• **Key Symbols**

- $\mu_1 - \mu_2 \rightarrow$ Difference between population means
- $\mu_d \rightarrow$ Paired difference in population means
- $p_1 - p_2 \rightarrow$ Difference between population proportions
- $\sigma_1^2 / \sigma_2^2 \rightarrow$ Ratio of population variances
- $D_0 \rightarrow$ Hypothesized value of difference
- $\bar{x}_1 - \bar{x}_2 \rightarrow$ Difference between sample means
- $\bar{d} \rightarrow$ Mean of sample differences
- $\hat{p}_1 - \hat{p}_2 \rightarrow$ Difference between sample proportions
- $s_1^2 / s_2^2 \rightarrow$ Ratio of sample variances
- $\sigma_{(\bar{x}_1 - \bar{x}_2)} \rightarrow$ Standard error for $\bar{x}_1 - \bar{x}_2$
- $\sigma_d \rightarrow$ Standard error for $\bar{d}$
- $\sigma_{(\hat{p}_1 - \hat{p}_2)} \rightarrow$ Standard error for $\hat{p}_1 - \hat{p}_2$
- $F_a \rightarrow$ Critical value for $F$-distribution
- $\nu_1 \rightarrow$ Numerator degrees of freedom for $F$-distribution
- $\nu_2 \rightarrow$ Denominator degrees of freedom for $F$-distribution
- $SE \rightarrow$ Sampling error in estimation

• **Formulas**

**Large Sample Confidence Interval for $\mu_1 - \mu_2$**

\[
(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sigma_{(\bar{x}_1 - \bar{x}_2)} = (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}
\]

\[
\approx (\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}
\]

**Large-Sample Test of Hypothesis for $\mu_1 - \mu_2$**

<table>
<thead>
<tr>
<th>One-Tailed Test</th>
<th>Two-Tailed Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0 : \mu_1 - \mu_2 = D_0$</td>
<td>$H_0 : \mu_1 - \mu_2 = D_0$</td>
</tr>
<tr>
<td>$H_a : \mu_1 - \mu_2 &lt; D_0$</td>
<td>$H_a : \mu_1 - \mu_2 \neq D_0$</td>
</tr>
<tr>
<td>or $H_a : \mu_1 - \mu_2 &gt; D_0$</td>
<td></td>
</tr>
</tbody>
</table>

Where $D_0 =$ Hypothesized difference between the means (this difference is often hypothesized to be equal to 0)
**Test statistic:**

\[ z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sigma_{(\bar{x}_1 - \bar{x}_2)}} \]

where

\[ \sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \approx \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \]

**Rejection region:** \( z < -z_{\alpha} \)

[or \( z > z_{\alpha} \) when \( H_a : \mu_1 - \mu_2 > D_0 \)]

**Rejection region:** \( |z| > z_{\alpha/2} \)

---

**Small-Sample Confidence Interval for \((\mu_1 - \mu_2)\): Independent Samples**

\[ (\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \sqrt{s_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} \]

Where

\[ s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \]

and \( t_{\alpha/2} \) is based on \((n_1 + n_2 - 2)\) degrees of freedom

---

**Small-Sample Test of Hypothesis for \((\mu_1 - \mu_2)\): Independent Samples**

**One-Tailed Test**

\( H_0 : \mu_1 - \mu_2 = D_0 \)

\( H_a : \mu_1 - \mu_2 < D_0 \)

[or \( H_a : \mu_1 - \mu_2 > D_0 \)]

**Two-Tailed Test**

\( H_0 : \mu_1 - \mu_2 = D_0 \)

\( H_a : \mu_1 - \mu_2 \neq D_0 \)

**Test statistic:**

\[ t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \]

**Rejection region:**

\( t < -t_{\alpha} \)

[or \( t > t_{\alpha} \) when \( H_a : \mu_1 - \mu_2 > D_0 \)]

**Rejection region:** \( |t| > t_{\alpha/2} \)
Approximate Small Sample Procedures when $\sigma_1^2 \neq \sigma_2^2$

Equal sample sizes ($n_1 = n_i = n$)

Confidence interval: \[ (\bar{x}_1 - \bar{x}_2) \pm t_{a/2} \sqrt{\left(\frac{s_1^2}{n} + \frac{s_2^2}{n}\right)} / n \]

Test statistic for $H_0 : (\mu_1 - \mu_2) = 0$ : \[ t = (\bar{x}_1 - \bar{x}_2) / \sqrt{\left(\frac{s_1^2}{n} + \frac{s_2^2}{n}\right)} / n \]

Where $t$ is based on $\nu = n_1 + n_2 - 2 = 2(n - 1)$ degrees of freedom.

Unequal Sample Sizes ($n_1 \neq n_1$)

Confidence interval: \[ (\bar{x}_1 - \bar{x}_2) \pm t_{a/2} \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)} \]

Test statistic for $H_0 : (\mu_1 - \mu_2) = 0$ : \[ (\bar{x}_1 - \bar{x}_2) / \sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)} \]

Where $t$ is based on degrees of freedom equal to \[ \nu = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}} \]

Note: The value of $\nu$ will generally not be an integer. Round $\nu$ down to the nearest integer to use the $t$-table.

