Dean&Francis
ISSN 2959-6157

Galton Board Experiment: Proof of Central Limit Theorem

Junhe Gong

Beanstalk International Bilingual School, Beijing, China

*Corresponding author:
MaisyGong25sy@bibs.com.cn

Abstract:

This paper investigates the application of the Central Limit Theorem (CLT) using a Galton board experiment. The Galton board, which produces a binomial distribution of ball positions, serves as a practical model to test the CLT's assertion that the distribution of sample means approaches a normal distribution as the sample size increases. In this experiment, 30 independent trials were conducted, with 100 balls passing through 10 rows of pegs in each trial. The sample means for each trial were computed, and the resulting sampling distribution was analyzed. A normal distribution curve was fitted to the data, visually demonstrating alignment with CLT predictions. Additionally, statistical tests, including the Shapiro-Wilk test, were applied to rigorously evaluate the normality of the sample means, providing empirical validation for the theoretical model. The findings confirm the applicability of the CLT to the Galton board, as the sampling distribution closely follows a normal pattern, highlighting the theorem's generalizability even when the original data follows a binomial distribution.

Keywords: Central limit theorem; Bernoulli trial; Shapiro-Wilk test.

1. Introduction

The Central Limit Theorem (CLT) is one of the cornerstones of statistics. Standing along with key terms such as normal distribution, and sampling distribution, CLT is the first thing people should know when one enters the kingdom of statistics [1]. The CLT provides the foundation for many applications in probability, data analysis, and statistical inference, as it is an extremely important interdisciplinary concept. The theorem asserts as the sample size increases, the distribution of sample means will tend to follow a normal distribution, regardless of the shape of the population from which the samples are drawn. This powerful concept is an essential tool for researchers who seek to understand a glimmer of rationality in such a discipline filled with abstract symbols and uncertainty. To explore the practical implications of the CLT, the author will utilize the Galton board, a well-designed mathematics model that serves to visualize the CLT.

Understanding how the Central Limit Theorem manifests in real-world scenarios is critical, and the Galton board is the best example which serves as

ISSN 2959-6157

a strong indicator illustrating the statistical pattern. The Galton board is an upright board with evenly spaced nails driven into its upper half. The nails are arranged in staggered order. The lower half of the board is divided with vertical slats into a number of narrow rectangular slots [2]. Putting simply, the Galton board is a model that consists of three parts: one is the entry at the top, where one can put small beads into it; one is the pins in the placed middle that form in a triangular pattern, disturbing or deciding what position will the ball falls, and the last one is bins that settled in the bottom of the model, collecting every possible outcome. Also, one of the important rules assumed here is that the ball cannot move in any direction indefinitely without meeting a pin [3]. As balls accumulate, one can easily see the distribution of the balls as they are all demonstrated in the bins.

With such a useful tool and iron rule in statistics, this paper wants to verify the existence and usefulness of CLT by designing an experiment on the Galton board model to figure out if the CLT theorem is always present in statistics. The experiment will identify an experimental group and a control group based on statements included in the CLT and will test whether the result follows or refutes the CLT by using a statistical indicator.

To be specific, the paper will first focus on the sampling distribution of the sample means generated by the Galton board, which simulates a binomial distribution as balls randomly drop through a series of pegs. The importance of this research lies in showing how, despite the binomial nature of the Galton board, the sample means across multiple trials conform to a normal distribution as the number of trials and sample size increases. This has profound implications for statistical modeling, allowing researchers to use normal approximations in real-world situations even when the underlying distribution is not normal. After proceeding with the sample means collected from experiments, the tester will gather data and evaluate whether the CLT is applicable in this experiment.

2. Theorem involved in the Galton board

2.1 Bernoulli Trial and Binomial Distribution

Before some further explanation about how the Galton board testified the Central Limited theorem (CLT), it is extremely important to study the designation of the Galton board. Thus, the author will introduce two key concepts in the course of probability and statistics—the Bernoulli trial and binomial distribution. The definition of a Bernoulli trial is that a Bernoulli trial is a random experiment that satisfies three conditions. The first condition is there are only two possible outcomes which is success or failure; the second condition is that the probability constantly remains the same; and the third condition is independence, meaning the outcome of one trial does not affect the outcome of any other trials.

