Lecture 29: Chapter 11, Section 2 Categorical & Quantitative Variable Inference in Two-Sample Design

- Sampling Distribution of Difference between Means
- 2-sample *t* Statistic for Hypothesis Test
- Test with Software or by Hand
- 2-sample Confidence Interval
- Pooled 2-sample *t* Procedures

©2011 Brooks/Cole, Cengage Learning

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-4)
  - Displaying and Summarizing (Lectures 5-12)
  - Probability (discussed in Lectures 13-20)
  - Statistical Inference
    - 1 categorical (discussed in Lectures 21-23)
    - 1 quantitative (discussed in Lectures 24-27)
    - cat and quan: paired, 2-sample, several-sample
    - 2 categorical
    - 2 quantitative

Inference Methods for C→Q *(Review)*

- Paired: reduces to 1-sample *t* (already covered)
  - Focused on mean of differences
- Two-Sample: 2-sample *t* (similar to 1-sample *t*)
  - Focus on difference between means
- Several-Sample: need new distribution (*F*)

Display & Summary, 2-Sample Design *(Review)*

- **Display: Side-by-side boxplots:**
  - One boxplot for each categorical group
  - Both share same quantitative scale
- **Summarize:** Compare
  - Five Number Summaries (looking at boxplots)
  - Means and Standard Deviations

Looking Ahead: *Inference for population relationship will focus on means and standard deviations.*

©2011 Brooks/Cole, Cengage Learning

Notation

- Sample Sizes n_1, n_2
- **Sample**
  - Means $\bar{x}_1, \bar{x}_2$
  - Standard deviations s_1, s_2
- **Population**
  - Means μ_1, μ_2
  - Standard deviations σ_1, σ_2

Two-Sample Inference

Inference about $\mu_1 - \mu_2$

- Test: Is it zero? (Suggests categorical explanatory variable does *not* impact quantitative response)
- C.I.: If diff $\neq 0$, how different are pop means?

Estimate $\mu_1 - \mu_2$ with $\bar{x}_1 - \bar{x}_2$ …

(Probability background) As R.V., $\bar{X}_1 - \bar{X}_2$ has

- **Center:** mean (if samples are unbiased) $\mu_1 - \mu_2$
- **Spread:** s.d. (if independent) $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \approx \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
- **Shape:** (if sample means are normal) normal

Two-Sample Inference

Note: claiming that the difference between population means is zero (or not)

$$H_o : \mu_1 - \mu_2 = 0 \text{ vs. } H_a : \mu_1 - \mu_2 \neq 0$$

is equivalent to claiming the population means are equal (or not).

$$H_0 : \mu_1 = \mu_2 \text{ vs. } H_a : \mu_1 \neq \mu_2$$

Two-Sample *t* Statistic

Standardize difference between sample means

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(assuming Ho true)

- Mean 0 if $H_o : \mu_1 - \mu_2 = 0$ is true
- s.d. > 1 but close to 1 if samples are large
- Shape: bell-shaped, symmetric about 0

