Discussion 4

8/27/2024

Dan Chen

Two Sample t tests theory

KEY IDEA

The ***central limit theorem*** implies that the distribution of sample means is normal when the population distribution is normal, or is approximately normal when the sample size is large.

VALIDITY CONDITIONS

The quantitative variable should have a symmetric distribution in both groups or you should have at least 20 observations in each group and the sample distributions should not be strongly skewed.

What does strongly skewed mean? If the data is strongly skewed, how large does the sample size need to be to compensate for a skewed distribution?
See Discussion 4.Rmd

Normal distribution theory

X ~ N(u, V) means X is a normal or Gaussian random variable with mean equal to **u** and variance equal to **V** (variance is just the square of the standard deviation)

If **X ~ N(u1, V1)** and **Y ~ N(u2, V2)**, then **X+Y ~ N(u1 + u2, V1 + V2)** and also

X-Y ~ N(u1 - u2, V1 + V2)

The mean of the sum or difference of two random variables is equal to the sum or difference of their means.

The variance of the sum or difference of two **normal** random variables is equal to the sum of their variances

Two Sample t tests theory

Suppose we have two samples **y** = (y_1, y_2, y_3, …, y_{n1}) and **a** = (a_1, a_2, a_3, …, a_{n2}) that satisfy the validity conditions, then the sample mean of each sample follows a normal distribution

$\bar{y}$ ~ N(u1, V1), $\bar{a}$ ~ N(u2, V2)

We know that the standard error of **$\bar{y}$** should be equal to **$sd_y/\sqrt{(n1)}$**, so **V1 = $(sd_y)^2/n1$**

Standard error of **$\bar{a}$** is equal to **$sd_a/\sqrt{(n2)}$**, so **V2** = **$(sd_a)^2/n2$**

($\bar{y}$ - $\bar{a}$) ~ N(u1 - u2, V1 + V2)

The standard error of **($\bar{y}$ - $\bar{a}$)** is equal to the square root of **V1 + V2** which is equal to

$\sqrt{[(sd_y)^2/n1 + (sd_a)^2/n2]}$

Two Sample t tests theory

This is how why the validity conditions are important for using the theory based formula.

Only if the sample mean of both samples are normally distributed can you calculate the standard error.

$$t = \frac{statistic - hypothesized\ value\ under\ Ho}{SE} = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Two Sample t tests theory

Calculating the p value:

- If your sample size is large, treat your t score as a z score and use pnorm()
- If you assume the true standard deviation of both samples is identical you can use pt() with degrees of freedom equals n1 + n2 - 2
- If you make no assumptions about equal standard deviation, then you can approximate the degrees of freedom as

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

Practice Problem (6.3.21)

A psychology study (Rutchick, Slepian, and Ferris, 2010) investigated whether using a red pen causes people to assign lower scores than using a blue pen. A group of 128 students in an introductory undergraduate psychology class were asked to grade an eighth grader's essay on a scale of 0–100. Half of the student graders were randomly assigned to use a red pen while grading the essay, and the other half were randomly assigned to use a blue pen. The researchers reported that the group using a red pen gave an average score of 76.20, and the SD of the scores was 12.29. For the group using blue pens, the average score was 80.00, and the SD of the scores was 9.36.

a. State the null and alternative hypotheses to be tested.

b. Without having access to the 128 scores that were assigned to the essays, can you feel comfortable in applying a two-sample t-test using the summary statistics? Explain.

c. What additional information is needed to apply a two-sample t-test using the summary statistics?

d. Make a reasonable guess for the missing information and conduct a two-sample t-test. Report the values of the test statistic and p-value.

Practice Problem (6.3.21)

A psychology study (Rutchick, Slepian, and Ferris, 2010) investigated whether using a red pen causes people to assign lower scores than using a blue pen. A group of 128 students in an introductory undergraduate psychology class were asked to grade an eighth grader's essay on a scale of 0–100. Half of the student graders were randomly assigned to use a red pen while grading the essay, and the other half were randomly assigned to use a blue pen. The researchers reported that the group using a red pen gave an average score of 76.20, and the SD of the scores was 12.29. For the group using blue pens, the average score was 80.00, and the SD of the scores was 9.36.

f. Is it legitimate to draw a cause-and-effect conclusion based on these data and your test result? Explain why or why not.

g. What would you advise the researchers about how broadly they can generalize the conclusions of this study? For example, would you feel comfortable in generalizing the results of this study to eighth-grade teachers or to college professors? Explain.