UNITED STATES DEPARTMENT OF COMMERCE • Luther H. Hodges, *Secretary*
NATIONAL BUREAU OF STANDARDS • A. V. Astin, *Director*

Handbook of Mathematical Functions

With

Formulas, Graphs, and Mathematical Tables

Edited by
Milton Abramowitz and Irene A. Stegun

National Bureau of Standards
Applied Mathematics Series • 55

Issued June, 1964

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 20402 - - - - Price $6.50

26. Probability Functions

Marvin Zelen [1] and Norman C. Severo [2]

Contents

[1] National Bureau of Standards. (Presently, National Institutes of Health.)
[2] National Bureau of Standards. (Presently, University of Buffalo.)

26. Probability Functions

Mathematical Properties[3]

26.1. Probability Functions: Definitions and Properties

Univariate Cumulative Distribution Functions

A real-valued function $F(x)$ is termed a (univariate) cumulative distribution function (c.d.f.) or simply distribution function if

i) $F(x)$ is non-decreasing, i.e., $F(x_1) \leq F(x_2)$ for $x_1 \leq x_2$

ii) $F(x)$ is everywhere continuous from the right, i.e., $F(x) = \lim_{\epsilon \to 0+} F(x+\epsilon)$

iii) $F(-\infty)=0$, $F(\infty)=1$.

The function $F(x)$ signifies the probability of the event "$X \leq x$" where X is a random variable, i.e., $Pr\{X \leq x\} = F(x)$, and thus describes the c.d.f. of X. The two principal types of distribution functions are termed *discrete* and *continuous*.

Discrete Distributions: Discrete distributions are characterized by the random variable X taking on an enumerable number of values $\ldots, x_{-1}, x_0, x_1, \ldots$ with point probabilities

$$p_n = Pr\{X = x_n\} \geq 0$$

which need only be subject to the restriction

$$\sum_n p_n = 1.$$

The corresponding distribution function can then be written

26.1.1 $$F(x) = Pr\{X \leq x\} = \sum_{x_n \leq x} p_n$$

where the summation is over all values of x for which $x_n \leq x$. The set $\{x_n\}$ of values for which $p_n > 0$ is termed the domain of the random variable X. A discrete distribution of a random variable is called a *lattice distribution* if there exist numbers a and $b \neq 0$ such that every possible value of X can be represented in the form $a+bn$ where n takes on only integral values. A summary of some properties of certain discrete distributions is presented in **26.1.19–26.1.24.**

Continuous Distributions. Continuous distributions are characterized by $F(x)$ being absolutely continuous. Hence $F(x)$ possesses a derivative $F'(x) = f(x)$ and the c.d.f. can be written

26.1.2 $$F(x) = Pr\{X \leq x\} = \int_{-\infty}^{x} f(t)dt.$$

The derivative $f(x)$ is termed the *probability density function* (p.d.f.) or *frequency function*, and the values of x for which $f(x) > 0$ make up the domain of the random variable X. A summary of some properties of certain selected continuous distributions is presented in **26.1.25–26.1.34.**

Multivariate Probability Functions

The real-valued function $F(x_1, x_2, \ldots x_n)$ defines an n-variate cumulative distribution function if

i) $F(x_1, x_2, \ldots x_n)$ is a non-decreasing function for each x_i

ii) $F(x_1, x_2, \ldots x_n)$ is continuous from the right in each x_i; i.e., $F(x_1, x_2, \ldots x_n) = \lim_{\epsilon \to 0+} F(x_1, \ldots, x_i+\epsilon, \ldots, x_n)$

iii) $F(x_1, x_2, \ldots x_n) = 0$ when any $x_i = -\infty$; $F(\infty, \infty, \ldots, \infty) = 1$. The function $F(x_1, x_2, \ldots, x_n)$ signifies the probability of the event $X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n$ where $X_1, X_2, \ldots X_n$ is a set of n random variables.

