American Journal of Mathematics and Statistics 2016, 6(3): 115-121
DOI: 10.5923/j.ajms.20160603.05

Convergence of Binomial, Poisson, Negative-Binomial, and Gamma to Normal Distribution: Moment Generating Functions Technique

Subhash C. Bagui[1,*], K. L. Mehra[2]

[1]Department of Mathematics and Statistics, University of West Florida, Pensacola, USA
[2]Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, USA

Abstract In this article, we employ moment generating functions (mgf's) of Binomial, Poisson, Negative-binomial and gamma distributions to demonstrate their convergence to normality as one of their parameters increases indefinitely. The motivation behind this work is to emphasize a direct use of mgf's in the convergence proofs. These specific mgf proofs may not be all found together in a book or a single paper. Readers would find this article very informative and especially useful from the pedagogical stand point.

Keywords Binomial distribution, Central limit theorem, Gamma distribution, Moment generating function, Negative-Binomial distribution, Poisson distribution

1. Introduction

The basic Central Limit Theorem (CLT) tells us that, when appropriately normalised, sums of independent identically distributed (i.i.d.) random variables (r.v.'s) from any distribution, with finite mean and variance, would have their distributions converge to normality, as the sample size n tends to infinity. If we accept this CLT and are in knowledge of the fact that Binomial, Poisson, Negative-binomial and Gamma r.v.'s are themselves sums of i.i.d. r.v.'s, we can conclude the limiting normality of these distributions by applying this CLT. We must note, however, that the proof of this CLT is based on the use of Characteristic Functions theory involving Complex Analysis, the study of which primarily only advanced math majors in colleges and universities undertake. There are available, indeed, other methods of proof in specific cases, e.g., in case of Binomial and Poisson distributions through approximations of probability mass functions (pmf) by the corresponding normal probability density function (pdf) using Stirling's formula (cf., Stigler, S.M. 1986, pp.70-88, [8]; Bagui et al. 2013b, p. 115, [2]) or by simply approximating the ratios of successive pmf terms of the distribution one is dealing with (cf., Proschan, M.A. 2013, pp. 62-63, [6]). However, by using the parallel (to characteristic functions) methodology of mgf's, which does not involve Complex Analysis, we can also accomplish the same objective with relative ease. This is what we propose to explicitly demonstrate in this paper.

* Corresponding author:
sbagui@uwf.edu (Subhash C. Bagui)
Published online at http://journal.sapub.org/ajms

The structure of the paper is as follows. We provide some useful preliminary results in Section 2. These results will be used in section 3. In Section 3 we give all the details of convergence for all the above mentioned distributions to normal distribution. Section 4 contains some concluding remarks.

2. Preliminaries

In this section, we state some results that will be used in various proofs presented in section 3.

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 as*

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

assume to exist and be finite for all $|t| < h$ *for an* $h > 0$.

If X has a normal distribution with mean μ and variance σ^2 , then mgf of X is given by

$M_X(t) = e^{\mu t + \sigma^2 (t^2/2)}$, [3]. If $Z = (X - \mu)/\sigma$, then Z is said to have standard normal distribution (i.e., a normal distribution with mean zero and variance one). The mgf of Z is given by $M_Z(t) = e^{t^2/2}$.

Let F_X denote the cumulative distribution function (cdf) of the r.v. X.

Theorem 2.1. *Let F_X and F_Y be two cumulative distribution functions (cdf's) whose moments exist. If the mgf's exist for the r.v.'s X and Y and $M_X(t) = M_Y(t)$ for all t in $-h < t < h$, $h > 0$, then* $F_X(u) = F_Y(u)$ for all u (i.e., $f_X(u) = f_Y(u)$ for all u.)

A probability distribution is not always determined by its moments. Suppose X has cdf F_X and moments $E\left(X^r\right) = \mu_r'$ which exist for all $r = 1, 2, \cdots$. If $\sum_{r=1}^{\infty} \frac{\mu_r' t^r}{r!}$ has a positive radius of convergence for all $-h < t < h$, $h > 0$ (Billingsley 1995, Section 30, [4]; Serfling 1980, p. 46, [7]), then mgf exists in the interval $-h < t < h$, $h > 0$, and hence uniquely determines the probability distribution.

