

Archimedes, Gauss, and Stein

Jim Pitman and Nathan Ross

Characterizing the Gaussian Distribution

One of Archimedes' proudest accomplishments was a proof that the surface area of a sphere is equal to the surface area of the tube of the smallest cylinder containing it; see Figure 1. Legend has it that he was so pleased with this result that he arranged to have an image similar to Figure 1 inscribed on his tomb.

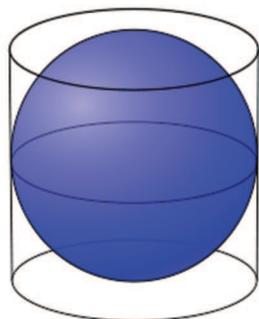


Figure 1. An illustration of the inscription on Archimedes' tomb.

More precisely, in the work *On the Sphere and Cylinder, Book I* as translated on page 1 of [2], Archimedes states that for every plane perpendicular to the axis of the tube of the cylinder, the surface areas lying above the plane on the sphere

Jim Pitman is professor of statistics at the University of California, Berkeley. His email address is pitman@stat.berkeley.edu.

Nathan Ross is a VIGRE postdoc in the statistics department at the University of California, Berkeley. His email address is ross@stat.berkeley.edu.

Research of the first author supported in part by NSF grant DMS-0806118.

DOI: <http://dx.doi.org/10.1090/noti905>

and on the tube are equal. See Figure 2 for an illustration and see also the discussion around Corollary 7 of [1].

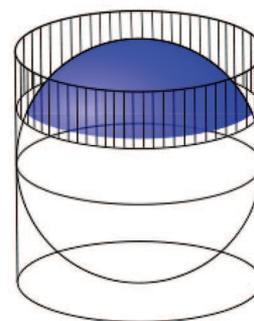


Figure 2. The surface area of the shaded "cap" of the sphere above a plane is equal to the striped surface area on the tube of the cylinder above the same plane.

In probabilistic terms, if a point is picked uniformly at random according to the surface area on the unit sphere in three dimensions, then its projection onto any given axis having origin at the center of the sphere is uniformly distributed on the interval $(-1, 1)$, independent of the angular component in the plane perpendicular to that axis. Formally, we have the following result.

Proposition 1. *If V is uniformly distributed on the interval $(-1, 1)$ and if Θ is uniformly distributed on the interval $(0, 2\pi)$ and independent of V , then*

$$\left(V, \sqrt{1 - V^2} \cos(\Theta), \sqrt{1 - V^2} \sin(\Theta) \right)$$

is uniformly distributed on the surface of the two-dimensional sphere of radius one.

In this article, we take Proposition 1 as a starting point for a discussion of characterizations of the centered Gaussian distribution which arise in Stein's method of distributional approximation. This discussion culminates in Theorem 7. We then generalize some of these results in the section on the beta-gamma algebra to obtain the characterization of the gamma distribution found in Proposition 9 and also mention an analog of Theorem 7 for the exponential distribution. We conclude in the last section with a discussion of some related literature.

To move from Archimedes' result above to characterizing the Gaussian distribution, we state the following result which was first realized by the astronomer Herschel and was made well known by the physicist Maxwell in his study of the velocities of a large number of gas particles in a container; see the introduction of [6].

Proposition 2. *Let $\mathbf{X} = (X_1, X_2, X_3)$ be a vector of independent and identically distributed (i.i.d.) random variables. Then X_1 has a mean zero Gaussian distribution if and only if, for all rotations $R : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $R\mathbf{X}$ has the same distribution as \mathbf{X} .*

Propositions 1 and 2 are related by the following observations. It is clear that, if \mathbf{X} is an $\mathbb{R}^3/\{0\}$ -valued random vector such that $R\mathbf{X}$ has the same distribution as \mathbf{X} for all rotations R , then $\mathbf{X}/\|\mathbf{X}\|$ is a rotation-invariant distribution on the surface of the two-dimensional unit sphere and is independent of $\|\mathbf{X}\| := \sqrt{X_1^2 + X_2^2 + X_3^2}$. Since the unique rotation-invariant distribution on the surface of a sphere of any dimension is the uniform distribution (Theorem 4.1.2 of [6]), the propositions of Archimedes and Herschel-Maxwell suggest the following characterization of mean-zero Gaussian distributions; we provide a proof and discussion of generalizations in the last section.

