STAT1201 - Summer Semester 2022

Lecture 7 - Comparing two populations

Dr. Wasanthi Thenuwara

20/12/2022

Comparing two population (group) means

- In lecture 5, we discussed how to perform hypothesis test and develop confidence intervals for a single population
- In scientific research, it is important, in many occassions to compare two populations or two groups
- Examples
  - Comparing effectiveness of two drugs.
  - Comparing the effect of caffeinated and decaffeinated drinks on pulse rate.
  - Comparing the nitrogen content of two lakes.
  - Comparing the tomato yields for two types of fertiliser.

Comparing two population (group) means - Learning objectives

In this lecture, you will learn how to

- Perform hypothesis test to compare means of the two independent populations using t-distribution (i.e: two sample t-test).
- Estimate Confidence Interval (CI) for the difference between the means of the two independent populations.
- Perform hypothesis test to compare proportions of the two independent populations using Z-distribution.

Two sample t-test

- This is a hypothesis test to compare the means of two independent populations using t-distribution.
- Following assumptions are made
  - The data follow normal distribution in each population.
  - The two samples are independent.
  - Each observation is a random sample from their respective populations.

Two sample t-test cont. . . (definitions and notations)

We will use the following notations and definitions in this lecture, unless otherwise stated.

Two populations –> Population 1 and Population 2

μ_1 –> Population 1 mean; μ_2 –> Population 2 mean

σ_1 –> Population 1 sd; σ_2 –> Population 2 sd

n_1 –> Sample 1 size; n_2 –> Sample 2 size

$\bar{x}_1$ –> Sample 1 mean; $\bar{x}_2$ –> Sample 2 mean

s_1 –> Sample 1 sd; s_2 –> Sample 2 sd

Two sample t-test cont. . . (null and alternative hypotheses)

To compare two population means, the null and alternative hypotheses is written as follows according to the nature of the experiment or study.

Case 1: to test whether means of the two populations are different

$H_0 : \mu_1 = \mu_2$ Vs $H_1 : \mu_1 \neq \mu_2$

Case 2: to test whether mean of the population 1 is greater than the mean of the population 2

$H_0 : \mu_1 = \mu_2$ Vs $H_1 : \mu_1 > \mu_2$

Case 3: to test whether mean of the population 1 is less than the mean of the population 2

$H_0 : \mu_1 = \mu_2$ Vs $H_1 : \mu_1 < \mu_2$

Two sample t-test cont. . . (test statistic)

To perform the hypothesis test, we need to calculate the test statistic like we did in one sample tests.

The test statistic follows the standard format in hypothesis testing.

$$t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{se(\bar{x}_1 - \bar{x}_2)}$$

This t-statistc follows a t-distribution.

Note that sample statistic here is the difference in the two sample means. Therefore, we need to find the standard error of the difference in the two sample means.Two different approaches will be used to calculate the standard error of the difference in the two sample means depending on the assumptions that we make about the unknown population standard deviations. We will discuss this in detail using examples.

Two sample t-test cont. . . (assumptions on unknown population standard deviations)

When we do the two sample t-test, we will consider two cases based on the assumptions made on unknown population standard deviations.

Case 1: Assume that unknown population standard deviations are NOT equal (i.e. $\sigma_1 \neq \sigma_2$).

Case 2: Assume that unknown population standard deviations are equal (i.e. $\sigma_1 = \sigma_2$). We use pooled t-test.

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

We use the following example to discuss the two sample *t*-test.

An experiment is conducted to test whether average breath holding time is different for young males and females. The breath holding times (in seconds) was measured for two groups of randomly selected male and female students aged between 18 to 20 years. The summary statistics are as follows.

Group	Sample_size	Sample_mean	Sample_SD
Male	10	50.19seconds	17.91seconds
Female	10	26.18seconds	7.29seconds

Note: It is not necessary to use the equal sample sizes. You can design the experiment with unequal sample sizes. However, the two sample sizes should not be substantially different.

