Lecture 7
Segment 2

NHST: A closer look

NHST: A closer look

- Important concepts/topics
  - Logic of NHST
  - NHST problems
  - NHST alternatives

NHST

- Null Hypothesis Significance Testing (NHST)
  - H_0 = null hypothesis: e.g., $r = 0$
  - H_A = alternative hypothesis: e.g., $r > 0$

NHST

- Null Hypothesis Significance Testing (NHST)
  - H_0 = null hypothesis: e.g., $B = 0$
  - H_A = alternative hypothesis: e.g., $B > 0$

NHST

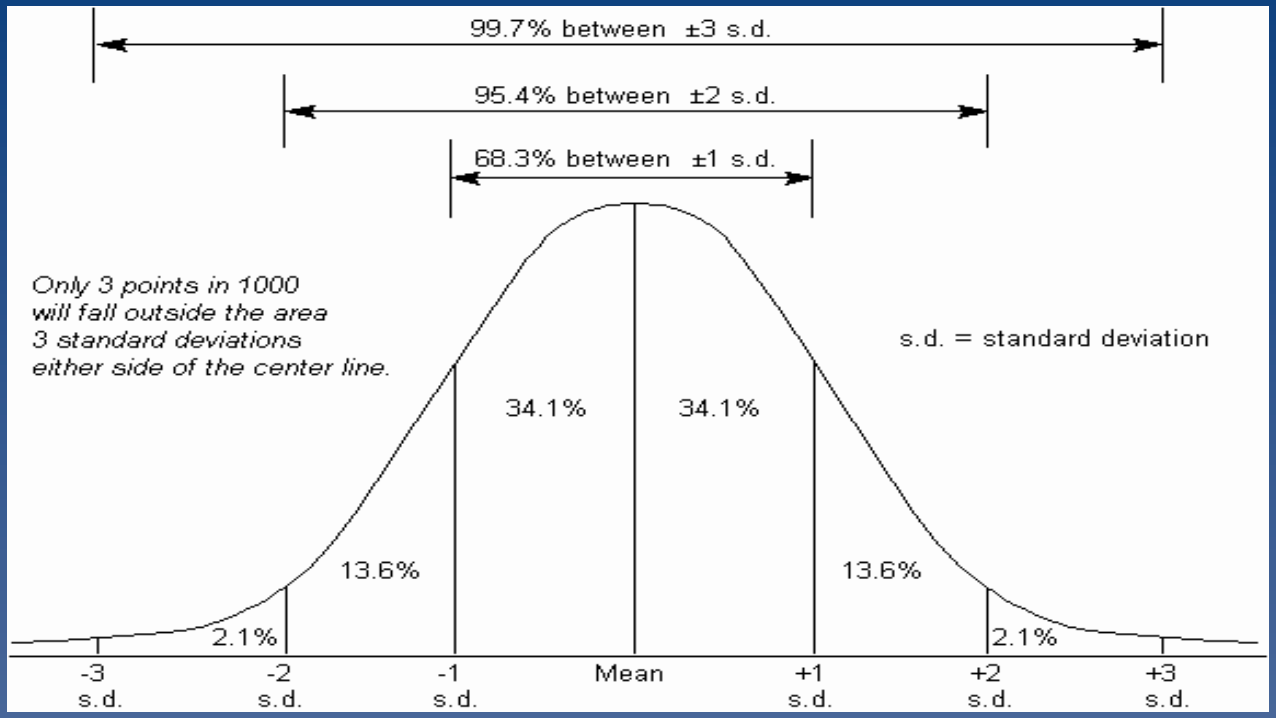
- Null Hypothesis Significance Testing (NHST)
  - Assume H_0 is true, then calculate the probability of observing data with these characteristics, given that H_0 is true
    - Thus, $p = P(D|H_0)$
    - If $p < \alpha$ then Reject H_0, else Retain H_0

NHST

Experimenter Decision

Truth	Retain H_0	Reject H_0
H_0 true	Correct Decision	Type I error (False alarm)
H_0 false	Type II error (Miss)	Correct Decision

