Derivation of Gaussian Probability Distribution: A New Approach

A. T. Adeniran[1], O. Faweya[2*], T. O. Ogunlade[3], K. O. Balogun[4]

[1]Department of Statistics, University of Ibadan, Ibadan, Nigeria
[2]Department of Statistics, Ekiti State University (EKSU), Ado-Ekiti, Nigeria
[3]Department of Mathematics, Ekiti State University (EKSU), Ado-Ekiti, Nigeria
[4]Department of Statistics, Federal School of Statistics, Ibadan, Nigeria
Email: at.adeniran@mail.ui.edu.ng, *lanrefaweya@gmail.com, topsmatic@gmail.com, ko.balogun1@yahoo.com

Abstract

The famous de Moivre's Laplace limit theorem proved the probability density function of Gaussian distribution from binomial probability mass function under specified conditions. De Moivre's Laplace approach is cumbersome as it relies heavily on many lemmas and theorems. This paper invented an alternative and less rigorous method of deriving Gaussian distribution from basic random experiment conditional on some assumptions.

Keywords

De Moivres Laplace Limit Theorem, Binomial Probability Mass Function, Gaussian Distribution, Random Experiment

1. Introduction

A well celebrated, fundamental probability distribution for the class of continuous functions is the classical Gaussian distribution named after the German Mathematician Karl Friedrich Gauss in 1809.

Definition 1.1 *Let* μ *and* σ *be constants with* $-\infty < \mu < \infty$ *and* $\sigma > 0$. *The function*

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; \quad \text{for } -\infty < x < \infty \tag{1}$$

is called the normal probability density function of a random variable X with parameters μ and σ.

Both in theories and applications, without element of equivocation, the Gaussian distribution function is the most essential and widely referencing distribu-

tion in statistics.

The well-known method of deriving this distribution first appeared in the second edition of the Doctrine of Chances by Abraham de Moivre (hence, **de Moivre's Laplace limit theorem**) published in 1738 ([1] [2] [3] [4] [5]). The mathematical statement of the popular de Moivre's theorem follows.

Theorem 1.1 (de Moivre's Laplace limit theorem) *As n grows large ($n \to \infty$), for x in the neighborhood of np, for moderate values of p ($p \neq 0$ and $p \neq 1$), we can approximate*

$$\binom{n}{x} p^x q^{n-x} \approx \frac{1}{\sqrt{2\pi npq}} \mathrm{e}^{-\frac{(x-np)^2}{2npq}} , \quad p+q=1, \ p,q>0. \tag{2}$$

Explicitly, the theorem asserts that suppose $n \in \mathbb{Z}^+$, and let p and q be probabilities, with $p+q=1$. The function

$$b(x;n,p) = \binom{n}{x} p^x (1-p)^{n-x} \quad \text{for } x = 0,1,2,\cdots,n \tag{3}$$

called the **binomial probability function** converges to the probability density function of the normal distribution as $n \to \infty$ with mean np and standard deviation $\sqrt{np(1-p)}$.

Although, De Moivre proved the result for $p = \frac{1}{2}$ ([6] [7]). [8] extended and generalized the proof to all values of p (probability of success in any trial) such that p is not too small and not too big. Feller result was expounded by [9]. [10] [11] [12] [13] used uniqueness property of moment generating function technique to proof the same theorem.

In this paper, we attempt to find an answer to the question: is there any alternative procedure to the derivation of Gaussian probability density function apart from de Moivre's Laplace limit theorem approach which relies heavily on many Lemmas and Theorems (Stirling approximation formula, Maclaurin series expansion etc.), as evidenced by the work of [8] and [9]?

2. Existing Technique

This section presents the summary proof of the existing de Moivre's Laplace limit theorem. First and foremost, the study state with proof, the most important lemma of the de-Moivre's Laplace limit theorem, Stirling approximation principle.

