A Modified Review and Proof of Central Limit Theorem in Relation with Law of Large Numbers

Casmir Chidiebere Onyeneke[1*]

[1] Department of Mathematics and Statistics, Hezekiah University, Umudi, Nigeria

Email Address

casdyx@yahoo.com(Casmir C. Onyeneke)
*Correspondence: casdyx@yahoo.com

Received: 24 August 2018**; Accepted:** 20 October 2018**; Published:** 16 November 2018

Abstract:

The study of the central limit theorem and law of large are paramount to mathematicians and statisticians because of the undisputable role they play in probability, estimation, decision making and analysis of confidence intervals. As the fundamental theorems underlying applied mathematics, it can never be overemphasized that the distribution of the sum of a large number of independent and identically distributed approximates to normal irrespective of the nature of the underlying distribution. This is the sole theorem of the study of central limit theorem and law of large numbers. The essence of this modified review and study of these powerful keys of mathematics is to elaborate the essential proofs and applications that uphold their position. Indeed, it is hard to overstate the importance of the central limit theorem due to the fact that it is the major reason for many statistical and mathematical procedures to work. When the distribution of statistical cumulative functions is viewed in the cause of this study, they all approximate to normal under repeated observation in the same condition for a long period of occurrences. The first thing to do in every given distribution is to standardize the function if it assumes non-normal. Next, is to normalize them by the assumptions of central limit theorem and law of large numbers. In this work, different approaches are used to tackle different distributions in order to get their approximation to normal by the application of central limit theorem and law of large numbers.

Keywords:

Central Limit Theorem, Law of Large Number, Discrete Random Variables, Cumulative, Moments, Probability Magnitude, Normal Distribution, Variance

1. Introduction

The central limit theorem and law of large numbers are applied in probability theory for conditions which the mean of an adequately large number of independent random variables, each with finite mean and variance, approximates to normal distributed [1]. One of the underlying assumptions is that the random variables must be identically distributed to depict the central limit theorem in its common form. In real life application, the societal quantities comprise most of the balanced sum of events that are not observed. As a result, the two theorems provide limited justification for the incidence of the normal probability distribution. It is not easy to separate central limit theorem from the law of large number because central limit theorem substantiates the approximation of large-sample statistics to the normal distribution in controlled experiments. This study examines the different approaches used to prove the central limit theorem and law of large numbers by some respective professional approaches. This work presented the proof which shows that the normal form variation in probabilities of independent and identically distributed functions have a limiting cumulative distribution function which approaches to normal distribution. The methods of cumulative and moment functions were used to back the evidence guiding the assumptions of the central limit theorem and laws of large numbers. A brief view of Martingale and Yuval central limit theorems were given. At the end, it was shown that central limit theorem which concludes $\lim_{n \to \infty} \emptyset_{Y_n}(t) = e^{-\frac{t^2}{2}}$ has a cumulative function of a standard normal distribution with mean 0 and Variance$\sigma^2 = 1$ with acontinuous limiting distribution.There are also some illustrations of the law of Large numbers both in discrete and continuous distributions. The central limit theorem and the law of large number have a relationship with Chebyshev's inequality which is stated that for random variables X with expected value μ and variance σ^2,$\rho\ (|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$for $\varepsilon > 0$. In thinking about a sequence $X_1, X_2, X_3, \ldots$ of an independent random variable with the same distribution,μ is the common expected value and σ^2 is the common variance. It is assumed that σ^2 is positive. LetS_n represents the normal sums denoted by $S_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} = \frac{\sum_i^n (x_i - \mu)}{\sigma\sqrt{n}}$.It is understood that the notion of standard normal distribution gives the understanding that the S_n have expected value 0 and variance 1. The central limit theorem now states that $\rho\ (S_n \leq x) \to \phi(x)$ for $n \to \infty$ for all $x \in R$, where ϕ is the distribution function of the standard normal distribution.

$$\phi(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-\frac{1}{2}t^2}\, dt \qquad (1)$$

The distribution function of the sums S_n converges to ϕ when n converges to infinity (∞).This is a quite amazing result and the absolute climax of probability theory. The surprising thing is that the limit distribution of the standard normal is independent of the distribution of theX_i.Consider a sequence of independent random variables $X_1, X_2 \ldots$ all having the same point probabilities $\rho\ (x_i = 0) = \rho\ (x_i = 1) = {}^1\!/_2$. The sums $S_n = X_1 + \cdots + X_n$ is binomially distributed with expected value $\mu = \frac{n}{2}$ and variance$\sigma^2 = \frac{n}{4}$. Thestandard normal thus becomes

$$S_n = \frac{X_1 + \cdots + X_n - \mu/2}{\sqrt{n}/2} \qquad (2)$$

The distribution of S_n is given by the distribution functionF_n. The Central Limit Theorem states that F_n converges to standard normal distributionϕ for $n \to \infty$ [2].

