THE USE OF TWO-SAMPLE t-TEST IN THE REAL DATA

Article *in* Advances and Applications in Statistics · September 2022

DOI: 10.17654/0972361722071

Mowafaq Muhammed Tawfeeq Al-kassab
Tishk International University

Aveen Hameed Majeed
Tishik International University

Advances and Applications in Statistics
© 2022 Pushpa Publishing House, Prayagraj, India
http://www.pphmj.com
http://dx.doi.org/10.17654/0972361722071
Volume 81, 2022, Pages 13-22 **P-ISSN: 0972-3617**

THE USE OF TWO-SAMPLE *t*-TEST IN THE REAL DATA

Mowafaq Muhammed Al-Kassab and Aveen Hameed Majeed

Department of Mathematics Education
Faculty of Education
Tishk International University
Ebril, Iraq
e-mail: mowafaq.muhammed@tiu.edu.iq
avin.hameed97@gmail.com

Abstract

The *t*-test is one of the most commonly used statistical methods. It was developed and accredited by William Gosset, Karl Pearson and R. Fisher in the 19th century. The test was further developed to the two-sample test (Snedecor and Cochran [10]) which is used to determine whether two populations are equal. A common application of the two-sample *t*-test is to test whether a process or treatment is superior to a current process or treatment. In this research, using the two-sample *t*-test, a comparison between the students in the three grades of the Department of Mathematics Education, Tishk International University, is made to see whether there is a significant difference between the grade point averages (GPAs) for the second, third and fourth grades,

Received: August 10, 2022; Accepted: September 15, 2022

2020 Mathematics Subject Classification: 62P99.

Keywords and phrases: grade points average (GPA), descriptive statistics, pooled variance, two-sample *t*-test, significant difference.

How to cite this article: Mowafaq Muhammed Al-Kassab and Aveen Hameed Majeed, The use of two-sample *t*-test in the real data, Advances and Applications in Statistics 81 (2022), 13-22. http://dx.doi.org/10.17654/0972361722071

Published Online: September 22, 2022

in addition to gender. The result showed that there is no significant difference on the average scores between grade 4 and grade 3, also no significant difference between grade 4 and grade 2, but there is a significant difference between grade 3 and grade 2. Similarly, there is no significant difference on the average scores due to gender.

1. Introduction

The two-sample *t*-test is a statistical test for comparing two groups' means. In pain research, it is one of the most commonly utilized statistical hypothesis tests (Yim et al. [12]). A test of significance is a formal technique for comparing observable facts to a claim, also known as a hypothesis, whose veracity is being determined. The claim is a declaration concerning a parameter, such as the population percentage P or the population mean μ. A significance test's results are given as a probability that indicates how well the data and the claim coincide. Xu et al. [11] mentioned that a significance test begins with a thorough description of the statements being compared. The null hypothesis is the assertion that is examined by a statistical test, denoted by (H_0) the purpose of the test is to determine how strong the evidence is against the null hypothesis. The null hypothesis is frequently stated as no difference. The alternative hypothesis denoted by (H_a) is when (H_0) is not true. The alternative is one-sided if a parameter is lower or higher than the null hypothesis value. If it specifies that the parameter differs from the null value, then it is two-sided (it could be either smaller or larger). The significant level (α) is therefore found at both ends of the curve. Half of (α) is on the upper end, while the other half is on the lower end. As a result, there is a low and a high cutoff value. The *p*-value is calculated by multiplying the area to the right of z by 2 in two-sided scenarios. The *p*-value may only be doubled after that. "A *t*-test is an inferential statistic that is used to see if there is a significant difference in the means of two groups that are connected in some way" (Maüll Miquel [7]). The two-sample *t*-test can be used to determine whether the means of two groups are equal. The *t*-test is a parametric test: it is founded on an assumption that the underlying population from which the samples are drawn is nearly normally distributed.

It is reasonably robust to failures in this assumption but should be treated with caution as the true distribution deviates from normal. The test considers means of two samples and tests the null hypothesis that the two samples are drawn from populations with the same mean [5, 6]. Variants are provided based on what is known about the underlying populations' variation. The *t*-test is a more cautious form the *z*-test, which relies on the Central Limit Theorem's confirmation that the sampling distribution of the mean is normal for large samples. The *t*-test can be used with fewer samples to compensate for the distribution distortion caused [5, 6].