Lemma 2.1 (Stirling Approximation Principle) *Given an integer $n; n > 0$, the factorial of a large number n can be replaced with the approximation*

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

Proof 2.1 *This lemma can be derived using the integral definition of the factorial,*

$$n! = \Gamma(n+1) = \int_0^\infty x^n \mathrm{e}^{-x} \mathrm{d}x \tag{4}$$

Note that the derivative of the logarithm of the integrand can be written

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln\left(x^n\mathrm{e}^{-x}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\left(n\ln x - x\right) = \frac{n}{x} - 1 \tag{5}$$

The integrand is sharply peaked with the contribution important only near $x = n$. Therefore, let $x = n + \delta$ where $\delta \le n$, and write

$$\ln\left(x^n\mathrm{e}^{-x}\right) = n\ln\left(n+\delta\right) - \left(n+\delta\right) = n\ln\left[n\left(1+\frac{\delta}{n}\right)\right] - \left(n+\delta\right)$$

$$= n\left[\ln\left(n\right) + \ln\left(1+\frac{\delta}{n}\right)\right] - \left(n+\delta\right) \tag{6}$$

Recall that the Maclaurin series of $f\left(x\right) = \ln\left(1+x\right) = x - \frac{1}{2}x^2 + 0\left(n\right)$. Therefore,

$$\ln\left(x^n\mathrm{e}^{-x}\right) = n\left[\ln\left(n\right) + \frac{\delta}{n} - \frac{1}{2}\left(\frac{\delta^2}{n^2}\right) + \cdots\right] - \left(n+\delta\right) = \ln\left(n^n\right) - n - \frac{\delta^2}{2n} + \cdots \tag{7}$$

Taking the exponential on both sides of the preceding Equation (7) gives

$$x^n\mathrm{e}^{-x} \approx \mathrm{e}^{\ln\left(n^n\right)-n-\frac{\delta^2}{2n}} = \mathrm{e}^{\ln\left(n^n\right)}\mathrm{e}^{-n}\mathrm{e}^{-\frac{\delta^2}{2n}} = \left(\frac{n}{\mathrm{e}}\right)^n\mathrm{e}^{-\frac{\delta^2}{2n}} \tag{8}$$

Plugging (8) into the integral expression for $n!$, that is, (4) gives

$$n! \approx \int_0^\infty\left(\frac{n}{\mathrm{e}}\right)^n\mathrm{e}^{-\frac{\delta^2}{2n}}\mathrm{d}\delta = \left(\frac{n}{\mathrm{e}}\right)^n\int_{-\infty}^\infty\mathrm{e}^{-\frac{\delta^2}{2n}}\mathrm{d}\delta \tag{9}$$

From (9), let $I = \int_0^\infty\mathrm{e}^{-\frac{\delta^2}{2n}}\mathrm{d}\delta$ and considering δ and κ as a dummy variable such that

$$I^2 = \int_0^\infty\mathrm{e}^{-\frac{\delta^2}{2\alpha}}\mathrm{d}\delta \times \int_0^\infty\mathrm{e}^{-\frac{\kappa^2}{2n}}\mathrm{d}\kappa = \int_0^\infty\int_0^\infty\mathrm{e}^{-\frac{1}{2n}\left(\delta^2+\kappa^2\right)}\mathrm{d}\delta\mathrm{d}\kappa \tag{10}$$

Transforming I^2 from algebra to polar coordinates yields $\delta = \rho\cos\left(\theta\right)$, $\kappa = \rho\sin\left(\theta\right)$ which implies $\delta^2 + \kappa^2 = \rho^2$ with Jacobian (J) of the transformation as

$$J = \begin{vmatrix} \dfrac{\partial\delta}{\partial\rho} & \dfrac{\partial\delta}{\partial\theta} \\ \dfrac{\partial\kappa}{\partial\rho} & \dfrac{\partial\kappa}{\partial\theta} \end{vmatrix} = \begin{vmatrix} \cos\left(\theta\right) & -\rho\sin\left(\theta\right) \\ \sin\left(\theta\right) & \rho\cos\left(\theta\right) \end{vmatrix} = \rho \tag{11}$$

Hence,

$$I^2 = \int_0^{2\pi}\int_0^\infty\mathrm{e}^{-\frac{\rho^2}{2n}}\left|J\right|\mathrm{d}\rho\mathrm{d}\theta = \int_0^{2\pi}\int_0^\infty\mathrm{e}^{-\frac{\rho^2}{2n}}\rho\mathrm{d}\rho\mathrm{d}\theta$$

$$= -n\int_0^{2\pi}\left[\mathrm{e}^{-u}\right]_0^\infty\mathrm{d}\theta = n\int_0^{2\pi}\mathrm{d}\theta = 2\pi n \tag{12}$$

Therefore, $I = I = \sqrt{2\pi n}$. Substituting for I in (9) gives

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n \quad \square \tag{13}$$

We now begin with proof of theorem (1.1) using the popular existing technique.