Thus $Pr\{X_1 \leq x_1, \quad X_2 \leq x_2, \quad \ldots, \quad X_n \leq x_n\} = F(x_1, x_2, \ldots x_n)$. The two principal types of n-variate distribution functions termed *discrete* and *continuous*, are defined in a manner similar to the corresponding cases for the univariate distribution function.

---

[3] *Comment on notation and conventions.*

a. We follow the customary convention of denoting a random variable by a capital letter, i.e., X, and using the corresponding lower case letter, i.e., x, for a particular value that the random variable assumes.

b. For statistical applications it is often convenient to have tabulated the "upper tail area," $1-F(x)$, or the c.d.f. for $|X|$, $F(x)-F(-x)$, instead of simply the c.d.f. $F(x)$. We use the notation P to indicate the c.d.f. of X, $Q=1-P$ to indicate the "upper tail area" and $A=P-Q$ to denote the c.d.f. of $|X|$. In particular we use $P(x)$, $Q(x)$, and $A(x)$ to denote the corresponding functions for the normal or Gaussian probability function, see **26.2.2–26.2.4.** When these distributions depend on other parameters, say θ_1 and θ_2, we indicate this by writing $P(x|\theta_1, \theta_2)$, $Q(x|\theta_1, \theta_2)$, or $A(x|\theta_1, \theta_2)$. For example the chi-square distribution **26.4** depends on the parameter ν and the tabulated function is written $Q(\chi^2|\nu)$.

Series Expansion

26.6.22

$$P(F'|\nu_1,\nu_2,\lambda)=e^{-\frac{\lambda}{2}(1-x)}\, x^{\frac{1}{2}(\nu_1+\nu_2-2)} \sum_{i=0}^{\frac{\nu_2}{2}-1} T_i \qquad (\nu_2 \text{ even})$$

where

$$T_0=1$$

$$T_1=\frac{1}{2}(\nu_1+\nu_2-2+\lambda x)\frac{1-x}{x}$$

$$T_i=\frac{1-x}{2i}[(\nu_1+\nu_2-2i+\lambda x)T_{i-1}+\lambda(1-x)T_{i-2}]$$

$$x=\frac{\nu_2}{\nu_1 F'+\nu_2}$$

Limiting Forms

26.6.23

$$\lim_{\nu_2\to\infty} P(F'|\nu_1,\nu_2,\lambda)=P(\chi'^2|\nu,\lambda), \qquad \chi'^2=\nu_1 F',\ \nu=\nu_1$$

26.6.24

$$\lim_{\nu_1\to\infty} P(F'|\nu_1,\nu_2,\lambda)=Q(\chi^2|\nu), \qquad \chi^2=\frac{\nu_2(1+c^2)}{F'}$$

where $\lambda/\nu_1\to c^2$ as $\nu_1\to\infty$.

Approximations to the Non-Central F-Distribution

26.6.25 $P(F'|\nu_1,\nu_2,\lambda)\approx P(x_1)$, ($\nu_1$ and ν_2 large)

where

$$x_1=\frac{F'-\frac{\nu_2(\nu_1+\lambda)}{\nu_1(\nu_2-2)}}{\frac{\nu_2}{\nu_1}\left[\frac{2}{(\nu_2-2)(\nu_2-4)}\left\{\frac{(\nu_1+\lambda)^2}{\nu_2-2}+\nu_1+2\lambda\right\}\right]^{\frac{1}{2}}}$$

26.6.26

$$P(F'|\nu_1,\nu_2,\lambda)\approx P(F|\nu_1^*,\nu_2),$$

$$F=\frac{\nu_1}{\nu_1+\lambda}F',\quad \nu_1^*=\frac{(\nu_1+\lambda)^2}{\nu_1+2\lambda}$$

26.6.27

$$P(F'|\nu_1,\nu_2,\lambda)\approx P(x_2),$$

$$x_2=\frac{\left[\frac{\nu_1 F'}{(\nu_1+\lambda)}\right]^{1/3}\left[1-\frac{2}{9\nu_2}\right]-\left[1-\frac{2(\nu_1+2\lambda)}{9(\nu_1+\lambda)^2}\right]}{\left[\frac{2}{9}\frac{\nu_1+2\lambda}{(\nu_1+\lambda)^2}+\frac{2}{9\nu_2}\left(\frac{\nu_1}{\nu_1+\lambda}F'\right)^{2/3}\right]^{\frac{1}{2}}}$$