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

Continue with the example.

We use F to denote females amd M to denote males. Then using standard notations,

$n_F = 10$; $n_M = 10$; $\bar{x}_F$ =26.18; $\bar{x}_M = 50.19$;

$s_F = 7.29$; $s_M = 17.91$

Looking at the sample statistics, it is very reasonable to assume that unknown population standard deviations are not equal.

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

Continue with the example.

The null and alternative hypotheses in this test can be written as follows.

$H_0 : \mu_M = \mu_F$ Vs $H_1 : \mu_M \neq \mu_F$

The t test statistic is given by,

$$t_{stat} = \frac{(\bar{x}_M - \bar{x}_F) - (\mu_M - \mu_F)}{se(\bar{x}_M - \bar{x}_F)}$$

This t-statistc follows a t-distribution. What are the degrees of freedom of the t-distribution?

df = min(n_M - 1, n_F - 1). This is a very conservative approximation to the real distribution (conservative means here that confidence intervals will probably be wider and hypothesis test will give a less significant p-value).

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

Continue with the example.

To calculate the test statistic we need to know the standard error of the sample statistic. That is, we need to find a value for $\text{se}(\bar{x}_M - \bar{x}_F)$.

Since the two populations are independent, we can include two separate sample variances to calculate the $\text{se}(\bar{x}_M - \bar{x}_F)$.

$$\text{se}(\bar{x}_M - \bar{x}_F) = \sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}$$

Substituting values,

$$\text{se}(\bar{x}_M - \bar{x}_F) = \sqrt{\frac{17.91^2}{10} + \frac{7.29^2}{10}}$$

$$\text{se}(\bar{x}_M - \bar{x}_F) = 6.114836$$

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

$$t_{stat} = \frac{(\bar{x}_M - \bar{x}_F) - (\mu_M - \mu_F)}{se(\bar{x}_M - \bar{x}_F)}$$

$$t_{stat} = \frac{(50.19 - 26.18) - (0)}{6.114836}$$

$$t_{stat} = 3.926516$$

We now need to find the p-value to make a conclusion.

since the df = min(n_M - 1, n_F - 1), the corresponding df of the t-distribution to find the p-value is (10-1) = 9.

Considering the t_9 distrbution, $\frac{p-value}{2}$ can be found considering the area beyond the t_{stat}.

Uisng R; $\frac{p-value}{2}$ = 1 - pt(3.926516, df=9).

$\frac{p-value}{2}$ = 1 - 0.9982618 = 0.0017382

p-value = 2*0.0017382 = 0.0034764

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

p-value $= 0.0034764 < 0.01$

We have strong evidence to conclude that the mean breath holding times are different between young males and females.

Poll Question 1

Consider the example we just discussed. Now the researcher uses 10 males and 9 females for her experiment. What is the degrees of freedom to test whether average breath holding time is different for young males and females?

a) 9

b) 10

c) 8

d) 18

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

The degrees of freedom for the t-distribution was based on the minimum of the two sample sizes. This was easy to use. However, this approximation is very coservative. The Welch (Welch 1936), gives a better degrees of freedom to the real t-distribution of the t-statistic.

We use R to perform the Welch t-test.

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

The Welch *t*-test

Consider the example again.

