International Mathematical Forum, Vol. 12, 2017, no. 9, 399 - 411
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/imf.2017.7118

Convergence of Binomial to Normal: Multiple Proofs

Subhash Bagui[1] and K. L. Mehra[2]

[1]Department of Mathematics and Statistics
The University of West Florida
Pensacola, FL 32514, USA

[2]Department of Mathematical and Statistical Sciences
University of Alberta
Edmonton, T6G 2G1, Canada

Abstract

This article presents four different proofs of the convergence of the Binomial $B(n, p)$ distribution to a limiting normal distribution, $n \to \infty$. These contrasting proofs may not be all found together in a single book or an article in the statistical literature. Readers of this article would find the presentation informative and especially useful from the pedagogical standpoint. This review of proofs should be of interest to teachers and students of senior undergraduate courses in probability and statistics.

Keywords: Binomial distribution, Central limit theorem, Moment generating function, Ratio method, Stirling's approximations

1. Introduction

The binomial distribution was first proposed by Jacob Bernoulli, a Swiss mathematician, in his book *Ars Conjectandi* published in 1713 – eight years after his death [7]. This distribution is probably most widely used discrete distribution in Statistics. Consider a series of n independent trials, each resulting in one of two possible outcomes, a success with probability p $(0 < p < 1)$ and failure with probability $q = 1 - p$. Let X_n denote the number of successes in these n trials.

Then the random variable (r.v.) X_n is said to have binomial distribution with parameters n and p, $b(n,p)$. The probability mass function (pmf) of X_n is given by (see [6], [7])

$$p_n(x) = P(X_n = x) = \frac{n!}{x!(n-x)!} p^x q^{n-x} \quad , \quad x = 0, 1, \cdots, n, \tag{1.1}$$

with $\sum_{x=0}^{n} p_n(x) = 1$. The mean and variance of the binomial r.v. X_n are given, respectively, by $\mu_n = np$ and $\sigma_n^2 = npq$. Most undergraduate elementary level statistics books list binomial probability tables (e.g., [6], [7]) for specified values of $n(\leq 30)$ and p. It is well known that (see [5]) if both np and nq are greater than 5, then the binomial probabilities given by the pmf (1.1) can be satisfactorily approximated by the corresponding normal probability density function (pdf). These approximations (see [5]) turn out to be fairly close for n as low as 10 when p is in a neighborhood of $1/2$. The French mathematician Abraham de Moivre (1738) (See Stigler 1986, pp.70-88) was the first to suggest approximating the binomial distribution with the normal when n is large. Here in this article, in addition to his proof based on the Stirling's formula, we shall present three other methods – the Ratio Method, the Method of Moment Generating Functions, and that of the Central Limit Theorems – for demonstrating the convergence of binomial to the limiting normal distribution. Under the first two methods, this is achieved by showing the convergence, as $n \to \infty$, of the "standardized" pmf of $b(n,p)$ to the standard normal probability density function (pdf). Under the latter two, this is achieved by showing the convergence, as $n \to \infty$, of the Laplace or Fourier transform of the Binomial distribution $b(n,p)$ to a Laplace or Fourier transform, from which then the standard normal distribution is identified as the limiting distribution.

The organization of the paper is described as follows. We state some useful preliminary results in Section 2. In Section 3, we provide the details of various proofs of the convergence of binomial distribution to the limiting normal. Section 4 contains some concluding remarks.

2. Preliminaries

In this section we state a few handy formulas, Lemmas, and Theorems which shall be needed in describing various methods of proofs presented in Section 3.

Formula 2.1. *When* n *is large, the Stirling's approximation formula* (*see* [1], [10]) *for approximating factorial function* $n! = n(n-1)(n-2)\cdots(2)(1)$ *is given by*

$$n! \approx \sqrt{2\pi n}(n)^n e^{-n} \text{ (i.e., } [n!/(\sqrt{2\pi n}n^n e^{-n})] \to 0 \text{ as } n \to \infty \text{).} \tag{2.1}$$

Formula 2.2. (i) *For* $-1 < x \le 1$, $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^k}{k}$.