A weaker sufficient condition for the moment sequence to determine a probability distribution uniquely is $\sum_{r=1}^{\infty} \frac{1}{(\mu_{2r}')^{1/(2r)}} = +\infty$. This sufficient condition is due to Carleman (Chung 1974, p. 82, [5]; Serfling 1980, p. 46, [7]).

Theorem 2.2. *Let $\{X_n, n \ge 1\}$ be a sequence of r.v.'s with the corresponding mgf sequence as $M_{X_n}(t)$, $n = 1, 2, \cdots$ and X be a r.v. with mgf $M_X(t)$ which are assumed exist for all $-h < t < h$, $h > 0$. If $\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$ for $-h < t < h$, 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 $\{\psi(n), n \ge 1\}$ be a sequence of reals. 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 an $\lim_{n\to\infty} \psi(n) = 0$.*

CLT *(See Bagui et al. 2013a, [1])***.** *Let $\{X_n : n \ge 1\}$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ, $-\infty < \mu < \infty$, and variance σ^2, $0 < \sigma^2 < \infty$, and set $S_n = \sum_{i=1}^{n} X_i$, $\bar{X}_n = [S_n/n]$ and*

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}.$$

Then $Z_n \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$, where $N(0,1)$ stands for a normal distribution with mean 0 *and variance* 1.

For Definition 2.1, Theorem 2.1, Theorem 2.2, and Lemma 2.1, see Casella and Berger, 2002, pp. 62-66, [4] and Bain and Engelhardt, 1992, p. 234, [3].

3. Congergence of Mgf's

3.1. Binomial

Binomial probabilities apply to situations involving a series of n independent and identical trials with two possible outcomes –a success with probability p and a failure with probability $q = 1 - p$ - on each trial. Let X_n be the number of successes in n trials, then X_n has binomial distribution with parameters n and p. The probability mass function of X_n is given by $f_{X_n}(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \cdots, n$. Thus the mean of X_n is $E(X_n) = np$ and the variance of X_n is Var $(X_n) = npq,$ $q = 1 - p$. The mgf of X_n is given by $M_{X_n}(t) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} = \left(q + pe^t\right)^n$. Let $Z_n = \left(X_n - np\right)/\sqrt{npq}$. With simplified notation $\sigma_n = \sqrt{npq}$, we have $Z_n = X_n/\sigma_n - np/\sigma_n$. Below we derive the mgf of Z_n. Now the mgf of Z_n is given by

$$M_{Z_n}(t)=E\left(e^{tZ_n}\right)=E\left(e^{t\left(X_n/\sigma_n-np/\sigma_n\right)}\right)=e^{-npt/\sigma_n}E\left(e^{(t/\sigma_n)X_n}\right)$$
$$=e^{-npt/\sigma_n}M_{X_n}(t/\sigma_n)=e^{-npt/\sigma_n}\left(q+pe^{t/\sigma_n}\right)^n$$
$$=\left(qe^{-pt/\sigma_n}+pe^{qt/\sigma_n}\right)^n. \quad (3.1)$$

Based on the Taylor's series expansion, there exists a number $\xi(n)$, between 0 and $qt/\sigma_n=t\sqrt{q/np}$ such that

$$e^{qt/\sigma_n}=1+\frac{qt}{\sigma_n}+\frac{q^2t^2}{(2!)\sigma_n^2}+\frac{q^3t^3}{(3!)\sigma_n^3}+\frac{q^4t^4}{(4!)\sigma_n^4}e^{\xi(n)}, \text{ where } \xi(n)\to 0 \text{ as } n\to\infty. \quad (3.2)$$

Similarly, based on the Taylor's series expansion, there exists a number $\varsigma(n)$, between 0 and $pt/\sigma_n=t\sqrt{p/nq}$ such that

$$e^{-pt/\sigma_n}=1-\frac{pt}{\sigma_n}+\frac{p^2t^2}{(2!)\sigma_n^2}-\frac{p^3t^3}{(3!)\sigma_n^3}+\frac{p^4t^4}{(4!)\sigma_n^4}e^{\varsigma(n)}, \text{ where } \varsigma(n)\to 0 \text{ as } n\to\infty. \quad (3.3)$$