Consider a sequence $X_1 X_2 X_3$ ... of independent random variables with the same distribution and let μ be the common expected value denoted byS_n,$S_n = X_1 + \cdots + X_n$. The law of large numbers then states that $\rho\ \left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \longrightarrow 0$ for $n \longrightarrow \infty$ for every $\varepsilon > 0$[3]In other words, the mean value of a sample from any given distribution conveys to the expected value of that distribution when the size n of the sample approaches ∞.At this point, examine the law of large numbers for discrete random variables. Let X_1 , $X_2, \dots, X_n$ be an independent trials process with finite expected value $\mu = E(X_i)$ and finite variance $\sigma^2 = V(x_j)$ [4]. Let $S_n = X_1 + X_2 + \cdots + X_n$ then for any $\varepsilon > 0$, $p\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \to 0$ as $n \to \infty$, equally, $p\left(\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right) \to 1$ as $n \to \infty$. Since X_1, $X_2, \dots, X_n$ are independent and have the same distributions, then $Var(S_n) = n\sigma^2$ and $V\left(\frac{S_n}{n}\right) = \frac{\sigma^2}{n}$ and $E\left(\frac{S_n}{n}\right) = \mu$. Chebyshev's inequality for $\varepsilon > 0$, $p\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$. Thefixed ε, gives $p\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \longrightarrow 0$ as $n \longrightarrow \infty$ or equivalently, $p\left(\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right) \longrightarrow 1$ as $n \longrightarrow \infty$.

The fact aboutlaw of large numbers for continuous probably distributions is that if X_1, $X_2 \dots X_n$ is an independent trials process with continuous density function f, finite expected value μ and variance σ^2. Then $S_n = X_1 + X_2 + \cdots + X_n$ is the sum of the X_i. Then for any real number $\varepsilon > 0$, $\lim_{n \to \infty} p\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0$. In other words, $\lim_{n \to \infty} p\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = 1$. However, this is not necessarily true if σ^2 is infinite.

2. Methodological Proofs

Let $X_1, X_2, \dots, X_N$ be a set of N independent random variables and each X_i has an arbitrary probability distribution $P(x_1, \dots, x_N)$ with mean μ_i and a finite variance σ_i^2 then the standard normal form of the variables is states as $x_{norm} = \frac{\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} \mu_i}{\sqrt{\sum_{i=1}^{N} \sigma_i^2}}$ has a limiting cumulative distribution function which approaches a normal distribution [5]. Under additional conditions on the distribution of the sums, the probability density itself is also normal with $\mu = 0$ and variance $\sigma^2 = 1$ [6].If conversion to normal form is not performed, then the variables $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ arenormally distributed with $\mu_X = \mu_x$ and $\sigma_X = \sigma_x/\sqrt{N}$

Consider the inverse Fourier transformation of $P_X(f)$ for more proof of the central limit theorem. According to [8]

$$F_f^{-1}[P_X(f)](x) = \int_{-\infty}^{\infty} e^{2\pi i f \chi} P(\chi) dx = \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{(2\pi i f X)^n}{n!} P(X) dx \tag{3}$$

$$= \sum_{-\infty}^{\infty} \frac{(2\pi i f)^n}{n!} \int_{-\infty}^{\infty} X^n P(X) dx = \sum_{n=0}^{\infty} \frac{(2\pi i f)^n}{n!} \langle X^n \rangle \tag{4}$$

$$\langle X^n \rangle = \langle N^{-n}(x_1 + x_2 + \cdots + x_N)^n \rangle = \int_{-\infty}^{\infty} N^{-n}(x_1 + \cdots + x_N)^n P(x_1) \dots P(x_N) dx_1 \dots d x_N \tag{5}$$

