# Lecture 6: Hypothesis Tests Applied to Means (Cont.)

## • 6. 1 Differences Between Population Means

- Hypotheses about single means are not very common

- It is more common to draw two samples from the same population, apply some experimental manipulation to one of them, and use the other as the control group

- The best available estimate of the difference between the means of two populations is the obtained difference between the means of the two samples…but this will be off every time

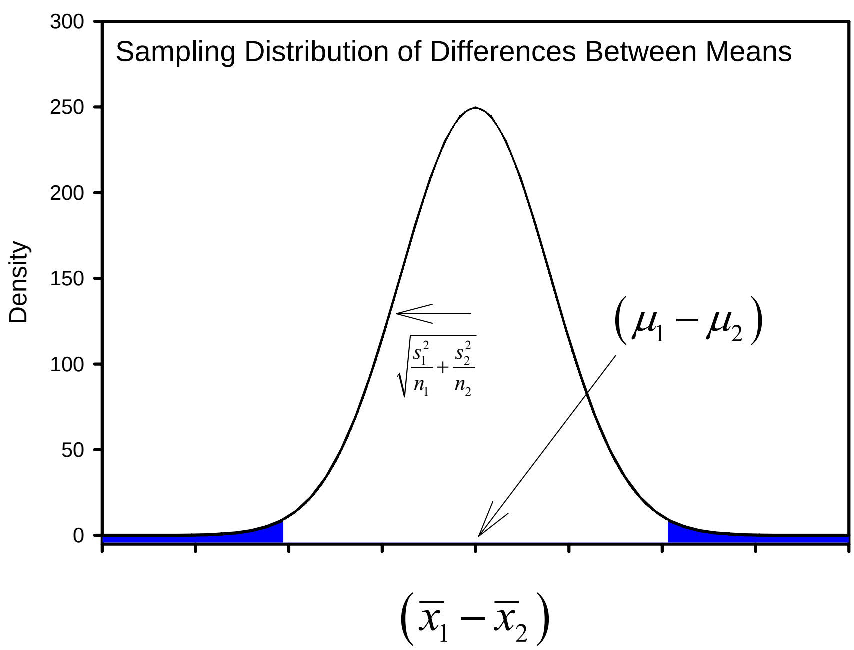
- We need a way to apply statistical inference to the differences between both means

  - $H_0$: $\mu_1 - \mu_2 = k$

  - $H_1$: $\mu_1 - \mu_2 \neq k$

  - We draw a sample $N_1$ from Population 1 and an independent sample $N_2$ from Population 2, and then consider the difference between the means $\bar{x}_1 - \bar{x}_2$

  - If we keep drawing pairs of samples and record the differences, we will obtain a sampling distribution of the sample differences




  - The resulting sampling distribution is: $\bar{x}_1 - \bar{x}_2 \sim N(\mu_1 - \mu_2, \sigma^2_{\bar{x}1 - \bar{x}2})$

  - For any pair of samples, we should expect

$E(\bar{x}_1 - \bar{x}_2) = E(\bar{x}_1) - E(\bar{x}_2) = \mu_1 - \mu_2$ (the expected difference between the two sample means is the difference between the population means)

Also, because of the independence of the two samples

$$\sigma^2_{\bar{x}1 - \bar{x}2} = \sigma^2_1 + \sigma^2_2 - 2\rho\sigma_1\sigma_2$$

$$\sigma^2_{\bar{x}1 - \bar{x}2} = \sigma^2_1 + \sigma^2_2 \ (\textit{Variance Sum Law})$$

- The variances of the means of both groups over repeated sampling are

$$\sigma^2_{\bar{x}1} = \frac{\sigma^2_1}{n_1} \text{ and } \sigma^2_{\bar{x}2} = \frac{\sigma^2_2}{n_2}\text{; or } \sigma_{\bar{x}1} = \frac{\sigma_1}{\sqrt{n_1}} \text{ and } \sigma_{\bar{x}2} = \frac{\sigma_2}{\sqrt{n_2}}$$

- Thus,

$$s^2_{\bar{x}1 - \bar{x}2} = \frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}$$