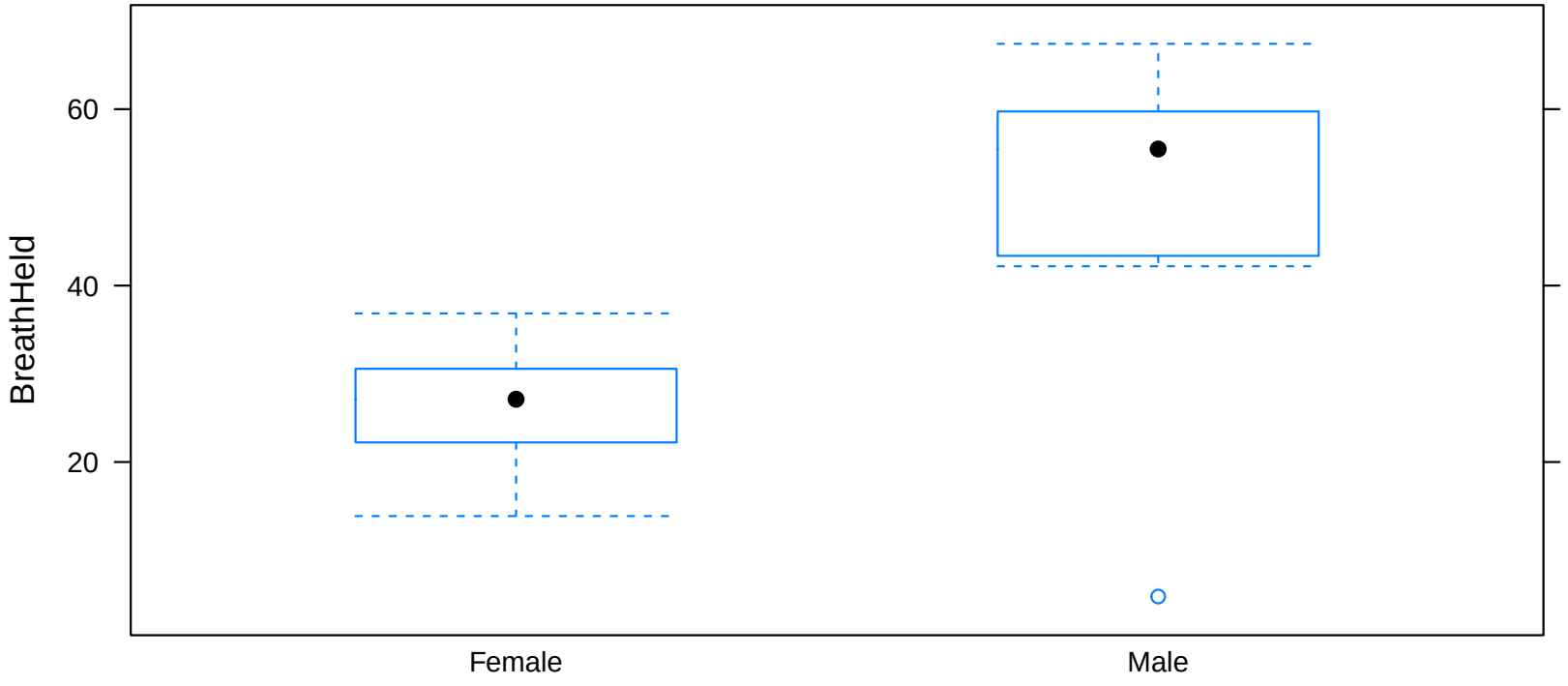
An experiment is conducted to test whether average breath holding time is different for young males and females. The breath holding times (in seconds) was measured for two groups of randomly selected male and female students aged between 18 to 20 years.The sample data is recorded in "M7Breath.csv" file.

First, we will draw a side-by-side box plot of breath holding time by sex for sample data to see the spread of the two distributions. This is a good initial step, if you are doing a two sample *t*-test for your research project.

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

The Welch t*-test*

```
breath = read.csv("M7Breath.csv")
library(lattice)
bwplot(BreathHeld ~ Sex, data=breath)
```



Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

The Welch t-test

We want to test whether the mean breath holding time is different for males and females.