Proposition 3. *Let $\mathbf{X} = (X_1, X_2, X_3)$ be a vector of i.i.d. random variables. Then X_1 has a mean zero Gaussian distribution if and only if, for V uniform on $(-1, 1)$ and independent of \mathbf{X} ,*

$$(1) \quad X_1 \stackrel{d}{=} V\sqrt{X_1^2 + X_2^2 + X_3^2}.$$

Here and in what follows, $\stackrel{d}{=}$ denotes equality in distribution of two random variables. The distribution of $\sqrt{X_1^2 + X_2^2 + X_3^2}$, where X_1, X_2, X_3 are independent standard normal variables, is referred to as the Maxwell or Maxwell-Boltzmann distribution; see page 453 of [12].

Proposition 3 characterizes centered Gaussian distributions as the one-parameter scale family of fixed points of the distributional transformation which takes the distribution of a random variable X to the distribution of $V\sqrt{X_1^2 + X_2^2 + X_3^2}$, where

X_1, X_2, X_3 are i.i.d. copies of X and V is uniform on $(-1, 1)$ independent of (X_1, X_2, X_3) . Such characterizations of distributions as the unique fixed point of a transformation also arise in Stein's method for distributional approximation.

We connect these concepts through the next lemma, which is the fundamental identity underlying Stein's method for Gaussian approximation. This lemma characterizes the Gaussian distribution of a random variable W by an identity relating expectations of functions of W . Stein's method is further developed by results which quantify the idea that if Stein's identity is approximately satisfied by W , then W is approximately Gaussian according to standard metrics on the space of distributions on the line. This framework and analogs for other distributions (e.g., Poisson and exponential) have proved to be very effective in establishing distributional limit theorems with rates of convergence in settings where other techniques (e.g., Fourier transforms and martingale methods) are difficult to apply. See [20] for a basic introduction and further references to the Stein's method literature.

Lemma 4 (Stein's Lemma [21]). *A random variable W has the mean zero, variance one Gaussian distribution if and only if, for all absolutely continuous functions f with bounded derivative,*

$$\mathbb{E}f'(W) = \mathbb{E}Wf(W).$$

We can relate the characterizations provided by Proposition 3 and Lemma 4, but first we need the following definition.

Definition 5. Let X be a random variable with distribution function F and such that $\mu_\alpha := \mathbb{E}|X|^\alpha < \infty$. We define $F^{(\alpha)}$, the α -power bias distribution of F , by the relation

$$dF^{(\alpha)}(x) = \frac{|x|^\alpha dF(x)}{\mu_\alpha},$$

and we write $X^{(\alpha)}$ for a random variable having this distribution. In other words, $X^{(\alpha)}$ has the α -power bias distribution of X if and only if for every measurable function f such that $\mathbb{E}|X|^\alpha|f(X)| < \infty$,

$$(2) \quad \mathbb{E}f(X^{(\alpha)}) = \frac{\mathbb{E}|X|^\alpha f(X)}{\mathbb{E}|X|^\alpha}.$$

Taking $\alpha = 1$ and $X \geq 0$, $X^{(1)}$ has the *size-biased* distribution of X , a notion which frequently arises in probability theory and applications [3], [5].

We can now state and prove the following result, which sheds some light on the relationship between Proposition 3 and Lemma 4.

Lemma 6. *If W is a random variable with finite second moment and f is an absolutely continuous*

function with bounded derivative then, for V uniform on the interval $(-1, 1)$ and independent of $W^{(2)}$,

$$2\mathbb{E}W^2\mathbb{E}f'(VW^{(2)}) = \mathbb{E}Wf(W) - \mathbb{E}Wf(-W).$$

Proof. The lemma is implied by the following calculation:

$$\begin{aligned}\mathbb{E}f'(VW^{(2)}) &= \frac{1}{2}\mathbb{E}\left[\int_{-1}^1 f'(uW^{(2)})du\right] \\ &= \frac{1}{2}\mathbb{E}\left[\frac{f(W^{(2)}) - f(-W^{(2)})}{W^{(2)}}\right] \\ &= \frac{\mathbb{E}Wf(W) - \mathbb{E}Wf(-W)}{2\mathbb{E}W^2},\end{aligned}$$

where in the final equality we use (2). \square

We now have the following main result for the Gaussian distribution, which is essentially a rephrasing of Proposition 2.3 on page 35 of [8].