26.7. Student's t-Distribution

If X is a random variable following a normal distribution with mean zero and variance unity, and χ^2 is a random variable following an independent chi-square distribution with ν degrees of freedom, then the distribution of the ratio $\frac{X}{\sqrt{\chi^2/\nu}}$ is called Student's t-distribution with ν degrees of freedom. The probability that $\frac{X}{\sqrt{\chi^2/\nu}}$ will be less in absolute value than a fixed constant t is

26.7.1

$$A(t|\nu)=P_r\left\{\left|\frac{X}{\sqrt{\chi^2/\nu}}\right|\le t\right\}$$

$$=\left[\sqrt{\nu}B\left(\frac{1}{2},\frac{\nu}{2}\right)\right]^{-1}\int_{-t}^{t}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}dx$$

$$=1-2I_x\left(\frac{\nu}{2},\frac{1}{2}\right), \qquad (0\le t<\infty)$$

where

$$x=\frac{\nu}{\nu+t^2}$$

Statistical Properties

26.7.2

mean: $m=0$

variance: $\sigma^2=\frac{\nu}{\nu-2}$ $\qquad(\nu>2)$

skewness: $\gamma_1=0$

excess: $\gamma_2=\frac{6}{\nu-4}$ $\qquad(\nu>4)$

moments:

$$\mu_{2n}=\frac{1\cdot 3\ \ldots\ (2n-1)\nu^n}{(\nu-2)(\nu-4)\ \ldots\ (\nu-2n)} \qquad (\nu>2n)$$

$$\mu_{2n+1}=0$$

characteristic function:

$$\phi(t)=E\left[\exp\left(it\frac{X}{\sqrt{\chi^2/\nu}}\right)\right]=\frac{\left(\frac{|t|}{2\sqrt{\nu}}\right)^{\nu/2}}{\pi\Gamma(\nu/2)}Y_{\frac{\nu}{2}}\left(\frac{|t|}{\sqrt{\nu}}\right)$$

Series Expansions

$$\left(\theta=\arctan\frac{t}{\sqrt{\nu}}\right)$$

26.7.3

$$A(t|\nu)=\begin{cases}\frac{2}{\pi}\left\{\theta+\sin\theta\left[\cos\theta+\frac{2}{3}\cos^3\theta+\ \ldots\right.\right.\\ \qquad\left.\left.+\frac{2\cdot 4\ \ldots\ (\nu-3)}{1\cdot 3(\nu-2)}\cos^{\nu-2}\theta\right]\right\} \\ \qquad\qquad (\nu>1 \text{ and odd}) \\ \frac{2}{\pi}\theta \qquad (\nu=1)\end{cases}$$

26.7.4

$$A(t|\nu)=\sin\theta\left\{1+\frac{1}{2}\cos^2\theta+\frac{1\cdot 3}{2\cdot 4}\cos^4\theta+\ \ldots\right.$$

$$\left.+\frac{1\cdot 3\cdot 5\ \ldots\ (\nu-3)}{2\cdot 4\cdot 6(\nu-2)}\cos^{\nu-2}\theta\right\} \qquad (\nu \text{ even})$$

Asymptotic Expansion for the Inverse Function

If $A(t_p|\nu)=1-2p$ and $Q(x_p)=p$, then

26.7.5

$$t_p \sim x_p+\frac{g_1(x_p)}{\nu}+\frac{g_2(x_p)}{\nu^2}+\frac{g_3(x_p)}{\nu^3}+\ldots$$

$$g_1(x)=\frac{1}{4}(x^3+x)$$

$$g_2(x)=\frac{1}{96}(5x^5+16x^3+3x)$$

$$g_3(x)=\frac{1}{384}(3x^7+19x^5+17x^3-15x)$$

$$g_4(x)=\frac{1}{92160}(79x^9+776x^7+1482x^5-1920x^3-945x)$$

Limiting Distribution

26.7.6

$$\lim_{\nu\to\infty} A(t|\nu)=\frac{1}{\sqrt{2\pi}}\int_{-t}^{t} e^{-x^2/2}dx=A(t)$$