Proof 2.2 *Using the result of lemma* (2.1), *Equation* (3) *can be rewritten as*

$$f(x;n,p) \approx \frac{\sqrt{2\pi n}\left(\dfrac{n}{e}\right)^n}{\sqrt{2\pi x}\left(\dfrac{x}{e}\right)^x \sqrt{2\pi(n-x)}\left(\dfrac{n-x}{e}\right)^{n-x}} p^x (1-p)^{n-x} \tag{14}$$

$$= \frac{1}{\sqrt{2\pi}} \frac{n^{n+\frac{1}{2}}}{x^{x+\frac{1}{2}}(n-x)^{n-x+\frac{1}{2}}} p^x (1-p)^{n-x}$$

Multiplying both numerator and denominator of Equation (14) by $\sqrt{n} \equiv n^{\frac{1}{2}}$ to get

$$f(x;n,p) \approx \frac{1}{\sqrt{2\pi n}} \frac{n^{n+\frac{1}{2}+\frac{1}{2}+x-x}}{x^{x+\frac{1}{2}}(n-x)^{n-x+\frac{1}{2}}} p^x (1-p)^{n-x} \tag{15}$$

$$= \frac{1}{\sqrt{2\pi n}}\left(\frac{x}{n}\right)^{-x-\frac{1}{2}}\left(\frac{n-x}{n}\right)^{-n+x-\frac{1}{2}} p^x (1-p)^{n-x}$$

Since x is in the neighborhood of np, change variables $x = np + \varepsilon$, where ε measures the distance from the mean, np, of the binomial and the measured quantity x. Re-write (15) in terms of ε and further simplify as follow

$$f(x;n,p) \approx \frac{1}{\sqrt{2\pi n}}\left(\frac{np+\varepsilon}{n}\right)^{-x-\frac{1}{2}}\left(\frac{n-np-\varepsilon}{n}\right)^{-n+x-\frac{1}{2}} p^x (1-p)^{n-x}$$

$$\approx \frac{1}{\sqrt{2\pi n}}\left[p\left(1+\frac{\varepsilon}{np}\right)\right]^{-x-\frac{1}{2}}\left[(1-p)\left(1-\frac{\varepsilon}{n(1-p)}\right)\right]^{-n+x-\frac{1}{2}} p^x (1-p)^{n-x}$$

$$\approx \frac{1}{\sqrt{2\pi n}}\left(1+\frac{\varepsilon}{np}\right)^{-x-\frac{1}{2}}\left(1-\frac{\varepsilon}{n(1-p)}\right)^{-n+x-\frac{1}{2}} p^{-\frac{1}{2}}(1-p)^{-\frac{1}{2}}$$

to get

$$f(x;n,p) \approx \frac{1}{\sqrt{2\pi np(1-p)}}\left(1+\frac{\varepsilon}{np}\right)^{-x-\frac{1}{2}}\left(1-\frac{\varepsilon}{n(1-p)}\right)^{-n+x-\frac{1}{2}} \tag{16}$$

Note that $x = \exp(\ln x)$. Therefore, rewriting (16) in exponential form to have

$$f(x;n,p) \approx \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left[\ln\left\{\left(1+\frac{\varepsilon}{np}\right)^{-x-\frac{1}{2}}\left(1-\frac{\varepsilon}{n(1-p)}\right)^{-n+x-\frac{1}{2}}\right\}\right]$$

$$\approx \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left[\left(-x-\frac{1}{2}\right)\ln\left(1+\frac{\varepsilon}{np}\right)+\left(-n+x-\frac{1}{2}\right)\ln\left(1-\frac{\varepsilon}{n(1-p)}\right)\right]$$

$$f(x;n,p) \approx \frac{1}{\sqrt{2\pi np(1-p)}}\exp\left[\left(-np-\varepsilon-\frac{1}{2}\right)\ln\left(1+\frac{\varepsilon}{np}\right)\right.$$
$$\left. +\left(-n(1-p)+\varepsilon-\frac{1}{2}\right)\ln\left(1-\frac{\varepsilon}{n(1-p)}\right)\right] \tag{17}$$