Theorem 7. *Let W be a random variable with finite second moment. The following are equivalent:*

- (i) W has the standard normal distribution.
- (ii) For all absolutely continuous functions f with bounded derivative,

$$\mathbb{E}f'(W) = \mathbb{E}Wf(W).$$

- (iii) $\mathbb{E}W^2 = 1$ and $W \stackrel{d}{=} VW^{(2)}$, where V is uniform on $(-1, 1)$ and independent of $W^{(2)}$.

Proof. The equivalence of the first two items of the proposition is (Stein's) Lemma 4 above.

The fact that (i) implies (iii) follows from Proposition 3 above coupled with the simple fact that for X_1, X_2, X_3 i.i.d. standard normal random variables, the density of $(X_1^2 + X_2^2 + X_3^2)^{1/2}$ is proportional to $x^2 e^{-x^2/2}$ (that is, $(X_1^2 + X_2^2 + X_3^2)^{1/2}$ has the same distribution as $X_1^{(2)}$).

Finally, we show that (ii) follows from (iii). If $W \stackrel{d}{=} VW^{(2)}$ and $\mathbb{E}W^2 = 1$, then using Lemma 6 we find that, for functions f with bounded derivative,

$$\begin{aligned}\mathbb{E}f'(W) &= \mathbb{E}f'(VW^{(2)}) \\ &= \frac{1}{2}(\mathbb{E}Wf(W) - \mathbb{E}Wf(-W)) = \mathbb{E}Wf(W),\end{aligned}$$

where in the last equality we use that $W \stackrel{d}{=} -W$, which follows in turn from the fact that $V \stackrel{d}{=} -V$ and hence $-VW^{(2)} \stackrel{d}{=} VW^{(2)}$. \square

Remark 8. The equivalence of (i) and (iii) is essentially the content of Proposition 2.3 on page 35 of [8], which uses the concept of the "zero-bias" transformation of Stein's method, first introduced in [11]. For a random variable W with mean zero and variance $\sigma^2 < \infty$, we say that W^* has the zero-bias distribution of W if, for all f with $\mathbb{E}|Wf(W)| < \infty$,

$$\sigma^2\mathbb{E}f'(W^*) = \mathbb{E}Wf(W).$$

We think of the zero-bias transformation acting on probability distributions with zero mean and finite variance, and Stein's lemma implies that this transformation has the centered Gaussian distribution as its unique fixed point. Proposition 2.3 on page 35 of [8] states that, for a random variable W with support symmetric about zero with unit variance, the transformation $W \rightarrow VW^{(2)}$ provides a representation of the zero-bias transformation. The equivalence of (i) and (iii) of theorem follows easily from these results.

Beta-Gamma Algebra

The equivalence between (i) and (iii) in Theorem 7 can be generalized as follows. For $r, s > 0$, let G_r and $B_{r,s}$ denote standard gamma and beta random variables having respective densities $\frac{1}{\Gamma(r)}x^{r-1}e^{-x}$, $x > 0$, and $\frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}y^{r-1}(1-y)^{s-1}$, $0 < y < 1$, where Γ denotes the gamma function.

Proposition 9. *Fix $p, r, s > 0$. A nonnegative random variable W has the distribution of cG_r^p for some constant $c > 0$ if and only if $W \stackrel{d}{=} B_{r,s}^p W^{(s/p)}$, where $B_{r,s}$ is independent of $W^{(s/p)}$.*

Remark 10. The equivalence in (i) and (iii) of Theorem 7 follows by taking $p = r = 1/2$, $s = 1$ in Proposition 9 and using the well-known fact that for Z having the standard normal distribution, $Z^2 \stackrel{d}{=} 2G_{1/2}$.

The proof of Proposition 9 uses the following result.

Lemma 11. *Let $\alpha, \beta > 0$. If $X \geq 0$ is a random variable such that $\mathbb{E}X^\alpha < \infty$, then*

$$(X^{(\alpha)})^\beta \stackrel{d}{=} (X^\beta)^{(\alpha/\beta)}.$$

Proof. By the definition of α/β -power biasing, we only need to show that

$$(3) \quad \mathbb{E}X^\alpha\mathbb{E}f((X^{(\alpha)})^\beta) = \mathbb{E}X^\alpha f(X^\beta)$$

for all f such that the expectation on the left-hand side exists. By the definition of α -power biasing, we have that, for $g(t) = f(t^\beta)$,