So,

$$F_f^{-1}[P_X(f)] = \sum_{n=0}^{\infty} \frac{(2\pi i f)^n}{n!} \langle X^n \rangle = \sum_{n=0}^{\infty} \frac{(2\pi i f)^n}{n!} \int_{-\infty}^{\infty} N^{-n}(x_1 + \cdots + x_N)^n P(x_1) \dots P(x_N) dx_1 \dots d x_N \tag{6}$$

$$= \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \left[\frac{2\pi i f (x_1 + \dots + x_N)}{N}\right]^n \frac{1}{n!} P(x_1) \dots P(x_N) dx_1 \dots dx_N \tag{7}$$

$$= \int_{-\infty}^{\infty} e^{2\pi i f (x_1 + \dots + x_N)/N} P(x_1) \dots P(x_N) dx_1 \dots dx_N \tag{8}$$

$$\left[\int_{-\infty}^{\infty} e^{2\pi i f x_1/N} P(x_1) dx_1\right] \times \ . \ . \ . \times \left[\int_{-\infty}^{\infty} e^{2\pi i f x_N/N} P(x_N) dx_N\right] = \left[\int_{-\infty}^{\infty} e^{2\pi i f x/N} P(x) dx\right]^N \quad (9)$$

$$= \left\{\int_{-\infty}^{\infty} \left[1 + \left(\frac{2\pi\, i\, f}{N}\right) x + \frac{1}{2}\left(\frac{2\pi\, i\, f}{N}\right)^2 x^2 + . . . \right] P(x) dx\right\}^N \quad (10)$$

$$= \left[1 + \frac{2\pi\, i\, f}{N}\langle x\rangle - \frac{(2\,\pi\, f)^2}{2N^2}\langle x^2\rangle + 0(N^{-3})\right]^N = exp\left\{N \ln\left[1 + \frac{2\pi\, i\, f}{N}\langle x\rangle - \frac{(2\,\pi\, f)^2}{2N^2}\langle x^2\rangle + 0(N^{-3})\right]\right\} \quad (11)$$

$$= exp\left\{N \ln\left[1 + \frac{2\pi\, i\, f}{N}\langle x\rangle - \frac{(2\,\pi\, f)^2}{2N^2}\langle x^2\rangle + 0(N^{-3})\right]\right\} \quad (12)$$

The expansion of$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + . . .$ gave

$$= F_f^{-1}[P_X(f)](x) = exp\left\{N\left[\frac{2\pi\, i\, f}{N}\langle x\rangle - \frac{(2\,\pi\, f)^2}{2N^2}\langle x^2\rangle + \frac{1}{2}\frac{(2\,\pi\, f)^2}{N^2}\langle x\rangle^2 + 0(N^{-3})\right]\right\} \quad (13)$$

$$= exp\left[2\pi\ i\ f\langle x\rangle - \frac{(2\pi\, f)^2(\langle x^2\rangle - \langle x\rangle^2)}{2N} + 0(N^2)\right] =\approx exp\left[2\pi\ i\ f\langle \mu_x\rangle - \frac{(2\pi\, f)^2\sigma_x^2}{2N}\right] \quad (14)$$

Since$\mu_x = \langle \chi\rangle$; $\sigma_x^2 = \langle x^2\rangle - \langle x\rangle$. Taking the Fourier transform,

$$P_X = \int_{-\infty}^{\infty} e^{-2\pi i f x} F^{-1}[P_X(f)] df \ = \int_{-\infty}^{\infty} e^{2\pi\, i\, f(\mu_x - \bar{x}) - (2\pi\, f)^2 \sigma_x^2/2N} df \quad (15)$$

It can be written as

$$\int_{-\infty}^{\infty} e^{-iaf - bf^2}\, df \quad (16)$$

In the formula, $a \equiv 2\pi(\mu_x - \bar{x})$ and $b \equiv (2\pi\sigma_x)^2/2N$ and it represents a Fouriertransform of a Gaussian function, so

$$\int_{-\infty}^{\infty} e^{iaf - bf^2}\, df = \ e^{-\frac{a^2}{4b}}\sqrt{\frac{\pi}{b}} \quad (17)$$

