Inferences About Differences Between Means
Edpsy 580

Carolyn J. Anderson
Department of Educational Psychology
University of Illinois at Urbana-Champaign
Outline

- General “template”
- Two Independent populations
  - Large sample
  - Any size sample
- Violations of Assumptions
  - Normality
  - Homogeneity of variance
  - Independence of observations
- SAS and t-tests
- Two Dependent Populations
- Power of t-test.
General Template

- **Statistical Hypotheses**: null & alternative

- **Select $\alpha$ level**: before you look at data.

- **Assumptions**: think critically

- **Test statistic**: choose appropriate one

- **Sampling distribution of test statistic**: needed for $p$-values, CI, power

- **Decision**: reject, retain or uncertain

- **Interpretation**: in context of research question(s)
Statistical Hypotheses

- Hypotheses of the form

\[ H_0 : \mu_1 = \mu_2 \]  “means are the same”

or equivalently,

\[ H_0 : \mu_1 - \mu_2 = 0 \]  “no difference between means”

- Different cases:
  - Sample size.
  - Assumptions about the population distribution.
Test Statistic

- General form of (many) test statistics:

\[
\text{test statistic} = \frac{(\text{sample statistic}) - (\text{null hypothesis value})}{\text{standard deviation of sample statistic}}
\]

- “t–ratio”

- “t–statistic”
Two Independent Populations

Independent or Dependent?

- Are the mean reading scores of high school seniors attending public schools the same as those students attending private schools?

- A group of patients suffering from high blood pressure are randomly assigned to either a treatment group or control group. Those assigned to the treatment group get a new drug while those assigned to the control receive the standard drug. Is the mean blood pressure lower with the new drug?
Independent vs Dependent Populations

Independent or Dependent?

- A sample of students are given a pre-test and then receive instruction on using SAS. To assess whether the instruction was effective, after completing the class, students take a second test. Are the test scores after the class better than those before the class.

- Students are randomly assigned to one of 2 instructional methods (eg, book & exercises or computer/multi-media presentation & practice problems). We want to assess whether the methods differ in terms of how much students learn. All students are given a pre-test and a post-test and are interested in whether the mean change scores are different for the two instructional methods.
Two Independent Populations

For both large and small sample cases . . .

Statistical Hypotheses:

\[ H_0 : \text{population 1 mean} = \text{population 2 mean} \]

\[ H_0 : \mu_1 = \mu_2 \]

Or equivalently,

\[ H_0 : \text{there's no difference between means} \]

\[ H_0 : \mu_1 - \mu_2 = 0 \]

Also need to specify the alternative hypotheses.
Two Independent Populations

Assumptions for LARGE samples:

- Observations are *independent*
  - Between populations, i.e., $Y_{i1}$ and $Y_{i2}$.
  - Within populations, i.e. $Y_{ij}$ and $Y_{kj}$ for $i \neq k$ and both $j = 1, 2$.

- Since sizes $n_1$ and $n_2$ are both *large*, by the CLT
  - The sampling distribution of $\bar{Y}_1 \approx \mathcal{N}(\mu_1, \sigma_1^2/n_1)$
  - The sampling distribution of $\bar{Y}_2 \approx \mathcal{N}(\mu_2, \sigma_2^2/n_2)$
General Template

Two Independent Populations

LARGE samples

Test Statistic:

- The parameters of interest: \((\mu_1 - \mu_2)\).

- Sample statistic \((\bar{Y}_1 - \bar{Y}_2)\) is an estimate of \((\mu_1 - \mu_2)\).

- If \(H_o\) is true, then \((\bar{Y}_1 - \bar{Y}_2)\) should be “close” to 0.

- Test statistic,

\[
t = \frac{(\bar{Y}_1 - \bar{Y}_2) - 0}{s(\bar{Y}_1 - \bar{Y}_2)} = \frac{(\bar{Y}_1 - \bar{Y}_2)}{s(\bar{Y}_1 - \bar{Y}_2)}
\]
Test Statistic

LARGE samples: What does \( s(\bar{Y}_1 - \bar{Y}_2) \) equal?

