Chapter 10: Inferences for two samples

Lianming Wang

University of South Carolina

January 4, 2023

Examples of studies involving two samples

- To compare the academic performance of between college male and female students.
- To evaluate whether our second mid-term exam is harder than the first one.
- To compare the popularity of two political candidates.
- To investigate whether men are more likely to work out regularly than women.

Identify the target parameter

- $\mu_1 - \mu_2$: mean difference; difference in average
- $p_1 - p_2$: difference between proportions, percentages, fractions, rates
- σ_1^2/σ_2^2: ratio of variances; difference in variability or spread

Sections 10.1-10.2. Comparing two-sample means: independent sampling

Part I: assuming large sample sizes in the two samples.

A $100(1-\alpha)\%$ confidence interval for $(\mu_1 - \mu_2)$ is

$$\begin{aligned} & (\bar{x}_1 - \bar{x}_2) \pm z_{\frac{\alpha}{2}}\, \sigma_{(\bar{x}_1 - \bar{x}_2)} \\ = \quad & (\bar{x}_1 - \bar{x}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \\ \approx \quad & (\bar{x}_1 - \bar{x}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}. \end{aligned}$$

Hypothesis Tests

One-Tailed Test

$H_0 : \mu_1 - \mu_2 = D_0$

$H_a : \mu_1 - \mu_2 > D_0$ (or $H_a : \mu_1 - \mu_2 < D_0$)

Test statistic: $z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \approx \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

Rejection Region: $z > z_\alpha$ (or $z < -z_\alpha$ when $H_a : \mu_1 - \mu_2 < D_0$)

Two-Tailed Test

$H_0 : \mu_1 - \mu_2 = D_0$

$H_a : \mu_1 - \mu_2 \neq D_0$

Test statistic: $z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \approx \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

Rejection Region: $z > z_{\frac{\alpha}{2}}$ or $z < -z_{\frac{\alpha}{2}}$

Conditions required for valid large-sample inferences about $(\mu_1 - \mu_2)$.

(1) The two samples are randomly selected in an independent manner from the two target populations.

(2) The sample sizes, n_1 and n_2 are both large. (That is, $n_1 \geq 30$ and $n_2 \geq 30$.)

Example 1. An investigation of repair times for two kinds of photocopying equipment produced the following results.

Type	Number of repairs	Mean	Standard deviation
1	60	84.2	19.4
2	60	91.6	18.8

a). At the .01 level of significance, do the mean repair times of these two types of equiptments appear to differ? Also calculate p-value for this test.

b). Construct 99% confidence interval for the mean difference of repair times between the two types of equipments. Interpret the result.

c). What assumptions are needed for the above inferences valid?

Part II: inference on two samples with small sample sizes

Assuming equal variances, i.e., $\sigma_1^2 = \sigma_2^2$.

A $100(1-\alpha)\%$ confidence interval for $(\mu_1 - \mu_2)$ is

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

where

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

and $t_{\frac{\alpha}{2}, n_1+n_2-2}$ is a critical value from the t distribution with $(n_1 + n_2 - 2)$ degrees of freedom.

Hypothesis Tests

One-Tailed Test

$H_0 : \mu_1 - \mu_2 = D_0$

$H_a : \mu_1 - \mu_2 > D_0$ (or $H_a : \mu_1 - \mu_2 < D_0$)

Test statistic: $t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

Rejection Region: $t > t_{\alpha, n_1+n_2-2}$ (or $t < -t_{\alpha, n_1+n_2-2}$ when $H_a : \mu_1 - \mu_2 < D_0$)

Two-Tailed Test

$H_0 : \mu_1 - \mu_2 = D_0$

$H_a : \mu_1 - \mu_2 \neq D_0$

Test statistic: $t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

Rejection Region: $|t| > t_{\frac{\alpha}{2}, n_1+n_2-2}$

Conditions required for valid small-sample inferences about $(\mu_1 - \mu_2)$.

(1) The two samples are randomly selected in an independent manner from the two target populations.

(2) Both sampled populations have distributions that are approximately normal.

(3) The population variances are equal. (That is, $\sigma_1^2 = \sigma_2^2$.)

Comments

- The two-sample t-test is a powerful tool for comparing population means when the assumptions are satisfied.

- When the sample sizes are equal (that is, $n_1 = n_2 = n$), the assumption of equal variances can be relaxed.

- Even when sample sizes are not equal and at the same time the variances are not equal, an approximate small-sample confidence interval or test can be obtained by modifying the number of degrees of freedom associated with t critical values or t test statistic. (Skip this part.)