Therefore,

$$P_X = \sqrt{\frac{\pi}{\frac{(2\pi\sigma_x)^2}{2N}}}\ \ exp\left\{\frac{-[2\pi(\mu_x - \bar{x})^2]}{4\frac{(2\pi\sigma_x)^2}{2N}}\right\} \quad (18)$$

$$= \sqrt{\frac{2\pi N}{4\pi^2\sigma_x^2}}\ exp\left[-\frac{4\pi^2(\mu_x - \bar{x})^2 2N}{4.4\pi^2\sigma_x^2}\right] \quad (19)$$

$$= \frac{\sqrt{N}}{\sigma_{x\sqrt{2\pi}}} e^{-(\mu_x - \bar{x})^2/2\sigma_x^2} \quad (20)$$

But $\sigma_x = \frac{\sigma_x}{\sqrt{N}}$and,$\mu_x = \mu$, so $P_X = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-(\mu_x - \bar{x})^2/2\sigma_x^2}$

Central limit theorem says that under general circumstances, if independent random variable are summed and normalized accordingly, then the limit approaches to a normal distribution [9]. Consider the density of the normal distribution $N\ (\mu, \sigma^2)$ with mean μ and varianceσ^2, that is $\frac{1}{\sqrt{2\ \pi\sigma^2}} e^{\frac{(x-\ \mu)^2}{2\sigma^2}}$. Suppose that X_i are independent, identically distributed random variables with zero mean and variance σ^2 , then$\frac{X_1 + \cdots + X_n}{\sqrt{n}} \to N(0, \sigma^2)$. If the variables do not have zero mean, they can easily be normalized by subtracting the expectation from them.The meaning of $Y_n \ \to \ Y$ is that each interval [a, b], $\rho_r[a \le Y_n \le b] \to \rho_r[a \ \le \ Y \le b]$.

The definition of the cumulative generating function of a random variable X is$K_x(t) = \log M_x(t)$ [10]. This implies that $M_x(t) = 1 + 0(t)$, so that it is logical to take its logarithm. Infact, using Taylor series of$\log(1 + x)$, $\log(1 + t) = t - \frac{t^2}{2} + \cdots$expands$K_x(t)$ as a power series, which begins as$K_x(t) = \left(E(x)t + \frac{E(x^2)}{2}t^2 + \cdots\right) - \frac{(E(x)t+\cdots)^2}{2} + \cdots = E(x) + \frac{E(x^2) - E(x)^2}{2}t^2 + \cdots$ The first two coefficient of $K_x(t)$ serves as the exponential generating function. Ignoring the $1/n$ factors, the expectation and the variance, the formula is called coefficient cumulative function$K_n(X)$as $K_n(X) = K_x^n(0)$. In particular, it shows that$K_0(X) = 0$, $K_1(X) = E(X)$;$K_2(X) = V(X)$. In general, using the Taylor series of$\log(1 + x)$,$K_n(X)$can be expressed as a polynomial in the moments. Conversely using the Taylor series of e^x can be used to show the moments as polynomials in the cumulative functions. This provides an example of Moebius inversion [11].

The formulas for balancing and simplifying extend to cumulative generating function as $K_{x+y}(t) = K_x(t) + K_y(t)$, $K_{cx}(t) = K_x(ct)$. Supposed that $X_1, X_2, \ldots$ are independent random variables with zero mean, then$\frac{K_{X_1+X_2+\cdots+X_n(t)}}{\sqrt{n}}$=$K_{X_1}\left(\frac{t}{\sqrt{n}}\right) + \cdots + K_{X_n}\left(\frac{t}{\sqrt{n}}\right)$. It can be written interms of the cumulative function as $K_m\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right) = \frac{K_m(X_1)+\cdots+K_m(X_n)}{n^{\mu/2}}$. It implies that $K_1(X_k) = 0$, so the first cumulative is not propelled. The second cumulative, the variance, is simply averages. What happened to all the higher cumulative if the cumulative are bonded by some constant C, then for M > 2, $K_m\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right) \leq \frac{nc}{\mu/2} \rightarrow 0$

In other words, all the higher cumulative disappear in the limit. The cumulative of the normalized sums tend to the cumulative of some fixed distribution, which must be the normal distribution. The limit cumulative generating function, which can be expressed as$\frac{\sigma^2 t^2}{2}$indeed corresponds to normal variable. In particular, if the normal distributed X_i are used, then $(X_1 + \cdots + X_n)/\sqrt{n}$ becomes normal distribution of the form $N(0, \sigma^2)$.