(2.2a)

In this series if we replace x by $-x$, we get

(ii) *for* $-1 \le x < 1$, $\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots = -\sum_{k=1}^{\infty} \frac{x^k}{k}$. (2.2b)

Definition 2.1. *Let* X *be a r.v. with probability mass function (pmf) or probability density function (pdf)* $f_X(x)$, $-\infty < x < \infty$. *Then the moment generating function (mgf) of the r.v.* X *is defined to the function*

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} f_X(x), & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x)dx, & \text{if } X \text{ is continuous} \end{cases}$$

for all $|t| < h$, $h > 0$.

If the mgf exists (i.e., if it is finite), there is only one unique distribution with this mgf. That is, there is a one-to-one correspondence between the r.v.'s and the mgf's if they exist. Consequently, by recognizing the form of the mgf of a r.v X , one can identify the distribution of this r.v.

Theorem 2.1. *Let* $\{M_{X_n}(t), n = 1, 2, \cdots\}$ *denote the sequence of mgf's corresponding to the sequence* $\{X_n\}$ *of r.v.'s* $n = 1, 2, \cdots$ *and* $M_X(t)$ *the mgf of a r.v.* X *, which are all assumed to exist for* $|t| < h$, $h > 0$. *If* $\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$ *for* $|t| < h$ *then* $X_n \xrightarrow{d} X$.

The notation $X_n \xrightarrow{d} X$ means that, as $n \to \infty$, the distribution of the r.v. X_n converges to the distribution of the r.v. X .

Lemma 2.1. *Let* $\{\psi(n), n \geq 1\}$ *be a sequence of real numbers such that* $\lim_{n\to\infty} \psi(n) = 0$. *Then* $\lim_{n\to\infty}\left(1+\frac{a}{n}+\frac{\psi(n)}{n}\right)^{bn} = e^{ab}$, *provided* a *and* b *do not depend on* n.

Theorem 2.2. *Let* $X_1, X_2, \cdots$ *be a sequence of independent and identically distributed (i.i.d.) r.v.'s with finite mean* μ *and variance* $\sigma^2 > 0$. *Set* $S_n = \sum_{i=1}^{n} X_i$ *and* $\bar{X}_n = (S_n/n)$. *Then,* $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1)$ *as* $n \to \infty$.

Theorem 2.2 is known as the *Central Limit Theorem* (CLT). The symbol $Z \sim N(0,1)$ denotes that the r.v. Z follows $N(0,1)$, which notation stands for the normal distribution with mean 0 and variance 1, referred to as the standard normal distribution.

Lemma 2.2. *Let* $Z \sim N(0,1)$, *whose pdf is given by* $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$, $-\infty < z < \infty$; *then* $M_Z(t) = e^{t^2/2}$.

Proof. Since $Z \sim N(0,1)$, it follows that

$$M_Z(t) = E(e^{tZ}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tz} e^{-z^2/2} dz = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(1/2)(z-t)^2} dz = e^{t^2/2}. \qquad \square$$

Big O and Small o Notations:

The Big-O equation $g(n) = O(f(n))$ symbolizes that ratio $[g(n)/f(n)]$ remains bounded, as $n \to \infty$; i.e., there exists a positive constant M such that $|g(n)/f(n)| < M$ for all n. On the other hand, Small-o equation $g(n) = o(f(n))$ symbolizes that the ratio $|g(n)/f(n)| \to 0$, as $n \to \infty$; i.e., given $\varepsilon > 0$, there is an $n_0 = n_0(\varepsilon)$ such that $|g(n)/f(n)| < \varepsilon$ for all $n \geq n_0$.

Thus, for example, if $|f(n)| \to 0$, as $n \to \infty$, $g(n) = O(f(n))$ implies that $|g(n)| \to 0$ at the same or higher rate than that of $|f(n)|$; whereas $g(n) = o(f(n))$ implies that $|g(n)| \to 0$ at a higher rate than that of $|f(n)|$.

For Definition 2.1, Theorem 2.1, Theorem 2.2, and Lemma 2.1, see Bain and Engelhardt 1992, pp. 78, 236, 238, and 234, respectively [3].

3. Proofs of Various Methods

In this section, we present four different proofs of the convergence of binomial $b(n, p)$ distribution to a limiting normal distribution, as $n \to \infty$.