Suppose $f(x) = \ln(1+x)$, using Maclaurin series $f(x) = x - \frac{1}{2}x^2 + 0(n)$

and similarly $f(x) = \ln(1-x) = -x - \frac{1}{2}x^2 + 0(n)$. So that,

$$\ln\left(1+\frac{\varepsilon}{np}\right) \approx \frac{\varepsilon}{np} - \frac{1}{2}\left(\frac{\varepsilon}{np}\right)^2 \quad \text{and} \quad \ln\left(1-\frac{\varepsilon}{n(1-p)}\right) \approx -\frac{\varepsilon}{n(1-p)} - \frac{1}{2}\left(\frac{\varepsilon}{n(1-p)}\right)^2 .$$

As a result,

$$f(x;n,p)$$

$$\approx \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left[-\varepsilon + \frac{\varepsilon^2}{2np} - \frac{\varepsilon^2}{np} + \varepsilon + \frac{\varepsilon^2}{2n(1-p)} - \frac{\varepsilon^2}{n(1-p)}\right] \qquad (18)$$

$$= \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left[-\frac{1}{2}\left(\frac{\varepsilon^2}{np(1-p)}\right)\right]$$

Recall that $x = np + \varepsilon$ which implies that $\varepsilon^2 = (x-np)^2$. From binomial distribution $np = \mu$, and $np(1-p) = \sigma^2$ which implies that $\sqrt{np(1-p)} = \sigma$. Making appropriate substitution of these in the Equation (18) yields

$$f(x;n,p) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right] = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; \quad \text{for } -\infty < x < \infty \quad (19)$$

The theorem confirmed.

We recommend that readers interested in the detailed proof of the theorem to consult the study expounded by [9].

3. The Proposed Technique

Suppose a random experiment of throwing needle or any other dart related objects at the origin of the cartesian plane is performed with the aim of hitting the centre (see **Figure 1**).

Due to human nature of inconsistency or lack of perfection, varying results in the throwing generate random errors. To make the derivation possible and less rigorous, we make the following assumptions:

1) The errors are independent of the orientation of the coordinate system.

2) Errors in perpendicular directions are independent. This means that being too high doesn't alter the probability of being off to the right.
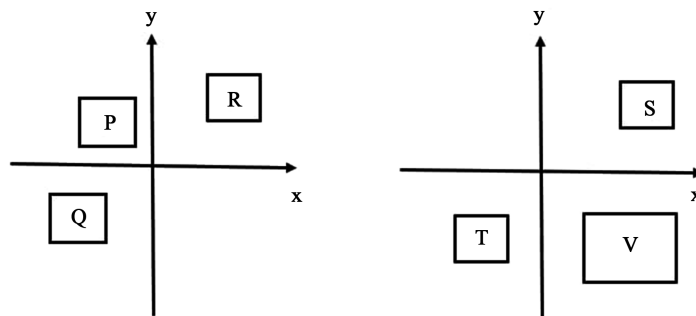

Figure 1. The possible results of the dart experiment.

3) Small errors are more likely than large errors. That is, throwings are more likely to land in region P than either Q or R, since region P is closer to the target (origin). Similarly, for the same reason, region Q is more likely than region R. Furthermore, there is higher possibility or tendency of hitting region V than either S or T, since V has the wider or bigger surface area and the distances from the origin are approximately the same.