Now substituting these two equations (3.2) and (3.3) in the last expression for $M_{Z_n}(t)$ in (3.1), we have

$$M_{Z_n}(t)=\left(1+\left(\frac{pqt}{\sigma_n}-\frac{pqt}{\sigma_n}\right)+\frac{pqt^2}{2!\sigma_n^2}(q+p)+\frac{pqt^3}{3!\sigma_n^3}(q^2-p^2)+\frac{pqt^4}{4!\sigma_n^4}(q^3e^{\xi(n)}-p^3e^{\varsigma(n)})\right)^n$$
$$=\left(1+\frac{t^2}{2n}+\frac{t^3(q-p)}{(n)(3!)(npq)^{1/2}}+\frac{t^4(q^3e^{\xi(n)}-p^3e^{\varsigma(n)})}{(n)(4!)(npq)}\right)^n. \quad (3.4)$$

The above equation (3.4) may be written 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 $n\to\infty$, then $\lim_{n\to\infty}\psi(n)=0$ for every fixed value of t. Thus based on Lemma 2.1 we have

$$\lim_{n\to\infty}M_{Z_n}(t)=e^{t^2/2}$$

for all real values of t. That is, in view of Theorems 2.1 and 2.2, we conclude that the r.v. $Z_n=\left(X_n-np\right)/\sqrt{npq}$ has the limiting standard normal distribution. Consequently, the binomial r.v. X_n has, for large n, an approximate normal distribution with mean $\mu_n=np$ and variance $\sigma_n^2=npq$.

3.2. Poisson

The Poisson distribution is appropriate for predicting rare events within a certain period of time. Let X_λ be a Poisson r.v. with parameter λ. The probability mass function of X_λ is given by $f_{X_\lambda}(x)=\frac{e^{-\lambda}\lambda^x}{x!}$, $x=0,1,2,\cdots$. Both the mean and variance of X_λ are λ. The mgf of X_λ is given by

$M_{X_\lambda}(t)=\sum_{x=0}^{\infty}e^{tx}\frac{e^{-\lambda}\lambda^x}{x!}=e^{\lambda(e^t-1)}$. For notational convenience let $\lambda=n$ and $Z_n=(X_n-n)/\sqrt{n}=(X_n/\sqrt{n}-\sqrt{n})$.

Below we derive the mgf of Z_n, which is given by

$$M_{Z_n}(t)=E\left(e^{tZ_n}\right)=E\left[e^{t(X_n/\sqrt{n}-\sqrt{n})}\right]=e^{-t\sqrt{n}}E\left[e^{\left(t/\sqrt{n}\right)X_n}\right]=e^{-t\sqrt{n}}M_{X_n}(t/\sqrt{n})$$

$$=e^{-t\sqrt{n}}e^{n\left(e^{t/\sqrt{n}}-1\right)}=\left[e^{-t}e^{\sqrt{n}\left(e^{t/\sqrt{n}}-1\right)}\right]^{\sqrt{n}}. \tag{3.5}$$

Now consider the simplification of the term $\sqrt{n}\left(e^{t/\sqrt{n}}-1\right)$ as

$\sqrt{n}\left(e^{t/\sqrt{n}}-1\right)=\sqrt{n}\left(1+\frac{t}{\sqrt{n}}+\frac{t^2}{(2!)n}+\frac{t^3}{(3!)n^{3/2}}+\frac{t^4}{(4!)n^2}e^{\varsigma(n)}-1\right)$, where $\varsigma(n)$ is number between 0 and $\frac{t}{\sqrt{n}}$ and converges to zero as $n\to\infty$. Further the above term $\sqrt{n}\left(e^{t/\sqrt{n}}-1\right)$ may be simplified as $\sqrt{n}\left(e^{t/\sqrt{n}}-1\right)$ $=t+\frac{t^2}{(2!)\sqrt{n}}+\frac{t^3}{(3!)n}+\frac{t^4}{(4!)n^{3/2}}\varsigma(n)$. Now substituting this in the last expression (3.5) for $M_{Z_n}(t)$, we have

$$M_{Z_n}(t)=\left[e^{-t}e^{t+t^2/[(2!)\sqrt{n}]+t^3/[(3!)n]+t^4\varsigma(n)/[(4!)n^{3/2}]}\right]^{\sqrt{n}}=e^{t^2/2}b(n),$$

where $b(n)=e^{t^3/[(3!)\sqrt{n}]}e^{[t^4\varsigma(n)]/[(4!)n^{3/2}]}$ which tends to 1 as $n\to\infty$. Hence, we have

$$\lim_{n\to\infty}M_{Z_n}(t)=e^{t^2/2}$$

for all real values of t. Using Theorems 2.1 and 2.2 we conclude that $Z_n=(X_n-n)/\sqrt{n}$ has the limiting standard normal distribution. Hence, the Poisson r.v. X_λ has also an approximate normal distribution with both mean and variance equal to $\lambda=n$, for large n.