- Hence, the standard error of the difference is

$$\sigma_{\bar{x}1 - \bar{x}2} = \sqrt{\sigma^2_{\bar{x}1} + \sigma^2_{\bar{x}2}}\text{, or}$$

$$s_{\bar{x}1 - \bar{x}2} = \sqrt{\frac{s^2_1}{n_1} + \frac{s^2_2}{n_2}}$$

- For two large samples drawn from populations with variances $\sigma^2_1$ and $\sigma^2_2$, the standard error of the difference is estimated as

$$\text{est. } \sigma_{\text{dif}} = \sqrt{(est.\sigma_{\bar{x}1})^2 + (est.\sigma_{\bar{x}2})^2}$$

even if $n_1 \neq n_2$, as long as $n_1$ and $n_2 \geq 1$

- Regardless of $N_1$ and $N_2$, and the form of the parent distribution, the expectation of the difference between two means is always the difference between their expectations and the variance of the difference between two independent means is the sum of their separate sampling variances. Moreover:

- If the distribution for each of the two populations is normal, the distribution of differences between sample means is normal

- As $N_1$ and $N_2$ become large, the sampling distribution of the difference between means approaches a normal distribution, regardless of the form of the original distributions (CLT)

- Thus, If $N_1$ and $N_2$ are very large, we can approximate the sampling distribution of the difference between means by a normal distribution, the test statistic is

$$z_{\text{diff}} = \frac{(\bar{x}_1 - \bar{x}_2) - E(\bar{x}_1 - \bar{x}_2)}{est.\sigma_{diff.}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{est.\sigma_{diff.}}$$

and this value can be referred to a normal distribution

- The hypothesis is of the form

  - $H_0$: $\mu_1 - \mu_2 = k$

  - $H_1$: $\mu_1 - \mu_2 \neq k$, (nondirectional) or

  - $H_1$: $\mu_1 - \mu_2 > k$, or $H_1$: $\mu_1 - \mu_2 < k$, directional

- Confidence intervals can be computed as for a single mean, as

  $(\bar{x}_1 - \bar{x}_2) \pm 1.96$ (est. $\sigma_{\text{dif}}$), for $\alpha = .05$

- **6. 2 Large-Sample Test for a Difference Between Means**

- Experiment to examine the effectiveness of a new drug therapy to treat depression

- Two random samples are taken from a clinical population of severely depressed individuals

1. We give a drug to one group (experimental group -- drug)

2. We give a placebo drug to the other group (control group -- placebo)

- Depression is measured after 3 weeks and it is found that Group 1 does better than Group 2. Is this a genuine effect from the drug or simply due to normal variability?

- $H_0$: $\mu_1 - \mu_2 = 0$

- $H_1$: $\mu_1 - \mu_2 \neq 0$

We want to know the difference between the sample means but also we want to know about their variability so we construct sampling distributions (in theory)

| | *Drug* | *Placebo* | *Difference* |
|---|---|---|---|
| | $\overline{y_{d1}}$ | $\overline{y_{p1}}$ | $\overline{y_{d1}} - \overline{y_{p1}}$ |
| | $\overline{y_{d2}}$ | $\overline{y_{p2}}$ | $\overline{y_{d2}} - \overline{y_{p2}}$ |
| | $\overline{y_{d3}}$ | $\overline{y_{p3}}$ | $\overline{y_{d3}} - \overline{y_{p3}}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ |
| | $\overline{y_{dN}}$ | $\overline{y_{pN}}$ | $\overline{y_{dN}} - \overline{y_{pN}}$ |
| $E$ | $\mu_d$ | $\mu_p$ | $\mu_d - \mu_p$ |

- In practice, we do the following:

*Sample* 1: $\bar{x}_1 = 1.82$, $s_1 = .7$, $N_1 = 125$

*Sample* 2: $\bar{x}_2 = 1.61$, $s_2 = .9$, $N_2 = 150$

- Even though the distributions of the two populations are unknown, the sample sizes are large and, thus, a normal approximation can be used to the sampling distribution of the differences

$$z_{\text{diff}} = \frac{(\bar{x}_1 - \bar{x}_2) - E(\bar{x}_1 - \bar{x}_2)}{est.\sigma_{diff.}}$$

- Under $H_0$, $E(\bar{x}_1 - \bar{x}_2) = (\mu_1 - \mu_2) = 0$

- Because $\sigma_1$ and $\sigma_2$ are unknown, the $\sigma_{\text{dif}}$ needs to be estimated. This is done by

$$\sigma^2_{\bar{x}1} = \frac{s^2{}_1}{n_1} = \frac{.49}{125} = .0039 \text{ and}$$

$$\sigma^2_{\bar{x}2} = \frac{s^2{}_2}{n_2} = \frac{.81}{125} = .0054$$

- Then, the estimated standard error of the mean difference is

$$\text{est. } \sigma_{\text{dif}} = \sqrt{(est.\sigma_{\bar{x}1})^2 + (est.\sigma_{\bar{x}2})^2} = \sqrt{.0039 + .0054} = .0964$$