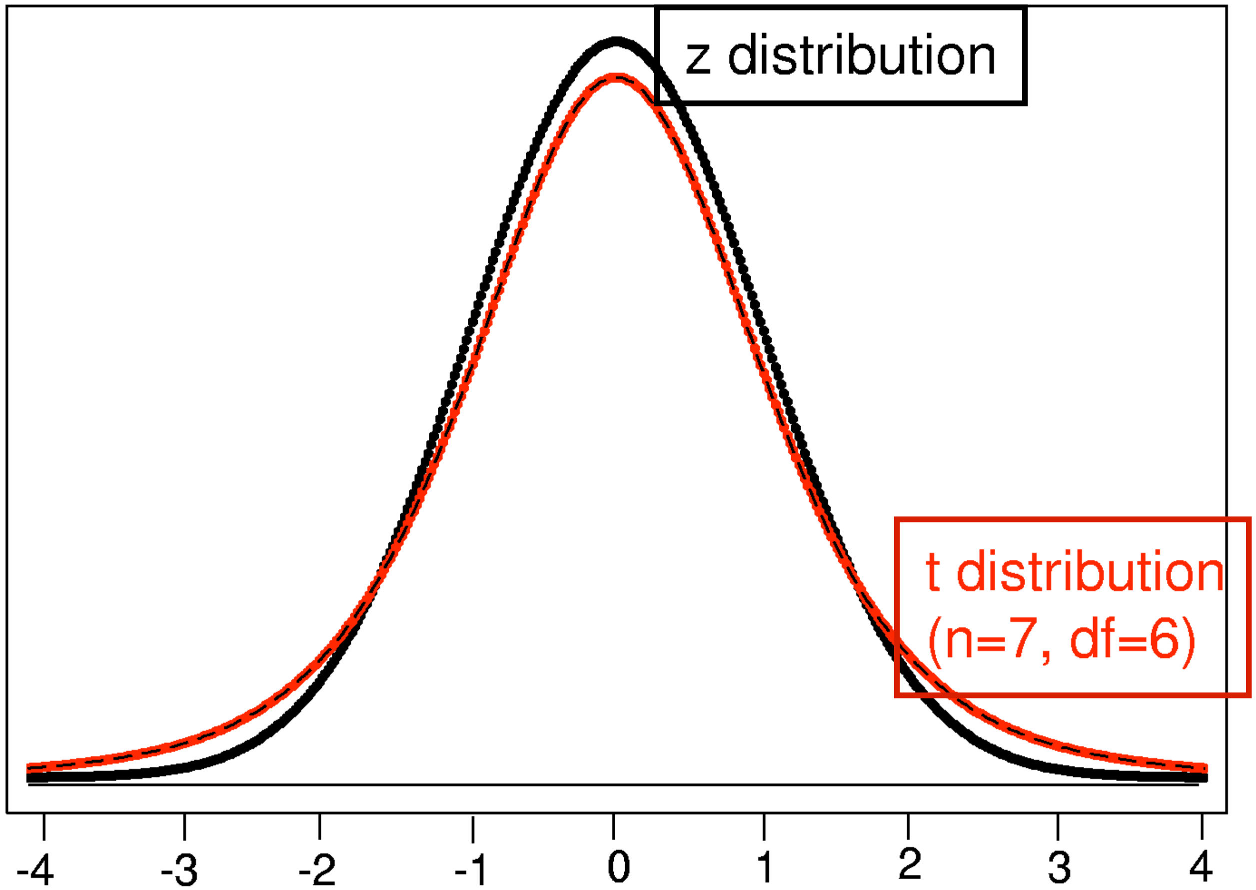
(but not quite the same as 1-sample t)

Shape of Two-Sample *t* Distribution

- ***t*** follows "two-sample *t*" dist *only if sample means are normal*
- 2-sample *t* like 1-sample *t*; df somewhere between smaller $n_i - 1$ and $n_1 + n_2 - 2$
- like *z* if sample sizes are large enough

