Lecture 19: Chapter 11 Sects. 1 & 2 Categorical & Quantitative Variables Inference in Paired & 2-Sample Design

- Cat→Quan Relations: Hypotheses, 3 Designs
- Inference for Paired Design
- Paired vs. Ordinary, *t* vs. *z*
- 2-Sample *t* Sampling Distribution and Statistic
- 2-Sample *t* test and CI, Ordinary and Pooled

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-3)
  - Displaying and Summarizing (Lectures 3-8)
  - Probability (discussed in Lectures 9-14)
  - Statistical Inference
    - 1 categorical (discussed in Lectures 14-16)
    - 1 quantitative (discussed in Lectures 16-18)
    - cat and quan: paired, 2-sample, several-sample
    - 2 categorical
    - 2 quantitative

Inference for Relationships: Two Approaches

- H_0 and H_a about variables: not related or related
  - Applies to all three C→Q, C→C, Q→Q
- H_0 and H_a about parameters: equality or not
  - C→Q: pop means equal? (mean diff=0? for paired)
  - C→C: pop proportions equal?
  - Q→Q: pop slope equals zero?

Either way, often do *test* before *confidence interval*.

1. Are variables related?
2. If so, quantify: how different are the parameters?

Example: C→Q *Test Relationship or Parameters*

- **Background**: Research question: "For all students at a university, are their Math SATs related to what year they're in?"
- **Question:** How can we formulate this in terms of parameters?
- **Response:**

Looking Ahead: *This is a several-sample design, to be discussed after paired and two-sample.*

  Practice: 11.2b p.525

Design for Cat→Quan Relationship *(Review)*

- Paired
- Two-Sample
- Several-Sample

Looking Ahead: *Inference procedures for population relationship will differ, depending on which of the three designs was used.*

Inference Methods for Cat→Quan Relationship

- Paired: reduces to 1-sample t (already covered)
- Two-Sample: 2-sample t (similar to 1-sample t)
- Several-Sample: need new distribution (F)

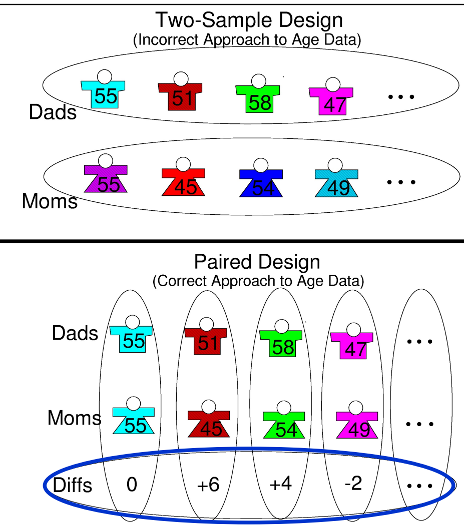
Example: *Paired vs. Two-Sample Data*

- **Background**: Research Question: "Are 'age of parent' and 'sex of parent' related for population of students at a university?"
- **Question:** How can this data set be used to answer the research question?
- **Response:**

DadAge	MomAge
55	55
51	45
58	54
47	49
. . .	. . .

Practice: 11.4a p.526

Paired Data: Incorrect vs. Correct Approach



Example: *Paired vs. Two-Sample Summary*

- **Background**: Research Question: "Are 'age of parent' and 'sex of parent' related for population of students at a university?"
- **Question:** Which output has enough info to do inference?

Descriptive Statistics: DadAge, MomAge

Variable	N	N*	Mean	Median	TrMean	StDev
DadAge	431	15	50.831	50.000	50.491	6.167
MomAge	441	5	48.406	48.000	48.166	5.511

Descriptive Statistics: AgeDiff

Variable	N	N*	Mean	Median	TrMean	StDev
AgeDiff	431	15	2.448	2.000	2.171	3.877

- **Response:**

***Looking Ahead:** We will standardize with the StDev of the differences, which cannot be found from the individual StDevs because of dependence.*

Example: *Consider Summaries in Paired Design*

- **Background**: To see if 'age of parent' and 'sex of parent' are related for population of students at a university, took sampled DadAge minus MomAge.

