African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 4, 2024 (pp. 153-161)
www.abjournals.org

# PERFECT HEMISPHERE TREND REALIZATION: A COMBINATORIAL MODIFICATION OF PROBABILITY DISTRIBUTIONS

**Mark Nwachukwu Ugo[1], JohnPaul Kenechukwu Iwuchukwu[2], and**

**Emwinloghosa Kenneth Guobadia[3]**

[1]Department of Statistics, School of Information Communication and Technology, Federal Polytechnic, Auchi, Nigeria.

[2,4]Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Nigeria.

[3]Department of Administration, Federal Medical Centre, Asaba, Delta State, Nigeria.

**Cite this article:**

Ugo, M. N., Iwuchukwu, K. J., Guobadia, E. K. (2024), Perfect Hemisphere Trend Realization: A Combinatorial Modification of Probability Distributions. African Journal of Mathematics and Statistics Studies 7(4), 153-161. DOI: 10.52589/AJMSS-K6FRZIGC

**Manuscript History**

Received: 14 Aug 2024

Accepted: 1 Oct 2024

Published: 11 Nov 2024



**ABSTRACT:** *The paper reviews creative ways to develop continuous probability models playing around the integration method and the concept of normalization. It further projects a probability distribution realized by combining two symmetric probability models that differ in shape, to produce a perfect hemisphere or half-sun trend. Normal and arcsine distributions are the root distributions used for this development. At some values of the parameter, the distribution can be right skewed; where other moments-related measures and estimation are studied as their properties, alongside simulation.*

**KEYWORDS:** Normal distribution, Arcsine distribution, Division arrangement, Hemisphere trend.


Article DOI: 10.52589/AJMSS-K6FRZIGC
DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 4, 2024 (pp. 153-161)
www.abjournals.org

## INTRODUCTION

Probability distributions are developed to model various life outcomes, making available the various possible simulated data prototypes of such phenomena. Many credits have been given to veteran researchers for the great advancement made in the provision of probability models, that characterize various shapes including increasing, decreasing, bathtub, inverted bathtub, uniform, skewed, and bell-curve shapes. These are various life trends and can be modelled by some renowned distributions including exponential, gamma, Weibull, Pareto, beta, Lomax, logistic, Burr XII, Gompertz, Gaussian, arcsine, uniform, Kumaraswamy, Lindley, Fréchet, and Gumbel distribution for the continuous category. More so, various forms of modifications and generalizations of many of these distributions can give weight to these trends expressed as light or heavy tail, and or light or heavy skewness, Buckland (1971).

Krithikadatta (2014) projected Normal distribution, which represents a probability model that mirrors a symmetric trend for the variable range $x \in R$. It is also known as the Gaussian distribution, which exhibits a bell curve when plotted, and the symmetry is centred about its mean with the width of the curve defined as the standard deviation. Its probability density function (pdf) is given as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

where $\mu \; and \; \sigma$ are the mean and standard deviation. However, the classic distribution has a standard form realized at the parameter values $\mu = 0 \; and \; \sigma = 1$ amidst various parametric possibilities. Other symmetric distributions that can be classified as bell-shape include Cauchy, student $t$ and logistic distribution, to mention a few. These continuous distributions share a similar variable range $x \in R$. It is worthy of note that the symmetricity of any probability models does not depend on whether the variable range is unbounded or not. Chattamvelli and Shanmugam (2021) studied another symmetric distribution with bounded support, the Arcsine distribution, generally defined as

$$f(x) = \frac{1}{\pi\sqrt{(x-a)(b-x)}} \tag{2}$$

where $x \in [a, b]$, and the standard form of the pdf is obtained at $a = 0 \; and \; b = 1$. It is known that arcsine distribution is a special case of the Pearson type 1 distribution, where the trend mirrors a perfect bathtub U-shape. Another distribution in this similitude is the Beta distribution, parametrically valued at $X \sim Beta\left(\frac{1}{2}, \frac{1}{2}\right)$. Uniform distribution is another symmetric distribution characterized by a straight line shape, which is also different in its own order.