©2011 Brooks/Cole, Cengage Learning

Shape of Two-Sample *t* Distribution



two-sample t with equal standard deviations and n1=n2=4 same as t with 6 df

What Makes Two-Sample *t* Large

Two-sample *t* statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

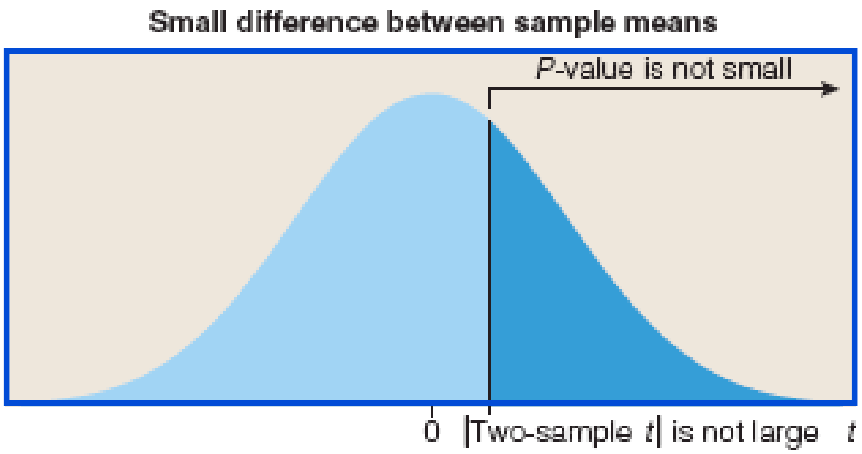
large in absolute value if…

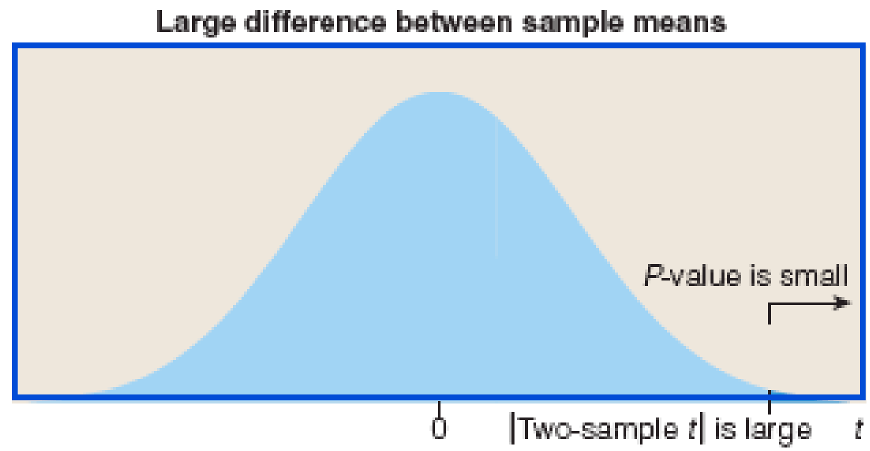
- $\bar{x}_1$ far from $\bar{x}_2$
- Sample sizes $n_1, \; n_2$ large
- Standard deviations $s_1, \; s_2$ small

©2011 Brooks/Cole, Cengage Learning

Example: *Sample Means' Effect on P-Value*

- **Background**: A two-sample t statistic has been computed to test $H_o: \mu_1 - \mu_2 = 0$ vs. $H_a: \mu_1 - \mu_2 > 0$.





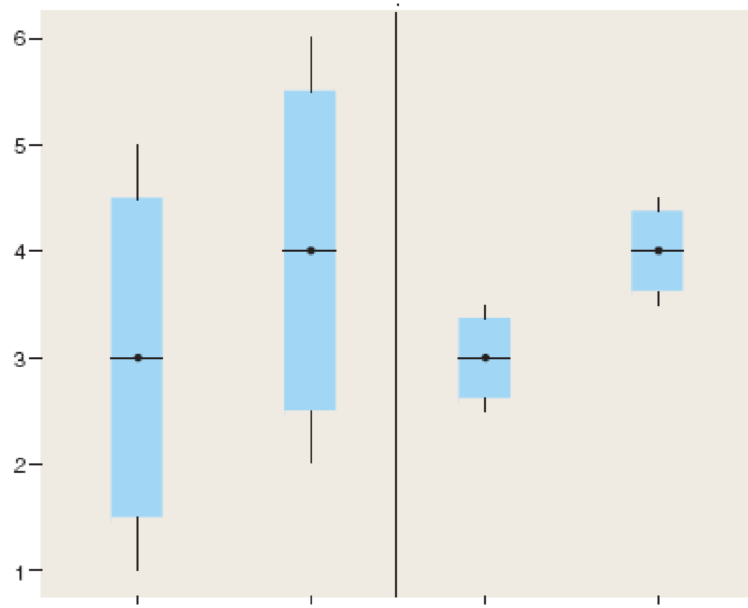
- **Question:** How does the size of the difference between sample means affect the *P*-value, in terms of area under the two-sample *t* curve?
- **Response:** If the difference isn't large, the *P*-value ____________

  As the difference becomes large, the *P*-value ____________

©2011 Brooks/Cole, Cengage Learning

Example: *Sample* *S.D.s' Effect on P-Value*

- **Background**: Boxplots with $\bar{x}_1 = 3, \bar{x}_2 = 4$ could appear as on left or right, depending on s.d.s.