- **FACT**: If \( \bar{Y}_1 \) and \( \bar{Y}_2 \) are independent, then

  \[
  \text{variance of a sum} = \text{sum of variances}
  \]

  \[
  \text{var}(\bar{Y}_1 - \bar{Y}_2) = \sigma^2_{Y_1 - Y_2} = \sigma^2_{Y_1} + \sigma^2_{Y_2}
  \]

- Unbiased estimators of \( \sigma^2_1 \) and \( \sigma^2_2 \),

  \[
  s^2_1 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (Y_{i1} - \bar{Y}_1)^2 \\
  s^2_2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_{i2} - \bar{Y}_2)^2
  \]
Estimating $s(\bar{Y}_1 - \bar{Y}_2)$

LARGE samples:

$$\text{var}(\bar{Y}_1 - \bar{Y}_2) = \sigma^2(\bar{Y}_1 - \bar{Y}_2) = \sigma^2_{\bar{Y}_1} + \sigma^2_{\bar{Y}_2}$$

- $s^2_1/n_1$ is an unbiased estimator of $\sigma^2_{\bar{Y}_1}$
- $s^2_2/n_2$ is an unbiased estimator of $\sigma^2_{\bar{Y}_2}$

Unbiased Estimator of $\sigma^2(\bar{Y}_1 - \bar{Y}_2)$

$$\left(\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}\right) = s^2_{\bar{Y}_1 - \bar{Y}_2}.$$
Test Statistic

LARGE samples:

- Estimator of $\sigma (\bar{Y}_1 - \bar{Y}_2)$

\[
S \bar{Y}_1 - \bar{Y}_2 = \sqrt{\left( \frac{s^2_1}{n_1} + \frac{s^2_2}{n_2} \right)}
\]

- Test statistic

\[
t = \frac{(\bar{Y}_1 - \bar{Y}_2) - 0}{S \bar{Y}_1 - \bar{Y}_2} = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\left( \frac{s^2_1}{n_1} + \frac{s^2_2}{n_2} \right)}}
\]
Sampling Distribution of Test Statistic

LARGE samples:

- When $H_o$ is true, $n_1$ and $n_2$ are large, (and our assumptions are valid), then
  
  test statistic: $t \approx \mathcal{N}(0, 1)$

- Alternatively by the CLT,
  
  - $\bar{Y}_1 \approx$ normal for large $n_1$
  - $\bar{Y}_2 \approx$ normal for large $n_2$
  
  $\Rightarrow (\bar{Y}_1 - \bar{Y}_2) \approx$ normal for large $n_1$ and $n_2$. 
Example: 2 Independent Populations

LARGE samples.

HSB data: Test whether mean Math T-scores of High School seniors attending public and private schools are the same.

<table>
<thead>
<tr>
<th>Public Schools</th>
<th>Private Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_1 = 506 )</td>
<td>( n_2 = 94 )</td>
</tr>
<tr>
<td>( \bar{Y}_1 = 51.45 )</td>
<td>( \bar{Y}_2 = 54.01 )</td>
</tr>
<tr>
<td>( s^2_1 = 91.74 )</td>
<td>( s^2_2 = 67.16 )</td>
</tr>
</tbody>
</table>
Example: 2 Independent Populations

- **Statistical Hypotheses:**
  \[ H_0 : \mu_1 - \mu_2 = 0 \quad \text{versus} \quad H_1 : \mu_1 - \mu_2 \neq 0 \]

