MONTE CARLO METHODS FOR COMPARING TWO POPULATION MEANS

N. Maglaperidze

Informatics and Engineering School, Georgian American University, Merab Aleksidze St. 10, 0160 Tbilisi, Georgia
nikoloz_maglaperidze@gau.edu.ge

(Received 25.05.2024; accepted 17.11.2024)

Abstract

Monte Carlo methods are widely used in modern scientific research and technology. Using Monte Carlo methods is the best way to deeply understand and study complex statistical methods. In this paper, we will use Monte Carlo methods to create sampling distributions, confidence intervals, and test hypotheses for comparing two population means for modified t-test allowing unequal variances.

Keywords and phrases: Two sample t test, P-Value, Confidence interval, Welch's test, stochastic simulation.

AMS subject classification (2010): 62F03, 62F25, 62J15, 68U20.

1 Introduction

Monte Carlo simulation is an effective method for approximating complex systems using random sampling. Monte Carlo methods generate random samples from a probability distribution related to the problem and use these samples to estimate the desired quantity or behavior. In this paper, we use Monte Carlo simulations [5] to sample from two normally distributed populations and compare their mean values. The first population has a mean of 20 and a standard deviation of 5, while the second has a mean of 20 and a standard deviation of 15. Figure 1 shows density plots of these distributions. The standard deviation of the second population are three times larger than the corresponding standard deviation of the first population. To compare the means of two populations during simulation, the Welch's test [1] for two samples employed. A sample of size ten was drawn from each normally distributed population using Monte Carlo simulation. During the calculations, 10,000 simulations were run. The sampling distribution of the Welch's statistic was developed using 10,000 simulations. We conducted statistical analyses and created visualizations using the R programming language [3].

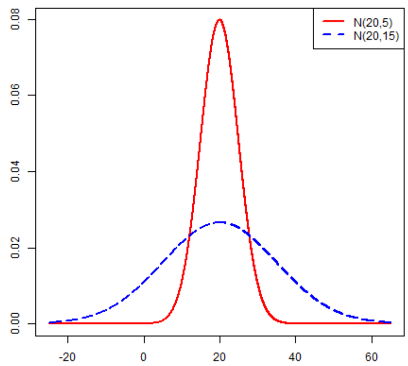

Figure 1: Two normal distributions with the same mean and with different standard deviation.

2 Welch's Test

Welch's t-test assumes that both groups of data are sampled from Gaussian populations, but does not assume those two populations have the same standard deviation [2].

Welch's test is based on the statistic

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}.$$

Where $\overline{x}_1$ and $\overline{x}_2$ are sample means, s_1 and s_2 are corrected sample standard deviations, and n_1 and n_2 are sample sizes in each group. Welch's test statistic follows a t-distribution. The degrees of freedom are equal to

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

3 Sampling Distribution of Welch's Test

The sampling distribution of a statistic represents the range of values that the statistic can take across all potential samples of identical size drawn from the same population. The sampling distribution depicts the theoretical pattern that would arise if we were to examine every conceivable sample of equivalent size from population [4].

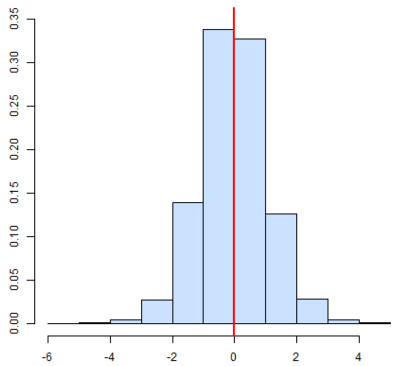

Figure 2: The histogram of the sampling distribution for Welch's statistic.

A set of ten observations was sampled from each population following a normal distribution (Figure 1) through Monte Carlo simulation. By selecting multiple samples from two populations that follow a normal distribution and computing Welch's statistic for each, we can generate the sampling distribution for Welch's statistic. Figure 2 displays the histogram of the sampling distribution for Welch's statistic [6]. The sampling distribution of a statistic is the distribution of values taken by the statistic in all possible samples of the same size from the same population. The values of Welch's statistic fluctuate across different samples, yet they are centered around the 0. While some samples exhibit Welch's statistic values below 0 and others above it, there is no consistent bias toward being consistently lower or higher. Figure 2 illustrates that the sampling distribution is symmetric.

4 P-Values and Confidence Intervals

For each of the 10,000 samples drawn from two normally distributed populations, we constructed 10,000 95 percent confidence intervals using Welch's test and calculated P-Values for the two-sided hypothesis. Out of the 10,000 confidence intervals, 9,507 (95%) encompassed the population parameter, and 493 (4.93%) of the P-Values were less than the 0.05 significance level [6]. Figure 3 displays a boxplot diagram of P-Values, clearly illustrating that a very small percentage of P-Values are less than 0.05.

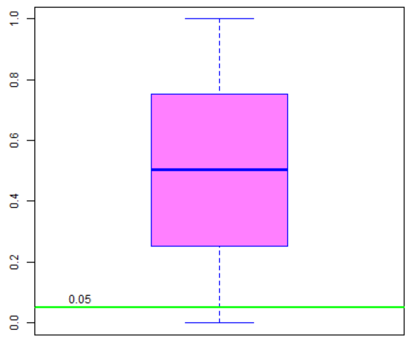

Figure 3: Boxplot diagram of P-Values.

References

1. Welch BL. The significance of the difference between two means when the population variances are unequal. *Biometrika* **29** (1938), no. 3/4 , 350-362.

2. Howell, D.C. *Statistical Methods for Psychology*. Wadsworth Publishing, 2012.

3. R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. URL https://www.R-project.org/.

4. Moore, D. S., McCabe, G. P., & Craig, B. A. *Introduction to the Practice of Statistics*. Sixth Edition. W. H. Freeman, New York, 2014.

5. Carsey, T. M., Harden, J. J. *Monte Carlo Simulation and Resampling Methods for Social Science*. Sage Publications, Inc, 2013.

6. Maglaperidze N. *Monte Carlo Methods for Comparing Two Population Means*. XXXVIII International Enlarged Sessions of the Seminar of Ilia Vekua Institute of Applied Mathematics, Tbilisi 2024.