In statistics, a Bernoulli trial can be represented using a random variable X, which takes the value 1 or 0, in which 1 always means success and 0 stands for failure. For instance, flipping coins is one of the typical examples that belong to the Bernoulli trial, as there are only two possible outcomes--head or tail (condition 1); the probability of fair coins that drop on one side always remains at 50 percent (condition 2); and each trial follows the rule of independence, since the probability a coin flip on one side does not influenced by previous trial(condition 3).

Likewise, the Galton board also follows the Bernoulli trial, since each pin in the Galton board can be seen as a binary, random decision point. With that point, the observer can see the event follows such three conditions: when a ball strikes a pin, it has two possible outcomes: left or right (condition1); the probability that the ball goes to right or left is constantly 50 percent for a symmetrical, fair device (condition 2); and each outcome occurs independently of previous pin (condition 3). However, in the Galton board, things get a little different because it is a vertical device and not as easy as flipping coins; for a single trial in the Galton board, the experimenter defines that failure in one trial is defined as the ball going left, and success defined as the ball goes left. In this case, the mathematical representation will be changed into X_i being 1 or 0, where 1 means the ball moves to the right; 0 means the ball moves to the left.

Once clarified each pin in the Galton board can be seen as a Bernoulli trial, it is not hard to understand the Galton board is a set of Bernoulli trials, since for each row, the bead will strike the pin once, leading people toward to the research of binomial distribution. The formula of the binomial distribution is given:

$P(x=k)=\binom{n}{k}p^k(1-p)^{n-k}$. Within the formula, K is the number of successes (the number that the ball goes to the right); n refers to the number of total trials (number of rows in this case); and p refers to the probability of success in a single trial($1-p$) refers to the probability of failure in a single trial.

The Bernoulli trial can be treated as a special scenario when $n=1, k=1$ .Thus, it is easy to understand that when n and k get higher, the Galton board becomes a set of Bernoulli trials that can be defined using the binomial distribution. For a single trial, noting n in the Galton board experiment stands for the number of rows because the number of rows means the number of decisions that a single ball has to make, and k can be seen as the number of balls chosen right-hand side within n rows (identical with what experimenter defined in a Bernoulli trial.

After knowing the Galton board follows the binomial distribution, one can calculate the probability of success by giving n times in k success. In other words, this formula can used to describe the probability of a bead in the Galton board experiment dropping into a specific bin.

Specifically, for example, when the number of rows (n)=4, and k=1, one can calculate the probability that the bead goes into bead 1 is $\binom{4}{1}0.5^1 0.5^3 = 25\%$. As the author observed, since both values of p an1-p are equal to 50%, $p^k(1-p)^{n-k}$ can be rewritten into p^n. In that case, this means for the same n, the probability that the ball falls into bins are all determined by the binomial coefficient, in which the observer can see the ratio between the probabilities of the bead falling into different bins patterns like $1:4:6:4:1(n=4)$ and $1:5:10:10:5:1(n=5)$.

2.2 Central Limit Theorem

The CLT is one of the most fundamental theorems in statistics, as it explains how the distribution of sample means approaches a normal distribution as the sample size becomes large. A normal distribution also known as a bell curve describes the shape of a univariate data set when it is symmetric, bell-shaped, and has a peak around the mean of the data set ("Measures of Shape", n.d.) [4]. In general, the CLT states that for a sufficiently large sample of independent and random variables, the sum (or average) of these variables will approach normality as the sample size increases [5].

To further explain this in the context of the CLT, the concept of a sampling distribution is needed. A sampling distribution refers to the probability distribution of a statistic, such as the sample mean or standard deviation. The CLT asserts that as the sample size increases, the sampling distribution of the sample mean tends to approximate a normal distribution, even when the population from which the samples are drawn is not normally distributed. For instance, in the context of the Galton board experiment, each bead's passage through the rows of pegs follows a binomial distribution. The CLT says that if testers repeatedly run the experiment and calculate the sample mean of ball positions from multiple trials, the distribution of these sample means—the sampling distribution—will increasingly resemble a normal distribution as the sample size and number of trials grow larger. That is, even though the original distribution of the sample in the Galton board follows precisely the binomial distribution, the sample mean, or the sampling distribution, will always stay in the pattern of normal distribution, regardless the original model is binomial.

