This time: 2 sample problem (quant.)

Next time: 2 sample problem (%)

Read: LN pp. 201 - 213 Today: LN L-186 → R-27 has info on matched pairs

Hwk 3 due next Tue; see web page for wed. disc. sec. next week

Matched Pairs: The difference between pairs is what's important. D = T − C

mean of differences = $\bar{d}$

sd of differences = S_d

types of pairs: control/treatment (note: has 2x subjects as column number b/c 2 groups)

before/after (aka longitudinal design)

2 diff. but comparable variables on an individual (ex: hind & fore leg length)

fun fact: gazelle/deer jumps called stotting, pronking, or pronging

HW 3 #4: it is some kind of paired comparison: held day constant, compared 2 things

not like C/T design or longitudinal design (before/after); is like deer data!

for deer example:

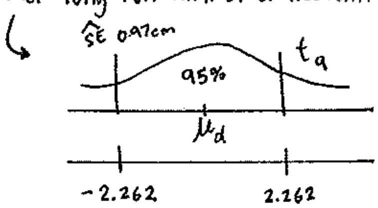
start with sample: the observed ___

then population: extend to broader universe

make inf. summary and fill out 1st two rows with P & S

layout imaginary data set

find long run hist. of $\bar{d}$ accounting for uncertainty in σ_d



$$\widehat{SE}_{IID}(\bar{d}) = \frac{\cancel{\sigma_d}\ S_d}{\sqrt{n}}$$ ← σ still a ? so estimate!

$$= \frac{3.06\text{cm}}{\sqrt{10}} \doteq 0.97\text{cm}$$

$$95\%\ \text{CI for } \mu_d = \bar{d} \pm \underbrace{(2.262)}_{t^{0.95}_{n-1}} \widehat{SE}(\bar{d})$$

$$= 3.30 \pm \underbrace{(2.262)(0.97)}_{2.19}$$

Daphnia example: an analysis of 2 independent samples, <u>not</u> paired!

ask: look at first two rows. Is there any connection between the two numbers? no = indep. sample

need to make a double model to include both groups!

Start as usual w/sample (the observed ___) and variable being measured. Use $\bar{y}$ & s because outcomes are quantitative continuous (use $\bar{y}_1$ and $\bar{y}_2$ to differentiate) continue as usual

note: book should say +0.0428 <u>days</u>

Keep focusing on differences (can subtract either group to pick pos. or neg.)

fill out top 2 rows of inferential summary

} Same but set up one diagram with y_1 and copy with y_2

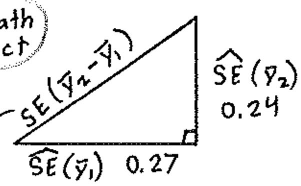
Q1: $SE(\bar{y}_2 - \bar{y}_1) = ?$

Q2: $SE(\bar{y}_2 - \bar{y}_1) \stackrel{?}{=} SE(\bar{y}_2) \; ? \; SE(\bar{x}_1)$

$\bar{y}_1$ unc. is $\widehat{SE}_1 = 0.27$ days

$\bar{y}_2$ unc. is $\widehat{SE}_2 = 0.24$ days

A2: math fact



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

$$= \sqrt{\left(\frac{s_1}{\sqrt{n_1}}\right)^2 + \left(\frac{s_2}{\sqrt{n_2}}\right)^2}$$

~~0.19~~ has to be larger than bigger side!

~~0.92~~ can't have leg longer than both legs together

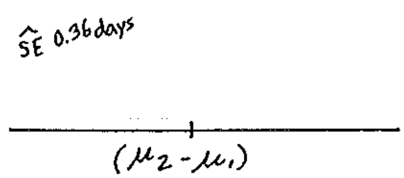
$$\boxed{\begin{matrix}\widehat{SE}(\bar{y}_2 - \bar{y}_1) \\ \text{or} \;\; {}'' \\ \widehat{SE}(\bar{y}_1 - \bar{y}_2)\end{matrix} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$0.27 < \widehat{SE}(\bar{y}_2 - \bar{y}_1) < (0.27 + 0.24)$

can fill 3rd row of inf. summary!

now do long run hist. of $(\bar{y}_2 - \bar{y}_1)$

(n doesn't need to be same in both groups)



$$95\%\ CI = (\bar{y}_2 - \bar{y}_1) \pm \underbrace{(2.179)}_{t^{95\%}_{n_1+n_2-2}} \widehat{SE}(\bar{y}_2 - \bar{y}_1)$$

Pg. 199 has full formula, but you don't need that (JMP knows it!)

2 versions of 2 independent story:

version A: unsure whether pop. SDs are the same (unpooled) = unequal variances for JMP

version B: suspect that they <u>are</u> the same (pooled)