From **Figure 2**, let the probability of the needle falling in the vertical strip from x to $x + \Delta x$ be denoted as $p(x)\Delta x$. Similarly, the probability of the needle falling in the horizontal strip from y to $y + \Delta y$ be $p(y)\Delta y$. Obviously, the function cannot be constant, due to the stochastic nature of the experiment. In this study, our interest is to know and obtain the form and characteristics of the function $p(x)$. From second assumption, the probability of the needle falling in the shaded region ABCD (see **Figure 2**) is

$$p(x)\Delta x \cdot p(y)\Delta y$$

Note that any regions r unit from the origin with area $\Delta x \Delta y$ has the same probability which is a consequence of the assumption that errors do not depend on the orientation. We can say that

$$p(x)\Delta x \cdot p(y)\Delta y = p(x)p(y)\Delta x \Delta y = g(r)\Delta x \Delta y \qquad (20)$$

where

$$g(r) = p(x)p(y) \qquad (21)$$

from fundamental rule of Calculus, differentiating (using product rule) both sides of Equation (21) with respect to θ gives

$$0 = p(x)\frac{\mathrm{d}}{\mathrm{d}\theta}p(y) + p(y)\frac{\mathrm{d}}{\mathrm{d}\theta}p(x) \qquad (22)$$

Here, $g'(r) = 0$ since $g(.)$ is independent of orientation. By transformation to polar coordinates, $x = r\cos\theta$ and $y = r\sin\theta$, we can rewrite the derivatives in Equation (22) as
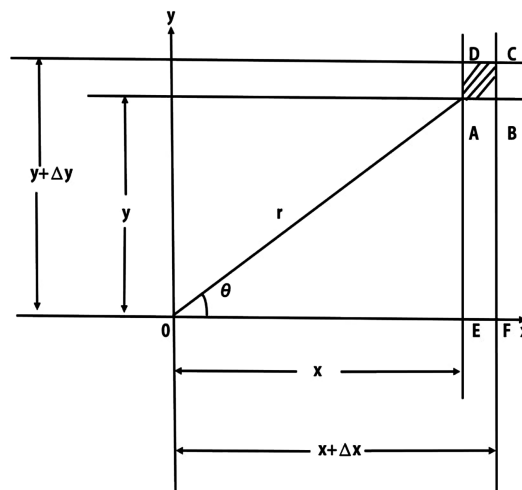

Figure 2. The typical example of the experiment.

$$0 = p(x)\frac{\mathrm{d}}{\mathrm{d}\theta}p(y = r\sin\theta) + p(y)\frac{\mathrm{d}}{\mathrm{d}\theta}p(x = r\cos\theta) \qquad (23)$$

Using chain rule of differentiation, (23) becomes

$$0 = p(x)p'(y)r\cos\theta - p(y)p'(x)r\sin\theta \qquad (24)$$

Rewriting Equation (24) again by replacing $r\cos\theta$ with x and $r\cos\theta$ with y yields

$$0 = p(x)p'(y)x - p(y)p'(x)y \qquad (25)$$

The above differential equation can be put in a form such that it can be solved using variable separable technique as

$$\frac{p'(x)}{p(x)x} = \frac{p'(y)}{p(y)y} \qquad (26)$$

This differential equation can only be true for any x and y, x and y are independent, if and only if the ratio $\dfrac{p'(x)}{p(x)x} = \dfrac{p'(y)}{p(y)y}$ defined by (26) is a constant. That is, if

$$\frac{p'(x)}{p(x)x} = \frac{p'(y)}{p(y)y} = c. \qquad (27)$$

Consider $\dfrac{p'(x)}{p(x)x} = c$ in (27) and rearrange to have

$$\frac{p'(x)}{p(x)} = cx. \qquad (28)$$

Integrating Equation (28) gives

$$\ln p(x) = \frac{cx^2}{2} + k_1 \quad \text{so that} \quad p(x) = ke^{\frac{cx^2}{2}}; \quad \text{where } k = e^{k_1}. \qquad (29)$$

By third assumption, c must be negative so that we write the probability function (29)

$$p(x) = ke^{-\frac{c}{2}x^2}; \quad \text{where } c \in \mathbb{R}^+ \qquad (30)$$

