Applied Mathematical Sciences, Vol. 14, 2020, no. 8, 349 - 359
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/ams.2020.914206

A Test of Normality from Allegorizing the Bell Curve or the Gaussian Probability Distribution as Memoryless and Depthless Like a Black Hole

Amitava Biswas

SHS School, University of Southern Mississippi
Hattiesburg, MS 39406, USA

Abhishek Bisaria

Spectral D&T Inc, Bridgewater, NJ 08807, USA

Abstract

A non-traditional formulation of the Gaussian probability distribution in two dimensions did not require the constant pi and mathematically resembled the exponential formula of radioactive decay. Shifting the position of reference did not affect the effective span of its remaining existence on the farther side of that reference position. Isolation prevailed like a black hole behind a circular boundary when observing outward from there. Consequently, any observer at any possible location, could as well view itself to be just in the center of that distribution, if the distribution was perfectly Gaussian. Utility of this viewpoint could include development of alternative tests of normality. Such tests might be useful for exploring random distribution of errors in unknown data sets in many areas of science from cosmology to virology.

Keywords: chance, expectation, invariance, simulation, statistics, visualization.

1 Introduction

This article stems from a few basic concepts such as memoryless probability [11]. Probabilistic as well as general behavior of many physical systems naturally depends on their past history of location, velocity, or other states. Therefore, such systems are said to have some memory. Whereas, some systems appear to be independent of their earlier conditions, and therefore, memoryless on that basis. For example, as shown in Fig 1, the cooling behavior of a hot object exposed to a cold environment may be considered memoryless under following situation. The cold environmental temperature is 0 C. The hot object is initially at 100 C. The characteristic time constant of the system is one hour. With these parameters, the initial cooling rate is 100 C per hour. So, the hot object appears to be reaching the environmental temperature precisely in one hour. After one hour has passed, the object is at 37 C above the environmental temperature and cooling at 37 C per hour. So, the object appears to be reaching the environmental temperature precisely in another one hour. After another one hour, the object is at 14 C above the environmental temperature and cooling at 14 C per hour. So, the object appears to be reaching the environmental temperature precisely in another one hour. This paradoxical memoryless tendency to reach the environmental temperature within a fixed interval every time, virtually forgets memory of past predictions or any influence of the passage of time. Such a memoryless characteristic has been found to be present and useful in many physical situations, such as the expected life of a light bulb [11] and radio-active decaying of some isotopes of carbon, plutonium, radium, thorium, and uranium [4]. This basic concept of memoryless behavior need not be restricted to time. It may be extended to spatial displacements or other abstract measures. Their common hallmark is a mathematical trend that is exponential of the first degree, presented in Appendix-A. In the following sections, we describe a mathematical perspective where a memoryless characterization of the Gaussian probability distribution could be observed. We also show a simple and sensitive mechanism to differentiate some other distributions from the Gaussian distribution, utilizing this perspective.

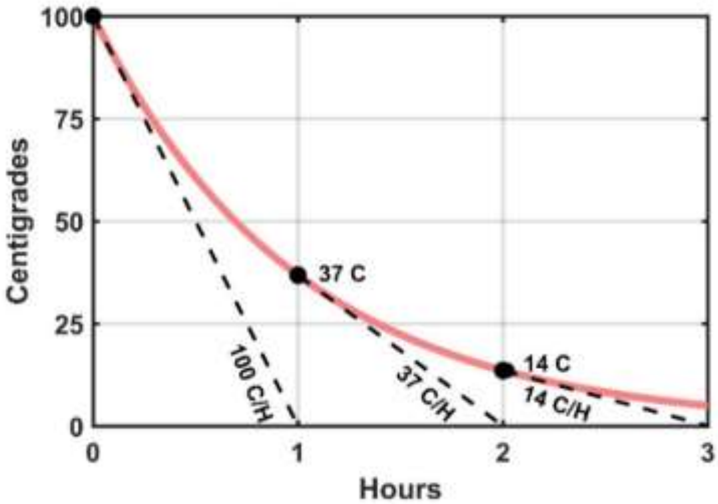

Fig 1. A simple example of memoryless cooling characteristic.