3.3. Negative Binomial

Consider an infinite series of independent trials, each having two possible outcomes, success or failure. Let $p=P(\text{success})$ and $q=P(\text{failure})=1-p$. Define the random variable X_n to be the number of failures before the n th success. Then X_n has negative binomial distribution with parameters n and p. Thus, the probability mass function of X_n is given by $f_{X_n}(x)=\binom{n+x-1}{x}p^nq^x$, $x=0,1,2,\cdots$. The mean of X_n is given by $E(X_n)=nq/p$ and the variance of X_n is given by $\text{Var}(X_n)=nq/p^2$. The mgf of X_n can be obtained as $M_{X_n}(t)=\sum_{x=0}^{\infty}e^{tx}\binom{n+x-1}{x}p^nq^x$ $=\left[p/(1-qe^t)\right]^n$. Let $Z_n=(X_n-nq/p)/(\sqrt{nq}/p)=\left[(pX_n)/\sqrt{nq}\right]-\sqrt{nq}$. Now the mgf of Z_n is given by

$$M_{Z_n}(t) = E\left(e^{tZ_n}\right) = E\left[e^{t\left(pX_n/\sqrt{nq}-\sqrt{nq}\right)}\right]$$

$$= e^{-(\sqrt{nq})t}E\left[e^{(p/\sqrt{nq})tX_n}\right] = e^{-(\sqrt{nq})t}M_{X_n}\left(\frac{pt}{\sqrt{nq}}\right)$$

$$= e^{-(\sqrt{nq})t}\left[p\Big/(1-qe^{pt/\sqrt{nq}})\right]^n = e^{-(q/\sqrt{nq})(t)n}\left(\tfrac{1}{p}-\tfrac{q}{p}e^{t/\sqrt{nq}}\right)^{-n}$$

$$= \left(\tfrac{1}{p}e^{(q/\sqrt{nq})t} - \tfrac{q}{p}e^{(t/\sqrt{nq})}\right)^{-n}. \tag{3.6}$$

According to Taylor's series expansion, there exists a number $\xi(n)$, between 0 and $\frac{q}{\sqrt{nq}}t$ such that

$$\frac{1}{p}e^{(q/\sqrt{nq})t} = \frac{1}{p}\left[1+\frac{qt}{\sqrt{nq}}+\frac{q^2t^2}{(2!)nq}+\frac{q^3t^3}{(3!)(nq)^{3/2}}e^{\xi(n)}\right]$$

$$= \frac{1}{p}+\frac{qt}{p\sqrt{nq}}+\frac{qt^2}{p(2n)}+\frac{(q)q^2t^3}{p(3!)(nq)^{3/2}}e^{\xi(n)}, \text{ where } \xi(n)\to 0 \text{ as } n\to\infty. \tag{3.7}$$

Similarly, there exists a number $\varsigma(n)$ between 0 and $\frac{t}{\sqrt{nq}}$ such that

$$\frac{q}{p}e^{t/\sqrt{nq}} = \frac{q}{p}\left[1+\frac{t}{\sqrt{nq}}+\frac{t^2}{(2!)nq}+\frac{t^3}{(3!)(nq)^{3/2}}e^{\varsigma(n)}\right]$$

$$= \frac{q}{p}+\frac{qt}{p\sqrt{nq}}+\frac{t^2}{p(2n)}+\frac{qt^3}{p(3!)(nq)^{3/2}}e^{\varsigma(n)}, \text{ where } \varsigma(n)\to 0 \text{ as } n\to\infty. \tag{3.8}$$

