AMS 5

5-17-05

Dana Morton

This time & Next time: Comparing 2 samples

reading: FPP ch. 27

* HW5 due Thursday May 26 (pgs 96-98)

Pitfalls of Significance Testing (continued)

(2) estimates + give/takes (interval estimation) are typically far more informative than P-values

Example - Case Study 12 - Flextime

H_0 (null): $\mu_0 = 6.3$ days

H_a (alt.): $\mu < 6.3$ days

Pvalue = 0.15% → so highly statistically significant

- By the logic of significance testing, the Null looks wrong.

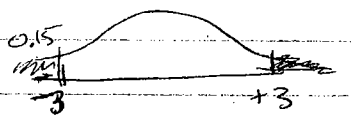
- But, can you work backwards from the Pvalue to look for practical significance?

- μ is < 6.3, but by how much?

Q: Does the Pvalue say anything about practical significance?

(How far backwards can we go?)

hist of Z under H_0



$P = 0.15\% \rightarrow z = -3 = \frac{\text{signal}}{\text{noise}}$

$\frac{\text{signal}}{\text{noise}} = \frac{(\text{obs. } \bar{y}) - (\text{exp. } \bar{y} \text{ under } H_0)}{SE}$

$= \frac{(\ ?\ ) - (6.3)}{?}$

- We know the ratio, but not the numerator and denominator of z.

A: No, because you can't work out the "signal" of the signal/noise ratio.

By contrast...

95% CI for μ

estimate of $\mu = \bar{y} = 5.4$ days

give/take: estimated $SE(\bar{y}) = 0.3$ days

95% CI for $\mu = \bar{y} \pm 2\widehat{SE}(\bar{y}) = (5.4 \pm 2(0.3))$ days

$= (5.4 \pm 0.6)$ days.

~~(((((((((~~ 95% CI for μ.

4.8 5.4 6.0

Practical significance: 5.4 vs 6.3

- a decline of 0.9 days per yr. per employee is large in real-world terms.

The Null value of 6.3 is not inside the 95% CI, so we reject it.

- Intervals answer both questions:
  - statistical significance?
  - practical significance?
- Whereas significance tests only answer:
  - statistical significance?

Case Study 14: Discount Pricing

Comparing 2 samples

Treatment (supposedly causal): discount strategy

Control: standard strategy.

Outcome (effect/response) variable: sales (# of cases of product sold over a 6 week period)

Basic design: controlled experiment.

Design 1 - RCT

: Randomize 120 stores: 60 to (T), 60 to (C)

- randomized controlled trial (RCT)
- good design: (but what is "good")
  1. validity: on avg., if the design were repeated, you would get the right answer
     - RCT is valid
  2. efficiency: how accurately the basic question can be answered.