While we appreciate the development of these symmetric distributions, it is pertinent to admit that they do not fit some lifetime symmetric data. Hence, we aim to combine two varying symmetric distributions of different shapes to realize a rare symmetric trend in literature, the *perfect hemisphere shape*. This "half-sun" model can suffice in applications where especially the bell-curve distributions may not be suitable. Other sections of this paper are arranged thus: materials and method, normal-arcsine distribution and properties, and simulation respectively.


Article DOI: 10.52589/AJMSS-K6FRZIGC
DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 4, 2024 (pp. 153-161)
www.abjournals.org

## MATERIALS AND METHOD

In this section, we examine some probability distribution development strategies. Amidst the various methods captured in Lai (2013), the mathematical combination of functions will be explored, using the integration and the concept of normalizing constant. This universal method of probability model development allows us access to exhaustively utilize any possible means to achieve convergence of the combined models (functions). It is worthy of note that no limit is placed on the mathematical operations to be adopted, with respect to the model combination. The selection of the models could be homogeneous, in the sense that a model can be used to form these combinations; and on the other hand non-homogeneous, such that different probability models can be adopted. More so, the choice of the models for these constructions cuts across either of the pdf(s) or the cumulative distribution function cdf(s).

Homogeneous Combinations include:

$$\blacksquare(x) = p(x)\,p(x), \quad \frac{p(x)}{p(x)}, \quad p(x) \pm p(x)$$

$$\blacksquare(x) = p(x)\,P(x), \quad \frac{p(x)}{P(x)}, \quad p(x) \pm P(x)$$

$$\blacksquare(x) = P(x)\,P(x), \quad \frac{P(x)}{P(x)}, \quad P(x) \pm P(x)$$

where Non-homogeneous Combinations include:

$$\blacksquare(x) = p(x)\,t(x), \quad \frac{p(x)}{t(x)}, \quad p(x) \pm t(x) \tag{3}$$

$$\blacksquare(x) = p(x)\,T(x), \quad \frac{p(x)}{T(x)}, \quad p(x) \pm T(x)$$

$$\blacksquare(x) = P(x)\,T(x), \quad \frac{P(x)}{T(x)}, \quad P(x) \pm T(x)$$

where $p(x), t(x),\ P(x)\ and\ T(x)$ are the pdfs and cdfs of probability models. These combinations, although probability models, can be treated like any other arbitrary mathematical functions, provided they are integrable. We would run into complex cases where some combinations may not converge upon integration application; however, with the idea of upper and or lower censoring, such integration difficulties are possibly handled. By censoring we imply the adjustment of either the lower or upper bound, or both of them. These are referred to as left, right and double censoring $(C_r,\ C_l\ and\ C_d)$ respectively; where $\int_{-\infty}^{\infty} \blacksquare(x)dx$

- $\int_{-\infty}^{x_{max}} \blacksquare(x)dx \rightarrow \ C_r$
- $\int_{x_{min}}^{\infty} \blacksquare(x)dx \ \rightarrow \ C_l$


Article DOI: 10.52589/AJMSS-K6FRZIGC
DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC

- $\int_{x_{min}}^{x_{max}} \blacksquare(x)dx \rightarrow \quad C_d$

**Normal-Arcsine Distribution and Properties**

Here, we develop a new model by combining two non-homogeneous probability models; recalling equations (1 and 2), where $p(x) \sim N(0, \sigma^2)$ and $t(x) \sim Arcsine[\pi]$. Of course, these two symmetric distributions exhibit different trends; however, their combination stirs curiosity about the possible trend(s) they may realize. If we take inference from the combinatorial division operation in equation (3), we realize that,

$$\blacksquare(x) = \frac{p(x)}{t(x)} = \frac{\{\pi\sqrt{(1-x)x}\}\ e^{-\frac{x^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \tag{4}$$