2.3 Formula of CLT

In this formula $\overline{X_n}\, N(\mu, \frac{\sigma^2}{n})$, $\overline{X_n}$ is the sample mean of n observations (sample size). In the context of the Galton board, this notion refers to the sample mean of the ball positions across multiple trials (since the tester is finding sampling distribution). After running several trials, the ball's positions are recorded, and the average of these positions across multiple trials forms the sample mean. $N(\mu, \frac{\sigma^2}{n})$ refers to the distribution of the sample mean is normally distributed with, μ represents the sample mean. This is the true population mean of the expected value of the ball's positions, σ^2 is the population variance. The variable measures the spread in the ball positions. In the Galton board the variance describes how spread out the ball position is, after they pass through the pins. n refers to the number of balls per trial in the Galton board experiment. For each trial, the tester collects a sample of ball positions, and n here represents the number of ball positions. As n increases, the variance of the sample means decreases, meaning the sample means becomes more stable and closely approximates the population mean. $\frac{\sigma^2}{n}$ is the variance of the sample mean, as the n gets larger, the variance of the ball gets smaller.

3. Experiment conducted to prove the CLT

3.1 Setting and Methodology

To investigate the application of the CLT in the context of the Galton board, the author conducted an experiment in which observers looked to identify the presence of the CLT in action. The primary aim of this experiment is to determine whether the CLT becomes evident as the number of trials increases, thereby testing its generalizability in this specific setup. To achieve this, the tester made a step-by-step experimental outline. The setting of the experiment involves designing an electronic Galton board with a fixed number of rows (10) and a fixed number of

ISSN 2959-6157

balls (100) per trial, and they are the controlled variables in this experiment.

The methodology involved in the experiment is presented. For each trial, the mean of the balls' final positions is calculated. As the CLT states as the number of trials increases, its sampling distribution will be more likely to conform to the pattern of normal distribution, the author will make a control group and a treatment group here. The experiment will repeat for 30 independent trials for the control group, resulting in 30 sample means. However, there is little to no documented evidence to support that a sample size of 30 is the magic number for non-normal distributions, so the author makes a treatment group that is larger in trial numbers [6]. The treatment group will be assigned 100 independent trials and collect its sample mean. In this case, the independent variable in the experiment is the number of trials, and the dependent variable in this experiment is the pattern of normal distribution. According to the CLT, the sampling distribution should approximate a normal distribution, even though the original Galton board distribution is binomial, and the treatment group will have a better resemblance to the pattern of a bell shape. If the sample means to follow a normal distribution, and the treatment group has a higher resemblance to the normal distribution, then the CLT's presence in this experiment is confirmed; otherwise, the CLT would not be considered applicable.

3.2 Data Processing

Table 1. 30 examples of different trials, with the mean shown in the last column.

	Bin0	1	2	3	4	5	6	7	8	9	10	Sample mean
Trial 1	0	1	1	13	16	22	31	10	3	3	0	5.23
Trial 2	0	0	5	15	24	27	20	5	3	1	0	4.74
Trial 3	0	1	7	8	19	25	24	13	3	0	0	4.99
Trial 4	1	0	4	12	23	27	22	10	1	0	0	4.81
Trial 5	0	0	3	11	21	23	28	7	6	1	0	5.12
Trial 6	0	2	4	9	24	22	17	13	7	1	1	5.11
Trial 7	1	1	8	15	17	19	29	7	1	2	0	4.74
Trial 8	0	1	4	9	24	23	24	12	3	0	0	4.99
Trial 9	1	0	4	8	17	32	23	10	5	0	0	5.08
Trial 10	0	1	4	9	23	33	21	4	4	1	0	4.88
Trial 11	0	1	6	5	19	30	20	15	2	1	1	5.14
Trial 12	2	0	5	8	23	20	27	9	5	1	0	5
Trial 13	0	1	4	10	20	20	23	16	6	0	0	5.17
Trial 14	0	0	2	12	18	29	27	9	3	0	0	5.06
Trial 15	0	1	3	11	20	24	20	12	9	0	0	5.16
Trial 16	0	2	6	12	16	28	20	12	3	1	0	4.91
Trial 17	0	1	3	20	13	27	20	9	6	1	0	4.94
Trial 18	0	2	6	14	19	24	17	14	4	0	0	4.84
Trial 19	0	1	2	11	20	29	19	11	7	0	0	5.1
Trial 20	0	0	3	15	23	27	20	4	6	2	0	4.92
Trial 21	0	2	6	14	23	23	16	7	6	3	0	4.83
Trial 22	0	0	5	14	16	22	18	17	4	4	0	5.21
Trial 23	0	0	1	15	21	24	23	14	2	0	0	5.03
Trial 24	0	2	3	9	15	30	21	12	7	1	0	5.2
Trial 25	0	1	3	9	16	34	22	9	6	0	0	5.11
Trial 26	1	1	3	12	20	25	18	12	5	3	0	5.07
Trial 27	0	0	5	8	23	22	24	9	7	2	0	5.17
Trial 28	0	1	3	11	22	28	18	10	6	1	0	5.03

