Internal Report SUF–PFY/96–01
Stockholm, 11 December 1996
1st revision, 31 October 1998

Hand-book on STATISTICAL DISTRIBUTIONS for experimentalists

by

Christian Walck

Particle Physics Group
Fysikum
University of Stockholm
(e-mail: walck@physto.se)

Contents

List of Tables

1 Introduction

In experimental work *e.g.* in physics one often encounters problems where a standard statistical probability density function is applicable. It is often of great help to be able to handle these in different ways such as calculating probability contents or generating random numbers.

For these purposes there are excellent text-books in statistics *e.g.* the classical work of Maurice G. Kendall and Alan Stuart [1,2] or more modern text-books as [3] and others. Some books are particularly aimed at experimental physics or even specifically at particle physics [4–7]. Concerning numerical methods a valuable references worth mentioning is [8] which has been surpassed by a new edition [9]. Also hand-books, especially [10], has been of great help throughout.

However, when it comes to actual applications it often turns out to be hard to find detailed explanations in the literature ready for implementation. This work has been collected over many years in parallel with actual experimental work. In this way some material may be "historical" and sometimes be naïve and have somewhat clumsy solutions not always made in the mathematically most stringent may. We apologize for this but still hope that it will be of interest and help for people who is struggling to find methods to solve their statistical problems in making real applications and not only learning statistics as a course. Even if one has the skill and may be able to find solutions it seems worthwhile to have easy and fast access to formulæ ready for application. Similar books and reports exist *e.g.* [11,12] but we hope the present work may compete in describing more distributions, being more complete, and including more explanations on relations given.

The material could most probably have been divided in a more logical way but we have chosen to present the distributions in alphabetic order. In this way it is more of a hand-book than a proper text-book.

After the first release the report has been modestly changed. Minor changes to correct misprints is made whenever found. In a few cases subsections and tables have been added. These alterations are described on page 184. In October 1998 the first somewhat bigger revision was made where in particular a lot of material on the non-central sampling distributions were added.

1.1 Random Number Generation

In modern computing Monte Carlo simulations are of vital importance and we give methods to achieve random numbers from the distributions. An earlier report dealt entirely with these matters [13]. Not all text-books on statistics include information on this subject which we find extremely useful. Large simulations are common in particle physics as well as in other areas but often it is also useful to make small "toy Monte Carlo programs" to investigate and study analysis tools developed on ideal, but statistically sound, random samples.

A related and important field which we will only mention briefly here, is how to get good basic generators for achieving random numbers uniformly distributed between zero and one. Those are the basis for all the methods described in order to get random numbers from specific distributions in this document. For a review see *e.g.* [14].

From older methods often using so called multiplicative congruential method or shift-generators G. Marsaglia et al [15] introduced in 1989 a new "universal generator" which became the new standard in many fields. We implemented this in our experiments at CERN and also made a package of routines for general use [16].

This method is still a very good choice but later alternatives, claimed to be even better, have turned up. These are based on on the same type of lagged Fibonacci sequences as is used in the universal generator and was originally proposed by the same authors [17]. An implementations of this method was proposed by F. James [14] and this version was further developed by M. Lüscher [18]. A similar package of routine as was prepared for the universal generator has been implemented for this method [19].

2 Probability Density Functions

2.1 Introduction

Probability density functions in one, discrete or continuous, variable are denoted $p(r)$ and $f(x)$, respectively. They are assumed to be properly normalized such that

$$\sum_r p(r) = 1 \quad \text{and} \quad \int_{-\infty}^{\infty} f(x)dx = 1$$

where the sum or the integral are taken over all relevant values for which the probability density function is defined.

Statisticians often use the *distribution function* or as physicists more often call it the *cumulative function* which is defined as

$$P(r) = \sum_{i=-\infty}^{r} p(i) \quad \text{and} \quad F(x) = \int_{-\infty}^{x} f(t)dt$$

2.2 Moments

Algebraic moments of order r are defined as the expectation value

$$\mu'_r = E(x^r) = \sum_k k^r p(k) \quad \text{or} \quad \int_{-\infty}^{\infty} x^r f(x)dx$$

Obviously $\mu'_0 = 1$ from the normalization condition and μ'_1 is equal to the mean, sometimes called the expectation value, of the distribution.

