May 24, 2017

This time:

Next time:

* no class monday (Holiday) *

- HW #3 due @ canvas by 11:59 pm Friday 26, 2017 (may)

★ Take home final given out last day of class + due 1 week later (there will be extra office hours all that week) ★

→ But discussion sections will take place tue - Fri. next week; monday people need to go to diff. discussion sec. next week.

Today: LN p. L-188 →

when working w/ differences

~~sample~~ pop. mean = μ_d SD: σ_d

~~pop~~ sample mean = $\bar{d}$ SD: S_d

Case Study (Daphenia Longispina)

↳ average age (days) at beginning of reproduction

* analysis of two independent samples

	(T) I.	(C) II.
	7.2	8.8
	7.1	7.5
	⋮	⋮
$\sum^n Y$	52.6	52.9
$\bar{Y}$	7.5143	7.5571
S^2	0.5047	0.4095

→ diff. in practical terms (large ?)

↳ look at it in comparrison to life span

→ (small in practical terms based on their two week life span)

→

* for 2 independent samples, you need to ~~so~~ set up a statistical model for each of the two samples

How ~~a~~ to distinguish the two

I.

you need ~~sob~~ the subscript 1 to distinguish ~~for~~ from experiment data 2

ex μ_1 , σ_1

$\bar{Y}_1$, S_1

$N_1 =$

$M_1 =$

II.

you need a subscript 2 in order to distinguish from experiment data 1

μ_2 , σ_1

$\bar{Y}_2$, S_2

$N_2 =$

$M_2 =$

Base inferential summary on the differences between sample I and Sample II

Paired comparisons formulas

R – 23 → 24

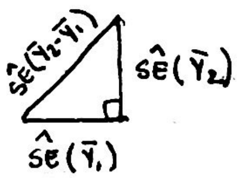
$\hat{SE}(\bar{Y}_1) = 0.2685$

$\hat{SE}(\bar{Y}_2) = 0.2419$

$\boxed{Q}$ $\hat{SE}(\bar{Y}_2 - \bar{Y}_1) = ?$

math fact: uncertainty in a sum of or difference between 2 independent sample means combines like the edges of a right triangle

Pythagorus



$$\hat{SE}(\bar{Y}_2-\bar{Y}_1) = \sqrt{[\hat{SE}(\bar{Y}_1)]^2 + [\hat{SE}(\bar{Y}_2)]^2}$$

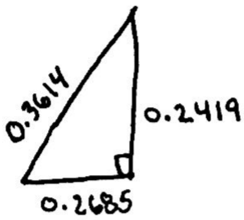
$$= \sqrt{\left(S_1/\sqrt{n_1}\right)^2 + \left(S_2/\sqrt{n_2}\right)^2}$$

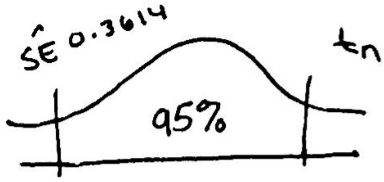
so

$$\hat{SE}(\bar{Y}_2-\bar{Y}_1) = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

R - 26
formula (11)

→ p. L-196





$$\hat{SE}(\bar{Y}_2-\bar{Y}_1) = 0.3614$$

2 - independent samples : dichotomous outcomes

Case Study: Redwood Trees
↳ randomly chosen trees / ~~their~~ + how they are effected by sudden oak disease in comparison to California + ~~Orgen~~ Oregon

* rate of infection in OR is higher than in CA
- for this case study, the means turns from $Y \to P$
p_1 + $\hat{p}_1$

$$\hat{SE}(\hat{p}_1-\hat{p}_2) = \sqrt{(\hat{SE}(\hat{p}_1))^2 + [\hat{SE}(\hat{p}_2)]^2}$$

$$= \sqrt{\left(\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}}\right)^2 + \left(\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\right)^2}$$

so $$\hat{SE}(\hat{p}_1-\hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

95% CI $(\hat{p}_1-\hat{p}_2) \pm 2\hat{SE}(\hat{p}_1-\hat{p}_2)$

$(-3.7\%) \pm \underbrace{2(1.9\%)}_{3.8\%}$