- **Assumptions:**
  - Public: \( Y_{11}, Y_{21}, \ldots, Y_{n_1 1} \) are independent and come from a population with mean \( \mu_1 \) and variance \( \sigma_1^2 \).
  - Private: \( Y_{12}, Y_{22}, \ldots, Y_{n_2 2} \) are independent and come from a population with mean \( \mu_2 \) and variance \( \sigma_2^2 \).
  - \( Y_{i1} \) and \( Y_{i2} \) are independent between populations.
Example: 2 Independent Populations

- **Test statistic:**

\[ t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \]

\[ = \frac{51.45 - 54.01}{\sqrt{\frac{91.74}{506} + \frac{67.16}{94}}} = \frac{-2.56}{.95} = -2.709 \]

- **Sampling distribution:** If \( H_0 \) is true and since \( n_1 \) and \( n_2 \) are both large, the test statistic is approximately distributed as a standard normal random variable.

- **Decision:** Since \( z = 1.960 \) is the critical value for a two-tailed test with \( \alpha = .05 \) (or \( p\)-value < .05 or \( p\)-value = .007), reject \( H_0 \).
Example: 2 Independent Populations

- **Interpretation/conclusion:** The data support the hypothesis that students attending public and private schools differ in terms of mean Math achievement test scores.

- **The 95% confidence interval for the difference,**

\[
\begin{align*}
&\left( \bar{Y}_1 - \bar{Y}_2 \right) \pm \frac{z_{\alpha/2}}{\hat{\sigma}} \left( \bar{Y}_1 - \bar{Y}_2 \right) \\
&-2.56 \pm 1.96(.95) \rightarrow (-4.42, -.70)
\end{align*}
\]

- **Do public or private schools have higher means on math achievement test scores?**
Any Size Sample

- **Hypotheses:**
  \[ H_o : \mu_1 = \mu_2 \quad \text{or} \quad \mu_1 - \mu_2 = 0 \]
  \[ H_a : \mu_1 \neq \mu_2 \quad \text{or} \quad \mu_1 - \mu_2 \neq 0 \] (< or >)

- **Assumptions:**
  - **Independent observations between and within populations.**
  - Population distributions of variables are **Normal**.
  - **Equal variances** or “homogeneous variances”.

\[ Y_{i1} \sim \mathcal{N}(\mu_1, \sigma^2) \quad \text{i.i.d.} \]
\[ Y_{i2} \sim \mathcal{N}(\mu_2, \sigma^2) \quad \text{i.i.d.} \]
Any Size Sample — Two Independent Populations

- **Statistical Hypotheses:**
  
  \[
  H_o : \mu_1 = \mu_2 \quad \text{or} \quad \mu_1 - \mu_2 = 0 \\
  H_a : \mu_1 \neq \mu_2 \quad \text{or} \quad \mu_1 - \mu_2 \neq 0 \quad (< \text{ or } >)
  \]

- **Assumptions:**
  
  - Independence between and within populations.
  - Population distributions of variables are **Normal**.
  - Equal variances or “homogeneous variances”, i.e.,
    
    \[
    Y_{i1} \sim \mathcal{N}(\mu_1, \sigma^2) \quad \text{i.i.d.} \\
    Y_{i2} \sim \mathcal{N}(\mu_2, \sigma^2) \quad \text{i.i.d.}
    \]
Any Size Sample

**Test Statistic:**

\[ t = \frac{\bar{Y}_1 - \bar{Y}_2}{s_{\bar{Y}_1 - \bar{Y}_2}} \]

**What does \( s_{\bar{Y}_1 - \bar{Y}_2} \) equal?**

- Variance of sum = sum of variances

\[ \sigma^2_{\bar{Y}_1 - \bar{Y}_2} = \frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2} \]

\[ = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} \]

- \( s^2_1 \) and \( s^2_2 \) are both estimators of \( \sigma^2 \)
- “Pool” information to get a better estimate.
Pooled (Within) Variance

\[
\begin{align*}
\bar{s}_{pool}^2 &= \frac{\sum_{i=1}^{n_1} (Y_{i1} - \bar{Y}_1)^2 + \sum_{i=1}^{n_2} (Y_{i2} - \bar{Y}_2)^2}{n_1 + n_2 - 2} \\
&= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \\
&= s_w^2 \quad \text{“within groups”}
\end{align*}
\]

