AMS 7-Lecture 5.15.18

THIS TIME: Hypothesis & Significance testing

NEXT TIME: Sample size determination

▷ Read: LN pg L-174 → L-185

Today: LN pg L-162 → 189

HWK #3 is due on canvas at 11:59pm on Sunday 5.27

L-139 Intertidal Crab Case Study

→ Confidence Interval (CI)

L-162 theory was wrong

* Null Hypothesis (H_0)
  - ↳ specific & single number prediction about the population mean that is being observed ($\mu_0 = 24.3°C$) → theory correct
  - ↳ the difference between $\mu_0 = 24.3°$ & $\bar{y} = 25.0°$ is due to unlucky random sampling (logical possibility)

95%

24.1 25.0 25.6

theory Probability 24.3° ∴ wrong b/c not in interval

* Alternative Hypothesis (H_A)
  - ↳ rejects null hypothesis ∴ $\mu \neq 24.3°C$ (2-sided Alt.)
  - ↳ theory wrong & the difference between 24.3 & 25.0 is real

▷ Neyman's (& Fisher's) logic:

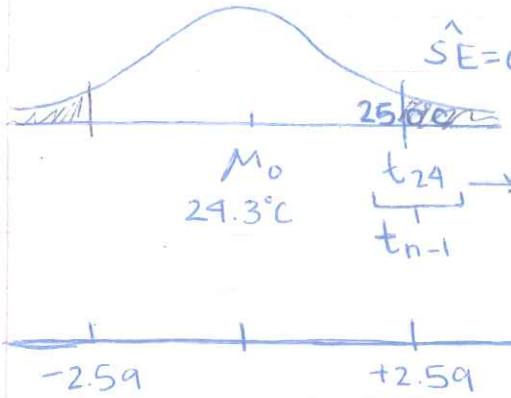
* try null on for size & see if discrepancy between how data came out vs. how data should have come out if null true
  - ↳ If large then favor alt. (reject null)
  - ↳ If not large, favor null (fails to reject null)

L-163 $\bar{y}$ vs. μ_0

→ For model, assume null is true ∴ Put in real value for μ ($\mu_0 = 24.3°C$) → temporary Pretend

→ EV of $\bar{y}$ bar is also equal to μ_0 due to temp. pretend.

The long run Histogram of $\bar{y}$ (if null was true)



* Accounting for uncertainty in σ

$$\frac{\bar{y} - \mu_0}{\hat{SE}(\bar{y})} = \frac{25.0°C - 24.3°C}{0.27°C} = \frac{0.7}{0.27} = +2.59$$

The t-statistic AKA t-ratio

Signal / noise ratio

→ Fisher (P-value)
Area under both tails
Chance if null was true, of getting data as extreme as, or more extreme than, what I got = numerical surprise measure = P-value

L-164

*P-value = measure of surprise of data (null true)
↳ NOT as good as confidence intervals

L-165

Alt #1- $M \neq \mu_0$: two sided alternative — two tailed test / two tailed p-value
Alt #2- $M > \mu_0$: one sided alternative — 1-tailed p-value / test
Alt #3- $M < \mu_0$: one sided alternative — left ↗

MATH FACT: when hyp. testing is done w/ 2-sided alt., its conclusion is identical to that of CI approach

Refer to L-167 for Pitfalls of hyp. testing

L-168 P-Hacking

L-169 We cannot work backwards from P-values but we can do this with CI (All info. we need)

L-171 Blood Pressure (example)
Statsig because too much data! (n = 1000)
↳ not clinically important

▷ Sample size Determination
↳ "How much data should I have?"

R-58 *Case study → arthropods L-174

$$n \doteq \left[t_{n-1}^{(1-\alpha)(2)}\right]^2 s^2 / (\mu_0 - \mu_A)^2$$ ← 2 tailed-test ← Built into JMP

▷ Neyman & Pearson conventions

	H_0 false	H_0 true	
Test says	Good	Type I error (α)	reject H_0
	Type II error (β)	Good	don't reject H_0

* Type I error : False rejection of null
* Type II error : False acceptance of null

L-178 → L-181

• Two-Sample Inferential Problems CRD

Population —like / IID→ exp. subj. → Treatment @random [ ]
→ control [ ]