Feb. 19, 2009

- Sample Size Determination w/ Hypothesis Testing -

H_0 (null) : $\mu = \mu_0$ (theory 1) $\mu_0 = 32$

H_A (alternative) : $\mu \neq \mu_0$ (2 sided alt.), (2 tailed test)

$\left.\begin{array}{l}\mu < \mu_0 \\ \mu > \mu_0\end{array}\right\}$ (1 sided alt.), (1 tailed test)

		Truth: H_0 false	H_0 true
Test says:	reject H_0	good	type 1 error
	don't reject H_0	type 2 error	good

type 1 : false rejection of null
type 2 : false acceptance of null

- If you get a random sample of n observations over and over again, sometimes by chance when H_0 is true, you would (without meaning to) fall into the type 1 error box. Other times when H_0 is false, you might fall into the type 2 error box. We want both of these error probabilities to be small.

- $P(\text{reject } H_0 \mid H_0 \text{ true}) = P(\text{type 1 error}) = \alpha$ = significance level of test
- $P(\text{don't reject } H_0 \mid H_0 \text{ false}) = P(\text{type 2 error}) = \beta$
  - we want this to be small
    → so we want $1-\beta$ to be big = power of the test

- you need at least: $$n = \frac{\left[t_{n-1}^{(1-\alpha)(1 \text{ or } 2)} + t_{n-1}^{(1-\beta)(1)}\right]^2 s^2}{(\mu_0 - \mu_A)^2}$$

- How to choose α, β?

Feb. 19, 2009

convention: $\left(\alpha = \begin{matrix} .05 \\ .01 \end{matrix}\right)$ $\left(\beta = \begin{matrix} .1 \\ .2 \end{matrix}\right)$

Ex: Marine arthropod calcium concentration

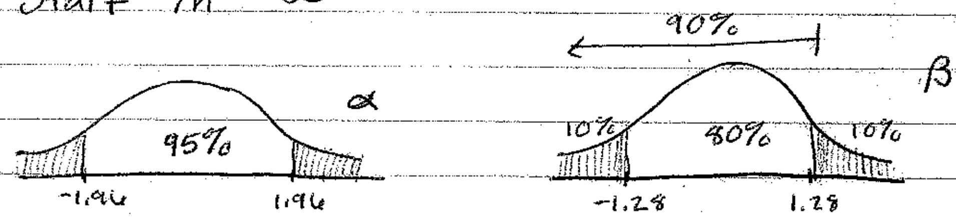
$s = 1.8$, null: $\mu = \mu_0 = 32$, alt: $\mu_A = 31.5$

so, $|\mu_0 - \mu_A| = .5$

2 tailed test: $\alpha = .05$; power $= 1 - \beta = .9$

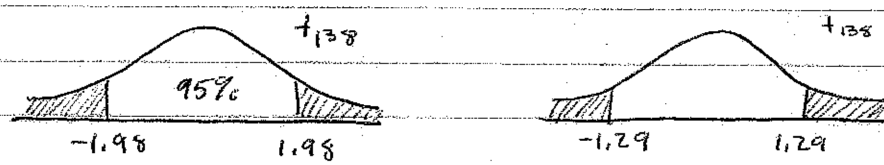
so, $\beta = .1$

$\boxed{1}$ Start w/ $n = \infty$



$$n = \frac{(1.96 + 1.28)^2 (1.8)^2}{(.5)^2} = 136$$

$\boxed{2}$



$$n = \frac{(1.98 + 1.29)^2 (1.8)^2}{(.5)^2} = 139$$

$\boxed{3}$ $n = 139$, so stop

Chapter 5: 2 Sample Inferential Problems

- How does you compare 2 samples?

  2 cases: 1) Paired comparisons

  2) Analysis of 2 independent samples

AR

Feb. 19, 2009

1) Paired Comparisons

- 3 ways for paired data to arise:

I. Matched Pairs

Ex: Cortex weight of rats

- The mean weight of the treatment group was 5.6% larger than the control group's mean weight.
  - → large in practical terms
- mean $\bar{y}$ of difference = $\bar{d}$ = 36.2
  SD of difference = s_d = 31.5

* attention in inference focuses on 1 sample, the column of the differences. This converts a 2 sample problem into a 1 sample problem.

II. Repeated Measures

- ~~same~~ measurement of same variable on n individuals at 2 different points in time.

Ex: Drug to lower blood pressure of hypertensive people

- measure BP before and after drug and calculate the difference
- mean BP before = 174
  mean BP after = 160
  > difference = −14 (A−B)
  - → difference is 8%, which is large in practical terms

III. Measurements on 2 different but comparible variables on n individuals

Ex: Hindleg and Foreleg lengths in deer.

- measure hindleg and foreleg in each deer and calculate difference in length for each deer.
- mean hindleg length = 144.7 cm
  mean foreleg length = 141.4 cm
  difference (mean dif.) = +3.30 cm

Feb. 19, 2009

- hindlegs of deer are only 2.3% longer than forelegs, so not practically significant (probably)

2) Analysis of 2 independent samples

ex: reproductive age in 2 groups of daphnia

Group 1:

Population	Sample	Imaginary
all clones from this genetic cross	the observed daphnia	possible $\bar{y}_1$'s

time — time til reproduction

$N = ?$ (big)

$n = 7$

mean $\bar{y}_1 = 7.5143$ days

SD $s_1 = .7104$

$n = 7$

ex mean = 7.54

7.5143
7.54

$M_1 = \infty$

mean $\mu_1 = ?$

SD $\sigma_1 = ?$

- long run mean: $E_{IID}(\bar{y}_1) = \mu_1 =$
- long run SD: $\hat{SE}_{IID}(\bar{y}_1) = \frac{s_1}{\sqrt{n}} = .2685$ days

Group 2:

Pop.	Sample	Imaginary
same as group 1	same	same

time — time

$n_2 = 7$

mean $\bar{y}_2 = 7.5571$ days

SD $s_2 = .6399$

7.5571
7.54

$M_2 = \infty$

$\mu_2 = ?$

SD $\sigma_2 = ?$

- long run mean $= \mu_2$
- long run SD $= .2419$

Feb 19, 2009

Inferential Summary

Unknown quantity of interest	$(\mu_2 - \mu_1)$ = pop. mean difference in time to reproduce between group 1 & 2
Estimate	$(\bar{y}_2 - \bar{y}_1) = 7.5571 - 7.5143$ $= +0.0428 \approx 62$ minutes
Give or take for $(\bar{y}_2 - \bar{y}_1)$	$\hat{SE}(\bar{y}_2 - \bar{y}_1) = .3414$ days
95% CI for $(\mu_2 - \mu_1)$	$(\bar{y}_2 - \bar{y}_1) \pm (2.179)\, SE(\bar{y}_2 - \bar{y}_1)$