If there is a horizontal shift of target from the origin to an arbitrary point μ which now mark the new center/target, then the probability function in (30) becomes

$$p(x) = ke^{-\frac{c}{2}(x-\mu)^2} \qquad (31)$$

Differentiating (31) and set the derivative equal to zero gives

$$p'(x) = -ck(x-\mu)e^{-\frac{c(x-\mu)^2}{2}} = 0 \qquad (32)$$

since $e^{-\frac{c(x-\mu)^2}{2}} \neq 0$ implies $x = \mu$. Therefore, Equation (31) has maximum value at $x = \mu$ and point of inflexion at $x = \mu \pm \dfrac{1}{\sqrt{k}}$. Obviously, (31) has given

us the basic form of the Gaussian distribution with constants k and c, and domain of X as $-\infty$ to ∞. Therefore, for Equation (31) to be regarded as a proper probability density function, the total area under the curve must be 1. That is

$$\int_{-\infty}^{\infty} k e^{-\frac{c}{2}(x-\mu)^2} dx = 1 \tag{33}$$

For a symmetric function $f(x)$, $\int_{-\infty}^{\infty} f(x) dx = 2\int_{0}^{\infty} f(x) dx$. Applying this property to Equation (33) yields

$$\int_{0}^{\infty} e^{-\frac{c}{2}(x-\mu)^2} dx = \frac{1}{2k}. \tag{34}$$

Squaring both sides of (34) to get

$$\int_{0}^{\infty} e^{-\frac{c}{2}(x-\mu)^2} dx \cdot \int_{0}^{\infty} e^{-\frac{c}{2}(y-\mu)^2} dy = \frac{1}{2k} \times \frac{1}{2k} \tag{35}$$

This is possible since x and y are just dummy variables. Recall that x and y are also independent, so we can write the product in LHS of (35) as a double integral to produce

$$\int_{0}^{\infty}\int_{0}^{\infty} e^{-\frac{c}{2}\left[(x-\mu)^2 + (y-\mu)^2\right]} dx dy = \frac{1}{4k^2}. \tag{36}$$

Putting $x - \mu = z \Rightarrow dx = dz$ and $y - \mu = w \Rightarrow dy = dw$ in the preceding Equation (36) gives

$$\int_{0}^{\infty}\int_{0}^{\infty} e^{-\frac{c}{2}\left[z^2 + w^2\right]} dz dw = \frac{1}{4k^2}. \tag{37}$$

The double integral (37) can be evaluated using polar coordinates as $z = r\cos\theta$ and $y = r\sin\theta$ with Jacobian (J) of the transformation as

$$J = \left| \frac{\partial(z,w)}{\partial(r,\theta)} \right| = \begin{vmatrix} \dfrac{dz}{dr} & \dfrac{dz}{d\theta} \\ \dfrac{dw}{dr} & \dfrac{dw}{d\theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r, \tag{38}$$

and

$$z^2 + w^2 = (r\cos\theta)^2 + (r\sin\theta)^2 = r^2. \tag{39}$$

So, Equation (37) now becomes

$$\int_{0}^{\frac{\pi}{2}}\int_{0}^{\infty} e^{-\frac{cr^2}{2}} |J| dr d\theta = \int_{0}^{\frac{\pi}{2}}\int_{0}^{\infty} e^{-\frac{cr^2}{2}} r dr d\theta = \frac{1}{4k^2}. \tag{40}$$

Evaluating the double integral $\int_{0}^{\frac{\pi}{2}}\int_{0}^{\infty} e^{-\frac{cr^2}{2}} r dr d\theta$ in Equation (40) by first letting $u = \dfrac{cr^2}{2}$, and solving for k in the resulting equation yields

$$k = \sqrt{\frac{c}{2\pi}} \tag{41}$$

Putting (41) in (31), the probability density function, $p(x)$, becomes

$$p(x) = \sqrt{\frac{c}{2\pi}} e^{-\frac{c}{2}(x-\mu)^2} \tag{42}$$

Again, integration of probability function over its domain gives 1. Therefore, from (42)

$$\int_{-\infty}^{\infty} p(x)\,dx = \int_{-\infty}^{\infty} \sqrt{\frac{c}{2\pi}} e^{-\frac{c}{2}(x-\mu)^2}\,dx = 2\int_0^{\infty} \sqrt{\frac{c}{2\pi}} e^{-\frac{c}{2}(x-\mu)^2}\,dx = 1 \tag{43}$$