Expected value of \(s_{pool}^2\): \(E(s_{pool}^2) = E(s_w^2) = \sigma^2\)

The estimator of \(\sigma^2_{\bar{Y}_1 - \bar{Y}_2}\),

\[
s_{\bar{Y}_1 - \bar{Y}_2}^2 = \frac{s_w^2}{n_1} + \frac{s_w^2}{n_2} = s_w^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)
\]
Any Size Sample

- So
\[
\sqrt{s_w^2/n_1 + s_w^2/n_2} = s_w \sqrt{\left( \frac{1}{n_1} + \frac{1}{n_2} \right)}
\]

- The test statistic,
\[
t = \frac{\bar{Y}_1 - \bar{Y}_2}{s_w \sqrt{\left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}
\]

- If \( H_o \) is true and assumptions are valid, then the Sampling Distribution of the test statistic Student’s \( t \) with degrees of freedom
\[
\nu = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2
\]
A random sample of \( n_1 = 10 \) students from private schools and \( n_2 = 10 \) from public schools from HSB data.

<table>
<thead>
<tr>
<th></th>
<th>Public Schools</th>
<th>Private Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size</td>
<td>( n_1 = 10 )</td>
<td>( n_2 = 10 )</td>
</tr>
<tr>
<td>Sample Mean</td>
<td>( \bar{Y}_1 = 53.99 )</td>
<td>( \bar{Y}_2 = 55.56 )</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>( s_1^2 = 119.15 )</td>
<td>( s_2^2 = 120.116 )</td>
</tr>
</tbody>
</table>
Example: Small Sample

- **Statistical hypotheses:**

  \[ H_0 : \mu_1 - \mu_2 = 0 \quad \text{versus} \quad H_a : \mu_1 - \mu_2 \neq 0 \]

- **Assumptions:**
  - Observations from public & private schools are independent within school type and between types.
  - Observations come from populations that are normal.
  - The variance of math test scores at public schools equals the variance of scores at private schools.

Statistical hypotheses:

\[ H_0 : \mu_1 - \mu_2 = 0 \quad \text{versus} \quad H_a : \mu_1 - \mu_2 \neq 0 \]

Assumptions:

- Observations from public & private schools are independent within school type and between types.
- Observations come from populations that are normal.
- The variance of math test scores at public schools equals the variance of scores at private schools.
Example: Small Sample

- **Test statistic:**

\[ s^2_w = \frac{(10 - 1)(119.15) + (10 - 1)(120.116)}{(10 + 10 - 2)} = 119.63 \]

\[ t = \frac{53.99 - 55.56}{\sqrt{119.63(1/10 + 1/10)}} = \frac{-1.57}{4.89} = -.321 \]

- **Sampling Distribution** of \( t \) is Students \( t \) distribution with degrees of freedom \( \nu = 10 + 10 - 2 = 18 \).

- **Decision:** Critical t-value for \( \alpha = .05 \) and \( \nu = 18 \) is 2.101 (or \( p\text{-value} = .75 \)); retain \( H_o \).

- **Conclusion:** The data do not support the conclusion that math scores for students attending public and private schools differ.
Example: Small Sample

Students $t$-distribution with df = 18

Reject $H_0$

Retain $H_0$

Reject $H_0$

$\alpha = 0.05/2 = 0.025$

$\alpha = 0.05/2 = 0.025$
Large vs Any Size Sample Test

Assumptions: There are fewer assumptions for large sample test.