The normal distribution

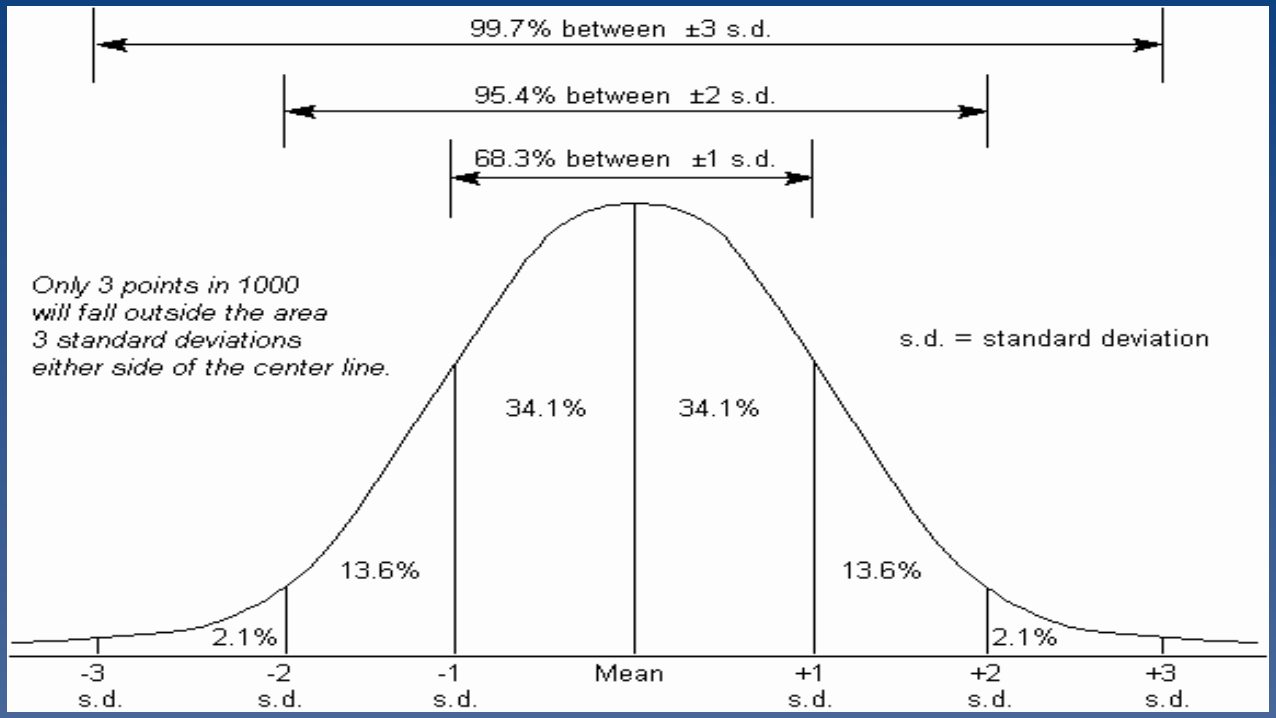


NHST

Experimenter Decision

Truth	Retain H_0	Reject H_0
H_0 true	$p = (1 - \alpha)$	$p = \alpha$
H_0 false	$p = \beta$ (1 - POWER)	$p = (1 - \beta)$ POWER

The normal distribution



NHST

- $p = P(D|H_0)$
- Given that the null hypothesis is true, the probability of these, or more extreme data, is p
  - NOT: The probability of the null hypothesis being true is p
  - In other words, $P(D|H_0) <> P(H_0|D)$

NHST can be applied to:

- *r*
  - Is the correlation significantly different from zero?
- B
  - Is the slope of the regression line for X significantly different from zero?