Descriptive Statistics: AgeDiff

Variable	N	N*	Mean	Median	TrMean	StDev
AgeDiff	431	15	2.448	2.000	2.171	3.877

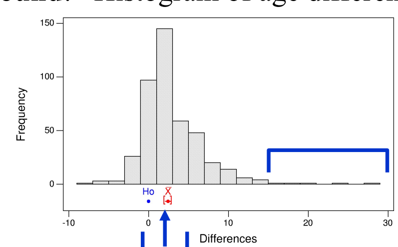
- **Question:** Which parent tended to be older in the sample?
- **Response:**

Example: *Display in Paired Design*

- **Background**: To see if 'age of parent' and 'sex of parent' are related for population of students at a university, took sampled DadAge minus MomAge.
- **Question:** How do we display the data?
- **Response:**

Example: *Display in Paired Design*

- **Background**: Histogram of age differences:



- **Question:** What does the histogram show?
- **Response:** Age differences have
  - Center: around _____ (dads tend to be about ____ yrs older)
  - Spread: most diffs within _________ yrs or mean)
  - Shape: ______________ (a few dads much older than wife)

Notation in Paired Study

- Differences have
  - Sample mean $\bar{x}_d$
  - Population mean μ_d
  - Sample standard deviation s_d
  - Population standard deviation σ_d

Test Statistic in Paired Study

- Start with ordinary 1-sample statistic $t = \frac{\bar{x}-\mu_o}{s/\sqrt{n}}$
- Substitute $\bar{x}_d$, s_d for ordinary summaries $\bar{x}$, s
- Substitute 0 for μ_o (H_0 will claim $\mu_d = 0$)
- Result is paired *t* statistic: $t = \frac{\bar{x}_d - 0}{s_d/\sqrt{n}}$

Example: *Paired t Test*

- **Background**: Paired test on students' parents' ages:

```
Paired T for DadAge - MomAge
                 N      Mean     StDev   SE Mean
DadAge         431    50.831     6.167     0.297
MomAge         431    48.383     5.258     0.253
Difference     431     2.448     3.877     0.187
95% CI for mean difference: (2.081, 2.815)
T-Test of mean difference = 0 (vs not = 0): T-Value = 13.11  P-Value = 0.000
```

- **Question:** What does output tell about formal test?
- **Response:** Testing
  - Unbiased? ____ *n*=431 large? ____ Pop≥10(431)? ____
  - $\bar{x}_d$= ________, *t* = ________ Large? ______
  - *P*-value = ______ Small? _____
  - Conclude pop mean diff =0? ____Sex and age related? ____

 Practice: 11.6d-f p.527

Example: *One- or Two-Sided H_a in Paired Test*

- **Background**: Paired test on students' parents' ages:

```
Paired T for DadAge - MomAge
                 N      Mean     StDev   SE Mean
DadAge         431    50.831     6.167     0.297
MomAge         431    48.383     5.258     0.253
Difference     431     2.448     3.877     0.187
95% CI for mean difference: (2.081, 2.815)
T-Test of mean difference = 0 (vs not = 0): T-Value = 13.11  P-Value = 0.000
```

- **Response:** Replace $H_a : \mu_d \neq 0$ with __________
  - *P*-value would be _________________
  - Conclude fathers in general are older? ______

 Practice: 11.45c p.572

Example: *Paired vs. Ordinary t vs. z*

- **Background**: Paired test on 431 students' parents' ages resulted in paired t-statistic +13.11.
- **Question:** What does this tell us about the P-value?
- **Response:**

Paired t same as ordinary t distribution

→ Ordinary t basically same as z for large n

→13.11 sds above mean unusual? ____ →P-val = ________

→Evidence that mean age diff is non-zero in pop.? ____

Note: for extreme t statistics, software not needed to estimate P-value.

Practice: 11.58b p.576

C.I. for Mean: σ Unknown *(Review)*

95% confidence interval for μ is

$$\bar{x} \pm \text{multiplier}\left(\frac{s}{\sqrt{n}}\right)$$

- multiplier from t distribution with n-1 degrees of freedom (df)
- multiplier at least 2, closer to 3 for *very* small n

Confidence Interval in Paired Design

Confidence interval for μ_d is

$$\bar{x}_d \pm \text{multiplier}\frac{s_d}{\sqrt{n}}$$

- Multiplier from t distribution with n-1 df
- Multiplier smaller for lower confidence
- Multiplier smaller for larger df

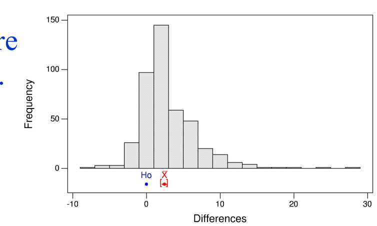
If n is small, diffs need to be approx. normal.

(Same guidelines as for 1-sample t)

Example: *Paired Confidence Interval*

- **Background**: Sample of 431 students' parents' age differences have mean +2.45, s.d. 3.88.
- **Question:** What is a 95% confidence interval for population mean age difference?
- **Response:** Since n is so large, t multiplier __________ for 95% confidence. (Also, skewed hist. OK.)

Pretty sure population of fathers are older by about ____ to ____ years.



