DOI: https://doi.org/10.24297/jam.v21i.9279

The Use of One Sample t-Test in the Real Data

Mowafaq Muhammed Al-Kassab

Department of Mathematics Education,Faculty of Education, Tishk International University, Erbil, Iraq

Mowafaq.muhammed@ tiu.edu.iq

Abstract:- The t-statistic is the ratio of the departure of the estimated value of a parameter from its hypothesized value to its standard error. The term "t-statistic" is abbreviated from "hypothesis test statistic". It was first derived as a posterior distribution in 1876 by Helmertand Lüroth. The purpose of this research is to study the t-test, especially the one sample t-test to determine if the sample data come from the same population. The grade points average (GPA) of the students for the second, third, and fourth grades of the Department of Mathematics Education, Tishk International University are used. The one sample t test is used to predict the GPA of the students for the second, third, and the fourth grades respectively, in addition to the overall average scores for the three grades. The 95% confidence interval for the true population average is also conducted.

Keywords: Great points average (GPA), Grades, Descriptive statistics, variance, One sample t-test, Significant difference, Confidence interval.

1. Introduction

The term "t-statistic" is abbreviated from "hypothesis test statistic". In statistics, the t-distribution was first derived as a posterior distribution, which is a type of conditional probability in Bayesian statistics. In common usage, the term posterior probability refers to the conditional probability of an event given which comes from an application of Bayes' theorem (Helmert, 1876). The t-distribution appeared in a more general form as Pearson Type IV distribution, the Pearson distribution is a family of continuous probability distributions (Pearson, 1894, Webb, Newton, & Cheng, 2013, Flannelly, Jankowski, & Flannelly, 2015), the t-distribution, also known as Student's t-distribution, gets its name from William Sealy Gosset (William, 1908), he pioneered small sample experimental design and analysis with an economic approach to the logic of uncertainty. Gosset published under the pen name Student and developed most famously Student's t-distribution originally, called Student's z and Student's test of statistical significance (William, 1908, Philip, 1984, Gerald, 2018).

The *t*-test is a statistical hypothesis test in which the test statistic follows a student's t-distribution under the null hypothesis, it includes many types. There is one sample t test, used to compare the mean of a sample with an assumed value of the population, the two samples independent t -test, which can be used when the two groups under comparison are independent of each other, and the paired t- test, which can be used when the two groups under comparison are dependent on each other (McDonald, 2009, Katherine, Flannelly, & Flannelly, 2018, Mishra, Pandey, Singh, Gupta, Sahu, Keshri , 2019).The one sample t-test will be treated in detail, it is used to compare the mean of a sample with an assumed value of the population, the population mean is not always known but it is sometimes hypothesized. For a valid test, we need data values that are: Independent (values are not related to one another), continuous, and obtained from simple random sample from the population (the population is assumed to be normally distributed) (Yim, Nahm, Han, & Park, 2010, Pagano, 2012, Jankowski, & Flannelly, 2015). The statistical hypotheses for one-sample t-tests take one of the following forms, depending on whether your research hypothesis is directional or not directional. (Pandey, 2015, Tae, 2015, Younis, 2015). In the example below μ refers to the population mean from which the study sample was drawn; μ_0 is replaced by the actual value of the population mean. The statistical hypotheses are identical to those used for one-sample z tests. The null hypothesis (H_0) and (two-tailed) alternative hypothesis (H_1) of the one sample t test can be expressed as (Heckert, Filliben, Croarkin, Hembree, Guthrie, Tobias, & Prinz, 2002): H_0: $\mu = \mu_0$ (the population mean is equal to the proposed population mean) versus H_1: $\mu \neq \mu_0$ (the population mean is not equal to the proposed population mean).

An important property of a test statistic is that its sampling distribution under the null hypothesis must be calculatable, either exactly or approximately, which allows p-value to be calculated (Weissman,2012, McDonald, 2014). We use the one-sample *t*-test when we collect data on a single sample drawn from a defined population (Sheynin, 1995). In this design, we have one group of subjects, collect data on these subjects and compare our sample statistic ($\bar{x}$) to the population parameter (μ). The population parameter tells us what to expect if our sample came from that population. If our sample statistic is very different (beyond what we would expect from sampling error), then our statistical test allows us to conclude that our sample came from a different population.