Central moments of order r are defined as

$$\mu_r = E((k - E(k))^r) \quad \text{or} \quad E((x - E(x))^r)$$

of which the most commonly used is μ_2 which is the variance of the distribution.

Instead of using the third and fourth central moments one often defines the coefficients of skewness γ_1 and kurtosis[1] γ_2 by

$$\gamma_1 = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} \quad \text{and} \quad \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

where the shift by 3 units in γ_2 assures that both measures are zero for a normal distribution. Distributions with positive kurtosis are called *leptokurtic*, those with kurtosis around zero *mesokurtic* and those with negative kurtosis *platykurtic*. Leptokurtic distributions are normally more peaked than the normal distribution while platykurtic distributions are more flat topped.

[1] From greek *kyrtosis* = curvature from *kyrt(ós)* = curved, arched, round, swelling, bulging. Sometimes, especially in older literature, γ_2 is called the coefficient of excess.

2.2.1 Errors of Moments

For a thorough presentation of how to estimate errors on moments we refer to the classical books by M. G. Kendall and A. Stuart [1] (pp 228–245). Below only a brief description is given. For a sample with n observations $x_1, x_2, \ldots, x_n$ we define the moment-statistics for the algebraic and central moments m'_r and m_r as

$$m'_r = \frac{1}{n}\sum_{r=0}^{n} x^r \quad \text{and} \quad m_r = \frac{1}{n}\sum_{r=0}^{n}(x - m'_1)^r$$

The notation m'_r and m_r are thus used for the statistics (sample values) while we denote the true, population, values by μ'_r and μ_r.

The mean value of the r:th and the sampling covariance between the q:th and r:th moment-statistic are given by.

$$\begin{aligned} E(m'_r) &= \mu'_r \\ Cov(m'_q, m'_r) &= \frac{1}{n}\left(\mu'_{q+r} - \mu'_q\mu'_r\right) \end{aligned}$$

These formula are exact. Formulæ for moments about the mean are not as simple since the mean itself is subject to sampling fluctuations.

$$\begin{aligned} E(m_r) &= \mu_r \\ Cov(m_q, m_r) &= \frac{1}{n}\left(\mu_{q+r} - \mu_q\mu_r + rq\mu_2\mu_{r-1}\mu_{q-1} - r\mu_{r-1}\mu_{q+1} - q\mu_{r+1}\mu_{q-1}\right) \end{aligned}$$

to order $1/\sqrt{n}$ and $1/n$, respectively. The covariance between an algebraic and a central moment is given by

$$Cov(m_r, m'_q) = \frac{1}{n}(\mu_{q+r} - \mu_q\mu_r - r\mu_{q+1}\mu_{r-1})$$

to order $1/n$. Note especially that

$$\begin{aligned} V(m'_r) &= \frac{1}{n}\left(\mu'_{2r} - \mu'^2_r\right) \\ V(m_r) &= \frac{1}{n}\left(\mu_{2r} - \mu_r^2 + r^2\mu_2\mu_{r-1}^2 - 2r\mu_{r-1}\mu_{r+1}\right) \\ Cov(m'_1, m_r) &= \frac{1}{n}\left(\mu_{r+1} - r\mu_2\mu_{r-1}\right) \end{aligned}$$

2.3 Characteristic Function

For a distribution in a continuous variable x the Fourier transform of the probability density function

$$\phi(t) = E(e^{\imath xt}) = \int_{-\infty}^{\infty} e^{\imath xt} f(x)dx$$

is called the *characteristic function*. It has the properties that $\phi(0) = 1$ and $|\phi(t)| \leq 1$ for all t. If the cumulative, distribution, function $F(x)$ is continuous everywhere and $dF(x) = f(x)dx$ then we reverse the transform such that