To form a new distribution from equation (4), we apply integration and subsequently, normalization. Since there are two possible support ranges to integrate with, we test-run both to see their convergence possibility, which will determine whether we will adopt integration by censoring. Using the support range from the normal distribution $-\infty < x < \infty$, we obtain a pdf

$$NASD_1 = \frac{4i\sqrt{2}e^{-\frac{x^2}{2b^2}}\pi^{\frac{3}{2}}\sqrt{(1-x)x}}{MeijerG\left[\left\{\left\{-\frac{1}{4},\frac{1}{4}\right\},\{.\}\right\},\left\{\left\{-1,-\frac{1}{2},0\right\},\{.\}\right\},\frac{1}{2b^2}\right]}$$

Similarly adopting the support range from Arcsine distribution $0 < x < 1$, we have

$$NASD_2 = \frac{8\ e^{-\frac{x^2}{2\sigma^2}}\sqrt{(1-x)x}}{\pi\ \aleph_{PFQ}} \tag{5}$$

where $\aleph_{PFQ} = hypergeometricPFQ\left[\left\{\frac{3}{4},\ \frac{5}{4}\right\},\ \left\{\frac{3}{2}, 2\right\},\ -\frac{1}{2\sigma^2}\right]$ is a generalized hyper-geometric function and $NASD$ implies Normal-Arcsine Distribution. While we may call both pdfs Normal-Arcsine distributions, we give preference to the latter for further exploration of pdf properties. This is primarily, to ensure relative mathematical efficiency, since we have a complex number "$i$" to deal with in the former.



Article DOI: 10.52589/AJMSS-K6FRZIGC

DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 7, Issue 4, 2024 (pp. 153-161) 

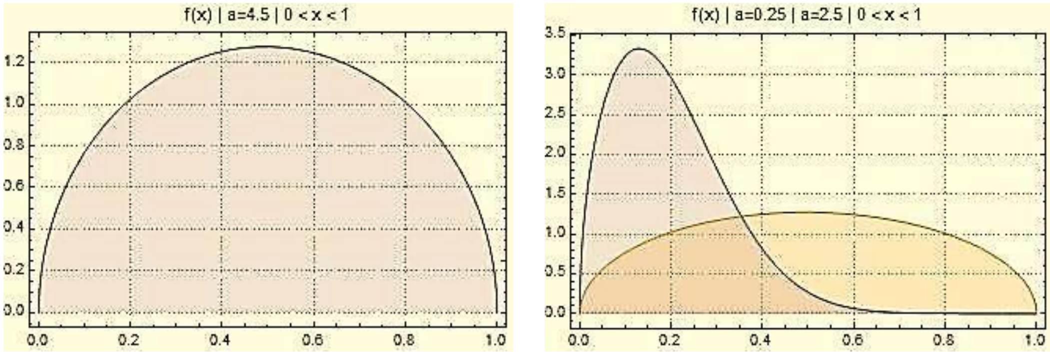


**Figure 1:** Graph plots of the $NASD_2$

From Figure 1 we can deduce that $NASD_2$ exhibits perfect hemisphere trend and can be right-skewed as well.

**Remark**

The shapes realized from this combination as clearly shown in Figure 1 are alien to the original shapes of the two distributions. This development has increased the modelling options of the combo; especially since that perfect hemisphere $NASD_2$ can be symmetric, and seems to be rare in literature.