2 Probability Distributions

Continuous random variables of single dimension in many areas of science, such as errors in measurement of the size of a galaxy or a virus, often follow a characteristic probability distribution. They may be mathematically modeled by a particular probability density function [4, 6, 11, 12]. One of the most commonly observed probability distributions is the normal distribution or the Gaussian distribution. According to the central limit theorem, this is the inevitable distribution of any variable that is a sum of sufficiently large number of random variables [11]. See Appendix-B, for the probability density function of the Gaussian distribution for single dimensional measurements. Albeit, sum of a very large number of random variables may be very close but may not be exactly this distribution. In Fig 2, the thick red bell curve represents the standard Gaussian probability density function. The closely overlapping density function in black dots is obtained from summing 12 random numbers with uniform probability over unit interval. With more summands that distribution becomes even closer to the Gaussian distribution, following the central limit theorem. Such a distribution may be considered in the category of Irwin-Hall distributions [5, 7]. Interestingly, there would be some difference, however small, if the number of summands for an Irwin-Hall distribution is a finite number. This fact will be utilized in the results section of this article. In this context, visually comparing Figures 1 and 2, neither Gaussian nor Irwin-Hall distributions appear to be memoryless. Moving the usual position of reference of these distributions from the mean would generally affect the expected measure of all possible samples or outcomes that are on one side of that new reference. In the following sections, we will develop a memoryless perspective for the Gaussian distribution and utilize that perspective to distinguish it from other distributions. Firstly, we will select a few distributions that are not easily distinguishable from the Gaussian distribution.

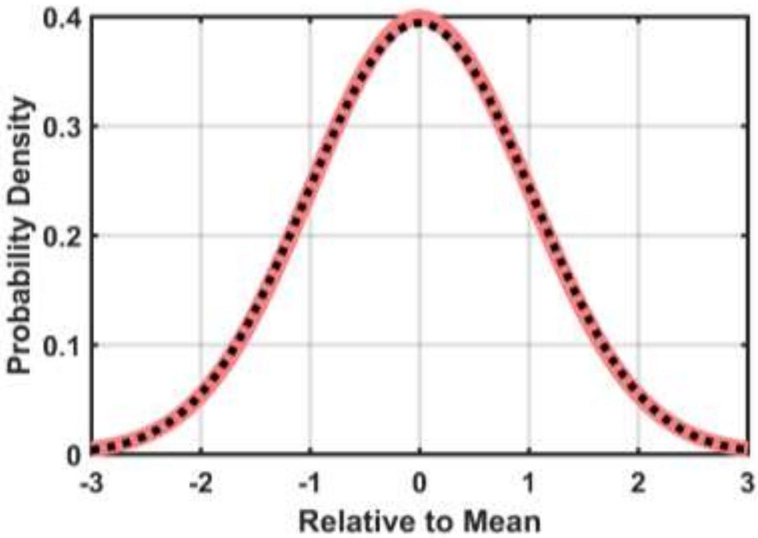

Fig 2. The Gaussian distribution (red), and Irwin-Hall (black dots).

4 Kolmogorov-Smirnov Test of Normality

The Kolmogorov-Smirnov (K-S) test is well known for checking whether a distribution is different from the Gaussian distribution [8, 13]. The Irwin-Hall distributions were subjected to this test. As shown in the bottom trace of Figure 3, the K-S test correctly rejected normality at the usual significance level ($p<0.05$) of all 100 trials involving sets of 10,000 sums of 2 random numbers with uniform probability over unit interval. The upper traces in Figure 3 are similarly obtained but used 3 or more summands, and the K-S test failed to reject normality in many trials. All of these distributions are supposed to differ, even by a small extent, from the Gaussian distribution unless the number of summands is infinitely large. Therefore, Irwin-Hall distributions would be good candidates for evaluating a new test of normality.

