USC School of Business

Data Driven Decision Making: Two-Sample Hypothesis Testing

GSBA 545, Fall 2021

Professor Dawn Porter

Two-Sample Hypothesis Testing

USC School of Business

- Independent Samples
  - CI for Mean Differences & Hypothesis Testing with Difference D_0
- Paired Samples
  - CI for Paired Mean Differences & Hypothesis Testing with Difference D_0
- Proportion Differences
  - CI for Proportion Differences & Hypothesis Testing with Difference D_0
- A/B Testing Applications

Independent Samples: Intervals

Random samples are taken from two independent populations; we assume the variances are approximately equal: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Population 1:	Population 2:
mean μ_1, unknown σ_1^2	mean μ_2, unknown σ_2^2
• random sample of size n_1	• random sample of size n_2
• mean $\bar{x}_1$ & variance s_1^2	• mean $\bar{x}_2$ & variance s_2^2

Pooled estimate of the *standard error* is:

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

and then the *t*-based CI is:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2} \times se(\bar{x}_1 - \bar{x}_2)$$

* $t_{\alpha/2}$ is based on $n_1 + n_2 - 2$ degrees of freedom

Independent Samples: Calcium & BP

Assume that two populations of men (one taking extra calcium and one isn't) have normally distributed blood pressures and their population variances are equal. At $\alpha = 0.10$, does the calcium supplement correspond to a *change* in BP?

<u>Population 1: Calcium supplement</u>

Random sample of 10 men taking calcium yields blood pressure changes with:

- sample mean of –5.0
- sample variance of 76.444

<u>Population 2: No calcium</u>

Random sample of 11 men receiving a placebo yields blood pressure changes with:

- sample mean of 0.273
- sample variance of 34.818

To create the *pooled standard error,*

$$\boxed{se(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{76.444}{10} + \frac{34.818}{11}} = 3.288}$$

Independent Samples: Calcium & BP

Assume that two populations of men (one taking extra calcium and one isn't) have normally distributed blood pressures and their population variances are approximately equal. At $\alpha = 0.10$, does the calcium supplement correspond to a *change* in BP?

$df = (n_1 + n_2 - 2) = (10 + 11 - 2) = 19$, so $t_{0.05,\, df=19} \rightarrow 1.729$

A 90% CI for $\mu_1 - \mu_2$ is :

$$[(\bar{x}_1 - \bar{x}_2) \pm t_{0.05} \times se(\bar{x}_1 - \bar{x}_2)] = [(-5 - 0.273) \pm 1.729(3.288)]$$

$$= [-5.273 \pm 5.685] = [-10.958, 0.412]$$

We can be 90% confident that the mean change in blood pressure by taking the calcium supplement is between -10.958 and 0.412 points.

On average, the difference is a decrease of 5.273 points.

Independent Samples: t- Statistic

- $D_0 = \mu_1 - \mu_2$ is the claimed difference between population means
- D_0 varies depending on the context of the situation
- Often $D_0 = 0$, so H_0 means that there is no difference between the population means
- This is analyzed using a t distribution with $df = (n_1 + n_2 - 2)$

Test statistic is:

$$t = \frac{(\bar{x}_1 - \bar{x}_2 - D_0)}{se(\bar{x}_1 - \bar{x}_2)}$$

If the differences are fairly normal, we can reject H_0: $\mu_d = D_0$ at the α level of significance if and only if the appropriate rejection point condition holds or, equivalently, if the corresponding p-value is less than α.

Example: Calcium & BP

Assume two populations of men (one taking extra calcium and one isn't) have normally distributed blood pressures and their population variances are equal.

<u>Population 1: Calcium supplement</u>

n_1 = 10 men with BP changes of:

- $\bar{x}_1 = -5.0$
- $s_1^2 = 76.444$

<u>Population 2: No calcium</u>

n_2 = 11 men with BP changes of:

- $\bar{x}_2 = 0.273$
- $s_2^2 = 34.818$

At α = 0.10, does the calcium supplement correspond to a significant *decrease* in BP?

Two-Sample Testing: Calcium & BP

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α**.
3. Calculate **test statistic** value.
4. Determine **critical value(s)**.
5. Decide whether to **reject H_0** and interpret the statistical result.