### 3.1 Moment

If X is a random variable from a continuous distribution and $g(x)$ is the density function, then the $r^{th}$ moment about the origin of X is defined by

$$E \;\; E(X^r) = \int_0^\infty x^r g(x) dx = \mu_r'.$$

Now, for the moments $NASD_2$, we recall equation (5) and obtain

$$E(X^r) = \frac{8}{\pi \; \aleph_{PFQ}} \; \int_0^1 x^r \, e^{-\frac{x^2}{2\sigma^2}} \sqrt{(1-x)x},$$

$$= \frac{2^{-r} Gamma\left[\frac{3}{2}+r\right] HypergeometricPFQRegularized\left[\left\{\frac{1}{4}(3+2r), \frac{1}{4}(5+2r)\right\}, \left\{\frac{3+r}{2}, \frac{4+r}{2}\right\}, -\frac{1}{2b^2}\right]}{HypergeometricPFQ\left[\left\{\frac{3}{4}, \frac{5}{4}\right\}, \left\{\frac{3}{2}, 2\right\}, -\frac{1}{2b^2}\right]}$$

Of course, the first four $r^{th}$ moments of the $NASD_2$ are further obtained at $r = 1,2,3 \; and \; 4$

$$\mu_1' = \frac{HypergeometricPFQ\left[\left\{\frac{5}{4}, \frac{7}{4}\right\}, \left\{2, \frac{5}{2}\right\}, -\frac{1}{2b^2}\right]}{2HypergeometricPFQ\left[\left\{\frac{3}{4}, \frac{5}{4}\right\}, \left\{\frac{3}{2}, 2\right\}, -\frac{1}{2b^2}\right]} = \mu, \qquad \mu_2' = \frac{5HypergeometricPFQ\left[\left\{\frac{7}{4}, \frac{9}{4}\right\}, \left\{\frac{5}{2}, 3\right\}, -\frac{1}{2b^2}\right]}{16HypergeometricPFQ\left[\left\{\frac{3}{4}, \frac{5}{4}\right\}, \left\{\frac{3}{2}, 2\right\}, -\frac{1}{2b^2}\right]},$$

$$\mu_3' = \frac{7HypergeometricPFQ\left[\left\{\frac{9}{4}, \frac{11}{4}\right\}, \left\{3, \frac{7}{2}\right\}, -\frac{1}{2b^2}\right]}{32HypergeometricPFQ\left[\left\{\frac{3}{4}, \frac{5}{4}\right\}, \left\{\frac{3}{2}, 2\right\}, -\frac{1}{2b^2}\right]} \qquad \mu_4' = \frac{21HypergeometricPFQ\left[\left\{\frac{11}{4}, \frac{13}{4}\right\}, \left\{\frac{7}{2}, 4\right\}, -\frac{1}{2b^2}\right]}{128HypergeometricPFQ\left[\left\{\frac{3}{4}, \frac{5}{4}\right\}, \left\{\frac{3}{2}, 2\right\}, -\frac{1}{2b^2}\right]}.$$


Article DOI: 10.52589/AJMSS-K6FRZIGC
DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC

Therefore, the variance, skewness and kurtosis of the $NASD_2$ can be obtained.

$$Variance(\mu_2) = \mu'_2 - \mu^2 = \sigma^2 \qquad Skewness\ (S_k) = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}} = \frac{\mu'_3 - 3\mu'_2\mu + 2\mu^3}{(\mu'_2 - \mu^2)^{\frac{3}{2}}},$$

$$Kurtosis\ (K_s) = \frac{\mu_4}{(\mu_2)^2} = \frac{\mu'_4 - 4\mu'_3\mu + 6\mu'_2\mu^2 - 3\mu^4}{(\mu'_2 - \mu^2)^2}$$