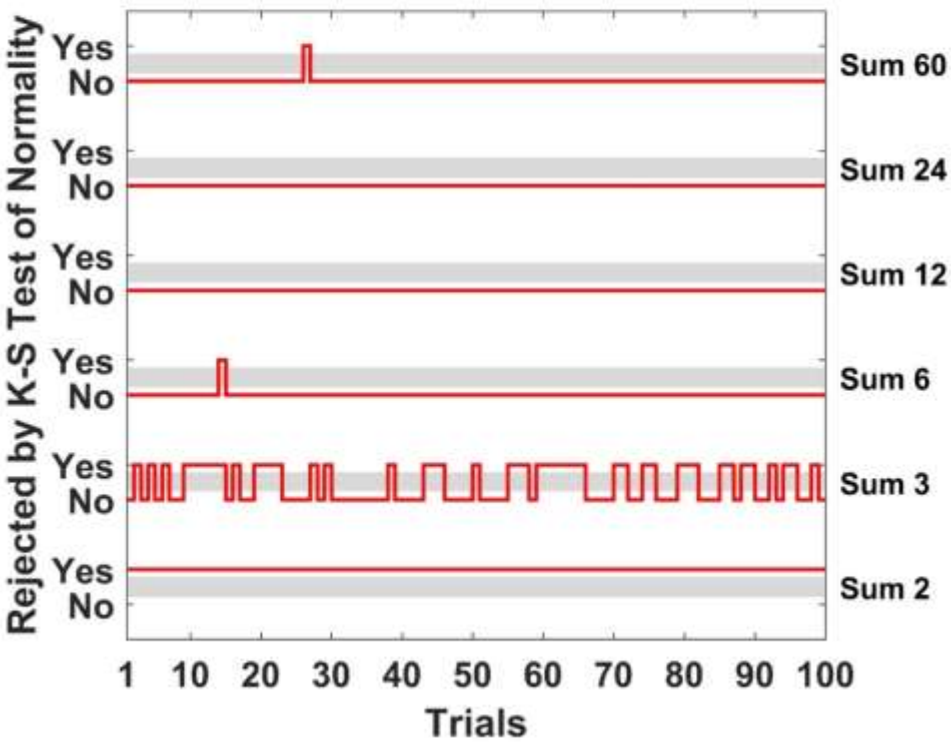

Fig 3. Results of the Kolmogorov-Smirnov (K-S) test.

5 From Univariate to Bivariate Perspective

To formulate the memory-less and pi-less expression, we will need a bivariate version of the one-dimensional distribution [3]. In Figure 4, the height of the bell shape along the z-axis from any location on the x-y plane represents the joint probability density of two independent samples x and y from a basically one-dimensional distribution. This resultant two-dimensional distribution is circular symmetric about the z-axis if the original one-dimensional distribution is Gaussian

[3, 4, 6, 11, 12]. See Appendix-C, for the probability density function of a two-dimensional distribution where the x and y coordinates of each possible outcome position is obtained independently from a single dimensional Gaussian distribution.

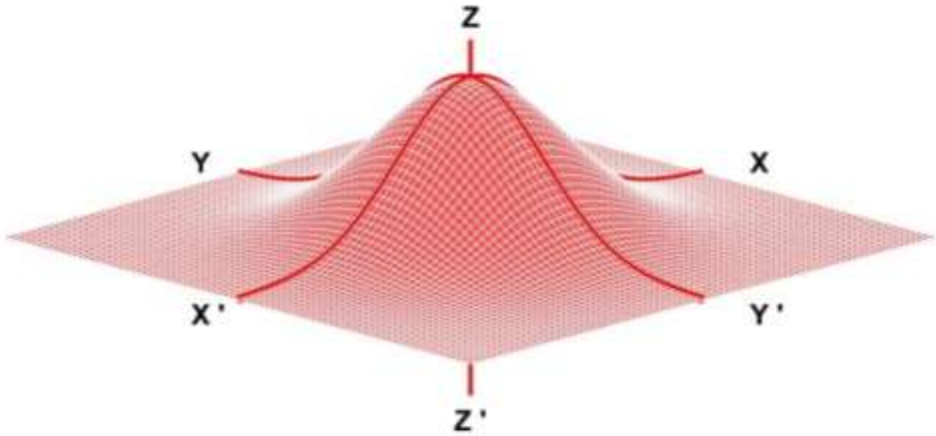

Fig 4. A bivariate Gaussian distribution

6 From Bivariate to Univariate Perspective

The x and y coordinates of each dot in Figure 5 are two independent samples randomly obtained from a one-dimensional original distribution. A circle has been drawn through each dot. The center of each of these circles coincides with the center of the entire two-dimensional distribution. The probability-weighted average area of all such circles will be used by the proposed basic concept, presented below. Thereby, we will practically consider a univariate system, using only the area of these circles for the proposed procedure [3].

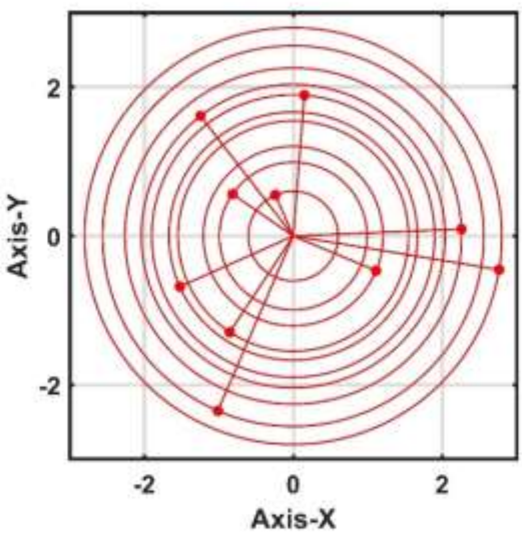

Fig 5. Every dot is associated with a circle.