Further simplification of the preceding Equation (43) gives

$$\int_0^{\infty} e^{-\frac{c}{2}(x-\mu)^2} = \sqrt{\frac{\pi}{2c}} \tag{44}$$

One of the important goals in mathematical theory of statistics is to obtain the mean and variance of any probability function under study. The mean, μ, is defined to be the value of the integral $\int_{-\infty}^{\infty} x p(x)\,dx$. The variance, σ^2, is the value of the integral $\int_{-\infty}^{\infty} (x-\mu)^2 p(x)\,dx$. Therefore, using Equation (42),

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \sqrt{\frac{c}{2\pi}} e^{-\frac{c}{2}(x-\mu)^2}\,dx = 2\int_0^{\infty} (x-\mu)^2 \sqrt{\frac{c}{2\pi}} e^{-\frac{c}{2}(x-\mu)^2}\,dx \tag{45}$$

or equivalently as

$$\sigma^2 = \sqrt{\frac{2c}{\pi}} \int_0^{\infty} (x-\mu)\left[(x-\mu) e^{-\frac{c}{2}(x-\mu)^2} \right] dx \tag{46}$$

consider Equation (46) and using integration by part ($\int u\,dv = uv - \int v\,du$) with

$u = x - \mu \Rightarrow du = dx$ and $dv = (x-\mu) e^{-\frac{c}{2}(x-\mu)^2}\,dx \Rightarrow v = -\frac{1}{c} e^{-\frac{c}{2}(x-\mu)^2}$, we have

$$\sigma^2 = \sqrt{\frac{2c}{\pi}} \left[-\frac{x-\mu}{c} e^{-\frac{c}{2}(x-\mu)^2} - \int_0^{\infty} -\frac{1}{c} e^{-\frac{c}{2}(x-\mu)^2}\,dx \right]$$

$$= \sqrt{\frac{2c}{\pi}} \left[-\frac{1}{c} \lim_{n\to\infty} (x-\mu) e^{-\frac{c}{2}(x-\mu)^2} \Big|_0^n + \frac{1}{c} \int_0^{\infty} e^{-\frac{c}{2}(x-\mu)^2}\,dx \right]$$

$$= \sqrt{\frac{2c}{\pi}} \left[-\frac{1}{c} \lim_{n\to\infty} (n-\mu) e^{-\frac{c}{2}(n-\mu)^2} + \frac{1}{c} \int_0^{\infty} e^{-\frac{c}{2}(x-\mu)^2}\,dx \right]$$

$$= \sqrt{\frac{2c}{\pi}} \left[0 + \frac{1}{c} \int_0^{\infty} e^{-\frac{c}{2}(x-\mu)^2}\,dx \right] = \frac{1}{c} \sqrt{\frac{2c}{\pi}} \int_0^{\infty} e^{-\frac{c}{2}(x-\mu)^2}\,dx$$

putting (44) in the preceding equation above, gives

$$\sigma^2 = \frac{1}{c} \times \sqrt{\frac{2c}{\pi}} \times \sqrt{\frac{\pi}{2c}} \Rightarrow c = \frac{1}{\sigma^2} \tag{47}$$

Substituting (47) in (42), the derived probability density function has form

$$p(x) = \sqrt{\frac{1}{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty. \tag{48}$$

Based on the three aforestated basic assumptions, we have easily derived Equation (48) famously known anywhere in the whole world as **Normal or Gaussian distribution function** with mean μ and standard deviation σ.

To verify that Equation (19) is a proper probability density function with parameters μ and σ is to show that the integral

$$I = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx$$

is equal to 1.

Change variables of integration by letting $z = \dfrac{x-\mu}{\sigma}$, which implies that $dx = \sigma dz$. Then

$$I = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma dz = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{z^2}{2}} dz = \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-\frac{z^2}{2}} dz$$

so that

$$I^2 = \left[\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-\frac{x^2}{2}} dx\right]\left[\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-\frac{y^2}{2}} dy\right] = \frac{2}{\pi} \int_0^{\infty}\int_0^{\infty} e^{-\frac{x^2+y^2}{2}} dxdy$$

Here x, y are dummy variables. Switching to polar coordinate by making the substitutions $x = r\cos\theta$, $y = r\sin\theta$ produces r as the Jacobian of the transformation. So

$$I^2 = \frac{2}{\pi} \int_0^{\frac{\pi}{2}}\int_0^{\infty} e^{-\frac{r^2}{2}} r dr d\theta$$