$H_0 : \mu_M = \mu_F$ Vs $H_1 : \mu_M \neq \mu_F$

```
t.test(BreathHeld ~ Sex, data=breath)
```

```
    Welch Two Sample t-test

data:  BreathHeld by Sex
t = -3.9287, df = 11.888, p-value = 0.002039
alternative hypothesis: true difference in means is not equ
95 percent confidence interval:
 -37.34272 -10.68128
sample estimates:
mean in group Female   mean in group Male
              26.181                50.193
```

Two sample t-test - Case 1 (Assume $\sigma_1 \neq \sigma_2$)

Poll Question 2

The appropriate null and alternative hypotheses to test whether average breath holding time for young males is longer than that of young females are:

a) $H_0 : \mu_M = \mu_F$ Vs $H_1 : \mu_M \neq \mu_F$

b) $H_0 : \overline{x}_M = \overline{x}_F$ Vs $H_1 : \overline{x}_M > \overline{x}_F$

c) $H_0 : \mu_M = \mu_F$ Vs $H_1 : \mu_M > \mu_F$

d) $H_0 : \mu_M = \mu_F$ Vs $H_1 : \mu_M < \mu_F$

Two sample t-test - Case 1 - (Assume $\sigma_1 \neq \sigma_2$)

Now suppose that you want to test whether the mean breath holding time for young males is longer than that of for young females.

$H_0 : \mu_M = \mu_F$ Vs $H_1 : \mu_M > \mu_F$

```
t.test(BreathHeld ~ Sex, data=breath,
       alternative = "less")
```

```
    Welch Two Sample t-test

data:  BreathHeld by Sex
t = -3.9287, df = 11.888, p-value = 0.001019
alternative hypothesis: true difference in means is less th
95 percent confidence interval:
     -Inf -13.1102
sample estimates:
mean in group Female   mean in group Male
              26.181                50.193
```

Two sample t-test - Case 1 - (Assume $\sigma_1 \neq \sigma_2$)

The Welch t-test

Notice the alternative hypothesis in the test results. Even though our alternative hypothesis was $H_1 : \mu_M > \mu_F$, in the t.test() function, we wrote alternative = "less". This is because, the R function treats females as the first sample and males as the second sample following the alphabetical order.

In other words, the alternative hypothesis can be written as

$H_1 : \mu_M > \mu_F$ or $H_1 : \mu_F < \mu_M$ to test whether the mean breath holding time for young males is longer than that of for young females.

Two sample t-test - Case 2 (Assume $\sigma_1 = \sigma_2$) - Pooled t-test

If we assume that the unknown population standard deviations of the two independent populations are equal (i.e.$\sigma_1 = \sigma_2$), we can combine the standard deviations and use a common standard deviation (i.e. a pooled standard deviation) in the two sample t-test. This is called the Pooled two sample t-test (or Pooled t- test).

We first need to calculate pooled variance to be used in the Pooled t-test. The Pooled variance is denoted by S_p^2.

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

Following the standard format for the test statistic,

$$t_{stat} = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{se(\bar{x}_1-\bar{x}_1)}$$

where

$$\text{se}(\bar{x}_1 - \bar{x}_2) = \sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}$$

Two sample t-test - Case 2 (Assume $\sigma_1 = \sigma_2$) - Pooled t-test

Then the t-test statistic can be written as;

$$t_{stat} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}}$$

This t_{stat} has a t-distribution with df $= (n_1 + n_2 - 2)$

Two sample t-test - Case 2 (Assume $\sigma_1 = \sigma_2$) - Pooled t-test - Example

A researcher is interested to test whether males aged between 18 - 20 years are taller, on average, compared to the females in the same age group. The random samples of males and females aged between 18 to 20 years were chosen and their heights in cm were measured. The summary measures are as follows.

Group	Sample_size	Sample_mean	Sample_SD
Male	25	172.26cm	6.3cm
Female	20	167.32cm	6.1cm

$n_F = 20$; $n_M = 25$; $\bar{x}_F$ =167.32; $\bar{x}_M = 172.26$;

$s_F = 6.1$; $s_M = 6.3$

Two sample t-test - Case 2 - Pooled t-test - Example

The null and alternative hypotheses are;

$H_0 : \mu_M = \mu_F$ Vs $H_1 : \mu_M > \mu_F$

$$t_{stat} = \frac{(\bar{x}_M - \bar{x}_F) - (\mu_M - \mu_F)}{\sqrt{S_p^2(\frac{1}{n_M} + \frac{1}{n_F})}}$$