2. Literature Review

In 1908, William Gosset, an Englishman publishing under the pseudonym student, developed the *t*-test. The *t*-test is a parametric test: it is founded on an assumption that the underlying population from which the samples are drawn is nearly normally distributed. The test is considered to be one of the most commonly used statistical methods. This method was developed and is accredited to Gosset, Karl Pearson, in the 19th century (Edwards and Fisher [2]). The method was further developed into the "two-sample test" (Snedecor and Cochran [10]). The key new statistical challenge for Gosset was that with such tiny samples, it was uncertain how well the sample standard deviation, s, represented the population standard deviation σ. This made it exceedingly difficult to detect if the two barleys differed much. A common application of the two-sample *t*-test is to test whether a new process or treatment is superior to a current process or treatment. Because comparing the means of two samples is such a typical experimental design, the student's *t*-test for two samples is mathematically equivalent to a one-way with two categories. McDonald [8] mentioned that "for the two-sample *t*-test, we need two variables. One variable defines the two groups. The second variable is the measurement of interest". Peck et al. [9] mentioned that two samples are deemed independent if the individuals or objects that make up one sample have no bearing on the individuals or subjects in the other sample. "The two-sample *t*-test should be used to compare the mean values of two samples in this scenario". He, also, mentioned that "if the observations in the first sample are coupled with some particular observations in the other sample, the samples are considered to be paired". The *t*-test is a useful statistical tool with moderate power for determining if there is a significant difference between two groups. The point at which *t* becomes significant is also influenced by sample size; the bigger the *n*, the lower the *t* necessary to become significant. Excessive repeating of the *t* test on the same dataset is a typical mistake in the usage of the *t*-test (Foster and Gerald [3]).

3. Derivation of the Two-sample *t*-test

Let $X_{1,1}, X_{1,2}, \ldots, X_{1,n1}$ and $X_{2,1}, X_{2,2}, \ldots, X_{2,n2}$ be two samples from two independent normal distributions with means μ_1, μ_2 and common variance σ^2. The unbiased estimators for the population means μ_1, μ_2 and the variance σ^2, respectively, are:

$$\bar{x}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_{i,j}, \tag{1}$$

$$s_j^2 = \frac{1}{n_j - 1}\sum_{i=1}^{n_j}(x_{i,j} - \bar{x}_j)^2, \quad j = 1, 2. \tag{2}$$

The maximum likelihood estimators can be written as (Heckert et al. [4]):

$$\hat{\mu}_1 = x_1, \quad \hat{\mu}_2 = x_2, \quad \hat{\sigma}^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}. \tag{3}$$

This yields the test statistic

$$t(x_1, x_2) = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)(\hat{\sigma}^2)}} \tag{4}$$

or, simply,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\hat{\sigma}^2}{n_1} + \frac{\hat{\sigma}^2}{n_2}}} \tag{5}$$

which has a *t*-distribution with $(n_1 + n_2 - 2)$ degrees of freedom (Chang and Pal [1]).

4. Application of the *t*-test

In this section, we present descriptive statistics for the data of this research, which is the GPA score of the students for three grades from Mathematics Education Department, Faculty of Education, Tishk International University, the second grade has nine students, seven female and two males, the third grade has sixteen students, ten of them females and six of them are males, and the fourth grade has nineteen students, ten females and nine males. A comparison between the students of the department is done to see whether there is a significant difference between the GPAs for the grades, and the gender. The two-sample *t*-test is used to see this significance. The data is given as in Table 1:

Table 1. The GPA scores of the students for the three grades

Grade 2	3.53	2.98	2.77	2.59	2.44	2.13	2.05	1.95	1.30
Grade 3	3.92	3.70	3.62	3.60	3.03	2.98	2.94	2.17	2.84
	2.72	2.70	2.66	2.60	2.53	2.44	2.32		
Grade 4	3.76	3.55	3.44	3.09	3.01	2.89	2.87	2.83	2.80
	2.78	2.76	2.67	2.58	2.31	2.30	2.04	1.81	1.23

The descriptive of the above data is given in Table 2:

Table 2. Descriptive of the scores of the students according to grades

Grades	*N*	Mean	SE mean	Minimum	Maximum	Median
Grade 2	9	2.416	0.217	1.300	3.530	2.440
Grade 3	16	2.923	0.131	2.170	3.920	2.780
Grade 4	18	2.724	0.138	1.230	3.760	2.800

