

TWO SAMPLE UNPAIRED T-TEST POWER CALCULATION USING SIMULATION

¹Prasanth Sambaraju

¹Independent Researcher, 500059, Hyderabad, INDIA,
e-mail: prashanth.kng1@gmail.com

Received: 22nd May; Accepted: 10th November

ABSTRACT

Statistical tests like t test are used to test the hypothesis for comparison of mean between two groups. The result from a t test is used to determine if there a significant difference between the two samples mean, which cannot be attributed to sampling error or to random occurrence. A t test is considered a parametric test, meaning test samples should meet assumptions of normality, equal variances and independence. The article aims to calculate the power of unpaired t test using simulation in Microsoft Excel.

Keywords: T test, Statistics, Simulation, Power, Effect size

1. INTRODUCTION

Statistical tests provide a mechanism for making quantitative decisions about a process. The goal is to determine whether the evidence is strong enough to reject a hypothesis about the process [1]. A t test is a statistical test, which is used to compare the means of two groups. T test are considered as parametric statistical methods, where probability distribution of probability variables are defined and inferences are made about the parameters of the distribution [2]. Unpaired t test is used to compare two population means. The test is performed to test the null hypothesis that the population means are equal. Power of a test is the probability that the test will reject the null hypothesis, when the null hypothesis is actually false. Power is given by $1 - \beta$, where β is the Type II error. [3]. Type II error is failure to reject a false null hypothesis. Excel doesn't output power value at the end of t test analysis. So to overcome this limitation, the t test is repeated many times after generating random numbers with specified mean and standard deviation. Power is obtained by the ratio of the number of times P value obtained from t test was less than alpha value (0.05) to the total number of times simulation was performed.

2. MATERIALS AND METHODS

Two groups, with five values per group were sampled from normal distribution with population mean value of 25 and 35, standard deviation value of 10 for both groups. The following formula was used in Excel to generate these values, where MEAN was 25 for one group, 35 for the second group and SD is 10 for both groups [4].

$$\text{NORM.INV}(\text{RAND}(), \text{MEAN}, \text{SD}) \quad (1)$$

2.1. Test Procedure in Excel

T.TEST function in Excel was used to perform unpaired t test in Excel. The function uses four parameters array1 (data for group1), array2 (data for group2), tails (1 for one-tailed and 2 for two-tailed distribution) and type (1 for paired, 2 for two-sample for equal variance and 3 for two-sample unequal variance) [5].

2.2. Equations Comparison of variances

F test in Excel was used to compare the variances between two groups. The test is performed with the null hypothesis that the variances of two population groups are equal. The function returns the two-tail test probability that the variances in both groups are significantly different. When the obtained *P* value is more than alpha level, two group variances are considered to be the same. The following formula was used in Excel, where array1 corresponds to group 1 and array2 corresponds to group 2 simulated values [6].

$$F. \text{TEST}(\text{array1}, \text{array2}) \quad (2)$$

The result from the *F* test is used to generate parameter value for the fourth argument in the *t* test procedure. If variances between both groups are same then a value of 2 is used, else value of 3 is used. The result obtained from *F* test, is subsequently used in *t* test calculation.

Table 1. Formulas used in simulation (Sheet name: simulation)

Cell address	Formula
G1	=25 (Mean 1)
G2	=35 (Mean 2)
F4 - F5	=10 (standard deviation value)
B3 - B7	=NORM.INV(RAND(),\$G\$1,\$G\$4) (random value from normal distribution with mean (25) and standard deviation (10)
C3 - C7	=NORM.INV(RAND(),\$G\$2,\$G\$5) (random value from normal distribution with mean (35) and standard deviation (10)
B8	=AVERAGE(B3:B7)
C8	=AVERAGE(C3:C7)
B9	=STDEV.S(B3:B7)
C9	=STDEV.S(C3:C7)
B11	=F.TEST(B3:B7,C3:C7) (compare variances for both groups)
M3	=0.05 (alpha value)
G7	=IF(M3>B11,3,2) (if alpha value is greater than P value, then outputs 3 else 2)
G8, K1	=T.TEST(B3:B7,C3:C7,2,G7) (two-tailed t test to compare mean between two groups, the value of fourth parameter value is dependent on F test result)
N3	=COUNTIF(K2:K10001,"<"&M3)/COUNT(K2:K10001) (count all the P values in K2:K10001 cells which are less than to 0.05 and then divide it by 10000)

2.3. Manual calculation for *F* test

F test is used to test the equality of variances between two population groups. The following assumptions must be satisfied before conducting the test