<table>
<thead>
<tr>
<th>Large Sample</th>
<th>Any size (small) sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Independence between and within groups</td>
<td>Independence between and within groups</td>
</tr>
<tr>
<td>2 Sampling distribution of <em>means</em> is approximately normal (by C.L.T.)</td>
<td>Population distributions of variable are normal</td>
</tr>
<tr>
<td>3 —</td>
<td>homogeneous variance ( \sigma_1^2 = \sigma_2^2 )</td>
</tr>
</tbody>
</table>
Sampling distribution of the test statistic differs.

- For large sample $\rightarrow$ standard normal.
- For any size sample $\rightarrow$ Students’s $t$ distribution.

Our example...
Any Size using Large Sample

Any size sample size method with the full sample (i.e, \( n = 600 \)),

\[
s_w^2 = s_{\text{pool}}^2 = \frac{505(91.73719) + 93(67.15709)}{506 + 94 - 2} = 87.9145
\]

\[
s_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{87.9145 \left( \frac{1}{506} + \frac{1}{94} \right)} = 1.05309
\]

\[
t = \frac{-2.56}{1.053} = -2.431 \quad \text{with} \quad \nu = 600 - 2 = 598
\]

Large sample method \( \rightarrow s_{\bar{Y}_1 - \bar{Y}_2} = .95 \) and \( t = -2.709 \).
SAS and $t$ tests

**SAS/ASSIST:**
- Input data into working memory.
- One the main toolbar: Solutions
  - ASSIST
    - Data Analysis
      - ANOVA
        - $t$-tests
          - Compare two group means:
            - Active Data Set (select one).
            - Dependent Variable.
            - Classification (the group or population).
- RUN
- Input data into working memory.
- One the main toolbar → Solutions → Analysis → Analyst

- File → Open by SAS name & select your data set
- On (Analyst) toolbar → Statistics → Hypothesis tests → Two-sample t-test for means → fill in boxes → OK
PROC TTEST data = HSB;
   CLASS sctyp;
   VAR math;
RUN;
Violations of Assumptions

- Mostly a concern with small samples (any size sample method).

- Effects and how to deal with it.

- The assumptions we’ll discuss:
  - Normality of sampling distribution of means
  - Homogeneity of variance
  - Independence within and between groups.
Violations of Normality

**Effects:**

- $t$-test is “robust”. The stated or “nominal” $\alpha$-level is very similar to the actual $\alpha$-level, so long as $n_1$ and $n_2$ are “moderately” large.

- Departure from normality can have more of an effect with 1-tailed test than a 2-tailed test.

**Remedies:**

- Get larger samples (so you can use any size test).

- Perform non-parametric test where you don’t make the normality assumptions (i.e., Mann-Whitney).
Homogeneity of Variance

Effects for three cases where $\sigma_1^2 \neq \sigma_2^2$:

- When $n_1 = n_2$ and $\sigma_1^2 > \sigma_2^2$, the $t$-test is robust.

- When $n_1 > n_2$ and $\sigma_1^2 > \sigma_2^2$, the $t$-test is “conservative” with respect to type I errors. (i.e., you make fewer type I errors than you would expect).

- When $n_1 < n_2$ and $\sigma_1^2 > \sigma_2^2$, the $t$-test is “liberal” with respect to type I errors. (i.e., you make more type I errors than you would expect).

- When $\sigma_1^2 \neq \sigma_2^2$:
  - The greater the difference between $\sigma_1^2$ and $\sigma_2^2$, the further off the $\alpha$-levels.
  - The greater the difference between $n_1$ and $n_2$, the further off the $\alpha$-levels.
Homogeneity of Variance

Remedies

- Take larger samples (& do large sample test).
- Use a non-parameteric test (i.e, Mann-Whitney).
- Perform a “Quasi” $t$-test or “Welch’s” $t$ test.