Example: *Paired Confidence Interval by Hand*

- **Background**: Mileage differences for 5 cars, city minus highway, had mean -5.40, s.d. 1.95.
- **Question:** What else is needed to set up a 95% confidence interval by hand for population mean difference?
- **Response:** Need ____________________________ (obtained from table before software was available)
  Interval is

Note: *n* very small→ *t* multiplier closer to 3 than to 2.

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-3)
  - Displaying and Summarizing (Lectures 3-8)
  - Probability (discussed in Lectures 9-14)
  - Statistical Inference
    - 1 categorical (discussed in Lectures 14-16)
    - 1 quantitative (discussed in Lectures 16-18)
    - 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.*

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?

Looking Back: *Estimated* μ *with* $\bar{x}$; *established the center, spread, and shape of* $\bar{X}$ *relative to* μ .

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

(Probability background) as R.V., $\bar{X}_1 - \bar{X}_2$ has what center, spread and shape?

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)

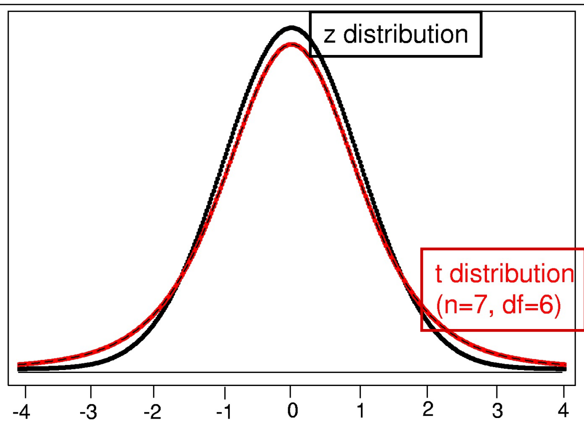
- Mean 0 if $H_o : \mu_1 - \mu_2 = 0$ is true
- s.d. > 1 but close to 1 if samples are large
- Shape: bell-shaped, symmetric about 0

  (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

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 One-Sample *t* Large *(Review)*

One-sample *t* statistic

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$$

t large in absolute value if…

- Sample mean far from μ_o
- Sample size *n* large
- Standard deviation *s* small

What Makes Two-Sample *t* Large

Two-sample *t* statistic

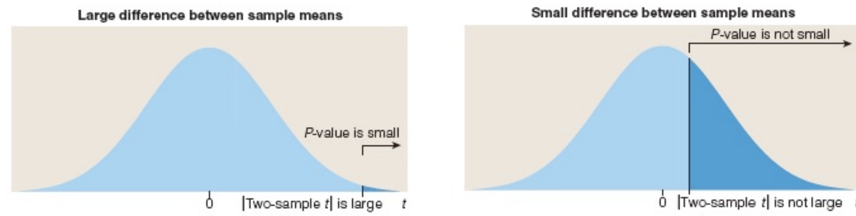
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

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

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. $H_a: \mu_1 - \mu_2 > 0$.



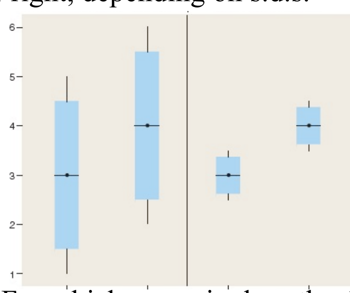
- **Question:** How does the size of the difference between sample means affect the *P*-value, in terms of area under the two-sample *t* curve?
- **Response:** If the difference isn't large, the *P*-value ________

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

 Practice: 11.14a p.540

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.



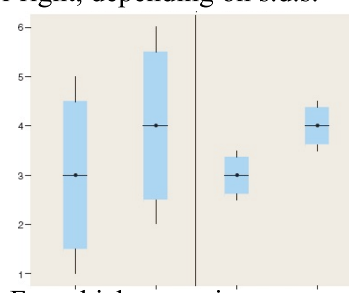
Context: *sample mean monthly pay (in \$1000s) for females (\$3000) vs. males (\$4000).*

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

 Practice: 11.14b p.540

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.



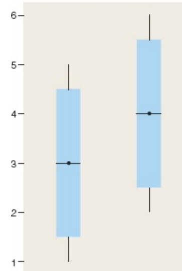
Context: *sample mean monthly pay (in \$1000s) for females (\$3000) vs. males (\$4000).*

- **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.

 Practice: 11.14b p.540

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.

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?________

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? ____

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.

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.

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.

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.

Lecture Summary
(Inference for Cat→Quan; Paired)

- Inference for relationships
  - Focus on variables
  - Focus on parameters
- cat→quan relationship: paired, 2- or several-sample
- Inference for paired design
  - Output
  - Display
  - Notation
  - Test statistic
  - Form of alternative
- Paired t vs. ordinary t vs. z
- Paired confidence interval vs. hypothesis test

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