- Thus, $z = \dfrac{(\bar{x}_1 - \bar{x}_2) - E(\bar{x}_1 - \bar{x}_2)}{est.\sigma_{diff.}} = \dfrac{1.82 - 1.61}{.0964} = 2.18$

- If $\alpha$ is selected as .05, the 5% critical value is $z = |1.96|$. The observed difference is larger than the critical value and thus it can be concluded that a difference exists between these two populations. It could be reasonable to assume, given proper randomization and everything else being equal, that such a difference is attributable to the drug

- 95%CI: $(\bar{x}_1 - \bar{x}_2) \pm 1.96 \cdot (.0964) = .21 \pm .1889 =$

(.021, .399). It does not contain 0.

## • 6. 3 Using The *t* Distribution To Test Hypotheses About Differences

- If the sample sizes are not very large and the variances of both populations can be assumed to be equal, the *t* distribution can be used to test for two (or more) independent samples

- In this case, the *t* test is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - k}{est.\sigma_{diff.}}, \text{ where } k = \mu_1 - \mu_2,$$

which follows the same form as for a single mean with degrees of freedom as

$\upsilon = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$

- If one assumes equal variances in the population ($\sigma_1 = \sigma_2 = \sigma$), then

$$\sigma_{diff} = \sqrt{\left(\frac{\sigma^2}{n_1}\right) + \left(\frac{\sigma^2}{n_2}\right)} = \sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

- Given that we now have two estimates of the same parameter, $\sigma^2$, the pooled estimate is better than either one taken separately. This *pooled variance estimate* is a weighted average of the two sample variances, weighted by the degrees of freedom of each sample, as

$$s^2{}_p = \frac{(df_1)s^2{}_1 + (df_2)s^2{}_2}{df_1 + df_2} = \frac{(n_1 - 1)s^2{}_1 + (n_2 - 1)s^2{}_2}{(n_1 - 1) + (n_2 - 1)}$$

$$s^2{}_p = \frac{(n_1 - 1)s^2{}_1 + (n_2 - 1)s^2{}_2}{n_1 + n_2 - 2}$$

- If $n_1 = n_2$, then $s^2{}_p = \dfrac{(n-1)(s^2{}_1 + s^2{}_2)}{2(n-1)} = \dfrac{(s^2{}_1 + s^2{}_2)}{2}$

- And the estimated standard error of the difference is

$$\text{est.}\sigma_{diff} = \sqrt{est.\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$= \sqrt{\frac{(n_1 - 1)s^2{}_1 + (n_2 - 1)s^2{}_2}{n_1 + n_2 - 2}\left(\frac{n_1 + n_2}{n_1 n_2}\right)}$$

- *Example* 1: Two random samples of subjects (experiment and control) are being compared on the basis of their scores on a motor learning task. The individuals in the experiment group receive feedback after each attempt, whereas the individuals in the control group do not. Does the "experiment" (i.e., performance feedback) affect the performance on the task?

  - $H_0$: $\mu_1 - \mu_2 = 0$
  - $H_1$: $\mu_1 - \mu_2 \neq 0$

- It is assumed that the population distributions of scores are normal and that the population variances are equal. $N_1 = 5$ and $N_2 = 7$

- $\alpha = .01$. $2Q = .01$ with $N_1 + N_2 - 2 = 10$ degrees of freedom, critical $t$ value = 3.169

  - If values $\geq |3.169|$, reject $H_0$

- Sample results:

  *Sample* 1: $\bar{x}_1 = 18$, $s^2{}_1 = 7.00$, $N_1 = 5$

  *Sample* 2: $\bar{x}_2 = 20$, $s^2{}_2 = 5.83$, $N_1 = 7$

- Estimated standard error of the difference is

$$\text{est.}\sigma_{diff} = \sqrt{\frac{(n_1-1)s^2{}_1 + (n_2-1)s^2{}_2}{n_1+n_2-2}\left(\frac{n_1+n_2}{n_1 n_2}\right)}$$

$$= \sqrt{\frac{(5-1)7 + (7-1)5.83}{5+7-2}\left(\frac{12}{35}\right)}$$

$$= 1.47$$

- The $t$ ratio is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - E(\bar{x}_1 - \bar{x}_2)}{est.\sigma_{diff.}} = \frac{-2}{1.47} = -1.36$$