In SAS, it’s called “approximate $t$ statistic for unequal variances”
Quasi $t$-test or Welch’s $t$-test

is same as the $t$-test (for any size), except

- When computing the test statistic, use

$$
S_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{s^2_{Y_1} + s^2_{Y_2}} = \sqrt{\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}}
$$

- Adjust the degrees of freedom

$$
\nu = \frac{s^2_{Y_1}}{(s^2_{Y_1})^2 \nu_1} + \frac{s^2_{Y_2}}{(s^2_{Y_2})^2 \nu_2}
$$

to nearest integer.
Quasi $t$-test or Welch’s $t$-test

- In SAS/PROC TTEST, this corresponds to $df$ for “Satterthwaite” or $df$ and $p$-value for unequal variances.

- The test statistic is compared to Student’s $t$ distribution with the degrees of freedom give above.

- SAS commands:

```sas
PROC TTEST data = HSB cochran;
   CLASS sctyp;
   VAR math;
```

- …look at output.
Violation of Independence Assumption

- Within groups $\implies$ big problem; beyond scope of this course.

- Between groups $\implies$
  - Paired comparison or dependent $t$-test (if appropriate)
  - If paired comparison test not applicable/appropriate, Big problem.
Two Dependent Populations

“Paired Comparisons $t$-test”

**Situation:** Have pairs of individuals (experimental units, subjects, etc) and you measure a variable on each “member” of the pair.

- Husband & wife and measure of marital satisfaction (are they the same?)
- Mother’s & father’s scores on a measure of authoritarian parenting style.
- Pre- and post-test scores to assess effectiveness of instruction or an intervention.
- Twins raised apart and a measure of cognitive ability.
Paired Comparisons $t$-test

- Note: by design $n_1 = n_2 = n$.

- Statistical Hypotheses:
  - Three equivalent Null hypotheses:
    
    $H_o : \mu_{diff} = 0$ where $\mu_{diff} = E(D_i)$ and $D_i = Y_{i1} - Y_{i2}$
    
    or
    
    $H_o : \mu_1 = \mu_2$
    
    or
    
    $H_o : \mu_1 - \mu_2 = 0$
    
    - $H_a$ can be 1– or 2–tailed. i.e.,
      
      $H_a : \mu_1 < \mu_2$ or $\mu_1 - \mu_2 < 0$
      
      $H_a : \mu_1 > \mu_2$ or $\mu_1 - \mu_2 > 0$
      
      $H_a : \mu_1 \neq \mu_2$ or $\mu_1 - \mu_2 \neq 0$
Paired Comparisons $t$-test

Assumptions:

- Independence of observations across pairs (within groups).
- Observations between groups (maybe) dependent.
- The variable is normally distributed in each population. ie.

$$Y_{i1} \sim \mathcal{N}(\mu_1, \sigma_1^2) \quad i.i.d.$$  
$$Y_{i2} \sim \mathcal{N}(\mu_2, \sigma_2^2) \quad i.i.d.$$  

$i = 1, 2, \ldots, n.$
Paired Comparisons $t$-test

Notes regarding the assumptions:

$\sigma_1^2$ does not have to equal $\sigma_2^2$.

The correlation between $Y_{i1}$ and $Y_{i2}$ is not necessarily 0.
Paired Comparisons $t$-test

Test Statistic:

- The $t$ statistic:
  \[ t = \frac{\overline{Y}_1 - \overline{Y}_2}{s_{\overline{Y}_1 - \overline{Y}_2}} = \frac{\overline{Y}_D}{s_{\overline{Y}_D}} \]

  The $t$ statistic on the far right is the same as the one population $t$–ratio/statistic.

- The standard error of the mean difference is the square root of
  \[ s_{\overline{Y}_D}^2 = s_{\overline{Y}_1}^2 + s_{\overline{Y}_2}^2 - 2\text{cov}(\overline{Y}_1, \overline{Y}_2) \]
  \[ = s_{\overline{Y}_1}^2 + s_{\overline{Y}_2}^2 - 2r_{12}s_{\overline{Y}_1}s_{\overline{Y}_2} \]

  where $r_{12}$ is the correlation between $Y_1$ and $Y_2$. 