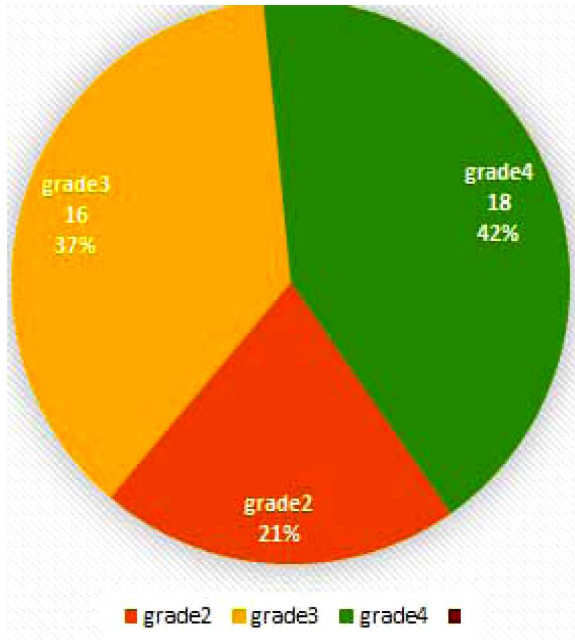

Figure 1. The percentage of students according to grades.

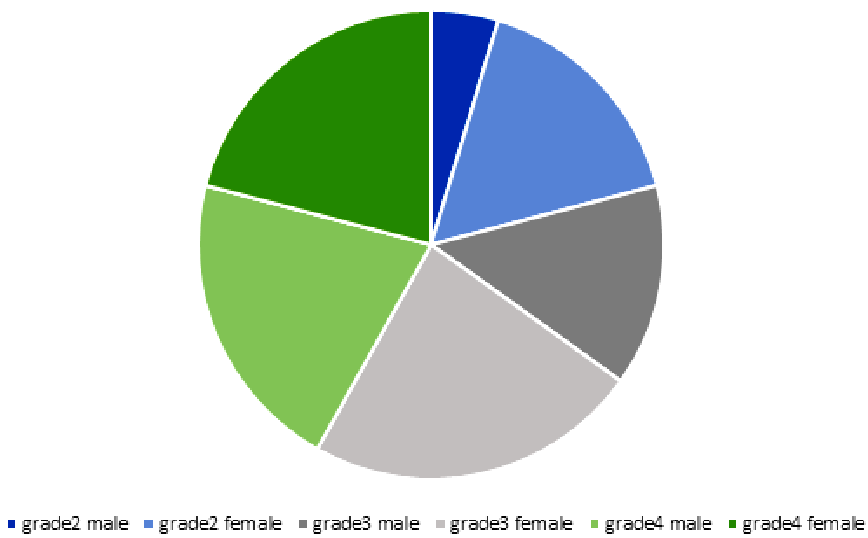

Figure 2. The percentage of students according to grades and gender.

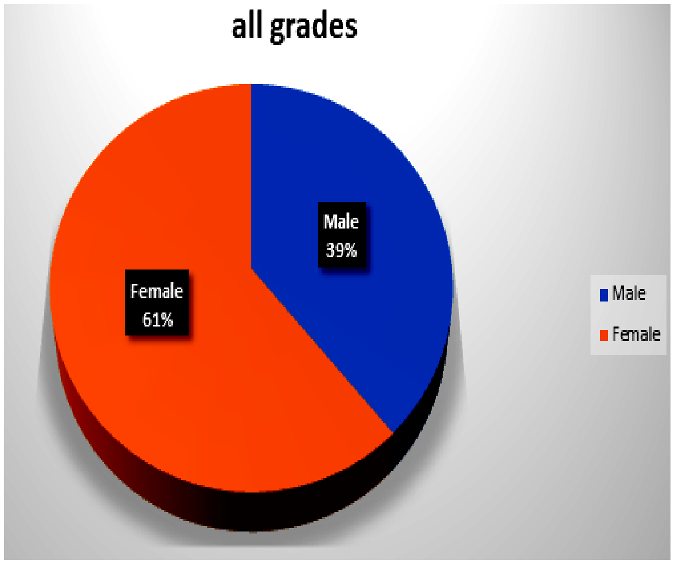

Figure 3. The percentage of students for all the grades according to gender.

Now, we will apply the two-sample *t*-test to see whether there is a significant difference between the GPA scores of the student according to their grades and according to their gender.

Two-sample *t*-test between grades 3 and 4

The null hypothesis is that there is no significant difference between the two grades 3 and 4, i.e., $H_0 : \mu_3 = \mu_4$.

Regarding to equations (1), (2), (3) and (5), we have the following table:

Table 3. Two-sample *t*-test for grades 3 and 4

Grade	*N*	Mean	Variance	Pooled variance	*t*-value	*p*-value
4th	19	2.724	0.36	0.324	1.03	0.310
3rd	16	2.923	0.28			

From the above table, we can see that there is no significant difference between the GPAs of the two grades. Therefore, we accept the null hypothesis.