3.1. Use of Stirling's Approximation Formula [4]

Using Stirling's formula given in Definition 2.1, the binomial pmf (1.1) can be approximated as

$$P(X_n = x) \approx \frac{\sqrt{2\pi n}(n)^n e^{-n}}{\sqrt{2\pi x}(x)^x e^{-x} \sqrt{2\pi(n-x)}(n-x)^{n-x} e^{-(n-x)}} p^x q^{n-x}$$

$$= \frac{\sqrt{n}}{\sqrt{2\pi}\sqrt{x(n-x)}} \left(\frac{np}{x}\right)^x \left(\frac{nq}{n-x}\right)^{n-x}$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{npq}\sqrt{\left(\frac{x}{np}\right)\left(\frac{n-x}{nq}\right)}} \left(\frac{x}{np}\right)^{-x} \left(\frac{n-x}{nq}\right)^{-(n-x)}$$

$$= C\left(\frac{x}{np}\right)^{-x-1/2} \left(\frac{n-x}{nq}\right)^{-(n-x)-1/2}, \tag{3.1}$$

where $C = 1/[\sqrt{2\pi}\sqrt{npq}]$. Now taking natural logarithms on both sides of (3.1), we have

$$\ln P(X_n = x) \approx \ln C - (x + 1/2)\ln\left(\frac{x}{np}\right) - (n - x - 1/2)\ln\left(\frac{n-x}{nq}\right). \tag{3.2}$$

Let $Z_n = \dfrac{X_n - np}{\sqrt{npq}}$ and $z = \dfrac{x - np}{\sqrt{npq}}$. The last equation leads to: $x = np + z\sqrt{npq}$

so that $\dfrac{x}{np} = \left(1 + z\sqrt{\dfrac{q}{np}}\right)$; and $n - x = nq - z\sqrt{npq}$ so that $\dfrac{n-x}{nq} = \left(1 - z\sqrt{\dfrac{p}{nq}}\right)$.

In view of the preceding equations and equations (2.2a) and (2.2b), we can rewrite (3.2) as

$$\ln P(Z_n = z) \approx \ln C - (np + z\sqrt{npq} - 1/2)\ln\left(1 + z\sqrt{\frac{q}{np}}\right)$$

$$-(nq - z\sqrt{npq} - 1/2)\ln\left(1 - z\sqrt{\frac{p}{nq}}\right)$$

$$=: \ln C + I_n(z,p,q) + I_n(-z,p,q) \qquad \text{(say)} \tag{3.3}$$

where, using Taylor series expansions and detailing of the expressions, we obtain

$$I_n(z,p,q) = -(np + z\sqrt{npq} - 1/2)\sum_{j=1}^{\infty}\frac{(-1)^{j-1}}{j}\left(\sqrt{\frac{q}{np}}\right)^j z^j$$

$$= -np\left[z\sqrt{\frac{q}{np}} - \frac{z^2}{2}\left(\frac{q}{np}\right) + \frac{z^3}{3}\left(\frac{q}{np}\right)^{3/2}\sum_{j=4}^{\infty}\frac{(-j)^{j-1}}{j}\left(\sqrt{\frac{q}{np}}\right)^{j-3} z^{j-1}\right]$$

$$-z\sqrt{npq}\left[z\sqrt{\frac{q}{np}} - \frac{z^2}{2}\left(\frac{q}{np}\right) + z^2\left(\frac{q}{np}\right)\sum_{j=3}^{\infty}\frac{(-j)^{j-1}}{j}\left(\sqrt{\frac{q}{np}}\right)^{j-2} z^{j-2}\right]$$

$$+\frac{1}{2}\left[z\sqrt{\frac{q}{np}} - z\left(\frac{q}{np}\right)^{1/2}\sum_{j=2}^{\infty}\frac{(-j)^{j-1}}{j}\left(\sqrt{\frac{q}{np}}\right)^{j-1} z^{j-1}\right]$$

$$= -z\sqrt{npq} + \frac{z^2}{2}q - \frac{z^3}{3}\frac{q^{3/2}}{\sqrt{np}} + \frac{z^3 q^{3/2}}{\sqrt{np}}\sum_{j=1}^{\infty}\frac{(-1)^{j-4}}{(j+3)}\left(\sqrt{\frac{q}{np}}\right)^j z^j$$

$$-z^2 q + \frac{z^3}{2}\frac{q^{3/2}}{\sqrt{np}} - \frac{z^3 q^{3/2}}{\sqrt{np}} - \frac{z^3 q^{3/2}}{\sqrt{np}}\sum_{j=1}^{\infty}\frac{(-1)^{j-3}}{(j+2)}\left(\sqrt{\frac{q}{np}}\right)^j z^j$$