Example 2: The following are scores on a standardized test for two independent random samples of students from two high schools.

School A: 78, 84, 81, 78, 76, 83, 79, 75, 85, 81
School B: 85, 75, 83, 87, 80, 79, 88, 94, 87, 82

a). At the .05 level, do it appear that school A has a lower mean score than school B? Also calculate p-value for this test.

b). Construct 95% confidence interval for the difference between the mean scores for the two schools. Interpret the result.

c). What assumptions are needed to ensure the validity of the above inference?

Section 10.4 Comparing Two Population Means: Paired Difference Experiments

A $100(1-\alpha)\%$ confidence interval for a paired mean difference $(\mu_d = \mu_1 - \mu_2)$ is

Large Sample:

$$\bar{x}_d \pm z_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n_d}}$$

Small Sample:

$$\bar{x}_d \pm t_{\frac{\alpha}{2}} \frac{s_d}{\sqrt{n_d}}$$

where $t_{\frac{\alpha}{2}}$ is based on $(n_d - 1)$ degrees of freedom.

Hypothesis Tests

One-Tailed Test	Two-Tailed Test
$H_0: \mu_1 - \mu_2 = D_0$	$H_0: \mu_1 - \mu_2 = D_0$
$H_a: \mu_1 - \mu_2 > D_0$ (or $H_a: \mu_1 - \mu_2 < D_0$)	$H_a: \mu_1 - \mu_2 \neq D_0$
Large Sample	
Test statistic: $z = \frac{\bar{x}_d - D_0}{s_d/\sqrt{n_d}}$	Test statistic: Same as left
RR: $z > z_\alpha$ (or $z < -z_\alpha$ when $H_a: \mu_1 - \mu_2 < D_0$)	RR: $\lvert z \rvert > z_{\frac{\alpha}{2}}$
Small Sample	
Test statistic: $t = \frac{\bar{x}_d - D_0}{s_d/\sqrt{n_d}}$	Test statistic: Same as left
RR: $t > t_{\alpha, n_d - 1}$ (or $t < -t_{\alpha, n_d - 1}$ when $H_a: \mu_1 - \mu_2 < D_0$)	RR: $\lvert t \rvert > t_{\frac{\alpha}{2}, n_d - 1}$

Conditions required for valid **large-sample** inferences about $\mu_d = (\mu_1 - \mu_2)$.

(1) A random sample of differences is selected from the target population of differences.

(2) The sample size n_d is large (i.e. $n_d \geq 30$).

Conditions required for valid **small-sample** inferences about $\mu_d = (\mu_1 - \mu_2)$.

(1) A random sample of differences is selected from the target population of differences.

(2) The population of differences has a normal distribution.

Example 3: An experiment was conducted to examine whether blood-pressure levels can be consciously lowered by persons trained in biofeedback exercises. Data collected on six subjects trained in these exercises produced the following.

Subject	1	2	3	4	5	6
Before	137	201	167	150	173	169
After	130	181	150	153	163	160

Conduct an appropriate hypothesis test to address the research concen at the 0.05 level of significance.

Section 10.5: Comparing Two Population Proportions: Independent Sampling

A $100(1-\alpha)\%$ confidence interval for a proportion difference $(p_1 - p_2)$ is

$$\begin{aligned} & (\hat{p}_1 - \hat{p}_2) \pm z_{\frac{\alpha}{2}}\, \sigma_{(\hat{p}_1 - \hat{p}_2)} \\ = \quad & (\hat{p}_1 - \hat{p}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \\ \approx \quad & (\hat{p}_1 - \hat{p}_2) \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}. \end{aligned}$$

Hypothesis Tests

One-Tailed Test	Two-Tailed Test
$H_0 : p_1 - p_2 = 0$	$H_0 : p_1 - p_2 = 0$
$H_a : p_1 - p_2 > 0$ (or $H_a : p_1 - p_2 < 0$)	$H_a : p_1 - p_2 \neq 0$
Test statistic: $z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1}+\frac{1}{n_2})}}$, where $\hat{p} = \frac{x_1+x_2}{n_1+n_2}$.	Test statistic: Same
RR: $z > z_\alpha$ (or $z < -z_\alpha$ when $H_a : p_1 - p_2 < 0$)	RR: $\lvert z \rvert > z_{\frac{\alpha}{2}}$

Conditions required for valid large-sample inferences about $(p_1 - p_2)$.

(1) The two samples are randomly selected in an independent manner from the two target populations.