Two-sample t-test between grades 2 and 3

The null hypothesis is that there is no significant difference between the two grades 2 and 3, i.e., $H_0 : \mu_2 = \mu_3$.

Regarding to equations (1), (2), (3) and (5), we have the following table:

Table 4. Two-sample t-test for grades 2 and 3

Grade	N	Mean	Variance	Pooled variance	t-value	p-value
2nd	9	2.416	0.424	0.327	2.135	0.0427
3rd	16	2.923	0.276			

From the above table, we can see that there is a significant difference between the GPAs of the two grades. Therefore, we reject the null hypothesis.

Two-sample t-test between grades 2 and 4

The null hypothesis is that there is no significant difference between the two grades 2 and 4, i.e., $H_0 : \mu_2 = \mu_4$.

Regarding to equations (1), (2), (3) and (5), we have the following table:

Table 5. Two-sample t-test for grades 2 and 4

Grade	N	Mean	Variance	Pooled variance	t-value	p-value
2nd	9	2.416	0.424	0.382	1.23	0.229
4th	19	2.724	0.36			

From the above table, we can see that there is no significant difference between the GPAs of the two grades. That is, we accept the null hypothesis.

Two samples t-test between females and males

The null hypothesis is that there is no significant difference between females and males, i.e., $H_0 : \mu_f = \mu_m$.

Regarding to equations (1), (2), (3) and (5), we have the following table:

Table 6. Two samples t-test for the females and males

Grade	N	Mean	Variance	Pooled variance	t-value	p-value
Male	17	2.803	0.246	0.368	0.607	0.547
Female	27	2.689	0.443			

From the above table, we can see that there is no significant difference between the GPAs of the two genders. That is, we accept the null hypothesis.

5. Conclusions

According to the GPA scores, there is no significant difference between grades 3 and 4, there is no significant difference between grades 2 and 4. There is a significant difference between grades 2 and 3. According to the GPA scores, there is no significant difference between males and females for the three grades altogether. The maximum GPA score is for grade 3, and the minimum GPA score is for grade 2.

References

[1] C.-H. Chang and N. Pal, A revisit to the Behrens-Fisher problem: comparison of five test methods, Comm. Statist. Simulation Comput. 37(6) (2008), 1064-1085.

[2] A. W. F. Edwards and R. A. Fisher, Statistical methods for research workers, Landmark Writings in Western Mathematics: Case Studies, I. Grattan-Guinness, ed., Elsevier, Amsterdam, 2005, pp. 1640-1940.
doi: 10.1016/B978-044450871-3/50148-0.

[3] S. Foster and K. Gerald, Review of the two sample t tests, Nurse Anesthesia 1(1) (1990), 38-40.

[4] N. A. Heckert, J. J. Filliben, C. M. Croarkin, B. Hembree, W. F. Guthrie, P. Tobias and J. Prinz, Handbook 151: NIST/SEMATECH e-Handbook of Statistical Methods, 2002.

[5] F. Bacchus, AIPS 2000 planning competition: The Fifth International Conference on Artificial Intelligence Planning and Scheduling Systems, AI Magazine 22(3) (2001), 47-56.

[6] F. Bacchus and M. Ady, Planning with resources and concurrency: a forward chaining approach, IJCAI, Vol. 1, 2001, pp. 417-424.

[7] A. Maüll Miquel, Functional stability of activated graphene-based electrodes after sterilization with ethylene oxide, Bachelor's thesis UPF, 2021.

[8] J. H. McDonald, Handbook of Biological Statistics, Vol. 2, Sparky House Publishing, Baltimore, MD, 2009.

[9] R. Peck, C. Olsen and J. Devore, Introduction to Statistics and Data Analysis, Cengage Learning, Boston, 2012.

[10] G. Snedecor and W. Cochran, Arc sine transformation for proportions, Statistical Methods, 8th ed., Iowa State University Press, Ames, 1989, pp. 289-290.

[11] M. Xu, D. Fralick, J. Z. Zheng, B. Wang, X. M. Tu and C. Feng, The differences and similarities between two-sample *t*-test and paired *t*-test, Shanghai Archives of Psychiatry 29(3) (2017), 184-188.

[12] K. H. Yim, F. S. Nahm, K. A. Han and S. Y. Park, Analysis of statistical methods and errors in the articles published in the Korean Journal of Pain, The Korean Journal of Pain 23(1) (2010), 35-41.