$$+\frac{1}{2}\left[z\sqrt{\frac{q}{np}} - z\left(\frac{q}{np}\right)^{1/2}\sum_{j=1}^{\infty}\frac{(-j)^{j-2}}{(j+1)}\left(\sqrt{\frac{q}{np}}\right)^j z^j\right]$$

$$= -z\sqrt{npq} - \frac{z^2}{2}q + O(1/\sqrt{n}), \tag{3.3a}$$

for all fixed z, as $n \to \infty$, the last equality following clearly from the preceding expression, since the three infinite sums in this expression are all (in absolute value) dominated – for each fixed z and sufficiently large n – by $\ln\left[1/(1-|z|\sqrt{q/np})\right]$, which converges to zero, as $n \to \infty$, for each fixed z. Similarly, using Taylor series expansions and appropriate detailing of terms as in (3.3a) (in fact, by just inter changing p and q and changing the sign of z in (3.3a)), we obtain

$$I_n(-z,q,p) = z\sqrt{npq} - \frac{z^2}{2}p + O(1/\sqrt{n}), \tag{3.3b}$$

for all fixed z, as $n \to \infty$. Combining (3.3), (3.3a), and (3.3b), we have

$$\ln P(Z=z) \approx \ln C - \frac{z^2}{2} + O(1/\sqrt{n}). \tag{3.4}$$

Hence, from (3.4) we get for large n,

$$P(Z_n = z) \approx Ce^{-z^2/2} = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz, \tag{3.4a}$$

where $dz \approx \Delta_n = \left(1/\sqrt{npq}\right)$. The above equation (3.4a) may also be written as

$$P(X_n = x) \approx \frac{1}{\sqrt{2\pi}\sqrt{npq}} e^{-\frac{(x-np)^2}{2(npq)}}. \tag{3.4b}$$

This completes the proof. □

Remark 3.1. *Uniform Approximation*. From equations (3.3) (a) - (b), it becomes apparent at once that the order term $O(1/\sqrt{n})$ in (3.4) holds uniformly in z over any given real compact interval $[z_1, z_2]$. Consequently, the approximate equality (3.4a) holds uniformly in z over any compact sub-interval on the real line. In fact, it follows from these equations that the preceding approximation does also hold uniformly, as $n \to \infty$, over the expanding sequence of compact intervals $[z_n^{(-)}, z_n^{(+)}]$, where $z_n^{(-)} = [(k_n^{(-)} - np)/\sqrt{npq}]$ and $z_n^{(+)} = [(k_n^{(+)} - np)/\sqrt{npq}]$, with $k_n^{(-)}$ and $k_n^{(+)}$ given by $k_n^{(-)} = \min\{k : [(z_n^{(k)})^3/\sqrt{n}] \uparrow 0\}$ and $k_n^{(+)} = \max\{k : [(z_n^{(k)})^3/\sqrt{n}] \downarrow 0\}$. We leave the details of this as an exercise for the readers (cf. Feller [4] Vol I; Theorem 1, pp. 184-186)

3.2. The Ratio Method [8]

The ratio method was introduced by Proschan (2008). The ratio of two successive probability terms of the binomial pmf stated in (1.1) is given by

$$\frac{P(X_n = x+1)}{P(X_n = x)} = \frac{x!(n-x)!}{(x+1)!(n-x-1)!}\frac{p}{q} = \frac{(n-x)}{(x+1)}\frac{p}{q} \tag{3.5}$$

Consider the transformation $z = z_n = (x-np)/\sqrt{npq}$. Then $x = np + z\sqrt{npq}$, so that substituting this value of x into (3.5) produces

$$\frac{P(X_n = np + z\sqrt{npq} + 1)}{P(X_n = np + z\sqrt{npq})} = \frac{(n - np - z\sqrt{npq})}{(np + z\sqrt{npq} + 1)}\left(\frac{p}{q}\right). \qquad (3.6)$$