$$\mathbb{E}X^\alpha\mathbb{E}g(X^{(\alpha)}) = \mathbb{E}X^\alpha g(X),$$

which is (3). \square

Proof of Proposition 9. The usual beta-gamma algebra (see [9]) implies that $G_r \stackrel{d}{=} B_{r,s}G_{r+s}$ where $B_{r,s}$ and G_{r+s} are independent. Using the elementary fact that $G_{r+s} \stackrel{d}{=} G_r^{(s)}$, we find that for fixed $r, s > 0$, G_r satisfies $G_r \stackrel{d}{=} B_{r,s}G_r^{(s)}$. Now applying Lemma 11 to G_r with $\alpha = s$ and $\beta = p$, we have that $W = G_r^p$ satisfies $W \stackrel{d}{=} B_{r,s}^p W^{(s/p)}$ and the forward implication now follows after noting that $(cX)^{(\alpha)} \stackrel{d}{=} cX^{(\alpha)}$.

Now assume that $W \stackrel{d}{=} B_{r,s}^p W^{(s/p)}$ for fixed p , $r, s > 0$. We show that $W \stackrel{d}{=} c G_r^p$ for some $c > 0$. First, note by Lemma 11 that, if $X = W^{1/p}$, then

$$(4) \quad X \stackrel{d}{=} B_{r,s} X^{(s)}$$

and we will be done if this implies that $X \stackrel{d}{=} G_r$. Observe that, by writing $X^{(s)}$, we have been tacitly assuming that $\mathbb{E}W^{s/p} = \mathbb{E}X^s < \infty$, which implies that $\mathbb{E}(B_{r,s}X^{(s)})^s < \infty$ so that using the definition of power biasing yields $\mathbb{E}X^{2s} < \infty$. Continuing in this way, we find that $\mathbb{E}X^{ks} < \infty$ for all $k = 1, 2, \dots$ and thus that $\mathbb{E}X^q < \infty$ for all $q \geq s$. Moreover, writing $a_k := \mathbb{E}X^{ks}$ and taking expectations in (4) after raising both sides to the power k , we have

$$a_k = \mathbb{E}B_{r,s}^{ks} \frac{a_{k+1}}{a_1},$$

where we have again used the definition of power biasing. We can solve this recursion after noting that, for $\alpha > -r$,

$$\mathbb{E}B_{r,s}^\alpha = \frac{\Gamma(r + \alpha)\Gamma(r + s)}{\Gamma(r + \alpha + s)\Gamma(r)},$$

to find that, for $k = 0, 1, \dots$,

$$a_k = \left(\frac{a_1 \Gamma(r)}{\Gamma(r + s)} \right)^k \frac{\Gamma(r + sk)}{\Gamma(r)}.$$

For any value of $a_1 > 0$, it is easy to see using Stirling's formula that the sequence $(a_k)_{k \geq 1}$ satisfies Carleman's condition

$$\sum_{k=1}^n a_{2k}^{-1/2k} \rightarrow \infty, \quad \text{as } n \rightarrow \infty.$$

Thus, for a given value of a_1 , there is exactly one probability distribution having moment sequence $(a_k)_{k \geq 1}$ (see the remark following Theorem (3.11) in Chapter 2 of [10]). Finally, it is easy to see that the random variable

$$X^s := \frac{a_1 \Gamma(r)}{\Gamma(r + s)} G_r^s$$

has moment sequence $(a_k)_{k \geq 1}$. \square

Exponential Distribution

The exponential distribution has many characterizing properties, many of which stem from its relation to Poisson processes. For example, by superimposing two independent Poisson processes into one, we easily find that, if Z_1 and Z_2 are independent rate-one exponential variables, then $2\min\{Z_1, Z_2\}$ is also a rate-one exponential (this is in fact characterizing as shown in Theorem 3.4.1 of [6]).

For our framework above, we use the memoryless property of the exponential distribution in the context of renewal theory. In greater detail, for any nonnegative random variable X , we define the renewal sequence generated from X as (S_1, S_2, \dots) , where $S_i = \sum_{k=1}^i X_k$ and the X_k are i.i.d. copies of

X . For a fixed $t > 0$, the distribution of the length of the interval $[S_{K_t}, S_{K_t+1}]$ containing t and the position of t in this interval depend on t and the distribution of X in some rather complicated way. We can remove this dependence on t by starting the sequence in "stationarity", meaning that we look instead at the sequence $(X', X' + S_1, \dots)$, where X' has the limiting distribution of $S_{K_t+1} - t$ as t goes to infinity; see Chapter 5, Sections 6 and 7.b of [13].