Paired Comparisons \( t \)-test

- **Sampling Distribution** of test statistic is

  Student’s \( t \) distribution with \( \nu = n - 1 \).

- **Decision & Conclusion**: … same as other tests.
Comments Regarding Dependent $t$ Test

- The paired comparison $t$-test is equivalent to computing difference scores for each pair and doing a one sample $t$-test on mean difference scores...

- Sample statistics,

$$\bar{D} = \frac{1}{n} \sum_{i=1}^{n} (Y_{i1} - Y_{i2}) = \frac{1}{n} \sum_{i=1}^{n} D_i$$

$$s_{\bar{D}} = \sqrt{\frac{s_D^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} (D_i - \bar{D})^2 / (n - 1)}{n}}$$
Dependent \( t \) Test = 1 Population \( t \) Test

- **Hypotheses:** \( H_o : \mu_D = 0 \) vs \( H_a : \mu_D \neq 0 \).

- **Test statistic,**
  \[
  t = \frac{\bar{D} - 0}{s_{\bar{D}}} = \frac{\bar{D}}{s_{\bar{D}}}
  \]

- **Sampling distribution of test statistic is Student’s \( t \) distribution with \( \nu = n - 1 \)**
Example: Paired Comparison $t$-test

The girls in the HSB data set and suppose we’re interested in whether the girls’ mean science score is the same as their math score.

There are $n = 327$ girls and

<table>
<thead>
<tr>
<th>Science</th>
<th>Math</th>
<th>Difference Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{Y}_1 = 50.5388$</td>
<td>$\bar{Y}_2 = 51.4346$</td>
<td>$\bar{D} = -0.8957$</td>
</tr>
<tr>
<td>$s_1 = 9.2166$</td>
<td>$s_2 = 9.0685$</td>
<td>$s_D = 7.6335$</td>
</tr>
<tr>
<td>$s_{\bar{Y}_1} = .5097$</td>
<td>$s_{\bar{Y}_2} = .5015$</td>
<td>$s_{\bar{D}} = .4221$</td>
</tr>
<tr>
<td>$r = .6516$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: Paired Comparison $t$-test

- For girls, $H_o : \mu_{sci} = \mu_{math}$ vs $H_a : \mu_{sci} \neq \mu_{math}$

- Assumptions are...

- $\alpha = .05$

- Using the writing and math statistics:

$$
\begin{align*}
\frac{s_{\bar{Y}_1 - \bar{Y}_2}^2}{s_{\bar{Y}_1}^2} &= \frac{s_{\bar{Y}_2}^2}{s_{\bar{Y}_2}^2} - 2r(Y_1, Y_2) s_{\bar{Y}_1} s_{\bar{Y}_2} \\
&= (0.5097)^2 + (0.5015)^2 - 2(0.6516)(0.5097)(0.5015) \\
&= 0.1782
\end{align*}
$$

- This is the same as $s_D = \sqrt{0.1782} = 0.4221$ given on previous slide.
Example: Paired Comparison $t$-test

- Test Statistic using sample statistics for the science and math scores,

$$ t = \frac{50.5388 - 51.4346}{\sqrt{.1782}} = \frac{-0.8958}{.4221} = -2.12 $$

- Test statistic using the difference score sample statistics,

$$ t = \frac{\overline{Y}_1 - \overline{Y}_2}{s_{\overline{Y}_1 - \overline{Y}_2}} = \frac{-0.8957}{.4221} = -2.12 $$

- With $\nu = 327 - 1 = 326$, $p$-value $P(|t| \geq 2.12) = .03$, or $t(326,.975) = 1.97$

(Note: $z_{.975} = 1.96$).
Example: Paired Comparison $t$-test

- Reject $H_0$.