The ratio on the left hand side of (3.6) can be expressed as

$$\frac{P\left[(X_n - np)\big/\sqrt{npq} = z + 1\big/\sqrt{npq}\right]}{P\left[(X_n - np)\big/\sqrt{npq} = z\right]} = \frac{P\left(Z_n = z + \Delta_n\right)}{P(Z_n = z)}, \quad (3.7)$$

where $Z_n = (X_n - np)\big/\sqrt{npq}$ and $\Delta_n = 1\big/\sqrt{npq}$. Note that as $n \to \infty,\ \Delta_n \to 0$. The ratio on the right hand side of (3.6) can be simplified as

$$\frac{(n - np - z\sqrt{npq})}{(np + z\sqrt{npq} + 1)}\left(\frac{p}{q}\right) = \frac{(q - z\sqrt{(pq)/n})}{(p + z\sqrt{(pq)/n} + 1/n)}\left(\frac{p}{q}\right)$$

$$= \frac{(1 - z\sqrt{p/(nq)})}{(1 + z\sqrt{q/(np)} + 1/np)} = \frac{(1 - zp\Delta_n)}{(1 + zq\Delta_n + q\Delta_n^2)}. \quad (3.8)$$

Combining (3.7) and (3.8), we get

$$\frac{P(Z_n = z + \Delta_n)}{P(Z_n = z)} = \frac{(1 - zp\Delta_n)}{(1 + zq\Delta_n + q\Delta_n^2)}. \qquad (3.9)$$

Now assume the existence of a "sufficiently" smooth pdf $f(z)$ such that for large n, the probability $P(Z_n = z)$ can be approximated by the differential $f(z)dz$. Hence, we can take $P(Z_n = z) \approx f(z)dz$ and $P(Z_n = z + \Delta_n) \approx f(z + \Delta_n)dz$ on the left hand side of (3.9), so that the probability ratio becomes equivalent to $f(z + \Delta_n)\big/ f(z)$. This, in conjunction with (3.9), implies that

$$\frac{f(z + \Delta_n)}{f(z)} \approx \frac{(1 - zp\Delta_n)}{(1 + zp\Delta_n + q\Delta_n^2)}. \qquad (3.10)$$

Now taking logarithms on both side of (3.10), we have

$$\ln f(z + \Delta_n) - \ln f(z) = \ln(1 - zp\Delta_n) - \ln(1 + zq\Delta_n + q\Delta_n^2), \qquad (3.11)$$

Dividing both sides of (3.11) by Δ_n and taking the limits as $n \to \infty$, or equivalently, as $\Delta_n \to 0$, the equation (3.11) takes the form

$$\lim_{n\to\infty}\left[\frac{\ln f(z+\Delta_n)-\ln f(z)}{\Delta_n}\right] \approx \lim_{\Delta_n\to 0}\left[\frac{\ln(1-zp\Delta_n)}{\Delta_n}-\frac{\ln\left(1+zq\Delta_n+q\Delta_n^2\right)}{\Delta_n^2}\right]. \tag{3.12}$$

The left hand side of (3.12) is nothing but the derivative of $\ln[f(z)]$. Applying L'Hopital's theorem, the limit on the right hand side of (3.12) takes the form

$$\lim_{\Delta_n\to 0}\left[\frac{\ln(1-zp\Delta_n)}{\Delta_n}-\frac{\ln\left(1+zq\Delta_n+q\Delta_n^2\right)}{\Delta_n}\right] = \lim_{\Delta_n\to 0}\left[\frac{-zp}{(1-zp\Delta_n)}\right]-\lim_{\Delta_n\to 0}\left[\frac{zq+2q\Delta_n}{(1+zq\Delta_n+q\Delta_n^2)}\right]$$
$$= -zp - zq = -z\ . \tag{3.13}$$

Now combining (3.12) and (3.13) we may conclude that $\dfrac{d\ln[f(z)]}{dz} \approx -z$. Integrating both sides of this equation with respect to z, we get $\ln[f(z)] \approx -\dfrac{z^2}{2}+c$, where c is the constant of integration. This in turn gives $f(z) \approx ke^{-z^2/2}$, where $k=e^c$ must be equal $1/\sqrt{2\pi}$ to make the RHS a valid $N(0,1)$ density. Thus, we conclude that, as $n \to \infty$, the r.v. $Z_n = (X_n - np)/\sqrt{npq}$ has the standard normal as its distribution in the limit; or equivalently, that the r.v. X_n follows approximately a normal distribution with mean $\mu_n = np$ and variance $\sigma_n^2 = npq$ when n is large.
□

