Statistics One

Lecture 13

Student's t-test

Two segments

- Overview
- Dependent & Independent t-tests

Overview

- From multiple regression to t-tests?!
  - This is an unusual progression for a stats course
  - So why take this approach?

Overview

- To re-iterate the lesson from Lecture 1
  - Nothing beats a simple controlled experiment!

Overview

- The examples discussed in multiple regression were complicated, considering the limitations placed on the final interpretations, e.g.,
    - The slope for X is B
    - But if you add another X then the slope changes!

Overview

- The examples discussed in multiple regression were complicated, considering the limitations placed on the final interpretations, e.g.,
    - X and Y are correlated
    - M fully mediates the relationship between X and Y
    - BUT correlation does not imply causation
    - So statistical mediation is not the same as true mediation

Overview

- The examples discussed in multiple regression were complicated, considering the limitations placed on the final interpretations, e.g.,
    - X and Y are correlated
    - Add a moderator variable
    - X and Y are not correlated!

Overview

- Let’s assume a simple experimental design
  - Independent variable
    - Vaccine
    - Placebo
  - Dependent variable
    - Rate of polio

Overview

- Two means can be compared using a t-test

Overview

- In this lecture, 4 tests, each compare means
  - z-test
  - t-test (single sample)
  - t-test (dependent)
  - t-test (independent)

Overview

- Why is it called Student's t-test?
- Developed by William Gossett in 1908
  - To monitor the quality of stout beer at the Guiness brewery in Dublin, Ireland
  - Management at Guiness considered their process a secret so they convinced Gossett to publish his work using the pen name "Student"

Overview

- z = (observed – expected) / SE
- t = (observed – expected) / SE

  – SE: Standard error

When to use z and t?

- z
  - When comparing a sample mean to a population mean and the standard deviation of the population is known
- Single sample t
  - When comparing a sample mean to a population mean and the standard deviation of the population is not known
- Dependent samples t
  - When evaluating the difference between two related samples
- Independent samples t
  - When evaluating the difference between two independent samples

observed, expected, and SE

	observed	expected	SE
z	Sample mean	Population mean	SE for a mean
t (single sample)	Sample mean	Population mean	SE for a mean
t (dependent)	Difference	Difference	SE for a difference
t (independent)	Difference	Difference	SE for a difference

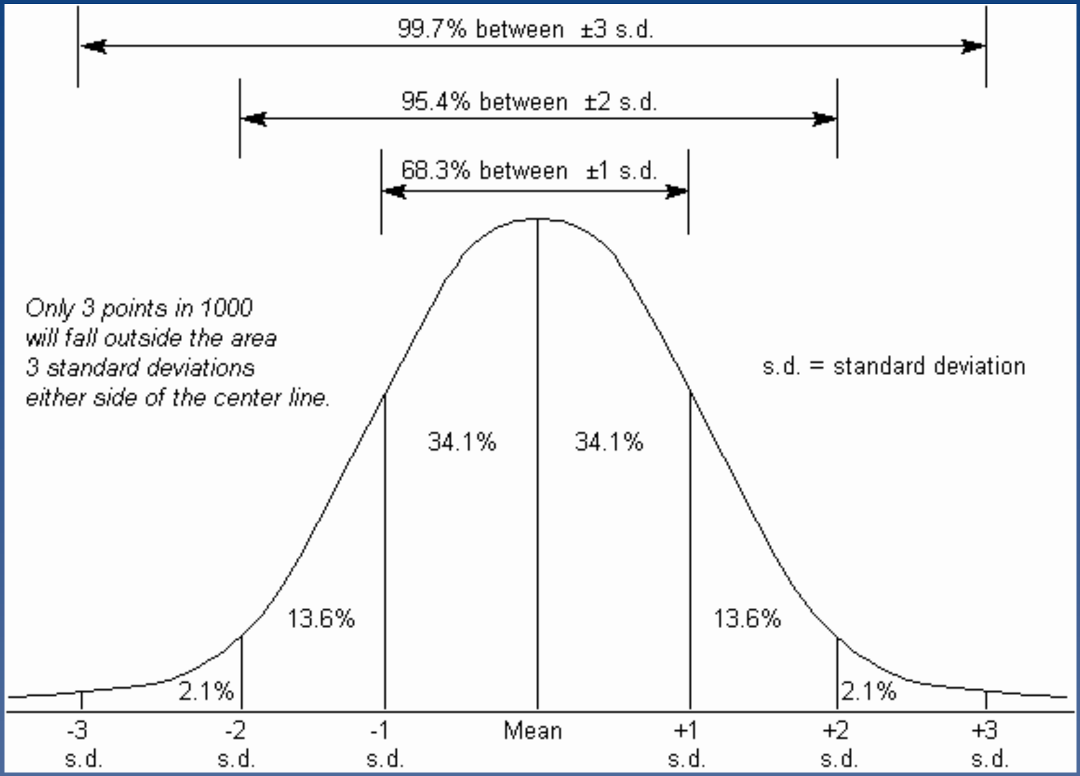
observed, expected, and SE

σ: population standard deviation
μ: population mean
SD: sample standard deviation
M: sample mean
SE: standard error
SE_M: standard error for a mean
SE_{MD}: standard error for a difference (dependent)
$SE_{Difference}$: standard error for a difference (independent)

p values for z and t

- Exact p value depends on:
  - Directional or non-directional test?
  - df
    - Different t-distributions for different sample sizes

z distribution



df

	df
z	NA
t (single sample)	N-1
t (dependent)	N-1
t (independent)	(N1 – 1) + (N2 – 1)

Single sample t

- Compare a sample mean to a population mean

$t = (M - \mu) / SE_M$

$SE^2_M = SD^2/N$

$SE_M = SD/SQRT(N)$

$SD^2 = \Sigma(X - M)^2 / (N - 1) = SS/df = MS$

Single sample t

- Suppose it takes rats just 2 trials to learn how to navigate a maze to receive a food reward

- A researcher surgically lesions part of the brain and then tests the rats in the maze. Is the number of trials to learn the maze significantly more than 2?

Single sample t

Rat #	X		
1	8		
2	6		
3	4		
4	9		
5	3		

Single sample t

Rat #	X	X-M	$(X\text{-}M)^2$
1	8	2	4
2	6	0	0
3	4	-2	4
4	9	3	9
5	3	-3	9
			26

Single sample t

$SD^2 = \Sigma(X - M)^2 / (N - 1) = SS/df = 26 / 4 = 6.5$

$SE^2_M = SD^2/N = 6.5 / 5 = 1.3$

$SE_M = 1.14$

$t = (M - \mu) / SE_M = (6 - 2) / 1.14 = 3.51$

For a directional test with alpha = .05, df = 4, p = .012

$\therefore$ Reject H_0

Single sample t

- Effect size (Cohen's *d*)
  - $d = (M - \mu) / SD = (6 - 2) / 2.55 = 1.57$