If X has a continuous distribution with finite mean, then the distribution of X' is the size-biased distribution of X times an independent variable which is uniform on $(0, 1)$ [13]. Heuristically, the memoryless property which characterizes the exponential distribution (Chapter 12 of [4]) implies that the renewal sequence generated by an exponential distribution is stationary (that is, X and X' have the same distribution) and vice versa. The following result implies that this intuition is correct.

Theorem 12 ([16]). *Let W be a nonnegative random variable with finite mean. The following are equivalent:*

- (i) W has the exponential distribution with mean one.
- (ii) For all absolutely continuous functions f with bounded derivative,

$$\mathbb{E}f'(W) = \mathbb{E}f(W) - f(0).$$

- (iii) $\mathbb{E}W = 1$ and $W \stackrel{d}{=} UW^{(1)}$, where U is uniform on $(0, 1)$ and independent of $W^{(1)}$.

Similar to the case of the normal distribution, the crucial link between (ii) and (iii) of Theorem 12 is provided by the following lemma; the proof is similar to that of Lemma 6.

Lemma 13. *If W is a nonnegative random variable with finite mean and f is an absolutely continuous function with bounded derivative, then*

$$\mathbb{E}W\mathbb{E}f'(UW^{(1)}) = \mathbb{E}f(W) - f(0).$$

Proof of Theorem 12. The equivalence of (i) and (iii) is a special case of Proposition 9 with $r = s = p = 1$, and the equivalence of (ii) and (iii) can be read from Lemma 13 (note in particular that (ii) with $f(x) = 1$ implies that $\mathbb{E}W = 1$). \square

Remark 14. For a nonnegative random variable W with finite mean, the transformation $W \rightarrow UW^{(1)}$ is referred to in the Stein's method literature as the "equilibrium" transformation, first defined in this context in [16], where Theorem 12 is also proved.

Due to the close relationship between the exponential and geometric distributions, it is not surprising that there is a discrete analog of Theorem 12 with the exponential distribution replaced by

the geometric; see [17] for this discussion in the context of Stein's method.

Proof of Proposition 3 and Discussion

Proof of Proposition 3. We will show that, for $n \geq 2$ and Y_1, \dots, Y_n nonnegative i.i.d. random variables, $Y_1 \stackrel{d}{=} cG_{1/(n-1)}$ for some $c \geq 0$ if and only if

$$(5) \quad Y_1 \stackrel{d}{=} B_{1/(n-1),1}(Y_1 + \dots + Y_n),$$

where $B_{1/(n-1),1}$ is independent of (Y_1, \dots, Y_n) and $G_a, B_{a,b}$ are gamma and beta variables as defined above. From this point, (5) with $n = 3$ yields a characterization of the distribution of X^2 after noting that $V^2 \stackrel{d}{=} B_{1/2,1}$ and if X has a mean zero and variance one normal distribution, then $X^2 \stackrel{d}{=} 2G_{1/2}$. The proposition then follows since the distribution of X is determined by the distribution of X^2 through the representation (1).

The forward implication is a consequence of Proposition 9 coupled with the fact that $G_{a+b} \stackrel{d}{=} G_a + G_b$, where G_a and G_b are independent. To establish the result, we assume (5) and show that $Y_1 \stackrel{d}{=} cG_{1/(n-1)}$. Since we assume that Y_1 is nonnegative, we define the Laplace transform

$$\varphi(\lambda) = \mathbb{E}e^{-\lambda Y_1}, \quad \lambda \geq 0.$$

By conditioning on the value of $B_{1/(n-1),1}$ in (5), we find, for $\lambda > 0$,

$$\begin{aligned} \varphi(\lambda) &= \mathbb{E}\varphi(B_{1/(n-1),1}\lambda)^n \\ &= \frac{1}{n-1} \int_0^1 u^{-(n-2)/(n-1)} \varphi(u\lambda)^n du \\ &= \frac{1}{(n-1)\lambda^{1/(n-1)}} \int_0^\lambda t^{-(n-2)/(n-1)} \varphi(t)^n dt, \end{aligned}$$

where we have made the change of variable $t = u\lambda$ in the last equality. We can differentiate the equation above with respect to λ , which yields

$$\begin{aligned} \varphi'(\lambda) &= -\frac{\lambda^{-n/(n-1)}}{(n-1)^2} \int_0^\lambda t^{-(n-2)/(n-1)} \varphi(t)^n dt \\ &\quad + \frac{1}{(n-1)\lambda} \varphi(\lambda)^n; \end{aligned}$$

thus

$$(6) \quad \varphi'(\lambda) = \frac{-\varphi(\lambda) + \varphi(\lambda)^n}{(n-1)\lambda}.$$

Consequently we find that φ satisfies the differential equation (6) with boundary condition $\varphi(0) = 1$.