3.3. The MGF Method [2], [3]

Let X_n be a binomial r.v. with parameters n and p. Then the mgf of the r.v. X_n evaluates to $M_{X_n}(t) = \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x q^{n-x} = (q+pe^t)^n$. Let $Z_n = \left(X_n - np\right)/\sqrt{npq}$, so that upon setting $\sigma_n = \sqrt{npq}$, we can rewrite Z_n as $Z_n = X_n/\sigma_n - np/\sigma_n$. Below we derive the mgf of Z_n, which is given by

$$M_{Z_n}(t) = E(e^{tZ_n}) = E(e^{t(X_n/\sigma_n - np/\sigma_n)}) = e^{-npt/\sigma_n} E(e^{(t/\sigma_n)X_n})$$

$$= e^{-npt/\sigma_n} M_{X_n}(t/\sigma_n) = e^{-npt/\sigma_n}\left(q + pe^{t/\sigma_n}\right)^n$$

$$= \left(qe^{-pt/\sigma_n} + pe^{qt/\sigma_n}\right)^n . \qquad (3.14)$$

The Taylor series expansion for e^{qt/σ_n} may be written as

$$e^{qt/\sigma_n} = 1 + \frac{qt}{\sigma_n} + \frac{q^2t^2}{(2!)\sigma_n^2} + \frac{q^3t^3}{(3!)\sigma_n^3} + \frac{q^4t^4}{(4!)\sigma_n^4} e^{\xi(n)} , \qquad (3.15)$$

where $\xi(n)$ is a number between 0 and $qt/\sigma_n = t\sqrt{q/np}$, and $\xi(n) \to 0$ as $n \to \infty$. Similarly, the Taylor's series expansion for e^{-pt/σ_n} may be written as

$$e^{-pt/\sigma_n} = 1 - \frac{pt}{\sigma_n} + \frac{p^2t^2}{(2!)\sigma_n^2} - \frac{p^3t^3}{(3!)\sigma_n^3} + \frac{p^4t^4}{(4!)\sigma_n^4} e^{\varsigma(n)} , \qquad (3.16)$$

where $\varsigma(n)$ is a number between 0 and $pt/\sigma_n = t\sqrt{p/nq}$, and $\varsigma(n) \to 0$ as $n \to \infty$.

Now substituting these two equations (3.15) and (3.16) in the last expression for $M_{Z_n}(t)$ in (3.14), we have

$$M_{Z_n}(t) = \left(1 + \left(\frac{pqt}{\sigma_n} - \frac{pqt}{\sigma_n}\right) + \frac{pqt^2}{2!\sigma_n^2}(q+p) + \frac{pqt^3}{3!\sigma_n^3}(q^2 - p^2) + \frac{pqt^4}{4!\sigma_n^4}(q^3e^{\xi(n)} - p^3e^{\varsigma(n)})\right)^n$$

$$= \left(1 + \frac{t^2}{2n} + \frac{t^3(q-p)}{(n)(3!)(npq)^{1/2}} + \frac{t^4(q^3e^{\xi(n)} - p^3e^{\varsigma(n)})}{(n)(4!)(npq)}\right)^n . \qquad (3.17)$$

The above equation (3.17) may be rewritten as

$$M_{Z_n}(t) = \left(1 + \frac{t^2}{2n} + \frac{\psi(n)}{n}\right)^n, \text{ where } \psi(n) = \frac{t^3(q-p)}{\sqrt{n}(\sqrt{pq})(3!)} + \frac{t^4(q^3e^{\xi(n)} - p^3e^{\varsigma(n)})}{n(pq)(4!)} .$$

Since $\xi(n)$, $\varsigma(n) \to 0$ as $n \to \infty$, it follows that $\lim_{n\to\infty} \psi(n) = 0$ for every fixed value of t. Thus, by Lemma 2.1, we have

$$\lim_{n\to\infty} M_{Z_n}(t) = e^{t^2/2} = M_Z(t)\text{, where } Z \sim N(0,1)\text{,}$$

for all real values of t. In view of Theorem 2.1, thus, we can conclude that, as $n \to \infty$, the r.v. $Z_n = (X_n - np)/\sqrt{npq}$ has the standard normal as its limiting distribution; or equivalently, that the binomial r.v. X_n has, for large n, an approximate normal distribution with mean $\mu_n = np$ and variance $\sigma_n^2 = npq$. □