First calculate S_p^2.

$$S_p^2 = \frac{(n_M - 1)s_M^2 + (n_F - 1)s_F^2}{(n_M - 1) + (n_F - 1)}$$

$$S_p^2 = \frac{(25 - 1)6.3^2 + (20 - 1)6.1^2}{(25 - 1) + (20 - 1)}$$

$$S_p^2 = 38.59419$$

$$t_{stat} = \frac{(172.26 - 167.32) - (0)}{\sqrt{38.59419(\frac{1}{25} + \frac{1}{20})}}$$

$$t_{stat} = \frac{4.94}{1.863727}$$

$$t_{stat} = 2.650603$$

Two sample t-test - Case 2 - Pooled t-test - Example cont. . .

The corresponding df $= (25 + 20 - 2) = 43$

This is a one tail (upper tail) test. The p-value is the area beyond the test statistics in the t distribution with 43 degrees of freedom.

We can find p-value using R,

p-value = 1 - pt(2.6506, df = 43) or

p-value = pt(-2.6506, df=43)

p-value = 0.00561

Thus, p-value < 0.01

Coclusion: We have strong evidence to conclude that mean height of males aged 18-20 years is greater than the mean height of females in the same age group. That is, males aged 18-20 years are taller, on average than that of females.

Confidence Interval (CI) for the difference between the means of the two independent populations (i.e. CI for (μ_1 - μ_2)

We can calculate confidence intervals for $(\mu_1 - \mu_2)$

The standard formula for estimating CIs is as follows.

sample statistic $\pm$ MOE

In the case of the difference in two population means $(\mu_1 - \mu_2)$, CI is wriiten as

$(\bar{x}_1 - \bar{x}_2) \pm t^* \ .\ \text{se}(\bar{x}_1 - \bar{x}_2)$ —- (1)

t^* depends on the level of confidence and the degrees of freedom in the t distribution

Confidence Interval (CI) for the difference between the means of the two independent populations [i.e. CI for (μ_1 - μ_2)]

Again we consider two cases like we did in hypothesis tests before.

Case 1: $\sigma_1 \neq \sigma_2$

Case 2: $\sigma_1 = \sigma_2$

Standard errors and degrees of freedom of the t distribution depend on Case 1 and Case 2.

CI for the difference in the means of two populations

Case 1: $\sigma_1 \neq \sigma_2$

Consider breath holding time example again. Estimate 95% CI interval for the difference in the mean breath holding times between males and females. Consider the minimum of the two sample sizes to find the degrees of freedom.

$n_F = 10$; $n_M = 10$; $\bar{x}_F$ =26.18; $\bar{X}_M = 50.19$; $s_F = 7.29$; $s_M =$ 17.91

Recall that $\text{se}(\bar{x}_M - \bar{x}_F) = \sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}$

$\text{se}(\bar{x}_M - \bar{x}_F) = \sqrt{\frac{17.91^2}{10} + \frac{7.29^2}{10}} = 6.114836$

df = min(n_M - 1, n_F - 1). Thus df = 10 - 1 = 9

Uisng R, t^* = qt(0.975, df=9). Thus, $t^* = 2.262157$

95% CI for (μ_M - μ_F) is; (50.19 -26.18) $\pm$ 2.262157*6.114836

24.01 $\pm$ 13.83272 —> (10.18, 37.84)cm

CI for the difference in the means of two populations

Case 1: $\sigma_1 \neq \sigma_2$

95% CI for (μ_M - μ_F) is (10.18, 37.84)

That is, we are 95% confident that the difference in the population mean breath holding time of males and females lies between 10.18 seconds and 37.84 seconds.

Try yourself

What is the value of t^* to estimate the 90% CI for (μ_M - μ_F)?

.

.

.

.