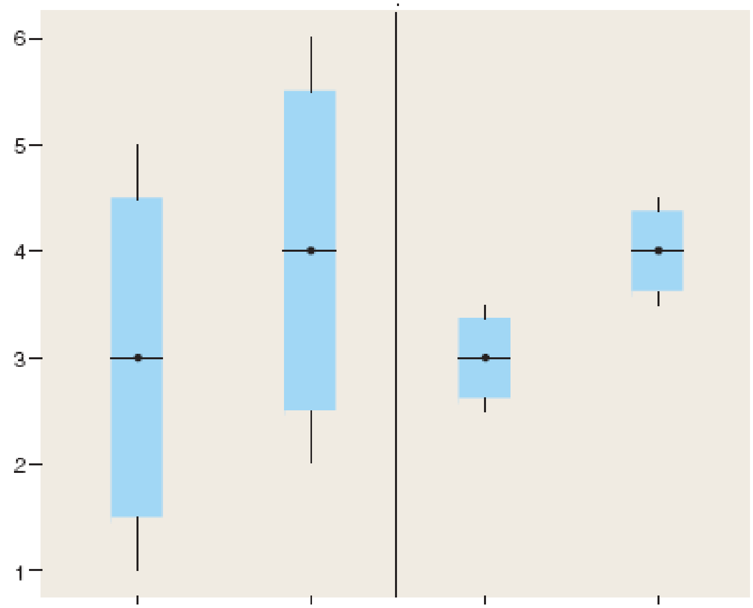


Context: *sample mean monthly pay (in $1000s) for females ($3000) vs. males ($4000).*

- **Question:** For which scenario does the difference between means appear more significant?
- **Response:** Difference between means appears more significant on

Example: *Sample S.D.s' Effect on P-Value*

- **Background**: Boxplots with $\bar{x}_1 = 3, \bar{x}_2 = 4$ could appear as on left or right, depending on s.d.s.



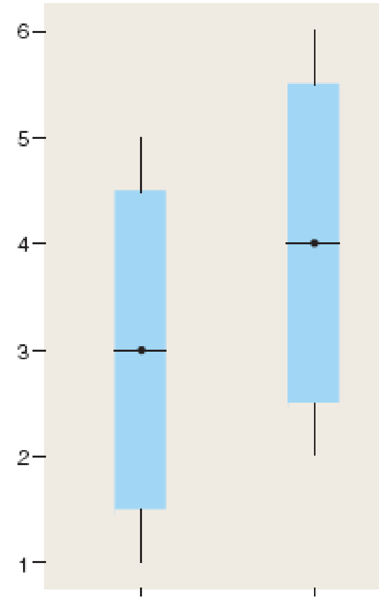
Context: *sample mean monthly pay (in $1000s) for females ($3000) vs. males ($4000).*

- **Question:** For which scenario are we more likely to reject $H_o: \mu_1 - \mu_2 = 0$?
- **Response:** On ______ ________ s.d.s → ________ two-sample *t* → ________ *P*-value → rejecting H_0 is more likely.

©2011 Brooks/Cole, Cengage Learning

Example: *Sample Sizes' Effect on Conclusion*

- **Background**: Boxplot has $\bar{x}_1 = 3, \bar{x}_2 = 4$.



Context: *sample mean monthly pay (in \$1000s) for females (\$3000) vs. males (\$4000).*

- **Question:** Which would provide more evidence to reject H_0 and conclude population means differ: if the sample sizes were each 5 or each 12?
- **Response:** sample size ( ) provides more evidence to reject H_0 .