Now substituting these two expressions (3.7) and (3.8) in the last expression for $M_{Z_n}(t)$ in (3.6), we have

$$M_{Z_n}(t) = \left[\left(\frac{1}{p}-\frac{q}{p}\right)-\frac{t^2}{2n}\left(\frac{1}{p}-\frac{q}{p}\right)+\frac{q}{p}\frac{t^3}{(3!)(nq)^{3/2}}\left(q^2e^{\xi(n)}-e^{\varsigma(n)}\right)\right]^{-n}$$

$$= \left[\frac{1-q}{p}-\frac{t^2}{2n}\frac{1-q}{p}+\frac{t^3}{(n)p\sqrt{nq}}\left(q^2e^{\xi(n)}-e^{\varsigma(n)}\right)\right]^{-n}$$

$$= \left[1-\frac{t^2}{2n}+\frac{1}{n}\frac{t^3}{(3!)\sqrt{nq}}\left(q^2e^{\xi(n)}-e^{\varsigma(n)}\right)\right]^{-n}. \tag{3.9}$$

The above equation (3.9) can be written as $M_{Z_n}(t) = \left(1-\frac{t^2}{2n}+\frac{\psi(n)}{n}\right)^{-n}$, where $\psi(n) = \frac{t^3}{\sqrt{nq}} \cdot (q^2e^{\xi(n)}-e^{\varsigma(n)})$.

Since both $\xi(n), \varsigma(n)\to 0$ as $n\to\infty$, $\lim_{n\to\infty}\psi(n)=0$ for every fixed value of t.

Hence by lemma 2.1 we have

$$M_{Z_n}(t) = e^{t^2/2}$$

for all real values of t. Hence, by Theorems 2.1 and 2.2, we conclude the r.v.

$Z_n = (X_n - nq/p)/(\sqrt{nq}/p)$ has the limiting standard normal distribution. Accordingly, the negative-Binomial r.v. X_n has approximately a normal distribution with mean $\mu_n = \frac{nq}{p}$ and variance $\sigma_n^2 = \frac{nq}{p^2}$, for large n.

3.4. Gamma

The Gamma distribution is appropriate for modeling waiting times for events. Let X be a Gamma r.v. with pdf $f_X(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$, $\alpha, \beta > 0$ and $x > 0$. The α is called the shape parameter of the distribution and β is called the scale parameter of the distribution. For convenience let us denote α by $\alpha = n$. It is well known that the mean of X is $E(X) = n\beta$ and the variance of X is $\mathrm{Var}(X) = n\beta^2$. The mgf of X is given by

$$M_X(t) = \frac{1}{\Gamma(n)\beta^n} \int_0^\infty e^{tx} x^{n-1} e^{-x/\beta} dx = (1-\beta t)^{-n}, \; t < 1/\beta.$$

Let $Z_n = (X - n\beta)/\beta\sqrt{n} = X/\beta\sqrt{n} - \sqrt{n}$. The mgf of Z_n is given by

$$M_{Z_n}(t) = E(e^{tZ_n}) = E\left[e^{t\left(X/\beta\sqrt{n} - \sqrt{n}\right)}\right] = e^{-\sqrt{n}(t)} E\left[e^{(t/\beta\sqrt{n})X}\right] = e^{-\sqrt{n}(t)} M_X(t/(\beta\sqrt{n}))$$

$$= e^{-(t/\sqrt{n})n}\left(1 - \frac{t}{\sqrt{n}}\right)^{-n} = \left[e^{t/\sqrt{n}} - \frac{t}{\sqrt{n}} e^{t/\sqrt{n}}\right]^{-n}, \; t < \sqrt{n}. \tag{3.10}$$

Observe that $e^{t/\sqrt{n}} = 1 + \frac{t}{\sqrt{n}} + \frac{t^2}{2n} + \frac{t^3}{(3!)n^{3/2}} e^{\xi(n)}$, where $\xi(n)$ is a number between 0 and $\frac{1}{\sqrt{n}}$ and tends to zero as $n \to \infty$, and $\frac{t}{\sqrt{n}} e^{t/\sqrt{n}} = \frac{t}{\sqrt{n}} + \frac{t^2}{n} + \frac{t^3}{(2!)n^{3/2}} + \frac{t^4}{(3!)n^2} e^{\xi(n)}$. Now substituting these two in the last expression of $M_{Z_n}(t)$ in (3.10), we have $M_{Z_n}(t) = (1 - \frac{t^2}{2n} + \frac{t^3 e^{\xi(n)}}{(3!)n^{3/2}} - \frac{t^3}{(2!)n^{3/2}} - \frac{t^4 e^{\xi(n)}}{(3!)n})^{-n}$. This can be written as $M_{Z_n}(t) = \left(1 - \frac{t^2}{2n} + \frac{\psi(n)}{n}\right)^{-n}$, where $\psi(n) = \frac{t^3 e^{\xi(n)}}{(3!)\sqrt{n}} - \frac{t^3}{2\sqrt{n}} - \frac{t^4 e^{\xi(n)}}{(3!)n}$. Since $\xi(n) \to 0$ as $n \to \infty$, $\lim_{n\to\infty} \psi(n) = 0$ for every fixed value of t. Hence by Lemma 2.1 we have