There are some advantages of providing a conditional explanation for values of the moments of a normal distribution. In other words, it is important to calculate these very moments. Note that the moment of a normal distribution with zero mean and unit variance can be written as

$$\int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} dx = \frac{x^{n+1}}{n+1} e^{-\frac{x^2}{2}} \int_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{x^{n+1}}{n+1}\left(-xe^{-\frac{x^2}{2}}\right) dx = \frac{1}{n+1}\int_{-\infty}^{\infty} x^{n+2} e^{-\frac{x^2}{2}} dx \quad (21)$$

The repeated relation is $M_r = \frac{M_{n+2}}{n+1} = M_{n+2} = (n+1)M_n$ [12].Since $M_0 = 1$ and $M_1 = 0$ where all odd numbersare zero because the distribution is symmetric and the event moments are $M_n = (n-1)M_{n-2} = \cdots (n-1)(n-3) \ldots 1$. Proof using moment briefly computes the limit of the moments of$Y_n = (X_1 + \cdots + X_n)/\sqrt{n}$. It is assumed that the variables X_i are independent with 0 mean, unit variance and bonded moments. The proof can be adapted to the case of varying variances.It implies that $M_1(Y_n) = 0$ and the second moment $M_2(Y_n) = F\left[\frac{(X_1+\cdots+X_n)^2}{n}\right] = \frac{\sum_i E\,(X_i^2)}{n} + \frac{\sum_{i\neq j} E\,(X_i X_j)}{n} = 1$since $E\,(X_i^2) = 1$,where as $E(X_i X_j) = E(X_i)E(X_j) = 0$

VOLUME 2, 2018
DOI: 10.31058/j.data.2018.24006

Then, the third moment assumes $M_3(X_i) \leq C_3$, $M_3(Y_n) = E\left[\frac{(X_1+\cdots+X_n)^3}{n^{3/2}}\right] = \frac{\sum_i E(X_c^3)}{n^{3/2}} + 3\frac{\sum_{i\neq j} E(X_c^{-2}X_j)}{n^{3/2}} + \frac{\sum_{i\neq j\neq k} E(X_i X_j X_k)}{n^{3/2}} \leq \frac{nC_3}{n^{3/2}} = \frac{C_3}{\sqrt{n}}$ which tends to zero.The fourth moment brings forth a more interesting calculation where$M_3(Y_n) = E\left[\frac{(X_1+\cdots+X_n)^4}{n^2}\right] = \frac{\sum_i E\, X_i^4}{n^2} + 4\frac{\sum_{i\neq j} E\,(X_i^3 X_j)}{n^2} + 3\frac{\sum_{i\neq j} E\,(X_i^2 X_j^2)}{n^2} = 6\frac{\sum_{i\neq j\neq k} E\,(X_i^2 X_j X_k)}{n^2} + \frac{\sum_{i\neq j\neq k\neq l} E\,(X_i X_j X_k X_l)}{n^2} = 0\,(n^2) + 3\frac{n(n-1)}{n^2} \rightarrow 3.$ The distribution of term with t variables with multiplicitiesM_1 ... M_i is at most $\frac{n^{m_1+\cdots+m_t}}{n^{m/2}} C_{m_1} \ldots C_{m_t}$ where C is a bound on $E(X_i^s)$. Thus the term is asymptotically zero if $m_1+\cdots+m_t < {}^{m}/_{2}$. If $m_1+\cdots+m_t \geq {}^{m}/_{2}$ then six $m_i \geq 2$otherwise the term is identically zero such that $t = {}^{m}/_{2}$and $m_i = 2$. In this case, the contribution of the term is $\frac{n\,(n-1)\ldots n-{}^{m}/_{2}+r}{n^{m/2}} \rightarrow 1$.Since the random variables as unit variance, the m^{th} moment converges to the number of such terms.

3. Results and Discussion

In distinguishing the facts in the law of large numbers and central limit theorem, the law of large numbers enables the investigation of the conditions under which a random sample of mean $\bar{X}_n$ tends to the population mean μ from which the sample is drawn whilethe central limit theory assists in findingthe probability magnitude of the discrepancy $|\bar{X}_n - \mu|$ or the size of the sample that can give a reliable estimate.

