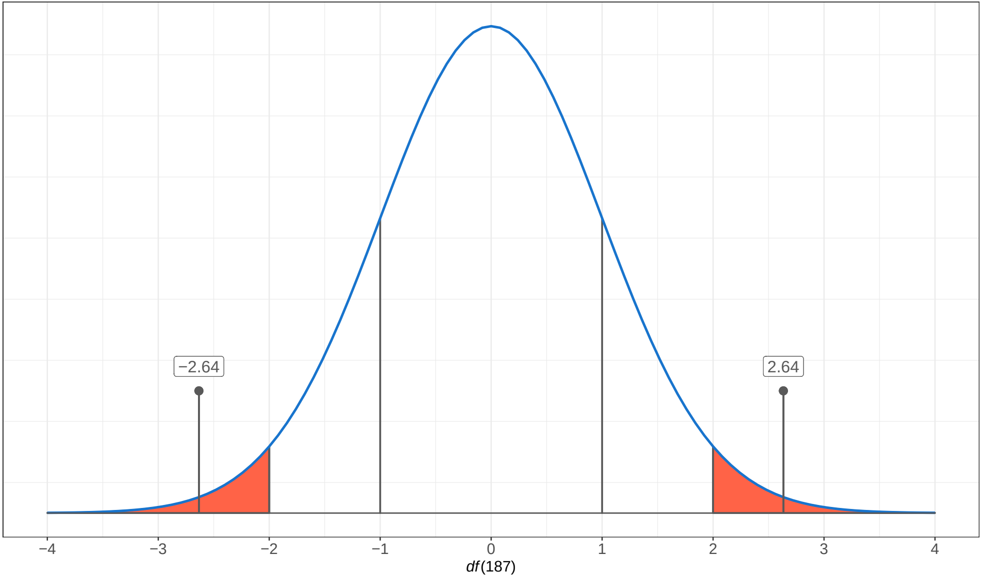
Results

- Assuming the null hypothesis, in reality, is true, the probability of obtaining a mean difference in birth weight ≥ 0.62 lbs. is 0.009 (0.90%)
- Birth weights appear to differ statistically significantly

Results



Results



Results

- But, is the difference of $M = 0.62$ lbs. meaningful?

Results

- But, is the difference of $M = 0.62$ lbs. meaningful?
- A meaningful difference implies practicality or usefulness in the real world

Results

- But, is the difference of $M = 0.62$ lbs. meaningful?
- A meaningful difference implies practicality or usefulness in the real world
- Effect size (ω^2): Proportion of variance explained in the model

Results

- But, is the difference of $M = 0.62$ lbs. meaningful?
- A meaningful difference implies practicality or usefulness in the real world
- Effect size (ω^2): Proportion of variance explained in the model
- Smoking status explains 0.009 (0.90%) of the variance in baby birth weight

Results

- But, is the difference of $M = 0.62$ lbs. meaningful?
- A meaningful difference implies practicality or usefulness in the real world
- Effect size (ω^2): Proportion of variance explained in the model
- Smoking status explains 0.009 (0.90%) of the variance in baby birth weight
- Thus, 100% - 0.991% = 99.10% of the variance in baby birth weight is left unexplained

Recap

Recap

- Descriptive statistics allow us to summarize data from a sample
- Inferential statistics allow us to predict and generalize about a population
- Hypothesis testing allows us to construct a sense of meaning about the world

Next Time

- Making decisions using hypothesis testing and prediction
  - Statistical variables
  - Multiple group comparisons (ANOVA)
  - Predicting outcomes (Regression)