Put $a = \dfrac{r^2}{2} \Rightarrow dr = \dfrac{da}{r}$. Therefore,

$$I^2 = \frac{2}{\pi} \int_0^{\frac{\pi}{2}}\int_0^{\infty} e^{-a} r \frac{da}{r} d\theta = -\frac{2}{\pi} \int_0^{\frac{\pi}{2}} \left[e^{-a}\right]_0^{\infty} d\theta = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} d\theta = 1$$

Thus $I = 1$, indicating that (48) is a proper probability density function. Other properties of the distribution such as; moments, moments generating function, cumulant generating function, characteristics function, parameter estimation and the likes can be found in [14] [15] [16].

4. Conclusion

While working with the outlined objective, we are able to establish that there exists an approach that is not only serving as an alternative proof of derivation of the Gaussian probability density function but also free from rigorous mathematical analysis and independent of Lemmas and Theorems. This paper can be classified as a theoretical study of Gaussian distribution and can serve as an excellent teaching reference in probability and statistics classes where only basic calculus and skills to deal with algebraic expressions, Maclaurin series expansion and Euler distribution of second kind (gamma function) are the only background requirements.

Acknowledgements

the improvement of the revised version of the paper.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Van der Vaart, A.W. (1998) Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511802256

[2] Blume, J.D. and Royall, R.M. (2003) Illustrating the Law of Large Numbers (And Confidence Intervals). *The American Statistician*, **57**, 51-57. https://doi.org/10.1198/0003130031081

[3] Lesigne, E. (2005) Heads or Tails: An Introduction to Limit Theorems in Probability. American Mathematical Society, Providence, Volume 28 of Student Mathematical Library. https://doi.org/10.1090/stml/028

[4] Walck, C. (2007) Handbook on Statistical Distributions for Experimentalists. Particle Physics Group, Fysikum University of Stockholm, Stockholm.

[5] Proschan, M.A. (2008) The Normal Approximation to the Binomial. *The American Statistician*, **62**, 62-63. https://doi.org/10.1198/000313008X267848

[6] Shao, J. (1999) Mathematical Statistics. Springer Texts in Statistics, Springer Verlag, New York.

[7] Soong, T.T. (2004) Fundamentals of Probability and Statistics for Engineers. John Wiley & Sons Ltd., Chichester.

[8] Feller, W. (1973) An Introduction to Probability Theory and Its Applications. Volume 1, Third Edition, John Wiley and Sons, Hoboken.

[9] Adeniran, A.T., Ojo, J.F. and Olilima, J.O. (2018) A Note on the Asymptotic Convergence of Bernoulli Distribution. *Research & Reviews: Journal of Statistics and Mathematical Sciences*, **4**, 19-32.

[10] Inlow, M. (2010) A Moment Generating Function Proof of the Lindeberg-Lévy Central Limit Theorem. *The American Statistician*, **64**, 228-230. https://doi.org/10.1198/tast.2010.09159

[11] Bagui, S.C., Bhaumik, D.K. and Mehra, K.L. (2013) A Few Counter Examples Useful in Teaching Central Limit Theorem. *The American Statistician*, **67**, 49-56. https://doi.org/10.1080/00031305.2012.755361

[12] Bagui, S.C., Bagui, S.S. and Hemasinha, R. (2013) Non-Rigorous Proof's Stirling's Formula. *Mathematics and Computer Education*, **47**, 115-125.

[13] Bagui, S.C. and Mehra, K.L. (2016) Convergence of Binomial, Poisson, Negative-Binomial, and Gamma to Normal Distribution: Moment Generating Functions Technique. *American Journal of Mathematics and Statistics*, **6**, 115-121.

[14] Casella, G. and Berger, R.L. (2002) Statistical Inference. Second Edition, Duxbury Thomson Learning: Integre Technical Publishing Co., Albuquerque.

[15] Young, G.A. and Smith, R.L. (2005) Essentials of Statistical Inference. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511755392

[16] Podgorski, K. (2009) Lecture Notes on Statistical Inference. Department of Mathematics and Statistics, University of Limerick, Limerick.