7 Gaussian Distribution without Pi

If the original one-dimensional distribution is Gaussian then the corresponding two-dimensional distribution is circular symmetric about the z-axis [3]. The probability-weighted average of squares of radial distances of all positions of the probability space from the center of the two-dimensional distribution is equal to sum of the two variances along directions x and y. That average times the constant pi, will be the probability-weighted average area of circles through all positions of the entire probability space. With that average area as a standard unit of measuring area, compact one-dimensional equations for the two-dimensional Gaussian distribution can be derived [3]. Interestingly, these equations do not need the constant pi, as shown in Appendix-E. These exponential equations of first degree bear the hallmark of memoryless behavior, as described in earlier sections.

8 Gaussian Distribution with a Swallower

It appears from the last section, the probability-weighted average area of annuli corresponding to all possible outcome positions beyond any circle of arbitrary size will be just a constant and always equal to the standard unit area, defined earlier, if and only if the original distribution is Gaussian, as shown in Appendix-D. Figure 6 shows a few sample positions from a two-dimensional distribution, grouped here by their distances from the center of the entire distribution.

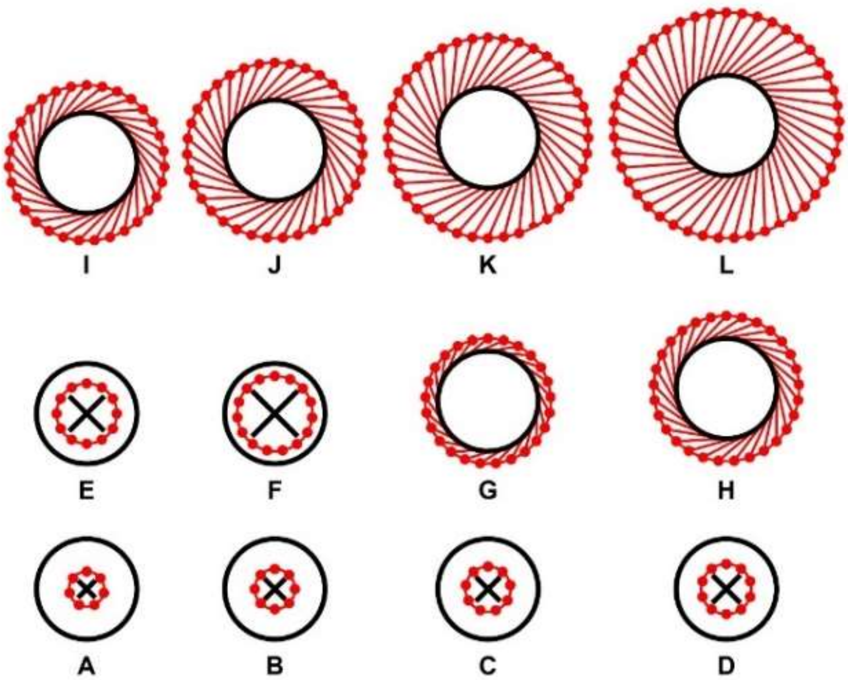

Fig 6. Contribution of selected red dots after purging by the black circle.

All sample positions of any group are on a common circle. Instead of concentric mapping, the selected groups are labeled A-L and mapped here side by side for explaining our basic concept. These groups have concentric black circles, called swallowers. These swallowers have the same radius which may be chosen arbitrarily at a time. As shown by cross marks for groups A-F, any sample position encased by the swallower may be ignored for our basic concept. The red tangential lines for groups G-L indicate the effective distance of each remaining sample position for our basic concept. The tangential lines are appropriate here, rather than radial, due to orthogonality of random independent components and the Pythagorean theorem [1, 2, 9]. This basic concept was subjected to a few simulation tests and the preliminary results are summarized in the next section.

9 Results & Conclusions

According to the equations presented in Appendix-D, an observer on the periphery of an arbitrary swallowing circle, concentric to the center of a two-dimensional perfectly Gaussian distribution, would expect an invariant probability-weighted average area of all possible annular rings, as described in previous sections. In Figure 7, the horizontal axis represents the size of the swallowing circle, enclosing from 0% to 90% of 1,000,000 sample positions from a two-dimensional distribution.