CI for the difference in the means of two populations

Case 1: $\sigma_1 \neq \sigma_2$

Poll Question 3

What is the R function to be used to find the t^* to estimate 99% CI in the example we just discussed?

a) qt(0.995, df=9)

b) qt(0.99, df=9)

c) qt(0.975, df=9)

d) qt(0.995, df=10)

CI for the difference in the means of two populations

Case 2: $\sigma_1 = \sigma_2$. Consider the height example again.

Estimate 90% CI for the difference in the mean heights between males and females.

$n_F = 20$; $n_M = 25$; $\bar{x}_F$ =167.32; $\bar{x}_M = 172.26$; $s_M = 6.3$; $s_F = 6.1$

We should use pooled variance to calculate the standard error here. Recall that

$$S_p^2 = \frac{(n_M-1)s_M^2+(n_F-1)s_F^2}{(n_M-1)+(n_F-1)} = 38.59419$$

$$\text{se}(\bar{x}_M - \bar{x}_F) = \sqrt{S_p^2(\frac{1}{n_M} + \frac{1}{n_F})}$$

$$\text{se}(\bar{x}_M - \bar{x}_F) = 1.863727$$

CI for the difference in the means of two populations cont. . .

Case 2: $\sigma_1 = \sigma_2$

df $= 25 + 20 - 2 = 43$

Using R: t^* = qt(0.95, df=43) —> $t^* = 1.681071$

90% CI for(μ_M - μ_F)

(172.26 - 167.32) $\pm$ 1.681071*1.863727 —> 4.94 $\pm$ 3.133057

(1.81cm, 8.07cm)

Interpret yourself.

CI for the difference in the means of two populations cont. . .

Poll Question 4

The 95% confidence interval for the difference between mean heights of males and females aged 18-20 years lies between 1.18cm and 8.70cm. The margin of error used to estimate this CI is:

a) 7.52

b) 3.76

c) 1.18

d) 4.35

Paired t - test

The hypothesis testing procedures discussed previously enabled you compare the differences in the means of two independent populations. However, you come across situations such that populations are related. In such situations, paired t-test should be used.

See Week 4 tutorial Part A question.

Comparing two independent population (group) proportions

Assumptions

- Each observation in the sample are randomly selected from their respective populations
- Populations follow binomial distributions
- Both *np* and *n(1-p)* are greater than 5 as to use the normal approximation for binomial distribution

Comparing two independent population (group) proportions - Example

A researcher is interested to test the effectiveness of nicotine inhaler on smoking reduction. The following table shows sustained reduction in smoking for two groups of smokers after four months of inhaler used. The first group was given nicotine inhaler, and the second group was given placebo inhaler.

Smoking_Status	Nicotine_Inhaler	Placebo_Inhaler
Reduction	32	18
No Reduction	68	82
Total	100	100

Comparing two independent population (group) proportions - Example cont. . .

Define:

$\hat{p}_1$ = Proportion of smokers in the sample who sustain a reduction in smoking using nicotine inhaler

$\hat{p}_2$ = Proportion of smokers in the sample who sustain a reduction in smoking using placebo inhaler

p_1 = Population proportion of all smokers who sustain a reduction in smoking using nicotine inhaler

p_2 = Population proportion of all smokers who sustain a reduction in smoking using placebo inhaler

Notes

p_1 and p_2 are unknown population parameters.

$\hat{p}_1$ and $\hat{p}_2$ are sample statistics that we can find using the given data.

Comparing two independent population (group) proportions - Example cont...

Smoking_Status	Nicotine_Inhaler	Placebo_Inhaler
Reduction	32	18
No Reduction	68	82
Total	100	100

$\hat{p_1} = \frac{32}{100} = 0.32$

$\hat{p_2} = \frac{18}{100} = 0.18$

The researcher is interested to test whether a nicotine inhaler is more effective in sustaining a reduction in smoking over a placebo inhaler. Thus the null and alternative hypotheses are

$H_0 : p_1 = p_2$ Vs

$H_1 : p_1 > p_2$

We use to z-test to test this hypothesis.

Comparing two independent population (group) proportions - Example cont...