2. Method and Results

In statistical tests, the probability distribution of the statistic is important. When a sample $x_1, x_2, x_3, ..., x_n$ of size n is drawn from normal population with mean μ and variance σ^2, that is N (μ, σ^2), the distribution of the sample mean $\bar{X}$ has normal distribution N (μ, σ^2/n). Under the null hypothesis $\mu = \mu_0$, the distribution of statistic

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

should be standardized as a normal distribution N (0,1). When the variance of the population is not known, replacement with the sample variance s^2 is possible. In this case the statistic

$$\frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

follows a t distribution with *(n-1)* degrees of freedom (In statistics, the degrees of freedom associated with some quantity is the number of values that are free to vary), where

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

is the unbiased estimator of the population variance σ^2 and we write $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

To apply this test, we present the data and description of this data, which is the greater point average (GPA) for three grades of the department of mathematics education, faculty of education, Tishk International University. The second grade has nine students. The third grade has sixteen. The fourth grade has nineteen. A comparison between the students of the department is done to see whether there is a significant difference between the average GPA and proposed GPA for the grades, and the gender. The one sample t-test is used to see this significance. The data is given as in table (1) below:

Table - 1 : The GPA scores of the students for the three grades

Grade 2	3.53	2.98	2.77	2.59	2.13	2.05	1.95	1.30	
Grade 3	3.92	3.70	3.62	3.60	3.03	2.98	2.94	2.84	
	2.72	2.70	2.66	2.60	2.53	2.44	2.32	2.17	
Grade 4	3.76	3.55	3.44	3.09	3.03	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								

Table -2: Descriptive of the GPA scores of the students according to grades

grades	N	Mean	SE Mean	Minimum	Maximum	Median
G2	9	2.416	0.217	1.300	3.530	2.440
G3	16	2.923	0.131	2.170	3.920	2.780
G4	18	2.724	0.138	1.230	3.760	2.800

We see that the minimum GPA is for grade 2, and the maximum GPA is for grade 3. The percentages of the students according to their grades and gender are given in fig. 1 and fig. 2 respectively.

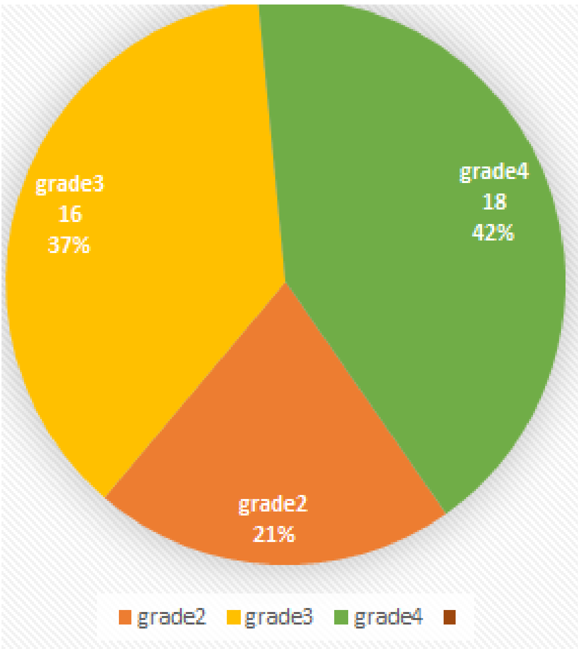

Figure-1: The percentage of students according to grades

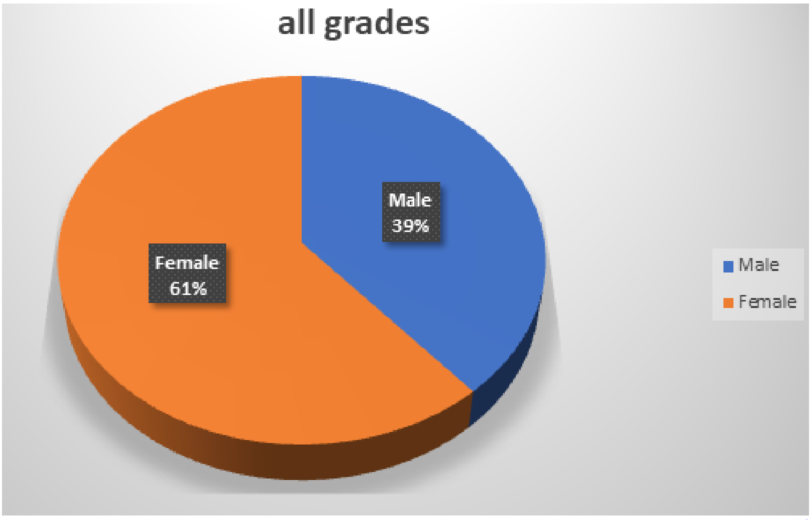

Figure-2: The percentage of students for all the grades according to gender

One sample t-test for Grade 2

We will apply the one sample t-test statistic to see whether there is a significant difference between the sample mean and the proposed population mean for grade 2 (2.5).

H_0. μ= 2.5

H_1. μ≠ 2.5

Table - 3: One sample t-test for grades 2

N	mean	Standard. Dev.	CI for μ	t-test	p-value
9	2.416	0.651	(1.915, 2.916)	-o.39	0.707

We see from table 3 that the null hypothesis H_0.μ= 2.5 is accepted, which means that the GPA score for grade 2 is 2.5.