- i) Both populations are normally distributed
- ii) Samples are independent from each other
- iii) Samples are randomly drawn from the population

P value can be obtained by manually by first calculating the F value as the ratio of variances of two samples and then using F value as first parameter for F.DIST function in Excel. Each group has five samples, so degrees of freedom (df1 and df2) is four [7].

$$F = \frac{s_1^2}{s_2^2} \tag{3}$$

$$P \text{ value} = 2 * F.DIST(F \text{ value}, 4, 4, TRUE)$$

2.4. Manual calculation for unpaired t test

The test is performed to test the null hypothesis that the two population means (μ_1 and μ_2) are equal. The t test for two groups was performed as described below

- a) Five samples for both were generated from normal distribution in Excel using mean and standard deviation
- b) The difference between two sample means were calculated ($\bar{x}_1 - \bar{x}_2$)
- c) Pooled standard deviation (S_p) was calculated using the formula, where n_1, n_2 are the sample size and s_1 and s_2 are the sample standard deviation for group1 and group2 respectively. Standard error (SE) for difference for mean differences is then calculated using pooled standard deviation.

$$S_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \tag{4}$$

$$SE(\bar{x}_1 - \bar{x}_2) = S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

d) T -statistic is calculated using the following equation. This statistic follows a t -distribution with $n_1 + n_2 - 2$ degrees of freedom (df). Using the Excel function TDIST, P value is obtained [7].

$$T = \frac{\bar{x}_1 - \bar{x}_2}{SE(\bar{x}_1 - \bar{x}_2)} \tag{5}$$

$$P \text{ value} = TDIST(t \text{ value}, df, 2)$$

2.5. Simulation using Data table

The result obtained (Cell G8) corresponds to p -value obtained after one t test. The test can be repeated by using the data table function in Excel. Using fill series (Home → Fill → Series → Series in Columns → Step value 1 and Stop value 10000) numbers from 1 to 10000 are added in cells J2:J10001. After selecting cells J1 to K10001, the data table tab (Home → Data → What-If Analysis → Data Table) was clicked. In the Data Table dialog box, the row input cell was left empty whereas for column input, an empty cell in the current worksheet was selected. Different P values are filled in cells K2 to K10001. Each value corresponds to a two-tailed t test P value obtained from two simulated arrays values, which are randomly selected from normal distribution. The power of unpaired t test is calculated in cell N3. To prevent automatic recalculation of formulas in worksheet, the calculation option is set to manual (Home → Formula → Calculation Options → Manual). This will make sure that the formulas will calculate only when the F9 key (Function + F9) is pressed. Tab. 3 give the F test two-tailed probability that the variances of two groups was same or different at alpha level 5%. Tab. 3 shows the t test results for same and different variances.

Table 2. P value below and above α level at the end of 10000 simulations

F test	$\alpha = 5\%$	
P value (less than α)	5.09%	509
P value (greater than α)	94.91%	9492

Table 3. t test results from same and different variances

	t test		
3 (Different variances)	1.10%	3.95% of 10000	21.61%
2 (Same variances)	26.72%	96.05% of 10000	28.15%
Total	2782	100%	

2.6. 95% Confidence Intervals

The formula to estimate confidence intervals using normal approximation, where z is the $100 \cdot (1 - \alpha, \alpha = 5\%)^{\text{th}}$ percentile from a standard normal distribution [8], n is the number of simulations (10000) and \hat{p} is the power of unpaired t test obtained from simulation (cell N3). It can be observed that as the number of simulations increases, tighter confidence intervals are obtained.

$$\hat{p} - \frac{z_{\alpha}}{2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} < p < \hat{p} + \frac{z_{\alpha}}{2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \tag{6}$$

Table 4. Reported vs obtained - power and 95% confidence intervals

Power value	n	Upper Limit	Lower limit	Remarks
0.275	1000	0.304	0.248	Reported values [4]
0.287	10000	0.294	0.276	Obtained values (subject to change when repeated)

2.7. Factors affecting power

- a) Alpha level: The simulation was repeated using an alpha value of 0.1. It was observed that the power of the test increased to 0.420 (Sheet name: alpha).
- b) Sample size: The simulation was repeated using ten simulated values for each group. It was observed that the power of test the increased to 0.571 (Sheet name: sample size).
- c) Effect size: Effect size is the magnitude of difference between the two populations [9]. Cohen's d is used as measure of effect size for unpaired t test [3]. One approach to increase effect size was by increasing the mean difference value between the two groups (Sheet name: effect size)

$$d = \frac{|\bar{x}_1 - \bar{x}_2|}{S_{pooled}} \tag{7}$$

$$s_{pooled} = \sqrt{\frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{n_1 + n_2 - 2}}$$

\bar{x}_1 = Mean of group1 values

\bar{x}_2 = Mean of group2 values

n_1 = count of values group1

n_2 = count of values group2

s_1 = standard deviation of group1 values

s_2 = standard deviation of group2 values

Effect size as measured by Cohen's d was increased by decreasing the Mean1 value (cell: G2) to 20. After simulation it was observed that the power of unpaired t test increased to 0.549.