$H_0 : p_1 = p_2$ Vs

$H_1 : p_1 > p_2$

The z-test statistic is given by,

$$z_{stat} = \frac{(\hat{p_1} - \hat{p_2}) - (p_1 - p_2)}{se(\hat{p_1} - \hat{p_2})}$$

$$z_{stat} = \frac{(\hat{p_1} - \hat{p_2}) - (p_1 - p_2)}{\sqrt{(\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2})}}$$

That is, the z-test statistic follows a standard normal distribution (or Z-distribution).

$$z_{stat} = \frac{(0.32 - 0.18) - (0)}{\sqrt{(\frac{0.32(1-0.32)}{100} + \frac{0.18(1-0.18)}{100})}}$$

$$z_{stat} = \frac{0.14 - 0}{0.06043178}$$

$$z_{stat} = 2.316662$$

Comparing two independent population (group) proportions - Example cont. . .

$H_0 : p_1 = p_2$ Vs $H_1 : p_1 > p_2$

$z_{stat} = 2.3166662$

Find the p-value now.

Since this is an upper tail test, the p-value is the area beyond the test statistic.

Using R:

p-value = 1 - pnorm(2.316662)

p-value = 1- 0.9897389

p-value = $0.01026 < 0.05$

Conclusion: We have moderate evidence to conclude that nicotine inhaler is more effective in sustaining a reduction in smoking over a placebo inhaler.

Comparing two independent population (group) proportions

Poll Question 5

If the alternative hypothesis in our nicotine inhaler example was $H_1 : p_1 \neq p_2$ what would be the p-value? (Hint: When the alternative hypothesis was $H_1 : p_1 > p_2$, the p-value was 0.01026).

a) 0.01026

b) 0.00513

c) 0.02052

d) Cannot be determined

Confidence Interval (CI) for the difference between the proportions of the two independent populations (i.e. CI for $(p_1 - p_2)$

We can calculate confidence intervals for $(p_1 - p_2)$

The standard formula for estimating CIs using Z-distribution is as follows.

sample statistic $\pm$ Z*.SE(sample statistic)

In the case of the difference in two population proportions $(p_1 - p_2)$, CI is written as

$(\hat{p}_1 - \hat{p}_2) \pm Z^* \ . \ \text{se}(\hat{p}_1 - \hat{p}_2)$

Z^* depends on the level of confidence.

CI for $(p_1 - p_2)$ cont. . .

Consider the example that we used for hypothesis testing. Estimate 95% CI for the difference between the population proportions (i.e. 95% CI for $p_1 - p_2$)

$(\hat{p}_1 - \hat{p}_2) \pm Z^* \cdot \text{se}(\hat{p}_1 - \hat{p}_2)$

$(\hat{p}_1 - \hat{p}_2) \pm Z^* \sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}$

Z^* can be found using R function qnorm(0.975)

$(0.32 - 0.18) \pm 1.959964 \sqrt{(\frac{0.32(1-0.32)}{100} + \frac{0.18(1-0.18)}{100})}$

$0.14 \pm 1.959964 * 0.06043178$

0.14 ± 0.1184441 —-> (0.0216, 0.2584)

We are 95% confident that the difference in the population proportions of sustaining in reduction in smoking for smokers who were given a nicotine inhlaler and placebo inhaler is between 2.16% and 25.84%.

CI for $(p_1 - p_2)$ cont. . .

Test yourself.

What is the value for Z^* if you estimate the 90% CI for $(p_1 - p_2)$?

.

.

CI for $(p_1 - p_2)$ cont. . .

Poll Question 6

Consider the nicotine inhaler example. Without doing any calculation, the 99% CI for $p_1 - p_2$ will be

a) wider than 95% CI

b) narrower than 95% CI

c) same as the 95%

d) between 95% CI and 90% CI

Next . . .

Module 8 - Statistical Models

Thursday, 22 December 2022 at 12:00

Zoom meeting ID: 818 1453 7986