- Because the obtained value is smaller (in absolute terms) than the critical value, given $\alpha = .01$, $H_0$ cannot be rejected

- Confidence intervals can be estimated as

$$99\%\ \text{CI} = (\bar{x}_1 - \bar{x}_2) \pm t_{(\alpha/2;\upsilon)} \cdot (\text{est.}\sigma_{diff})$$

$$= -2 \pm (3.169)(1.47) = (-6.66, 2.66).$$

The $p$ is .99 that the true difference between $\mu_1 - \mu_2$ is contained in this interval.

## • 6. 4 Paired Samples

- Examples:

    - same subjects, same variables (pretest, posttest)

    - same subjects, different variables (e.g., midterm, final)

    - related subjects, same variables (parent-child)

- Matching pairs reduces the variability

$\sigma^2_{\bar{x}1-\bar{x}2} = \sigma^2_1 + \sigma^2_2 - 2\text{cov}_{x1\text{-}x2}$ or

$\sigma^2_{\bar{x}1-\bar{x}2} = \sigma^2_1 + \sigma^2_2 - 2\rho\sigma_1\sigma_2$

the larger the covariance, the smaller the variability of the difference

- Dependent Samples *t* test - Example

1. Research question: Is there a difference between the verbal and performance scores of neurologically impaired children?

| | Verbal | Performance | Difference | |
|---|---|---|---|---|
| $n$ | $(X_1)$ | $(X_2)$ | $(X_d)$ | $(X_d - \overline{X_d})^2$ |
| 1 | 80 | 70 | 10 | 25 |
| 2 | 100 | 80 | 20 | 25 |
| 3 | 110 | 90 | 20 | 25 |
| 4 | 120 | 90 | 30 | 225 |
| 5 | 70 | 70 | 0 | 225 |
| 6 | 100 | 110 | -10 | 625 |
| 7 | 110 | 80 | 30 | 225 |
| 8 | 120 | 120 | 0 | 225 |
| 9 | 110 | 80 | 30 | 225 |
| 10 | 90 | 70 | 20 | 25 |
| | $\overline{X}_1 = 101$ | $\overline{X}_2 = 86$ | $\sum X_d = 150$ | $\sum (X_d - \overline{X}_d)^2 = 1850$ |
| | $(s_1 = 16.6)$ | $(s_2 = 17.1)$ | $\overline{X_d} = 15$ | |

- From the differences we perform a one sample *t* test.

2. Statistical Hypotheses.

$H_0$: $\mu_d = 0$

$H_1$: $\mu_d \neq 0$

Assumptions

$t$ follows a $t$ distribution with $n - 1$ degrees of freedom

3. Decision Rules

$\alpha = .05$, Critical $t = \pm\, 2.262$;

If $|t| \geq 2.262$, Reject $H_0$

4. Test Statistic (One sample $t$ test)

$$t = \frac{\overline{D} - \mu_d}{s_{\bar{d}}} \qquad s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$$

$$s^2{}_d = \frac{\sum (D - \overline{D})^2}{n-1} = \frac{1850}{9} = 205.6$$

$s_d = 14.34$

$$s_{\bar{d}} = \frac{s_d}{\sqrt{n}} = \frac{14.34}{\sqrt{10}} = 4.53$$

$$t = \frac{\overline{D} - \mu_d}{s_{\bar{d}}} = \frac{15}{4.53} = 3.31$$

5. Decision.

Observed $t >$ critical $t$; $3.31 > 2.26$, Reject $H_0$, $p < .05$

6. Confidence Intervals

$.95CI = 15 \pm (2.26) \cdot (4.53) = (4.75, 25.26)$

- Advantages of Dependent Sample Test (DST) over IST
    - DST is a more sensitive test because the sampling distribution is less variable
        - this is true when the matching of the pairs is related to the variable examined
        - if this is true, the correlation will be positive and the $s_{\bar{d}}$ will be smaller

- Suppose we compute the differences between the scores in the last example as independent samples.