3.4. The CLT Method [3], [4]
Let $Y_1, Y_2, \cdots, Y_n$ be a sequence of n independent and identical Bernoulli r.v.'s each with success probability p success, i.e.,

$$Y_i = \begin{cases} 1 & \text{with prob. } p \\ 0 & \text{with prob. } (1-p) = q \end{cases}$$

for $i = 1, 2, \cdots, n$. Then the sum $X_n = \sum_{i=1}^{n} Y_i$ can be shown to have a Binomial distribution with parameters n and p. To verify this assertion, we use the mgf methodology of Section 2: The mgf of each Y_k is $M_{Y_k}(t) = E(e^{tY_k}) = pe^t + qe^{(0)t}$ $= (q + pe^t)$, $k = 1, 2, \cdots, n$, so that the mgf of $X_n = \sum_{i=1}^{n} Y_i$ is given by $M_{X_n}(t) =$ $E(e^{tX_n}) = = E\left(e^{t\sum_{i=1}^{n} Y_i}\right) = (M_{Y_k}(t))^n = (q + pe^t)^n$. This is exactly the directly evaluated mgf of a binomial r.v. with parameters n and p in Section 3.3. Thus, X_n follows a $b(n, p)$ distribution. Since X_n is a sum of n independent and identically distributed r.v.'s with finite mean $\mu = p$ and variance $\sigma^2 = pq$, by the CLT Theorem 2.2, we can conclude at once that, $Z_n = \left[(X_n - n\mu)/(\sqrt{n}\sigma)\right]$ $= \left[(X_n - np)/(\sqrt{npq})\right] \xrightarrow{d} N(0,1)$, as $n \to \infty$; or equivalently, that the r.v. X_n follows approximately the normal distribution with mean np and variance npq when n is large. □

4. Concluding Remarks

This article presents four methods – a new elementary method and three other well-known ones – for deriving the convergence of the standard binomial $b(n, p)$ distribution to the limiting standard normal distribution, as $n \to \infty$.

The new elementary Ratio Method, of recent origin, is in fact just a modification of one the other three, viz., of the de Moivre's method based on Stirling's Approximation. The Ratio Method, however, circumvents the use of Stirling's approximation. The other two methods are, respectively, the Moment Generating Functions method based on Laplace transforms and the last one that of the Central Limit Theorems, based on Fourier transforms and complex Analysis.

While the last three methods utilize advanced theoretical tools and are broader in scope, the Ratio Method requires knowledge of only elementary calculus, namely, of derivatives, Taylor's series expansions, simple integration etc., along with that of basic probability concepts.

All four different types of methods have been employed here in one single place to demonstrate the convergence of the standardized binomial r.v. to a limiting normal $N(0,1)$ r.v.. This article may serve as a useful teaching reference paper. Its contents can be discussed profitably in senior level probability or mathematical statistics classes. The teachers may assign these different methods to students in senior level probability classes as class projects.

Acknowledgements. This article was presented in abstract and poster form at Florida Chapter of American Statistical Association (ASA) Annual meeting 2016, Tallahassee and ASA Joint Statistical Meeting 2016, Chicago.

References

[1] S.C. Bagui, S.S. Bagui, and R. Hemasinha, Nonrigorous proofs of Stirling's formula, *Mathematics and Computer Education*, **47** (2013), 115-125.

[2] S.C. Bagui and K. L. Mehra, Convergence of binomial, Poisson, negative-binomial, and gamma to normal distribution: Moment generating functions technique, *American Journal of Mathematics and Statistics*, **6** (2016), 115-121.

[3] L.J. Bain and M. Engelhardt, *Introduction to Probability and Mathematical Statistics*, Duxbury, Belmont, 1992.

[4] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, Wiley Eastern Limited, New Delhi, 1968.

[5] A. de Moivre, *The Doctrine of Chances*, H. Woodfall London, 1738.

[6] R. Khazanie, *Elementary Statistics in a World Applications*, 2nd ed., Scott, Foresman and Company, Glenview, 1986.

[7] J.T. McClave and T. Sincich, *Statistics*, Pearson, New Jersey, 2006.

[8] M.A. Proschan, The normal approximation to the binomial, *The American Statistician*, **62** (2008), 62-63. https://doi.org/10.1198/000313008x267848

[9] S.M. Stigler, S.M., *The History of Statistics: The Measurement of Uncertainty before 1900*, Cambridge, MA: The Belknap Press of Harvard University Press, 1986.

[10] J. Stirling, *Methodus Differentialis*, London, 1730.

Received: February 15, 2017; Published: March 19, 2017