$$M_{Z_n}(t) = e^{t^2/2}$$

for all real values of t. Hence, by Theorems 2.1 and 2.2, we conclude the r.v. $Z_n = (X - n\beta)/\beta\sqrt{n}$ has the limiting standard normal distribution. Accordingly, the Gamma r.v. X has approximately a normal distribution with mean $\mu_n = n\beta$ and varianc $\sigma_n^2 = n\beta^2$, for large n.

4. Concluding Remarks

It is well-known that a Binomial r.v. is the sum of i.i.d. Bernouli r.v.'s, a Poisson $P(\lambda)$ r.v., with $\lambda = n$ a positive integer, the sum of n i.i.d. $P(1)$ r. v.'s, a Negative-binomial r.v. the sum of i.i.d. geometric r.v.'s and a Gamma r.v. the sum of i.i.d. exponential r.v.'s. In view of these facts, one can easily conclude by applying the above stated general CLT that the above distributions, after proper normalizations, converge to a normal distribution as n, the number of terms in their respective sums, increases to infinity. But these facts may be beyond the knowledge of undergraduate students, especially those who are non-math majors. However, as demonstrated in the preceding Section 3 for the Binomial, Poisson, Negative-binomial and Gamma distributions, in dealing with distributional convergence problems where individual mgf's exist and are available, we can use the mgf technique effectively to formally deduce their limiting distributions. In our view, this latter technique is natural, equally instructive and at a more manageable level. In any case, it provides an alternative approach.

In the proof of general central limit theorem using mgf both Bain and Engelhardt (1992), [3] and Inlow (2010), [6a] use the mgf of sum of i.i.d r.v's. But we are using the existing mgf of all the above mentioned distributions without treating them as sums of i.i.d. r.v.'s. Bain and Engelhardt (1992), [3] discusses a proof of convergence of binomial to normal using mgf. But this paper formalizes mgf proofs of collection of distributions. The paper framed in this way can serve as an excellent teaching reference. The proofs are straightforward and require only an additional knowledge of Taylor series expansion, beyond the skills to handle algebraic equations and basic probabilistic concepts. The material should be of pedagogical interest, and can be discussed in classes where only basic calculus and skills to deal with algebraic expressions are the only background requirements. The article should also be of reading interest for senior undergraduate students in probability and statistics.

ACKNOWLEDGEMENTS

The authors are thankful to the Editor-in-Chief and an anonymous referee for their careful reading of the paper.

REFERENCES

[1] Bagui, S.C., Bhaumik, D.K., Mehra, K.L. (2013a). A few counter examples useful in teaching central limit theorem, *The American Statistician*, 67(1), 49-56.

[2] Bagui, S.C., Bagui, S.S., Hemasinha, R. (2013b). Nonrigourous proof's Stirling's formula, *Mathematics and Computer Education*, 47(2), 115-125.

[3] Bain, L.J. and Engelhardt, M. ((1992). *Introduction to Probability and Mathematical Statistics*, 2nd edition, Belmont: Duxbury Press.

[4] Billingsley, P. (1995). *Probability and Measure*, 3rd edition, New York: Wiley.

[5] Casella, G. and Berger, R.L. (2002). *Statistical Inference*, Pacific Grove: Duxbury.

[6] Chung, K.L (1974). *A Course in Probability Theory*, New York: Academic Press.

a. Inlow, Mark (2010). A moment generating function proof of the Lindeberg-Lévy central limit theorem, *The American Statistician*, 64(3), 228-230.

[7] Proschan, M.A. (2008). The Normal approximation to the binomial, *The American Statistician*, 62(1), 62-63.

[8] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*, New York; Wiley.

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