NHST for B

- $t = B / SE$
  - B is the unstandardized regression coefficient
  - SE = standard error
  - $SE = \sqrt{SS.RESIDUAL / (N - 2)}$

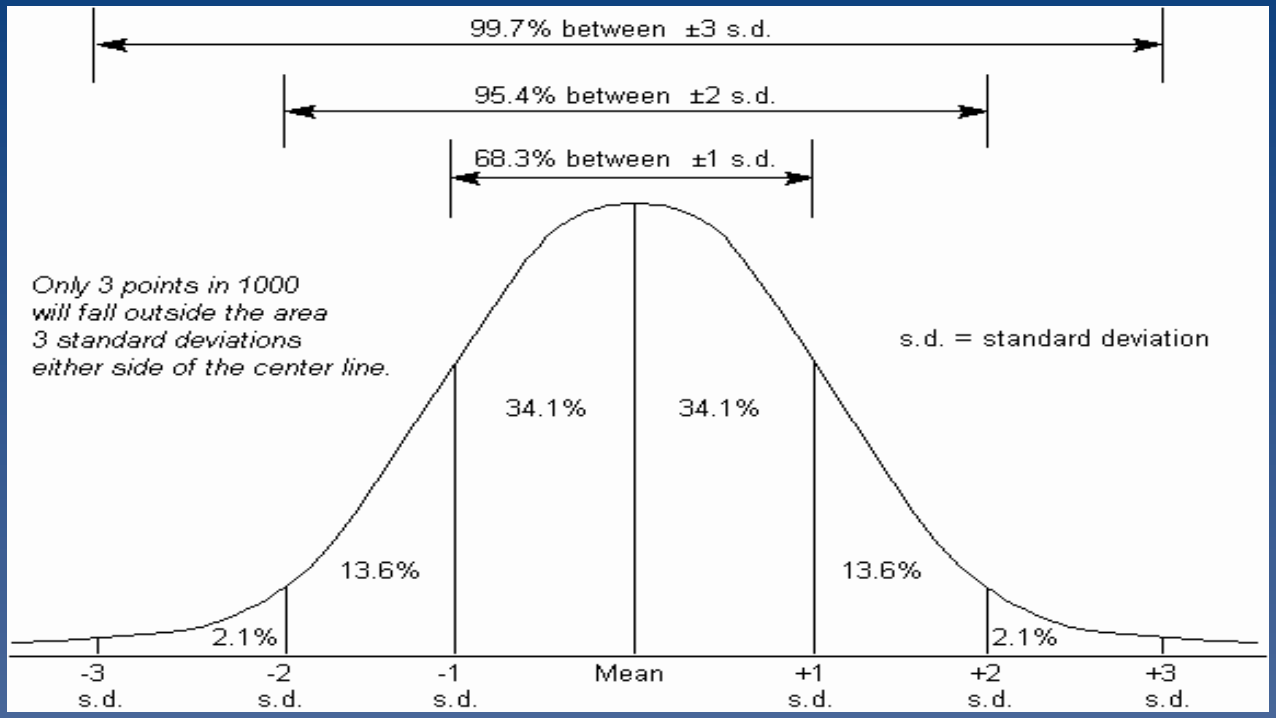
NHST: Problems!

- Biased by N
- Binary outcome
- Null “model” is a weak hypothesis

NHST: Problems!

- Biased by N
  - p-value is based on t-value
  - $t = B / SE$
  - $SE = \sqrt{SS.RESIDUAL / (N - 2)}$

The normal distribution



NHST: Problems!

- Binary outcome
  - Technically speaking, one must Reject or Retain the Null Hypothesis
  - What if $p = .06$?

NHST: Problems!

- Null “model” is a weak hypothesis
  - Demonstrating that your model does better than NOTHING is not very impressive

Alternatives to NHST

- Effect size
- Confidence intervals
- Model comparison

Alternatives to NHST

- Effect size
  - Correlation coefficient (r)
  - Standardized regression coefficient (B)
  - Model R^2

Alternatives to NHST

- Confidence intervals
  - Sample statistics are “point estimates”
    - Specific to the sample
    - Will vary as a function of sampling error

Alternatives to NHST

- Confidence intervals
  - Instead report “interval estimates”
    - Width of interval is a function of standard error

Alternatives to NHST

- Model comparison
  - Propose multiple models
    - Model A
    - Model B
  - Compare Model R^2

NHST: A closer look

- Important concepts/topics
  - Logic of NHST
  - NHST problems
  - NHST alternatives

Image in slides 7 and 9 was retrieved from
http://www.syque.com/quality_tools/toolbook/Variation/
Image375.gif