1. H_0: $\mu_1 - \mu_2 \geq 0$ vs H_a: $\mu_1 - \mu_2 < 0$
2. $\alpha = 0.10$
3. $$t = \frac{(-5.0-0.273)}{\sqrt{\frac{76.444}{10}+\frac{34.818}{11}}} = -1.604$$
4. Reject H_0 if $t < t_{0.10,19} < -1.328$
5. Since $t = -1.604 < -1.328$, we can **reject H_0** and conclude that taking calcium *does* correspond to a significant decrease in BP.

Paired Samples: Intervals

When the same units are used for both processes.

- Ex: using same people for 'before' and 'after' treatments
- Eliminates differences between sampled groups
- Only need to test differences between results

$\mu_d = \mu_1 - \mu_2$, where μ_1 and μ_2 are means from two populations

- $\bar{d}$ is the mean of the differences between pairs of values
- s_d is the standard deviation of the sample of paired differences

If the differences are fairly normally distributed, then a CI for μ_d is:

$$\left[\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}\right]$$

with $n - 1$ degrees of freedom.

* If population σ values are known, a *Z*-based interval can be used instead.

Paired Samples: Repair Costs

7 damaged cars are assessed for repairs by two different garages.

<u>Garage 1:</u>

- $\bar{x}_1 = \$932.90$

<u>Garage 2:</u>

- $\bar{x}_2 = \$1012.90$

<u>Differences:</u>

- $\bar{d}$ = 932.9 – 1012.9 = –80
- s_{d}^2 = 2533.11
- s_{d} = 50.33

$t_{0.025,\mathrm{df}=6}$ = 2.447. A 95% CI for μ_{d} with $n - 1$ = 6 degrees of freedom is:

$$\left[\bar{d} \pm t_{\alpha/2}\frac{s_{\mathrm{d}}}{\sqrt{n}}\right] \rightarrow \left[-80 \pm 2.447\frac{50.33}{\sqrt{7}}\right] \rightarrow [-126.55, -33.45]$$

- We are 95% sure that the mean difference of repair cost estimates is between –\$126.55 and –\$33.45.
- The mean of estimates at Garage 1 is between \$126.55 and \$33.45 *less* than that of Garage 2.

USC School of Business

Paired Samples: t- Statistic

- $D_0 = \mu_1 - \mu_2$ is the claimed difference between population means
- D_0 varies depending on the context of the situation
- Often $D_0 = 0$, so H_0 means that there is no difference between the population means
- This is analyzed using a t distribution with $df = (n - 1)$

Test statistic is:

$$t = \frac{\bar{d} - D_0}{s_{\mathrm{d}}/\sqrt{n}}$$

If the differences are fairly normal, we can reject H_0: $\mu_d = D_0$ at the α level of significance if and only if the appropriate rejection point condition holds or, equivalently, if the corresponding *p*-value is less than α.

* If population σ values are known, a *Z*-test can be used instead.

Paired Samples: Rejections

Alternative	Reject H_0 if:	p-value
H_a: $\mu_d > D_0$	$t > t_\alpha$	Area to the right of t
H_a: $\mu_d < D_0$	$t < -t_\alpha$	Area to the left of $-t$
H_a: $\mu_d \neq D_0$	$\|t\| > t_{\alpha/2}$*	Twice the area to the right of $\|t\|$

where t_α, $t_{\alpha/2}$, and p-values are based on $(n-1)$ degrees of freedom

* either $t > t_{\alpha/2}$ or $t < -t_{\alpha/2}$, and if population σ values are known, corresponding Z-values can be used instead.

Paired Samples: Repair Costs

Assume 7 damaged cars are assessed for repairs by two different garages to compare the costs of repairs. Garage 1 asserts that it is less expensive.

Garage 1:

- $\bar{x}_1 = \$932.90$

Garage 2:

- $\bar{x}_2 = \$1012.90$

Differences:

- $\bar{d}$ = 932.9 – 1012.9 = –80
- s^2_{d} = 2533.11
- s_{d} = 50.33

Is there enough evidence at α = 0.05 to believe the claim of Garage 1?

Paired Samples: Repair Costs

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α**.
3. Calculate **test statistic** value.
4. Determine **critical value(s)**.
5. Decide whether to **reject H_0** and interpret the statistical result.