3. RESULTS AND DISCUSSION

The results obtained for power were comparable to that of reported results. Fig. 1 show the plot of cumulative distribution of P values at the end of simulation. The power value at alpha level (5%) can be visualized in the plot. Power can be varied by changing either alpha level, sample size or effect size.

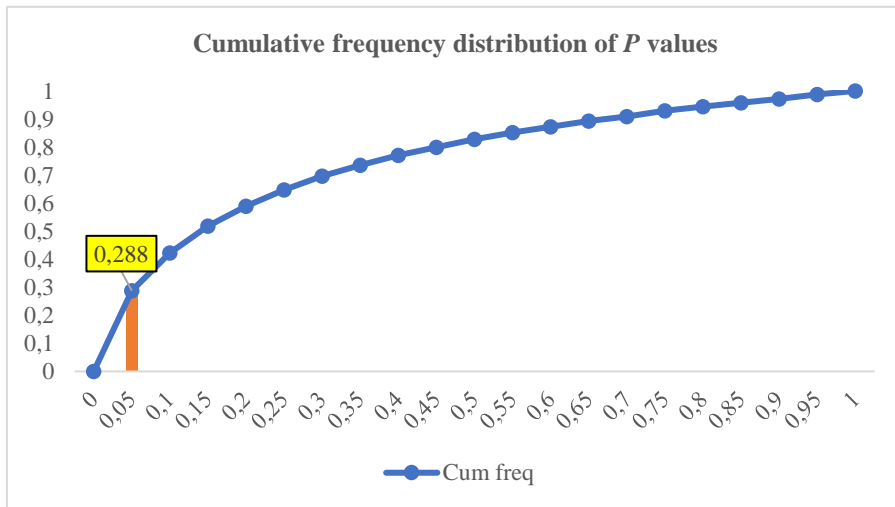


Figure 1. Plot of cumulative distribution of P values at end of simulation

4. CONCLUSIONS

Simulation using Microsoft Excel provides an easy way to understand power calculation for unpaired t test and also factors which affect it.

REFERENCES

- [1] What are statistical tests?
<https://www.itl.nist.gov/div898/handbook/prc/section1/prc13.htm#:~:text=A%20statistical%20test%20provides%20a,is%20called%20the%20null%20hypothesis> (2023.05.23)

- [2] Kim T.K. (2015): T test as a parametric statistic, *Korean Journal of Anesthesiology* 68(6), 540-546. <https://doi.org/10.4097/kjae.2015.68.6.540>
- [3] Nakagawa S., Cuthill I.C., (2007): Effect size, confidence interval and statistical significance: a practical guide for biologists, *Biological reviews* 82(4) 591-605. <https://doi.org/10.1111/j.1469-185X.2007.00027.x>
- [4] [GraphPad Prism 9 Statistics Guide - Monte Carlo example: Power of unpaired t test](#) (2023.05.22)
- [5] T.TEST function <https://support.microsoft.com/en-us/office/t-test-function-d4e08ec3-c545-485f-962e-276f7cbed055> (2023.05.22)
- [6] FTEST function <https://support.microsoft.com/en-us/office/f-test-function-4c9e1202-53fe-428c-a737-976f6fc3f9fd> (2023.05.22)
- [7] David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran, Michael J. Fry, Jeffrey W. Ohlmann, Essentials of Modern Business Statistics with Microsoft® Excel®, 8th edition, Cengage, Boston, 2019.
- [8] Morisette J.T., Khorram S. (1998): Exact Binomial Confidence Interval for Proportions. *Photogrammetric Engineering and Remote Sensing* 64(4), 281-283.
- [9] Hentschke H, Stüttgen M.C. (2011), Computation of measures of effect size for neuroscience data sets. *European Journal of Neuroscience* 34(12), 1887-1894. <https://doi.org/10.1111/j.1460-9568.2011.07902.x>