Let x_i ; $i = 1, 2 \ldots n$ be a sequence of independently and identically distributed random variables such that $E(x_i) = \mu$, $\mathrm{var}\,(x_i) = \sigma^2 < \infty$ [13]. Then $Y_n = \frac{(X_n - \mu)\sqrt{n}}{\sigma} \xrightarrow{L} Z\sim\ \phi_{\bar{x}-\mu}\,(t)$. Note that $Y_n = \frac{\bar{X}_n - \mu}{\sqrt{V(X_n)}} = \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} = \frac{(\bar{X}_n - \mu)\sqrt{n}}{\sigma}$ [13].Let $\phi_{X_k}(t)$ denote the cumulative function of X_K, the distribution function is $F(X_k)$ or F. Then, $\phi_{X_k}(t) = 1 + \mu it + u_2' \frac{(it)^2}{2!} + \cdots = \int_{-\infty}^{\infty} e^{itx}\ f(x)dx$. Since the first and the second moments of X_K exist and $\phi_{X_k - u}(t) = E\left(e^{it(X_{k}-\mu)}\right) = e^{-itu} E\left(e^{itX_K}\right) = e^{-itu}\phi_{X_K}(t) = e^{-itu}\left(1 + ui + u_2'\frac{(it)^2}{2!} + \cdots\right) = \left(1 - uit + u^2\frac{(it)^2}{2!} + \cdots\right)\left(1 + ui + u_2'\frac{(it)^2}{2!} + \cdots\right)$

$$= 1 + uit + u_2'\frac{(it)^2}{2!} - iut - u^2\,(it)^2 - uu_2'\frac{(it)^3}{2!} + \frac{u^2(it)^2}{2!} + \frac{u^3(it)^3}{2!} + u^2u_2'\frac{(it)^4}{2!} + \cdots$$

$= 1 + uit + u_2'\frac{(it)^2}{2!} - uit + u^2\frac{(it)^2}{2!} - u^2\,(it)^2 + \frac{u^3(it)^3}{2!} + u^2u_2'\frac{(it)^2}{2!} - u\ \ u_2'\frac{(it)^3}{2!} + \cdots$

$= 1 + u_2'\frac{(it)^2}{2!} + u^2\frac{(it)^2}{2!} - u^2\,(it)^2 + \cdots = 1 + u_2'\frac{(it)^2}{2!} + u^2\left(\frac{1}{2} - 1\right)(it)^2 + \cdots$; $\frac{1}{2} - 1 = -\frac{1}{2}$

$= 1 + u_2'\frac{(it)^2}{2!} - u^2\frac{(it)^2}{2!} + 0(t^2) = 1 + (u_2' - u^2)\frac{(it)^2}{2!} + 0(t^2) = 1 + \sigma^2\frac{(it)^2}{2!} + 0\,(t^2)$

$$= 1 - \sigma^2 \frac{it^2}{2!} + 0\,(t^2) \quad (22)$$

Consider the characteristic function of $Y_n = \frac{(\bar{X}_n - \mu)\sqrt{n}}{\sigma} = \frac{\frac{1}{n}\Sigma(X_n - \mu)\sqrt{n}}{\sigma} = \frac{\Sigma(X_n - \mu)}{\sigma\sqrt{n}}$, then

$\phi_{Y_n}(t) = \frac{\phi(X_n - \mu)\sqrt{n}}{\sigma}(t) = \phi_{X_k - u}\left(\frac{t\sqrt{n}}{\sigma}\right) = \sum_{i=1}^{\emptyset_n}(X_i - \mu)\left(\frac{t}{\sigma\sqrt{n}}\right)$
whereX_1 , X_2 ... X_n are identically distributed. Thus

$$\emptyset_{Y_n}(t) = \left\{\phi_{(X_j - u)}\left(\frac{t}{\sigma\sqrt{n}}\right)\right\}^n \quad (23)$$

The logarithm of both sides of equation (2) gives $\log\emptyset_{Y_n}(t) = n\log\emptyset_{(X_j-u)}\left(\frac{t}{\sigma\sqrt{n}}\right)$. By replacing t in equation (23) by $\frac{t}{\sigma\sqrt{n}}$ it gives $\emptyset_{x-u}\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 - \sigma^2\frac{(\frac{t}{\sigma\sqrt{n}})^2}{2!} + 0\,(t^2) = 1 - \frac{\sigma^2 t^2}{2!\,\sigma^2 n} + 0\,(t^2)$ and hence

$$\log\emptyset_{Y_n}(t) = n\log\left(1 - \frac{t^2}{2n} + 0\,(t^2)\right)) \quad (24)$$

