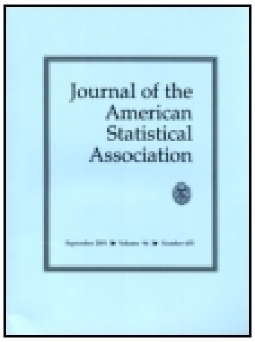

Taylor & Francis
Taylor & Francis Group

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

A New Probability Function and its Properties

Tseng-Tung Cheng

To cite this article: Tseng-Tung Cheng (1944) A New Probability Function and its Properties, Journal of the American Statistical Association, 39:226, 243-245

To link to this article: https://doi.org/10.1080/01621459.1944.10500682

Published online: 11 Apr 2012.

A NEW PROBABILITY FUNCTION AND ITS PROPERTIES

By Tseng-Tung Cheng
National Amoy University, China

Let $x_1, x_2, \cdots, x_n$ be n quantities which vary within the interval $(0, 1)$ and have the same probability function $\phi(x)$. Let $f(t)dt$ be the probability that the sum of these n quantities has a value between t and $t+dt$. Then

$$f(t) = \frac{d}{dt}\int\int \cdots \int_V \phi(x_1)\phi(x_2) \cdots \phi(x_n)dx_1dx_2 \cdots dx_n, \qquad (1)$$

where the region V, over which the n-tuple integral is extended, is defined by the inequalities

$$0 \leqq x_1 \leqq 1,\ 0 \leqq x_2 \leqq 1, \cdots, 0 \leqq x_n \leqq 1,\ x_1+x_2+ \cdots +x_n \leqq t. \qquad (2)$$

Let V' be the region defined by

$$0 \leqq x_1,\ 0 \leqq x_2, \cdots, 0 \leqq x_n,\ x_1 + x_2 + \cdots + x_n \leqq t,$$

and V_s be that part of V' in which s and only s of the x's are greater than 1, so that when $k < t \leqq k+1$, we may write symbolically

$$V' = V + V_1 + \cdots + V_k. \qquad (3)$$

Again, let $V^{(s)}$ be the region defined by

$$x_1 \geqq 1, \cdots, x_s \geqq 1,\ x_{s+1} \geqq 0, \cdots, x_n \geqq 0,\ x_1+x_2+ \cdots +x_n \leqq t \qquad (4)$$

where, of course, s is less than t. Furthermore, we will denote the values of the n-tuple integral

$$\int\int \cdots \int \phi(x_1)\phi(x_2) \cdots \phi(x_n)dx_1dx_2 \cdots dx_n$$

extended over the regions V, V', V_s and $V^{(s)}$ respectively[1] by I, I', I_s and $I^{(s)}$.

By the definitions given above, we may derive the following relations:

$$\left.\begin{aligned}
C_1^n I^{(1)} &= I_1 + C_1^2 I_2 + C_1^3 I_3 + \cdots + C_1^k I_k \\
C_2^n I^{(2)} &= \qquad\quad\ \ I_2 + C_2^3 I_3 + \cdots + C_2^k I_k \\
C_3^n I^{(3)} &= \qquad\qquad\qquad\quad\ I_3 + \cdots + C_3^k I_k \\
&\ \ \vdots \\
C_k^n I^{(k)} &= \qquad\qquad\qquad\qquad\qquad\qquad I_k
\end{aligned}\right\} \qquad (5)$$

[1] It is assumed here that the function $\phi(x)$ has also been defined for values of x lying outside the interval $(0, 1)$.

where C_s^n denotes, as usual, the number of combinations of n things taken s at a time.

With the aid of the relation

$$I' = I + I_1 + \cdots + I_k,$$

which follows immediately from equation (3), we find from equations (5) that

$$I = I' - C_1^n I^{(1)} + C_2^n I^{(2)} - \cdots + (-1)^k C_k^n I^{(k)}. \tag{6}$$

The transformation

$$x_1 = z_1 + 1, \cdots, x_s = z_s + 1, x_{s+1} = z_{s+1}, \cdots, x_n = z_n \tag{7}$$

carries the region $V^{(s)}$ into a new region $R^{(s)}$ defined by the inequalities

$$z_1 \geqq 0, z_2 \geqq 0, \cdots, z_n \geqq 0, z_1 + z_2 + \cdots + z_n \leqq t - s. \tag{8}$$

Hence

$$I^{(s)} = \int\int \cdots \int_{R^{(s)}} \phi(z_1 + 1) \cdots \phi(z_s + 1)\phi(z_{s+1}) \cdots \phi(z_n) dz_1 dz_2 \cdots dz_n. \tag{9}$$

If $\phi(x)$ be a polynomial in x, then each of the integrals I' and $I^{(s)}$ is the sum of a number of Dirichlet's integrals, and may therefore be expressed in terms of Γ functions. In particular, when $\phi(x) \equiv 1$,

$$I' = \frac{[\Gamma(1)]^n}{\Gamma(n + 1)} t^n = \frac{1}{n!} t^n, \tag{10}$$

$$I^{(s)} = \frac{1}{n!} (t - s)^n, \quad s = 1, 2, \cdots, k. \tag{11}$$

Substituting (10) and (11) into (6) gives

$$I = \frac{1}{n!} [t^n - C_1^n (t-1)^n + \cdots + (-1)^k C_k^n (t-k)^n], \quad k < t \leqq k+1. \tag{12}$$

Differentiating (12) we obtain finally

$$f(t) = \frac{dI}{dt} = \frac{1}{(n-1)!} [t^{n-1} - C_1^n (t-1)^{n-1} + \cdots + (-1)^k C_k^n (t-k)^{n-1}], \quad k < t \leqq k + 1. \tag{13}$$

As shown elsewhere,[2] the function $f(t)$ has the following properties:

[2] T. T. Cheng, Collected Scientific and Engineering Papers, National Amoy University.

(a) $f(t)$, together with its first $n-2$ derivatives, is continuous throughout the interval $(0, n)$.

(b) $f(t)$ is symmetric with respect to the point $t=\frac{1}{2}n$.

(c) $f(t)$ increases in the t-interval $(0, \frac{1}{2}n)$ and decreases in $(\frac{1}{2}n, n)$.

(d) The moments of $f(t)$ with respect to its arithmetic average $t=\frac{1}{2}n$ are given by the following identity:

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} \mu_k = \left(\frac{e^t - 1}{te^{\frac{1}{2}t}}\right)^n, \tag{14}$$

where μ_k denotes the kth moment. Hence

$$\mu_0 = 1, \mu_1 = 0, \mu_2 = n/12, \mu_3 = 0, \mu_4 = (5n^2 - 2n)/240, \cdots . \tag{15}$$

In general, $\mu_{2k+1}=0$ and μ_{2k} is a polynomial of the kth degree in n which may be written in the form

$$\mu_{2k} = \frac{(2k!)}{k!}\left[\frac{1}{3!}\frac{1}{2^2}\right]^k n^k[1 + \text{terms involving } 1/n]. \tag{16}$$

Neglecting the terms involving $1/n$ and denoting the standard deviation by σ, this equation gives

$$\sigma^2 = \mu_2 = n/12, \tag{17}$$

whereupon it follows that

$$\mu_{2k} = (2k - 1)(2k - 3) \cdots 3.1\sigma^{2k}, \tag{18}$$

which is a property characteristic of the normal curve

$$y = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2}.$$