By computing the derivative using (6) and the fact that $0 < \varphi(\lambda) \leq 1$ for $\lambda > 0$, we find that, for some constant $c \geq 0$,

$$\frac{1 - \varphi(\lambda)^{n-1}}{\lambda \varphi(\lambda)^{n-1}} = c, \quad \lambda > 0.$$

Solving this equation for $\varphi(\lambda)$ implies that

$$\varphi(\lambda) = (1 + c\lambda)^{-1/(n-1)},$$

which is the Laplace transform of $cG_{1/(n-1)}$, as desired. \square

The proof of Proposition 3 and the beta-gamma algebra suggest the following proposition and conjecture.

Proposition 15. Fix $n \geq 2$, $a > 0$ and let $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$ be a vector of nondegenerate i.i.d. random variables such that $Y_1 \geq 0$ and $\mathbb{E}Y_1^k < \infty$ for all $k > 1$. Then Y_1 is equal in distribution to cG_a for some constant $c > 0$ if and only if, for $V = B_{a,(n-1)a}$ independent of \mathbf{Y} ,

$$(7) \quad Y_1 \stackrel{d}{=} V(Y_1 + Y_2 + \dots + Y_n).$$

Proof. The forward implication of the proposition is an easy consequence of the following beta-gamma algebra facts: for G_a, G_b , and $B_{a,b}$ independent, $B_{a,b}G_{a+b} \stackrel{d}{=} G_a$ and $G_a + G_b \stackrel{d}{=} G_{a+b}$.

The remaining implication follows easily after using (7) to obtain a recursion relation for the moments of Y_1 which, up to the scale factor, determines those of a gamma distribution with the appropriate parameter. \square

The assumption in Proposition 15 that all moments of Y_1 are finite is likely unnecessary; we make the following conjecture.

Conjecture 16. Proposition 15 holds for each $n \geq 2$ and $a > 0$ with no moment assumption on Y_1 .

The forward implication of Proposition 15 is clearly unaffected by removing the moment assumption on Y_1 . Conversely, assuming (7), the proof of Proposition 3 above implies the conjecture in the special case where $a = 1/(n-1)$. In the general case, it is possible to follow the proof of Proposition 3, which leads to an integral equation for the Laplace transform of Y_1 and it is only a matter of showing that this integral equation has a unique scale family of solutions. In the case $a = 1/(n-1)$, the integral equation has a simple form from which we are able to deduce the required uniqueness, but in the general case we do not have the analogous argument.

Conjecture 16 is very similar to Lukacs's characterization of the gamma distribution [14] that positive, nondegenerate, independent variables X, Y have the gamma distribution if and only if $X + Y$ and $X/(X + Y)$ are independent. However, it does not appear that this result can be used to show the difficult implication of the conjecture. Note additionally that Lukacs's result also characterizes beta distributions as the only distributions which can be written as $X/(X + Y)$ independent of $X + Y$ for positive, nondegenerate, independent

variables X, Y . Thus, a question related to our conjecture is that, if (7) holds for independent variables Y_1, \dots, Y_n and V , does this imply that V has a beta distribution?

Conjecture 16 is connected to the observation of Poincaré (see the introduction of [15]) that the coordinates of a point uniformly chosen on the $(n - 1)$ -dimensional sphere of radius \sqrt{n} are asymptotically distributed as independent standard Gaussians. Analogous to the discussion in the introduction, we can realize these uniformly distributed points as $\sqrt{n}R^{-1}(X_1, \dots, X_n)$, where X_1, \dots, X_n are independent standard normal variables and $R = (X_1^2 + \dots + X_n^2)^{1/2}$. Squaring these coordinates, Poincaré's result implies that $nX_1^2/(X_1^2 + \dots + X_n^2)$ is asymptotically distributed as X_1^2 . Since $X_1^2 \stackrel{d}{=} 2G_{1/2}$, taking the limit as $n \rightarrow \infty$ on the right side of (7) with $a = 1/2$ yields a related fact.