Consider the identity $\log(1 - Z) = -Z - \frac{Z^2}{2} - \frac{Z^3}{3} - \ldots$.. where $Z < 1$. Application of this identity to (24) gives $\log\emptyset_{Y_n}(t) = n\,(-\frac{t^2}{2n} - \frac{t^4}{8n^2} - \frac{t^6}{24n^3}\ldots) = \frac{t^2}{2} - \frac{t^4}{8n} - \frac{t^6}{24n^2} = -\frac{t^2}{2}$ as $n \to \infty$,

$$\lim_{n \to \infty}\emptyset_{Y_n}(t) = e^{-\frac{t^2}{2}} \quad (25)$$

Equation 4 has a characteristics function of a standard normal distribution with mean 0 and Variance $\sigma = 1$. The limiting distribution is continuous. When a die is rolled 420 times, the central limit theorem can be used to determine the probability that the sum of the rolls lies between 1400 and 1550. The sum is a random variable $S_{420} = X_1 + X_2 + \cdots + X_{420}$ where each X_j has a distribution $M_x = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$, we have seen that $\mu = E(X) = 7/2$ $and\ \sigma^2 = V(X) = 35/12$thus $E(S_{420}) = 420 \times 7/2 = 1470$, $\sigma^2(S_{420}) = 420 \times 35/12 = 1225$ and $\sigma(S_{420}) = 35$. Therefore, $p\,(1400 \le S_{420} \le 1550) = p\left(\frac{1399.5 - 1470}{35} \le S_{420} \le \frac{1550.5 - 1470}{35}\right) = p\,(-2.01 \le S_{420} \le 2.30) =$
$NA\,(-2.01, 2.30) = 0.9670$

Again, let X_1 , $X_2, \ldots, X_n$ be a Bernouli trials process with probability 0.3 for success and 0.7 for failure. Let $X_{ij} = 1$ if jth outcome is a success and 0 otherwise. Then $E\left(X_{ij}\right) = 0.3$ and $V\left(X_{ij}\right) = (0.3)(0.7) = 0.21$ if $A_n = \frac{S_n}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$is the average of X_i. Then $E(A_n) = 0.3$ and $V(A_n) = var\frac{(S_n)}{n^2} = 0.21/n$. Example, let $\varepsilon = 0.1$,$p\,(|A_n - 0.3| \ge 0.1) \le \frac{0.21}{n\,(0.1)^2} = \frac{21}{n}$. if $n = 100$, $p\,(|A_{100} - 0.3| \ge 0.1) \le 0.21$ or if $n = 100$; $p\,(|A_{1000} - 0.3| \ge 0.1) \le 0.021$. These can be rewritten as$p(0.2 < A_{100} < 0.4) \ge 0.79$, $p(0.2 <$

$A_{1000} < \ 0.4) \ \geq 0.979$. These can be compared with the actual values, which are (to six decimal places)$p(0.2 < A_{100} < \ 0.4) \ \doteq 0.962549, p(0.2 < A_{100} < 0.4) \doteq 1$.

Suppose we choose at random n numbers from the interval [0, 1] with uniform distribution. Then if $X_i$describes the i^{th} choice, it gives$\mu = E(X_i) = \int_0^1 x \ dx = {}^1/_2$; $\sigma^2 = var(X_i) = \int_0^1 x^2 \, dx - u^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$. Hence $E\left(\frac{S_n}{n}\right) = \frac{1}{12n}$and for any $\varepsilon > 0$; $p\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right) \leq \frac{1}{12n\varepsilon^2}$. This says that we choose n numbers at random from$[0, \ \ 1]$, and then the numbers are better than, $-\frac{1}{12n\varepsilon^2}$ that the difference $\left|\frac{S_n}{n} - \frac{1}{2}\right|$ is less that ε.Suppose we choose n real numbers at random, using a normal distribution with mean 0 and variance 1. Then $\mu = E(X_i) = 0, \ \ \sigma^2 = v(X_i) = 1$ Hence $E\left(\frac{S_n}{n}\right) = 0var\left(\frac{S_n}{n}\right) = \frac{1}{n}$ and $\left(\left|\frac{S_n}{n} - 0\right| \geq \varepsilon\right) \leq \frac{1}{n\varepsilon^2}$.