**Estimation**

In this section, we obtain the parameter values of the probability function that maximizes the likelihood function. Kinaci et al. (2014) and Fang et al. (2015) studied maximum likelihood estimation (MLE) under different data conditions including uncensored data and censoring. Censoring is a concept that describes the timing, in which data are recorded during a procedural observation; in the sense that truncated events are not treated like exhaustive investigations. In general, the likelihood functions for both data conditions are respectively given as:

$$(x, \theta) = \prod_{i=1}^{n}[g(x_i)]$$

$$L(x, \theta) = \frac{n!}{(N-n)!}\{\prod_{i=1}^{n} g(x_i)\}\{1 - G(x_T)\}^{N-n} \qquad (6)$$

where $g(x_i)\ and\ G(x_i)$ are the PDF and CDF of a distribution with independent random observations $x_i, i = 1, 2, \dots, n$; N is the number of specimens being investigated or the number of trials. Now, if the fixed time or cycle or count to an event (say failure time) is $x_0$, then for Type 1 censoring according to equation (6), the time of termination $x_T = x_0$; and $x_T = x_n$ for Type 2 case. However, for this study, we have our emphasis only on the complete or uncensored investigations. Let $x_i, i = 1,2, \dots, n$ be a vector of observations from $NASD_2$, and then the log-likelihood for the complete data is defined by:

$$(x, \sigma) = \log L(x, \sigma) = \sum_{i=1}^{n} \log\{f(x, \sigma)\}$$

$$l(x, \sigma) = \sum_{i=1}^{n} \log\left\{\frac{8\, e^{-\frac{y^2}{2\sigma^2}}\sqrt{(1-x)x}}{\pi\ \aleph_{\text{PFQ}}}\right\}$$

$$= \log\left[\frac{8}{\pi\ \aleph_{\text{PFQ}}}\right]^n + \sum_{i=1}^{n} \log\left[\left((1-x)x\right)^{\frac{1}{2}}\right] - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2$$

$$= n\log 8 - n log(\pi) - n\log(\aleph_{\text{PFQ})}) + \sum_{i=1}^{n} \log\left[\left((1-x)x\right)^{\frac{1}{2}}\right] - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 \qquad (7)$$

The score function for equation (7) is defined by


Article DOI: 10.52589/AJMSS-K6FRZIGC
DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC

$$\frac{\partial l}{\partial \sigma} = \frac{n\{\beth_{PFQ}\}}{\aleph_{\mathrm{PFQ})}} + \frac{1}{\sigma^3}\sum_{i=1}^{n} x_i^2$$

$$\frac{n\{\beth_{PFQ}\}}{\aleph_{\mathrm{PFQ})}} + \frac{1}{\sigma^3}\sum_{i=1}^{n} x_i^2 = 0$$

Where $\beth_{PFQ} = \frac{5}{16e^3} HypergeometricPFQ\left[\left\{\frac{7}{4},\frac{9}{4}\right\},\left\{\frac{5}{2},3\right\},-\frac{1}{2e^2}\right]$

A numerical analysis like the Newton-Raphson iterative method, which is a root-finding algorithm, can be used to obtain the MLEs of $\hat{\sigma}$. This scheme is given by

$$\hat{\sigma} = \sigma - H^{-1}(\sigma)\, S(\sigma)$$

where $S(\sigma)$ is the score function and $H^{-1}(\sigma)$ is the second derivative of the log-likelihood function termed the Hessian matrix, Bayal et al. 2022. Finally, it is expected that the different methods show efficiency concerning the selected size of data samples under consideration.

**Simulation Study**

Furthermore, the asymptotic character of the maximum likelihood estimates of the parameters of Normal-Arcsine distributions is investigated, through a Monte Carlo simulation study. For different sample sizes $n = 20,\ 50,\ 75,\ 100\ \&\ 250$, a $10000$ times trials are carried out; and the steps are given by the algorithm:

i) Choose a value M (which represents the number of Monte Carlo trials).

ii) Select the values $\sigma_0$ within the domain of their parameter supports.

iii) Simulate a sample of size *n* from the derived distributions.