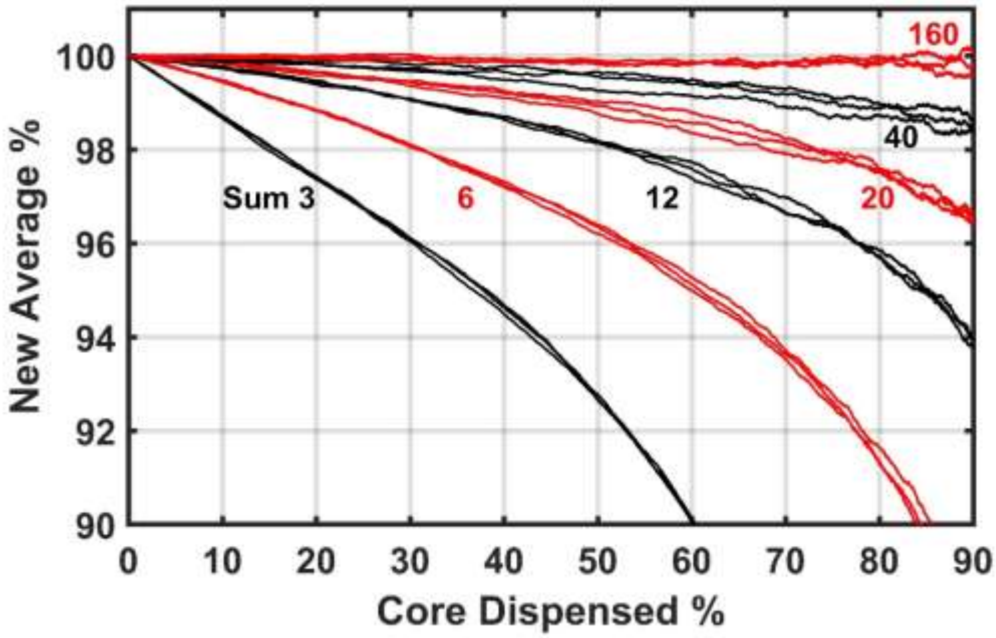

Fig 7. Sums of 160 uniform random numbers behaved almost as Gaussian.

The vertical axis represents the average area of all available annular rings from the data set. The inner circle of each ring coincides with the swallowing circle. The outer circle passes through a sample position. Each group of three adjacent traces in one color shows three trials using one particular two-dimensional distribution.

The x and y coordinates of each selected position on that two-dimensional distribution are two random outcomes x and y from an one-dimensional Irwin-Hall distribution involving a sum of a fixed number (3, 6, 12, 20, 40, or 160) of random numbers with uniform probability over unit interval. Figure 7 demonstrates that the Irwin-Hall distribution involving a large sum of random numbers approaches the Gaussian distribution, as predicted by the central limit theorem [11]. For sums of 160 random numbers, the average area of the annular rings remained close to the 100% level, even if the swallowing circle removed 90% of the samples. This behavior is consistent for distributions that are almost indistinguishable from the Gaussian distribution, according to this concept.

Acknowledgements. This work has benefited from faculty start-up funds granted by the University of Southern Mississippi to Dr. Amitava Biswas.

References

[1] A. Biswas, and A. Bisaria, An Amusing Emergence of the Pythagorean Theorem from Circular Planetary Motions, *International Journal of Technology*, **9** (2019), 43 - 44.

[2] A. Biswas, and A. Bisaria, A simple drill for the Pythagorean theorem with scissors on circles, *International Mathematical Forum*, **15** (2020) 145 - 149. https://doi.org/10.12988/imf.2020.91254

[3] A. Biswas, and A. Bisaria, A non-traditional concise formulation of the Gaussian distribution from computer simulation and analysis, *International Journal of Recent Scientific Research*, **11** (2020), 37710 - 37711.

[4] G.W. Castellan, *Physical Chemistry*, Addison-Wesley, Boston, 1983.

[5] P. Hall, The Distribution of Means for Samples of Size N Drawn from a Population in which the Variate Takes Values Between 0 and 1, All Such Values Being Equally Probable. *Biometrika*, **19** (1927), 240 - 245. https://doi.org/10.2307/2331961

[6] R.W. Hamming, *Methods of Mathematics Applied to Calculus, Probability, and Statistics*, Dover Publications, New York, 2004.