4. Conclusions

Without getting into the mathematical details, the Central Limit Theorem states that if you take a lot of samples from a certain probability distribution, the distribution of their sum (and therefore their mean) will be approximately normal, even if the original distribution was not normal. Furthermore, it gives you the standard deviation of the mean distribution which is $\frac{\sigma}{\sqrt{n}}$. When testing a statistical hypothesis or calculating a confidence interval, we generally take the mean of a certain number of samples from a population, and assume that this mean is a value from a normal distribution. The Central Limit Theorem tells us that this assumption is approximately correct, for large samples, and tells us the standard deviation to use.

The Martingale central limit theorem generalizes the result of central limit for random variables to martingales, which are stochastic processes where the change in the value of the process from time t to time t + 1 has expectation zero, even conditioned on previous outcomes [14]. Let $X_1, X_2 \ldots$ be a martingale with bounded increments. That is, suppose$E\,(X_{t+1} - X_t | X_1 \ldots, X_t) = 0$ And $|X_{t+1} - X_t| \leq K$ Almost surely for some fixed bound K and all t. also assume that $|X_1| \ \leq K$ almost surely. Define$\sigma_t^2 = E\{(X_{t+1} - X_t)^2 | X_1 \ldots, X_t\}$ and let$T_v = Min\,\{t\colon \sum_{l=1}^{t} \sigma_i^2 \geq V\}$[15]. Then$\frac{X_{TV}}{\sqrt{V}}$converges the distribution to the normal distribution with mean 0 and variance 1 as $V \rightarrow +\infty$. More explicitly, $\lim_{V \rightarrow +\infty} P\left(\frac{X_{TV}}{\sqrt{V}} < x\right) = \emptyset(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(\frac{-u^2}{2}\right) du, \quad x \in R.$

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

References

[1] Abramowitz, M.; Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing. National Bureau of Standards Applied Mathematics Series - 55. Dover, New York, 1972, Chapter 3, 53-60.

[2] Trotter, H. F. An Elementary Proof of the Central Limit Theorem. *Architectural Mathematics*, 1959, 10, 226-234.

[3] David B. Essentials of Statistics. David Brinks and Ventures Publishing ApS. 2010, Chapter 5, 26 -27.

[4] Lindeberg, J. W. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Bulletin of the American Mathematical Society*, 1922, 15, 211-225.

[5] Feller, W. The Fundamental Limit Theorems in Probability. *Bulletin of the American Mathematical Society*, 1945, 51, 800-832.

[6] Feller, W. An Introduction to Probability Theory and Its Applications 3rd Edition, Volume 2, Wiley Publishing, New York, 1971, Chapter 2, 31-98.

[7] Kallenberg, O. Foundations of Modern Probability. Springer-Verlag, New York, 1997, Chapters 3, 254-618.

[8] Feller, W. An Introduction to Probability Theory and Its Applications, 3rd Edition, Volume 1, Wiley Publishing, New York, 1968, Chapter 5, 229.

[9] Yuval F. Two Proves of the Central Limit Theorem. *Architectural Mathematics* 2010, Chapter 101, 567-934.

[10] Mlodinow, L. The Drunkard's Walk: How Randomness Rules Our Lives. Random House, New York, 2008, Chapter 2, 45-50.

[11] Spiegel, M. R. Theory and Problems of Probability and Statistics. McGraw-Hill, New York, 1992, Chapter 6, 112-113.

[12] Zabell, S. L. Alan Turing and the Central Limit Theorem. *Bulletin of the American Mathematical Society*, 1995, Chapter 102, 483-494.

[13] Artstein, S.; Ball, K.; Barthe F.; Naor A. Solution of Shannon's Problem on the Monotonicity of Entropy". *Journal of the American Mathematical Society*, 2004, 17(4), 4-10.

[14] Hall, P.; Heyde, C.C. Martingale Limit Theory and its application. Academic Press, New York 1980, Chapter 2, 17-25.

[15] Bradley, R. On the same results of M. I. Gordin: a classification of a misunderstanding. *Journal of Theoretical Probability (Springer)*, 1988, Chapter 1, 115-119.