(2) The sample sizes, n_1 and n_2 are both large. (That is, $n_1\hat{p}_1 \geq 15$, $n_1(1-\hat{p}_1) \geq 15$, $n_2\hat{p}_2 \geq 15$, and $n_2(1-\hat{p}_2) \geq 15$.)

Example 4: With the rapid growth in legalized gambling in the United States, there is concern that the involvement of youth in gambling activities has changed. University of Minnesota professor Randy Stinchfield compared the rates of gambling among Minnesota public school students. Based on survey data, the following table shows the percentage of ninth-grade boys who gambled weekly or daily on any game (e.g., cards, sports betting, lotteries) for the two years:

	1992	1998
Number of ninth-grade boys in survey	21,484	23,199
Number who gambled weekly/daily	4,684	5,313

Are the percentages of ninth-grade boys who gambled weekly or daily on any game in 1992 and 1998 significantly different?

Section 10.6. Comparing Two Population Variances: Independent Sampling

A $100(1-\alpha)\%$ confidence interval for σ_1^2/σ_2^2 is

$$\left(\left(\frac{s_1^2}{s_2^2}\right)\left(\frac{1}{F_{\alpha/2,df_n=n_1-1,df_d=n_2-1}}\right),\left(\frac{s_1^2}{s_2^2}\right)F_{\alpha/2,df_n=n_2-1,df_d=n_1-1}\right),$$

where $F_{\alpha/2,df_n=n_1-1,df_d=n_2-1}$ is the upper $\alpha/2$ quantile of the F distribution with numerator degree of freedom (df) (n_1-1) and denominator df (n_2-1), and $F_{\alpha/2,df_n=n_2-1,df_d=n_1-1}$ is the upper $\alpha/2$ quantile of the F distribution with numerator df (n_2-1) and denominator df (n_1-1).

Background knowledge I

- When $X_1, X_2, \dots, X_n$ are randomly sampled from a normal distribution $N(\mu, \sigma^2)$,

$$\sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi^2_{df=n-1}.$$

- That is,

$$\frac{n-1}{\sigma^2} s^2 \sim \chi^2_{df=n-1},$$

where s^2 is the sample variance of X's.

Background knowledge II

- **Definition of F distribution**: If X_x^2 and X_y^2 are independent χ^2 random variables with n_x and n_y as degrees of freedom, then $\frac{X_x^2/n_x}{X_y^2/n_y}$ follows the F distribution with $df_n = n_x$ and $df_d = n_y$. That is,

$$\frac{X_x^2/n_x}{X_y^2/n_y} \sim F_{df_n=n_x, df_d=n_y}.$$

- For two independent normal random samples: $X_1, X_2, \ldots, X_{n_x}$ and $Y_1, Y_2, \ldots, Y_{n_y}$,

$$\frac{s_x^2/s_y^2}{\sigma_x^2/\sigma_y^2} \sim F_{df_n=n_x-1, df_d=n_y-1}.$$

F-test for equal population variance

One-Tailed Test	Two-Tailed Test
$H_0 : \sigma_1^2 = \sigma_2^2$ $H_a : \sigma_1^2 > \sigma_2^2$ (or $H_a : \sigma_1^2 < \sigma_2^2$)	$H_0 : \sigma_1^2 = \sigma_2^2$ $H_a : \sigma_1^2 \neq \sigma_2^2$
Test statistic: $F = \frac{s_1^2}{s_2^2}$ (or $F = \frac{s_2^2}{s_1^2}$ when $H_a : \sigma_1^2 < \sigma_2^2$)	Test statistic: $F = \frac{\text{Larger sample variance}}{\text{Smaller sample variance}}$
RR: $F > F_\alpha$	RR: $F > F_{\alpha/2}$

Conditions required for valid inferences about σ_1^2/σ_2^2.

(1) The two samples are random and independent.

(2) Both sampled populations are normally distributed.

Note: the F test is sensitive to the normality assumption.

Example 5: Two automated filling processes are used in the production of automobile paint. The target weight of each process is 128.0 fluid oz (1 gallon). There is little concern about the process population mean fill amounts (no complaints about under/overfilling on average). However, there is concern that the population variation levels between the two processes are different. To test this claim, industrial engineers took independent random samples of n1 = 24 and n2 = 24 gallons of paint and observed the fill amounts. The sample standard deviations were reported $s_1 = 0.109$ and $s_2 = 0.127$, respectively.

Conduct an appropriate test for equal variances for the two populations of fill amounts from the two method at 0.10 level of significance.