1. H_0: $\mu_1 - \mu_2 \geq 0$ vs H_a: $\mu_1 - \mu_2 < 0$
2. $\alpha = 0.05$
3. $t = \dfrac{\bar{d} - D_0}{s_d/\sqrt{n}} = \dfrac{-80-0}{50.33/\sqrt{7}} = -4.2053$
4. Reject H_0 if $t < t_{0.05,6} = -1.943$
5. Since $t = -4.2053 < -1.943$, we can **reject H_0** and conclude the mean repair cost at Garage 1 is less than Garage 2.

* From JMP, for $t = -4.2053$, the p-value is 0.003, indicating very strong evidence that μ_1 is actually less than μ_2.

Population Proportions

Testing for a difference in proportions between populations.

Population 1:

- sample size n_1
- sample proportion $\hat{p}_1$ in category of interest
- $n_1\,\hat{p}_1 \geq 5$, $n_1\,(1-\hat{p}_1) \geq 5$

Population 2:

- sample size n_2
- sample proportion $\hat{p}_2$ in category of interest
- $n_2\,\hat{p}_2 \geq 5$, $n_2\,(1-\hat{p}_2) \geq 5$

Differences: $\hat{p}_1 - \hat{p}_2$

- Assume n_1 and n_2 are large
- Difference is approximately normal

Pooled estimate of the *standard error* is:

$$se(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Proportions: Z-statistic

Used when difference between population proportions are approximately normal and sample sizes are large enough.

Confidence Interval:

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \times se(\hat{p}_1 - \hat{p}_2)$$

Test Statistic:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - D_0}{se(\hat{p}_1 - \hat{p}_2)}$$

$D_0 = p_1 - p_2$ is the claimed or hypothesized difference between the population proportions.

Proportions: Rejections

Alternative	Reject H_0 if:	*p*-value
$H_a: p_1 - p_2 > D_0$	$z > z_\alpha$	Area to the right of z
$H_a: p_1 - p_2 < D_0$	$z < -z_\alpha$	Area to the left of $-z$
$H_a: p_1 - p_2 \neq D_0$	$\lvert z \rvert > z_{\alpha/2}$*	Twice the area to the right of $\lvert z \rvert$

* either $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$

Proportion CI: Liver Transplants

Assume two hospitals routinely perform liver transplants. What is a 95% CI for the difference in the fatality rate between the two?

Hospital 1:

- $n_1 = 100$, $x_1 = 77$
- $\hat{p}_1 = \frac{77}{100} = 0.77$

Hospital 2:

- $n_1 = 200$, $x_1 = 120$
- $\hat{p}_2 = \frac{120}{200} = 0.60$

Pooled standard error:

$$se(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{0.77(0.23)}{100} + \frac{0.6(0.4)}{200}} = 0.0545$$

A 95% CI for $p_1 - p_2$ is :

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \times se(\hat{p}_1 - \hat{p}_2) = [(0.77 - 0.60) \pm 1.96(0.0545)]$$

$$= [0.17 \pm 0.1068] = [0.0632, 0.2768]$$

- We are 95% sure that the true proportion difference between the hospitals is between 6.32% and 27.68%
- The population fatality rate for Hospital 1 is between 6.32% and 27.68% *higher* than that of Hospital 2.

Proportion Testing: Liver Transplants

Assume two hospitals routinely perform liver transplants and Hospital 2 asserts that their patients have a *significantly lower* fatality rate. At a 5% significance level, do you believe the claim?

Hospital 1:

- $n_1 = 100, x_1 = 77$
- $\hat{p}_1 = \frac{77}{100} = 0.77$

Hospital 2:

- $n_1 = 200, x_1 = 120$
- $\hat{p}_2 = \frac{120}{200} = 0.60$

Pooled standard error:

$$se(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{0.77(0.23)}{100} + \frac{0.6(0.4)}{200}} = 0.0545$$

Proportion Testing: Liver Transplants

USC School of Business

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α**.
3. Calculate **test statistic** value.
4. Determine **critical value(s)**.
5. Decide whether to **reject H_0** and interpret the statistical result.

1. $H_0: p_1 - p_2 \leq 0$ vs $H_a: p_1 - p_2 > 0$
2. $\alpha = 0.05$
3. $z = \dfrac{(0.77 - 0.60) - 0}{\sqrt{\frac{0.77(0.23)}{100} + \frac{0.60(0.40)}{200}}} = 3.1193$
4. Reject H_0 if $z > z_{0.05} = 1.645$
5. Since $z = 3.1193 > 1.645$, we can **reject** H_0 and conclude Hospital 2 does have a lower rate.