iv) Compute the maximum likelihood estimates $\hat{\sigma}_k\ of\ \sigma_k$

v) Redo steps 3-4 for $N$ number of times

vi) The computations of the following measures are obtained:

$$Average\ Bias = \left[\ \frac{1}{M}\sum_{i=1}^{M}(\hat{\sigma}_i - \sigma)\right]\ and$$

$$MSE = \left[\ \frac{1}{M}\sum_{i=1}^{M}(\hat{\sigma}_i - \sigma)^2\ \right]$$

See Table 2-7 for the *Average Bias and MSE of the NASD*$_2$


Article DOI: 10.52589/AJMSS-K6FRZIGC
DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC

**Table 1:** *Average Bias and MSE of the (NASD) Estimator* $\hat{\sigma}$

| Parameter | $N$ | Average Bias ($\sigma$) | MSE ($\sigma$) |
|---|---|---|---|
| $\sigma = 0.1$ | 20 | -1.6e-16 | 0.00899 |
| | 50 | -0.3e-16 | 0.00576 |
| | 75 | -7.8e-16 | 0.00387 |
| | 100 | -8.0e-16 | 0.00312 |
| | 250 | -10.6e-16 | 0.00125 |
| $\sigma = 0.5$ | 20 | -2.6e-15 | 0.05451 |
| | 50 | -2.8e-15 | 0.04729 |
| | 75 | -3.2e-15 | 0.04154 |
| | 100 | -6.1e-15 | 0.03215 |
| | 250 | -8.2e-15 | 0.03012 |
| $\sigma = 1.5$ | 20 | -3.1e-15 | 0.06999 |
| | 50 | -3.7e-15 | 0.06865 |
| | 75 | -8.3e-15 | 0.05824 |
| | 100 | -7.9e-15 | 0.04234 |
| | 250 | -9.7e-15 | 0.01425 |
| $\sigma = 2.5$ | 20 | 6.2e-15 | 0.06917 |
| | 50 | -2.2e-16 | 0.05934 |
| | 75 | -3.9e-15 | 0.05239 |
| | 100 | -4.5e-15 | 0.04236 |
| | 250 | -1.2e-15 | 0.03046 |

The estimates for the average bias and mean square error are presented in Table 1, and at different selected values of the parameter. Apparently, from the Table, we deduce that the estimates of the average bias and mean square error decrease as the sample size $n$ increases. This simply indicates that the estimators of the derived distributions are consistent and asymptotically stable.

## CONCLUSION

The paper presents some methodical approaches to the development of probability models; and a rare trend in the theory of probability, which features a perfect hemisphere shape which can also be termed the Normal ArcSine Distribution. This development stems from an integrative combination of two different symmetric trends, realized over the mathematical principle of normalization. The combination is a composite of the Normal distribution with a bell curve shape and the arcsine distribution with a U-shape; where the right skewed trend is another shape obtainable from the proposed model. Properties like moments and estimation were studied alongside simulation, underscoring the behaviour of the parameters. It is recommended that further studies be carried out on the application of this development.


Article DOI: 10.52589/AJMSS-K6FRZIGC
DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC

## REFERENCES

[1] Buckland, w. (1971). Continuous univariate distributions - Johnson, NL, and Kotz, S. 75-75

[2] Krithikadatta, J. (2014). Normal distribution. Journal of Conservative Dentistry and Endodontics, 17(1), 96-97.

[3] Chattamvelli, R., & Shanmugam, R. (2021). Arcsine Distribution. In Continuous Distributions in Engineering and the Applied Sciences-Part I (pp. 57-68). Cham: Springer International Publishing.

[4] Lai CD (2013) Constructions and applications of lifetime distributions. Applied Stochastic Models in Business and Industry, 29(2), 127-140.

[5] Bayat M, Koushki MM, Ghadimi AA, Tostado-Véliz M, Jurado F (2022) Comprehensive enhanced Newton Raphson approach for power flow analysis in droop-controlled islanded AC microgrids. Int. J. Electr. Power Energy Syst., 143, 108493.


Article DOI: 10.52589/AJMSS-K6FRZIGC
DOI URL: https://doi.org/10.52589/AJMSS-K6FRZIGC