Trial 29	0	4	4	9	19	27	23	10	4	0	0	4.9
Trial 30	0	4	5	5	22	23	28	10	2	1	0	4.95

The Table 1 is the data table that shows the results of the experiment on the Galton board simulation. The table displayed the number of balls that fell into each bin for each of the 30 trials (noting that there is a bin 0 because when several rows equal 10, there will be 11 possible outcomes for balls, and thus the tester placed that extra, naming bin 0). The table precisely shows how the binomial distribution governs the placement of the balls across the bins since the balls' final positions elaborate a binomial distribution that is formed by a sequence of independent Bernoulli trials, which is mentioned earlier in the text. As the data has shown, the distribution of balls concentrated in bins 4, 5, and 6, as these bins have higher binomial probabilities. This is because these bins correspond to the outcomes with higher binomial coefficients, making it more likely for the balls to fall into these areas.

Once the sample means are computed for all 100 trials, a sampling distribution of these means is formed. Then, the tester made a graph to illustrate the results of the sample mean better, see Fig. 1.

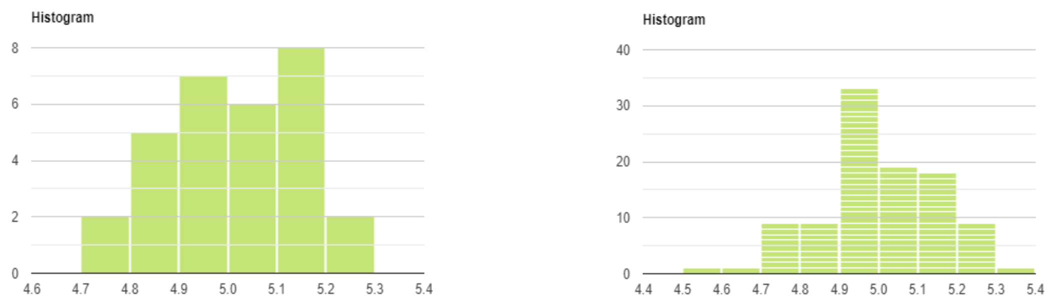

Fig. 1 The graph formed by the first (left) 30 sample mean and (right) 100 sample mean.

3.3 Analysis

The two graphs illustrate the distribution of balls in the control and treatment groups. As the graph has shown, both graphs have an x-axis standing for the sample mean, and the y-axis standing for the frequency that trial results in falls into an interval of 0.1. The author chose 0.1 for the graphs' interval because the sample size in this experiment is not overwhelmingly large enough to support the graph-making, as there could be situations in which no data is in a smaller interval. For this reason, the author chose 0.1, making graphs conform to a consistent pattern while still keeping the data as accurate and clear as possible. Specifically, by observing two charts, the finding is evident. The charts generated from these two datasets illustrate a clear distinction in their resemblance to a normal distribution. The chart representing the sample mean of 100 trials exhibits an evident bell-shaped curve, which indicates a high degree of symmetry and accumulation around the sample mean's mean. This graph is nearly identical to a theoretical normal distribution and therefore supports the assertion of the CLT that the bigger the sample mean, the more resemblance of the normal distribution. On the other hand, however, the chart constructed from smaller sample sizes revealed a less distinct normal distribution pattern. Though one can still see the accumulation of frequency around sample means of 5.0, the graph is inconsistent enough compared to a normal distribution curve in theory. This finding allows the researcher to conclude that smaller sample sizes do not reliably yield a distribution that resembles normality.