The forward implication of Proposition 3 is evidenced also by the creation of a three-dimensional Bessel process by conditioning a one-dimensional Brownian motion not to hit zero. Indeed, a process version of Proposition 3 is involved in the proof of the “ $2M - X$ ” theorem provided in [19]; see the section above on beta-gamma algebra and especially Section 2.3 of [7]. More generally, process analogs of the beta-gamma algebra can be found in Section 3 of [7].

Some extensions of the characterizations discussed in this article to more complicated distributions can be found in the recent work [18].

References

- [1] T. M. APOSTOL and M. A. MNATSAKIAN, A fresh look at the method of Archimedes, *Amer. Math. Monthly* **111**(6):496–508, 2004.
- [2] ARCHIMEDES and T. L. HEATH, *The Works of Archimedes*, Cambridge University Press, 1897. Translated by T. L. Heath.
- [3] R. ARRATIA and L. GOLDSTEIN, Size bias, sampling, the waiting time paradox, and infinite divisibility: when is the increment independent?, <http://arxiv.org/abs/1007.3910>, 2011.
- [4] N. BALAKRISHNAN and A. P. BASU, editors, *The Exponential Distribution: Theory, Methods and Applications*, Gordon and Breach Publishers, Amsterdam, 1995.
- [5] M. BROWN, Exploiting the waiting time paradox: Applications of the size-biasing transformation, *Probab. Engrg. Inform. Sci.* **20**(2):195–230, 2006.
- [6] W. BRYC, *The Normal Distribution: Characterizations with Applications*, Lecture Notes in Statistics, Vol. 100, Springer-Verlag, New York, 1995.
- [7] P. CARMONA, F. PETIT, and M. YOR, Beta-gamma random variables and intertwining relations between certain Markov processes, *Rev. Mat. Iberoamericana* **14**(2):311–367, 1998.
- [8] L. H. Y. CHEN, L. GOLDSTEIN, and Q.-M. SHAO, *Normal Approximation by Stein's Method*, Probability and Its Applications, Springer, 2010.
- [9] D. DUFRESNE, Algebraic properties of beta and gamma distributions, and applications, *Adv. in Appl. Math.* **20**(3):285–299, 1998.
- [10] R. DURRETT, *Probability: Theory and Examples*, second edition, Duxbury Press, Belmont, CA, 1996.
- [11] L. GOLDSTEIN and G. REINERT, Stein's method and the zero bias transformation with application to simple random sampling, *Ann. Appl. Probab.* **7**(4):935–952, 1997.
- [12] N. L. JOHNSON, S. KOTZ, and N. BALAKRISHNAN, *Continuous univariate distributions. Vol. 1*, second edition, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, A Wiley-Interscience Publication, John Wiley & Sons Inc., New York, 1994.
- [13] S. KARLIN and H. M. TAYLOR, *A First Course in Stochastic Processes*, second edition, Academic Press [a subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975.
- [14] E. LUKACS, A characterization of the gamma distribution, *Ann. Math. Statist.* **26**:319–324, 1955.
- [15] H. P. MCKEAN, Geometry of differential space, *Ann. Probab.* **1**:197–206, 1973.
- [16] E. PEKÖZ and A. RÖLLIN, New rates for exponential approximation and theorems of Rényi and Yaglom, *Ann. Probab.* **39**(2):587–608, 2011.
- [17] E. PEKÖZ, A. RÖLLIN, and N. ROSS, Total variation error bounds for geometric approximation, <http://arxiv.org/abs/1005.2774>, 2010, to appear in *Bernoulli*.
- [18] ———, Degree asymptotics with rates for preferential attachment random graphs, <http://arxiv.org/abs/1108.5236>, 2011, to appear in *Ann. Appl. Probab.*
- [19] L. C. G. ROGERS and J. W. PITMAN, Markov functions, *Ann. Probab.* **9**(4):573–582, 1981.
- [20] N. ROSS, Fundamentals of Stein's method, *Probability Surveys* **8**:210–293, 2011.
- [21] C. STEIN, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, CA., 1970/1971), Vol. II: Probability theory*, University of California Press, Berkeley, CA, 1972, pp. 583–602.