[7] J.O. Irwin, On the Frequency Distribution of the Means of Samples from a Population Having any Law of Frequency with Finite Moments, with Special Reference to Pearson's Type II, *Biometrika*, **19** (1927), 225 - 239. https://doi.org/10.2307/2331960

[8] F.J. Massey, The Kolmogorov-Smirnov Test for Goodness of Fit, *Journal of the American Statistical Association*, **46** (1951), 68 - 78.

[9] M. Mnatsakanian, Annular Rings of Equal Area, *Math. Horizons*, **5** (1997), 5 - 8. https://doi.org/10.1080/10724117.1997.11975029

[10] B. Ramachandran, On the Strong Memorylessness Property of the Exponential and Geometric Probability Laws, *Sankhyā*, **41** (1979), 244 - 251.

[11] S. Ross, *A First Course in Probability*, Pearson, New York, 2019.

[12] M.K. Simon, *Probability distributions involving Gaussian random variables: A handbook for engineers and scientists*, Springer, New York, 2007. https://doi.org/10.1007/978-0-387-47694-0

[13] S.F. Sufahani, and A. Ahmad, A Comparison between Normal and Non-Normal Data in Bootstrap, *Applied Math. Sciences*, **6** (2012), 4547 - 4560.

Appendix-A

See more details from references [4, 6, 11, 12].

τ = time constant of the system.

T = temperature of the hot body at any time t.

T_0 = initial temperature of the hot body at time t=0.

T_e = temperature of the environment.

The temperature of the hot body follows this *memoryless* exponential equation:

$$T = T_e + (T_0 - Te)e^{-\frac{t}{\tau}}$$

Whatever be that time t, the instantaneous rate of drop of temperature from T is

$$\frac{dT}{dt} = \frac{T - T_e}{\tau}$$

Whatever be that time t, the time to reach temperature T_e from T is always τ.

Appendix-B

See more details from references [3, 4, 6, 11, 12].

μ = mean of the distribution.

σ = standard deviation of the distribution.

x = value of the random variable.

$f(x)$ = Gaussian probability density per unit interval near the value x of the variable.

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

Appendix-C

See more details from references [3, 4, 6, 11, 12].

μ = mean of the distribution = 0 (assumed to simplify x-y coordinates).

σ = standard deviation of the *original* one-dimensional distribution *f(x)*.

x = value of the random variable along *x*-axis from the center (independent of *y*).

y = value of the random variable along *y*-axis from the center (independent of *x*).

g(x,y) = two-dimensional Gaussian probability density per unit area near (*x,y*).

$$g(x,y) = f(x)\, f(y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Appendix-D

See more details from reference [3].

Here, expected value = probability-weighted average value.

Since *g(x,y)* is circular symmetric around the z-axis, we have following polar form.

r = radial distance of any position *(x, y)* from the center = $\sqrt{(x^2 + y^2)}$

h(r) = two-dimensional Gaussian probability density per unit area at radius *r*.

$$h(r) = g(r,0) = f(r)\, f(0) = \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}}$$

$$\int_0^\infty \int_0^{2\pi} h(r) r \, d\theta \, dr = 1 \quad \text{[verify 100\% probability]}$$

A = expected area of any circle through position *(x,y)* with probability *g(x,y)*.

Radius of the circle of area *A* is the standard deviation of the distribution of (*x*+*y*).

$$A = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \pi(x^2 + y^2) g(x,y)\, dx\, dy = \int_0^\infty \int_0^{2\pi} \pi r^2 h(r) r \, d\theta \, dr = 2\pi\sigma^2$$

A_S = probability-weighted average area of annulus of all possible outer radius *r*.

s = radius of the swallowing circle, inner radius of the above annulus (*r*>*s*).

$$A_s = \frac{\int_s^\infty \int_0^{2\pi} \pi(r^2 - s^2) h(r) r \, d\theta \, dr}{\int_s^\infty \int_0^{2\pi} h(r) r \, d\theta \, dr} = 2\pi\sigma^2$$

Therefore, for Gaussian distributions, A_S = *A* (independent of *s*).

Appendix-E

See details on these π-less exponential equations in reference [3].

For circular symmetric two dimensional Gaussian probability distribution *g(x,y)*,

a = area of a circle through position *(x,y)* divided by the unit area *A* defined above.

PCA = periphery of the circle through position *(x,y)* for computing the ratio a.

$\hat{H}(a) = e^{-a}$ [Case 1 - probability density per unit area *A* near *PCA*.]

$\hat{H}(a) = e^{-a}$ [Case 2 - cumulative probability beyond *PCA*.]

Received: May 17, 2020; Published: June 3, 2020