- 95% confidence interval of plausible differences,

\[
\bar{Y}_{sci} - \bar{Y}_{math} \pm t(326,.975)S(\bar{Y}_{sci} - \bar{Y}_{math})
\]

\[-.8957 \pm 1.9672(.4221) \rightarrow (-1.73, -.07)\]

- The girls’ mean math scores are significantly (at the .05 level) from their science.

- The girls’ mean math scores are larger than their science scores... but is this important?

- SAS/ANALYST, ASSIST and commands.
Comments Regarding Dependent $t$ Test

- Under what circumstances can we relax the assumption that $Y_{i1}$ and $Y_{i2}$ come from normal distributions?

- If $Y_1$ and $Y_2$ are positively correlated, then the standard error (deviation) of ($\bar{Y}_1 - \bar{Y}_2$) is smaller, which leads to a larger $t$ statistic (more power) relative to the two independent group test. In the dependent groups test, the some variability is removed by the “matching” variable.

- Question: What if the correlation is negative?
If the matching variable is not related to what you’re measuring, then the dependent group test is bad... it has less power than independent group test & half the degrees of freedom.

- e.g. You randomly assign 20 students to one of two study methods. Your measure is a test score on material learned.
  - “Match” students in pairs based on their height?
  - “Match” students in pairs based on ability/IQ measure?
Power of $t$ Test

- For large samples (i.e., when use $\mathcal{N}(0, 1)$).

- Example

- For any size samples (i.e., when use Student’s $t$ distribution).
Power of $z$ Test: Null Hypothesis

For large samples when use $N(0,1)$
Power of $z$ Test: Alternative

For large samples when use $\mathcal{N}(0,1)$
Power of $z$ Test

For large samples when use $\mathcal{N}(01, )$.

- Basically, like what we did for 1 sample tests... for a 2-tailed alternative.

- Find the rejection region in terms of the original scale. The Critical mean differences

$$D_{crit} = (\bar{Y}_1 - \bar{Y}_2)_{crit}$$
$$= \mu_o \pm Z_{\alpha/2}s_{diff} = 0 \pm Z_{\alpha/2}s_{diff}$$

- Find the z-scores of $\pm D_{crit}$ using the “true” mean difference,

$$z_{alternative} = \frac{\pm D_{crit} - \mu_{true}}{s_{diff}}$$

and find the area.
Example: Power of \( \zeta \) Test

Dependent group \( t \)-test using female seniors from the HSB data set.

- Since \( n_{girls} = 327 \), we have a “large” sample so we’ll use the standard normal distribution.

- When performing the test,

![Graph showing the standard normal distribution with critical values for a \( \zeta \) test.](Image)
Example: Power of $z$ Test

Under the alternative hypothesis ($\mu = -0.8957$)

Paired Dependent Test: Alternative

Inferences About Differences Between Means
Example: Power of $z$ Test

Under the alternative and null together

$\text{Alternative and Null}$

$\text{Density, } f(x)$

$\text{Mean Differences}$

$-3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3$
Example: The Algebra

- What does $\bar{D}_{crit}$ equal?

$$\bar{D}_{crit} = \mu_o \pm 1.96s_{\bar{d}} = 0 \pm 1.96(0.4221) = \pm 0.8273$$

- Assume that the true difference (in the population) equals $-0.8957$ (i.e., $\mu_{diff} = -0.8957$).

$$z = \frac{-0.8273 - (-0.8957)}{0.4221} = 0.1620 \quad z = \frac{0.8273 + 0.8957}{0.4221} = 4.0820$$

- Power = $\text{Prob}(z \leq 0.1620) + \text{Prob}(z \geq 4.0820)$

$$\text{Prob}(z \leq 0.1620) = 0.56 \quad \text{Prob}(z \geq 4.0820) \sim 0.0000002$$
Power for Small Samples

- The logic is the same.
- Use Student’s $t$-distribution.

Possibilities:
- SAS/Analyst
- The cumulative density calculator at http://calculators.stat.ucla.edu/cdf/
- $p$–value program from course web-site.
- SAS program functions.