Approximation for Large Values of t and $\nu \leq 5$

26.7.7 $$A(t|\nu)\approx 1-2\left\{\frac{a_\nu}{t^\nu}+\frac{b_\nu}{t^{\nu+1}}\right\}$$

ν	1	2	3	4	5
a_ν	.3183	.4991	1.1094	3.0941	9.948
b_ν	.0000	.0518	−.0460	−2.756	−14.05

Approximation for Large ν

26.7.8 $$A(t|\nu)\approx 2P(x)-1, \qquad x=\frac{t\left(1-\frac{1}{4\nu}\right)}{\sqrt{1+\frac{t^2}{2\nu}}}$$

Non-Central t-Distribution

26.7.9

$$P(t'|\nu,\delta)=$$

$$\frac{1}{\sqrt{\nu}B\left(\frac{1}{2},\frac{\nu}{2}\right)}\int_{-\infty}^{t'}\left(\frac{\nu}{\nu+x^2}\right)^{\frac{\nu+1}{2}} e^{-\frac{1}{2}\frac{\nu\delta^2}{\nu+x^2}} Hh_\nu\left(\frac{-\delta x}{\sqrt{\nu+x^2}}\right)dx$$

$$=1-\sum_{j=0}^{\infty} e^{-\delta^2/2}\frac{(\delta^2/2)^j}{j!} I_x\left(\frac{1}{2}+j,\frac{\nu}{2}\right), \qquad x=\frac{\nu}{\nu+t'^2}$$

where δ is termed the non-centrality parameter.

Approximation to the Non-Central t-Distribution

26.7.10

$$P(t'|\nu,\delta)\approx P(x) \qquad \text{where } x=\frac{t'\left(1-\frac{1}{4\nu}\right)-\delta}{\left(1+\frac{t'^2}{2\nu}\right)^{\frac{1}{2}}}$$

Numerical Methods

26.8. Methods of Generating Random Numbers and Their A

$P(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt$

Normal (Gaussian) probability function

Random digits are digits generated by repeated independent drawings from the population 0, 1, 2, . . ., 9 where the probability of selecting any digit is one-tenth. This is equivalent to putting 10 balls, numbered from 0 to 9, into an urn and drawing one ball at a time, replacing the ball after each drawing. The recorded set of numbers forms a collection of random digits. Any group of n successive random digits is known as a *random number*.

Several lengthy tables of random digits are available (see references). However, the use of random numbers in electronic computers has resulted in a need for random numbers to be generated in a completely deterministic way. The numbers so generated are termed pseudo-random numbers. The quality of pseudo-random numbers is determined by subjecting the numbers to several statistical tests, see [26.55], [26.56]. The purpose of these statistical tests is to detect any properties of the pseudo-random numbers which are different from the (conceptual) properties of random numbers.

[9] The authors wish to express their appreciation to Professor J. W. Tukey who made many penetrating and helpful suggestions in this section.

Experience has sh
method is the most pre
random numbers on a c
of pseudo-random nur
$n=0, 1, 2, \ldots.$ Th
of generating pseudo-r

$$X_{n+1}=aX_n+b \pmod T$$

where b and T are relatively prime. The choice of T is determined by the capacity and base of the computer; a and b are chosen so that: (1) the resulting sequence $\{X_n\}$ possesses the desired statistical properties of random numbers, (2) the period of the sequence is as long as possible, and (3) the speed of generation is fast. A guide for choosing a and b is to make the correlation between the numbers be near zero, e.g., the correlation between X_n and X_{n+s} is

$$\rho_s=\frac{1-6\frac{b_s}{T}\left(1-\frac{b_s}{T}\right)}{a_s}+e$$

where

$a_s=a^s \pmod T$
$b_s=(1+a+a^2+\ldots+a^{s-1})b \pmod T$
$|e|<a_s/T$