Still, simply observing the chart is not convincing enough to conclude. Thus, the researcher used statistical tools to evaluate the resemblance of two graphs in terms of normal distribution. The author applied the Shapiro-Wilk Test, which proved to outperform the other three tests [7,8]. This statistical test is designed to assess the normality of a distribution by comparing the sample data to a perfectly normal distribution. An indicator called p-value is involved. Put simply, the higher the p-value is, the more resemblance the data follows the normal distribution. A p-value equal to or greater than 0.05 will be considered as the data do not significantly deviate from normality. Unsurprisingly, the Shapiro-Wilk Test results confirmed the visual observations: the p-value formed by the first 30 data is 0.2259, and the p-value formed by 100 data is 0.5695. As the latter one has a p-value greater than twice the previous one, the test is considered a success. Ultimately, the findings reinforce the conclusion that larger sample sizes, such as a sample mean of 100 trials, can lead to sampling

ISSN 2959-6157

distributions that conform more closely to a normal distribution, thereby validating the CLT and illustrating its generalizability in the Galton board experiment.

4. Conclusion

In conclusion, this experiment provides clear evidence for the existence and generalizability of the Central Limit Theorem (CLT) within the context of the Galton board problem. After identifying the concept of the Bernoulli trial and binomial, one can readily see the power of CLT since the original distribution of the Galton board is binomial, but the CLT changed the final distribution of the Galton board after collecting its sampling distribution. By adhering to the principles of the CLT and conducting a methodologically sound experiment, the tester successfully straightforwardly demonstrates the existence of CLT, which is an abstract and complex statistical concept and hard to illustrate without visualization. The results indicate a notable distinction between the graphs formed by the sample means of 30 and 100, with the latter more closely resembling a normal distribution, confirming that a larger sample size corrects the data towards a normal pattern. However, even the graph formed by 30 data points adheres to the normal distribution, as the p-value exceeds 0.05, suggesting that the sampling distribution method itself regulates the data within the bounds of normality, despite a smaller sample size. Through a combination of visual analysis and quantitative statistical evaluation, the author offers strong empirical support for the generalizability and applicability of the CLT, even when the underlying data is binomially distributed, as in the case of the Galton board experiment. Overall, the author provides visualized results based on the Galton board model and simplifies the process of proving the CLT, or at least transferring such a concept into a comprehensive manner.

References

[1] Daud, Auni Aslah Mat. “Mathematical modelling and symbolic dynamics analysis of three new galton board models.” Communications in Nonlinear Science and Numerical Simulation 2014, 19(10): 3476-3491.

[2] Conners, Abigail C., and John Cody. “Galton’s Quincunx: Creating an Accessible Model of the Central Limit Theorem.” (2024).

[3] Kwak, Sang Gyu, and Jong Hae Kim. “Central limit theorem: the cornerstone of modern statistics.” Korean journal of anesthesiology 2017, 70(2): 144-156.

[4] Kozlov, Valery Vasil’evich, and M. Yu Mitrofanova. “Galton board.” arXiv:nlin/0503024 (2005).

[5] W. L. Hays, Statistics, 5th ed., New York: Holt, Rinehart and Winston, 1994.

[6] Islam, Mohammad Rafiqul. “Sample size and its role in Central Limit Theorem (CLT).” Computational and Applied Mathematics Journal 2018, 4(1): 1-7.

[7] González-Estrada, Elizabeth, and Waldenia Cosmes. “Shapiro–Wilk test for skew normal distributions based on data transformations.” Journal of Statistical Computation and Simulation 2019, 89(17): 3258-3272.

[8] Razali, Nornadiah Mohd, and Yap Bee Wah. “Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests.” Journal of statistical modeling and analytics 2011, 2(1): 21-33.