    $\alpha = .05$. $2Q = .05$ with $N_1 + N_2 - 2 = 18$ degrees of freedom, critical $t$ value = 2.101

    - Estimated standard error of the difference is

$$\text{est.}\sigma_{diff} = \sqrt{\frac{(n_1-1)s^2{}_1 + (n_2-1)s^2{}_2}{n_1+n_2-2}\left(\frac{n_1+n_2}{n_1 n_2}\right)}$$

$$= \sqrt{\frac{(10-1)(16.6)^2 + (10-1)(17.1)^2}{10+10-2}\left(\frac{20}{100}\right)}$$

$$= 7.55$$

- The $t$ ratio is

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - E(\bar{x}_1 - \bar{x}_2)}{est.\sigma_{diff.}} = \frac{15}{7.55} = 1.99$$

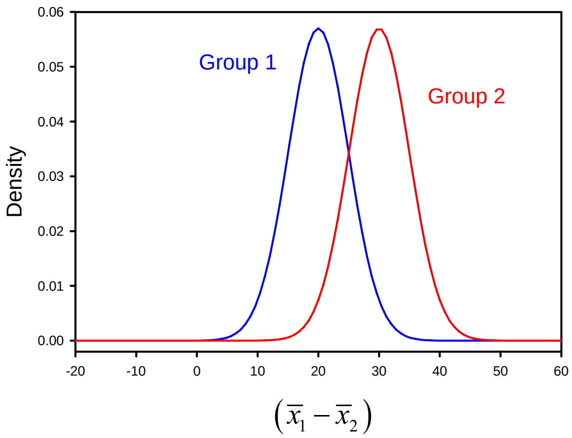
- Observed $t <$ critical $t$; $1.99 < 2.10$, Retain $H_0$, $p > .05$

## 6. 5 Heterogeneity of Variance

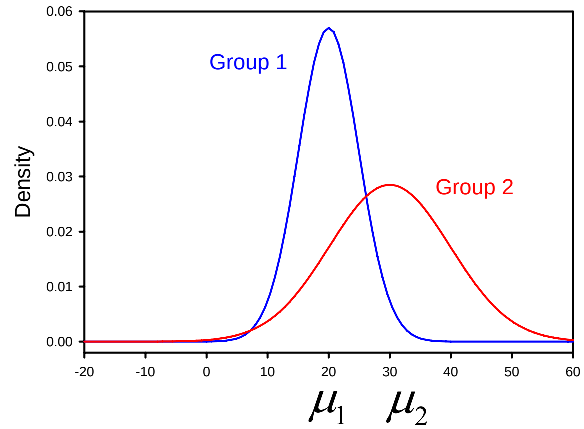
- To use the independent-samples *t*-test, we assume that the populations from which our samples were drawn have equal variances

- This is called the *homogeneity of variance* assumption

- What happens if this assumption is violated?




*equal variances*




*unequal variances*

- Sample variances will always differ at least by chance. But are they different enough that we should conclude something more than chance is to blame?

- We don't have to *assume* homogeneity of variance – we can test for it. It is an empirical question

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

- Rule of thumb – if one group's variance is at least 4 times the size of the other group's variance, we have a problem

- What to do?

  - Don't pool the variances

  - Use the smaller of *df*1 and *df*2 as if it were the total *df*

- **6. 6 Strength of Association**

- Scientific purpose is to predict or explain variation. The variable *Y* has some variance that we would like to account for. There are statistical indexes of how well our IV accounts for variance in the DV. These are measures of how strongly or closely associated our IVs and DVs are

- Variance accounted for (does knowing *X* reduce our uncertainty about *Y*)?

$$\omega^2 = \frac{\sigma_Y^2 - \sigma_{Y|X}^2}{\sigma_Y^2}$$

- We can say that a statistical relation exists when the variability of *Y* given *X* is smaller than the variability of *Y*

- For two treatment populations with equal variances, the strength of association is

$$\omega^2 = \frac{(\mu_1 - \mu_2)^2}{4\sigma_Y^2}$$

- How much of variance on *Y* is associated with the IV?

  - In each case, how much of the variance in *Y* is associated with the IV, group membership? As a mean difference gets bigger, so does the variance accounted for

- Effect size:

$$d = \frac{\left|(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)\right|}{s_p}$$ , or simplified $$d = \frac{\left|(\bar{x}_1 - \bar{x}_2)\right|}{s_p}$$

this represents the difference between means in standard deviation units

where .2 = small, .5 = medium, and .8 = large effect