©2011 Brooks/Cole, Cengage Learning

Practice: 11.14e p.540

Example: *Two-Sample t with Software*

- **Background**: Two-sample *t* procedure output based on survey data of students' age and sex.

```
Two-sample T for Age
Sex            N      Mean     StDev   SE Mean
female       281     20.28      3.34      0.20
male         163     20.53      1.96      0.15
Difference = mu (female) - mu (male  )
Estimate for difference:  -0.250
95% CI for difference: (-0.745, 0.245)
T-Test of difference = 0 (vs not =):
T-Value = -0.99  P-Value = 0.321  DF = 441
```

- **Questions:** Does a student's sex tell us something about age? If so, how do ages of male & female students differ in general?
- **Responses:** *P*-val=0.321 small? Age and sex related? Sample means "close"?_____ Diff. between pop means=0?_____

©2011 Brooks/Cole, Cengage Learning

Practice: 11.9 p.538

Example: *Two-Sample t by Hand*

- **Background**: Students' age and sex summaries:

281 females: mean 20.28 sd 3.34; 163 males: mean 20.53 sd 1.96

- **Question:** Are students' sex and age related?
- **Response:** Testing for relationship same as testing

  H_o : ______ vs. H_a : ______

  Standardized diff between sample mean ages is

  Samples are large → 2-sample t ______ z distribution.

  $|t|$ *is* just under 1→P-val for 2-sided H_a is ______

  Small? ______ Evidence that sex and age are related? ______

©2011 Brooks/Cole, Cengage Learning

Practice: 11.13c-e p.539

Two-Sample Confidence Interval

Confidence interval for diff between population means is

$$(\bar{x}_1 - \bar{x}_2) \pm \text{multiplier}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- Multiplier from two-sample *t* distribution
- Multiplier smaller for lower confidence
- Multiplier smaller for larger df

If samples are large, multiplier for 95% confidence is 2, as for *z* distribution.

©2011 Brooks/Cole, Cengage Learning

Example: *Two-Sample Confidence Interval*

- **Background**: Students' age and sex summaries:

281 females: mean 20.28 sd 3.34; 163 males: mean 20.53 sd 1.96.

- **Question:** What interval should contain the difference between population mean ages?
- **Response:** For this large a sample size, 2-sample *t* multiplier

  We're 95% sure that females are between years younger and years older than males, on average.

  Thus, is a plausible age difference, consistent with test not rejecting Ho.

©2011 Brooks/Cole, Cengage Learning

Practice: 11.18 p.541

Example: *Interpreting Confidence Interval*

- **Background**: A 95% confidence interval for difference between population mean hts, in inches, females minus males, is (-6.4, -5.3).
- **Question:** What does the interval tell us?
- **Response:** We're 95% sure that, on average, females are shorter by ______ to ______ inches. We would reject the null hypothesis of equal population means.

©2011 Brooks/Cole, Cengage Learning

Practice: 11.19d-e p.542

Example: *Changing Order of Subtraction*

- **Background**: A 95% confidence interval for difference between population mean hts, in inches, females minus males, is (-6.4, -5.3).
- **Question:** What would the interval for the difference be, if we took males minus females?
- **Response:** Interval for males minus females would be

Pooled Two-Sample *t* Procedure

If we can assume $\sigma_1 = \sigma_2$, standardized difference between sample means follows an actual *t* distribution with $df = n_1 + n_2 - 2$

- Higher df→narrower C.I., easier to reject H_0
- Some apply Rule of Thumb: use pooled *t* if larger sample s.d. not more than twice smaller.

Example: *Checking Rule for Pooled t*

- **Background**: Consider use of pooled *t* procedure.
- **Question:** Does Rule of Thumb allow use of pooled *t* in each of the following?
  - Male and female ages have sample s.d.s 3.34 and 1.96.
  - 1-bedroom apartment rents downtown and near campus have sample s.d.s $258 and $89.
- **Response:** We check if larger s.d. is more than twice smaller in each case.
  - 3.34 > 2(1.96)? so pooled *t* OK.
  - 258 > 2(89)? so pooled *t* OK.

©2011 Brooks/Cole, Cengage Learning

Practice: 11.20b p.542

Lecture Summary
(Inference for Cat & Quan; Two-Sample)

- Inference for 2-sample design
  - Notation
  - Test
  - Confidence interval
- Sampling distribution of diff between means
- 2-sample *t* statistic (role of diff between sample means, standard deviation sizes, sample sizes)
- Test with software or by hand
- Confidence interval
- Pooled 2-sample *t* procedures