One sample t-test for grade 3

We will apply the one sample t-test statistic to see whether there is a significant difference between the sample mean and the proposed population mean for grade 3 (2.5).

H_0. μ= 2.5

H_1. μ≠ 2.5

Table - 4: One sample t-test for grades 3

N	mean	Standard. Dev.	CI for μ	t-test	p-value
16	2.923	0.526	(2.643, 3.203)	3.22	0.006

We see from table 4 that the null hypothesis H_0.μ= 2.5 is rejected which means that the GPA for grade 3 is more than 2.5, the 95% confidence interval GPA score for grade 3 is between 2.6 and 3.2.

One sample t-test for grade 4

Now, we will apply the one sample t-test statistic to see whether there is a significant difference between the sample mean and the proposed population mean for grade 4 (2.5).

H_0. μ= 2.5

H_1. μ≠ 2.5

Table - 5: One sample t-test for grades 4

N	mean	Standard. Dev.	CI for μ	t-test	p-value
19	2.724	0.604	(2.433, 3.015)	1.64	0.124

We see from table 5 that the null hypothesis H_0.μ= 2.5 is accepted, which means that the GPA score for grade 4 is 2.5.

One sample t-test for all grades

We will apply the one sample t-test statistic to see whether there is a significant difference between the sample mean and the proposed population mean for all grades.

H_0. μ= 2.5

H_1. μ≠ 2.5

Table - 6: One sample t-test for all grades

N	mean	Standard. Dev	CI for μ	t-test	p-value
44	2.7332	0.6020	2.5501, 2.9162	3.57	0.014

We see from table 5 hat the null hypothesis H_0.μ= 2.5 is rejected, which means that the GPA score for mathematics department is more than 2.5. The GPA score 95% is located between 2.6 and 2.9.

3. Conclusion:

According to the GPA scores, we concluded that the GPA scores for grades 2 and 4 is 2.5. The GPA score for grade 3 is more than 2.5, its location is 95% between 2.6 and 3.2. The GPA score for all the grades together is 95% between 2.6 and 2.9.

References

1. Flannelly, K. J., Jankowski, K. R. B., & Flannelly, L. T. (2015). "Measures of variability in chaplaincy, health care, and related research". Journal of Health Care Chaplaincy, 21(3), 122–130. doi:10.1080/08854726.2015.105467

2. Gerald, B. (2018) "A Brief Review of Independent, Dependent and One Sample t-test" International Journal of Applied Mathematics and Theoretical Physics Vol. (4), No. 2, 50-54.

3. Gosset, W. S. (1908). "William Sealy Gosset." Biographical Encyclopedia of Mathematicians 1: 239.

4. Helmert, F. (1876). "Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers director Beobachtungen gleicher Genauigkeit." Astronomische Nachrichten 88: 113.

5. Heckert, N. A., Filliben, J. J., Croarkin, C. M., Hembree, B., Guthrie, W. F., Tobias, P., & Prinz, J. (2002). Handbook 151: NIST/SEMATECH e-Handbook of Statistical Methods.

6. Jankowski, K. R. B., & Flannelly, K. J. (2015). " Measures of central tendency in chaplaincy, healthcare, and related research". Journal of Health Care Chaplaincy, 21(1), 39–49. doi:10.1080/08854726.2014.989799.

7. Katherine R. B. Jankowski,Kevin J. Flannelly &Laura T. Flannelly. (2018). " The t-test: An Influential Inferential Tool in Chaplaincy and Other Healthcare Research" Journal of Health Care Chaplaincy. Vol. 24, No 1, https://doi.org/10.1080/08854726.2017.1335050

8. McDonald, J. H. (2009). Handbook of biological statistics. Vol. (2): sparky house publishing Baltimore, MD.

9. McDonald, J.H. (2014). "Handbook of Biological Statistics". Third Edition. University of Delaware. Baltimore, Maryland, U.S.A.: Sparky House Publishing

10. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. (2019) "Descriptive statistics and normality tests for statistical data". Ann Card Anaesth , Vol. 22, No.1, 67-72.

11. Pagano, R. P. (2012). "Understanding statistics in the behavioral sciences" (10th ed.). Belmont, CA: Wadsworth.

12. Pandey, R. M (2015). " Commonly Used t-tests in Medical Research". Journal of the Practice of Cardiovascular Sciences. Vol. 1. No. 2.

13. Pearson, K. (1894). "Contributions to the mathematical theory of evolution." Philosophical Transactions of the Royal Society of London. A 185: 71-110.

14. Philip J. B. (1984) “A Biographical Glimpse of William Sealy Gosset”. The American errors in the articles". The Korean journal of pain, 23(1), 35.