Paired Difference Confidence Interval for $\mu_d - \mu_1 - \mu_2$

Large Sample

\[ \bar{x}_d \pm z_{a/2} \frac{\sigma_d}{\sqrt{n}} \approx \bar{x}_d \pm z_{a/2} \frac{s_d}{\sqrt{n}} \]

Small Sample

\[ \bar{x}_d \pm t_{a/2} \frac{s_d}{\sqrt{n}} \]

where $t_{a/2}$ is based on $(n_d - 1)$ degrees of freedom
Paired Difference Test of Hypothesis for $\mu_d - \mu_1 - \mu_2$

**One-Tailed Test**

$H_0 : \mu_d = D_0$

$H_a : \mu_d < D_0$

[or $H_a : \mu_d > D_0$]

**Two-Tailed Test**

$H_0 : \mu_d = D_0$

$H_a : \mu_d \neq D_0$

**Large Sample**

Test statistic: $z = \frac{x_d - D_0}{\sigma_d} \approx \frac{x_d - D_0}{s_d}$

Rejection region: $z < -z_{\alpha}$

[or $z > z_{\alpha}$ when $H_a : \mu_d > D_0$]

**Small Sample**

Test statistic: $t = \frac{x_d - D_0}{s_d}$

Rejection region: $t < -t_{\alpha}$

[or $t > t_{\alpha}$ when $H_a : \mu_d > D_0$]

**Large Sample 100(1-\alpha)\% Confidence Interval for (p_1 - p_2)**

$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sigma_{(\hat{p}_1 - \hat{p}_2)} = (\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$

$\approx (\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$

**Large Sample Test of Hypothesis about (p_1 - p_2)**

**One-Tailed Test**

$H_0 : (p_1 - p_2) = D_0$

$H_a : (p_1 - p_2) < D_0$

[or $H_a : (p_1 - p_2) > D_0$]

Test statistic: $z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sigma_{(\hat{p}_1 - \hat{p}_2)}}$

Rejection region: $z < -z_{\alpha}$

[or $z > z_{\alpha}$ when $H_a : (p_1 - p_2) > D_0$]

**Two-Tailed Test**

$H_0 : (p_1 - p_2) = D_0$

$H_a : (p_1 - p_2) \neq D_0$

Rejection region: $|z| > z_{\alpha/2}$
Formulas

Chapter Nine: Inferences Based on Two Samples; Confidence Intervals and Tests of Hypothesis

Randall Miller

\[ A(1 - \alpha) \] 100% Confidence Interval for \( \frac{\sigma_1^2}{\sigma_2^2} \)

\[
\left( \frac{s_1^2}{s_2^2} \right) \left( \frac{1}{F_{L,\alpha/2}} \right) < \left( \frac{\sigma_1^2}{\sigma_2^2} \right) < \left( \frac{s_1^2}{s_2^2} \right) F_{U,\alpha/2}
\]

Where \( F_{L,\alpha/2} \) is the value of \( F \) that places an area \( \alpha / 2 \) in the upper tail of an \( F \)-distribution with

\[ \nu_1 = (n_1 - 1) \] numerator and \( \nu_2 = (n_2 - 1) \) denominator degrees of freedom, and \( F_{U,\alpha/2} \) is the value of \( F \) that places an area \( \alpha / 2 \) in the upper tail of an \( F \)-distribution with \( \nu_1 = (n_2 - 1) \) numerator and \( \nu_2 = (n_1 - 1) \) denominator degrees of freedom.

\( F \)-Test for Equal Population Variances

**One-Tailed Test**

\[ H_0 : \sigma_1^2 = \sigma_2^2 \]

\[ H_a : \sigma_1^2 < \sigma_2^2 \) \quad \text{(or } H_a : \sigma_1^2 > \sigma_2^2 \)

**Test statistic:**

\[ F = \frac{s_1^2}{s_2^2} \]

\[ \quad \left( \text{or } F = \frac{s_2^2}{s_1^2} \text{ when } H_a : \sigma_1^2 > \sigma_2^2 \right) \]

**Rejection region:**

\[ F > F_{\alpha,\nu_1,\nu_2} \]

\( \quad \text{where } F_{\alpha,\nu_1,\nu_2} \) are based on \( \nu_1 \) numerator degrees of freedom and \( \nu_2 \) denominator degrees of freedom; and \( \nu_1 \) and \( \nu_2 \) are the degrees of freedom for the numerator and denominator sample variances, respectfully.

**Two-Tailed Test**

\[ H_0 : \sigma_1^2 = \sigma_2^2 \]

\[ H_a : \sigma_1^2 \neq \sigma_2^2 \]

**Test statistic:**

\[ F = \frac{\text{Larger sample variance}}{\text{Smaller sample variance}} \]

\[ = \frac{s_1^2}{s_2^2} \text{ when } s_1^2 > s_2^2 \]

\[ \quad \left( \text{or } \frac{s_2^2}{s_1^2} \text{ when } s_2^2 > s_1^2 \right) \]

**Rejection region:**

\[ F > F_{\alpha/2,\nu_1,\nu_2} \]