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(t) e^{-\imath xt} dt$$

The characteristic function is related to the moments of the distribution by

$$\phi_x(t) = E(e^{\imath tx}) = \sum_{n=0}^{\infty} \frac{(\imath t)^n E(x^n)}{n!} = \sum_{n=0}^{\infty} \frac{(\imath t)^n \mu'_n}{n!}$$

e.g. algebraic moments may be found by

$$\mu'_r = \frac{1}{\imath^r} \left(\frac{d}{dt} \right)^r \phi(t) \bigg|_{t=0}$$

To find central moments (about the mean μ) use

$$\phi_{x-\mu}(t) = E\left(e^{\imath t(x-\mu)} \right) = e^{-\imath t\mu} \phi_x(t)$$

and thus

$$\mu_r = \frac{1}{\imath^r} \left(\frac{d}{dt} \right)^r e^{-\imath t\mu} \phi(t) \bigg|_{t=0}$$

A very useful property of the characteristic function is that for *independent* variables x and y

$$\phi_{x+y}(t) = \phi_x(t) \cdot \phi_y(t)$$

As an example regard the sum $\sum a_i z_i$ where the z_i's are distributed according to normal distributions with means μ_i and variances σ_i^2. Then the linear combination will also be distributed according to the normal distribution with mean $\sum a_i \mu_i$ and variance $\sum a_i^2 \sigma_i^2$.

To show that the characteristic function in two variables factorizes is the best way to show *independence* between two variables. Remember that a vanishing correlation coefficient does not imply independence while the reversed is true.

2.4 Probability Generating Function

In the case of a distribution in a discrete variable r the characteristic function is given by

$$\phi(t) = E(e^{\imath tr}) = \sum p(r) e^{\imath tr}$$

In this case it is often convenient to write $z = e^{\imath t}$ and define the *probability generating function* as

$$G(z) = E(z^r) = \sum p(r) z^r$$

Derivatives of $G(z)$ evaluated at $z = 1$ are related to *factorial moments* of the distribution

$$
\begin{aligned}
G(1) &= 1 \quad \text{(normalization)} \\
G^1(1) = \left.\frac{d}{dz}G(z)\right|_{z=1} &= E(r) \\
G^2(1) = \left.\frac{d^2}{dz^2}G(z)\right|_{z=1} &= E(r(r-1)) \\
G^3(1) = \left.\frac{d^3}{dz^3}G(z)\right|_{z=1} &= E(r(r-1)(r-2)) \\
G^k(1) = \left.\frac{d^k}{dz^k}G(z)\right|_{z=1} &= E(r(r-1)(r-2)\cdots(r-k+1))
\end{aligned}
$$

Lower order algebraic moments are then given by

$$
\begin{aligned}
\mu'_1 &= G^1(1) \\
\mu'_2 &= G^2(1) + G^1(1) \\
\mu'_3 &= G^3(1) + 3G^2(1) + G^1(1) \\
\mu'_4 &= G^4(1) + 6G^3(1) + 7G^2(1) + G^1(1)
\end{aligned}
$$

while expression for central moments become more complicated.

A useful property of the probability generating function is for a branching process in n steps where

$$G(z) = G_1(G_2(\ldots G_{n-1}(G_n(z))\ldots))$$

with $G_k(z)$ the probability generating function for the distribution in the k:th step. As an example see section 29.4.4 on page 107.

2.5 Cumulants

Although not much used in physics the *cumulants*, κ_r, are of statistical interest. One reason for this is that they have some useful properties such as being invariant for a shift in scale (except the first cumulant which is equal to the mean and is shifted along with the scale). Multiplying the x-scale by a constant a has the same effect as for algebraic moments namely to multiply κ_r by a^r.

As the algebraic moment μ'_n is the coefficient of $(\imath t)^n/n!$ in the expansion of $\phi(t)$ the cumulant κ_n is the coefficient of $(\imath t)^n/n!$ in the expansion of the logarithm of $\phi(t)$ (sometimes called the *cumulant generating function*) *i.e.*

$$\ln \phi(t) = \sum_{n=1}^{\infty} \frac{(\imath t)^n}{n!}\kappa_n$$

and thus

$$\kappa_r = \frac{1}{\imath^r}\left(\frac{d}{dt}\right)^r \ln \phi(t)\bigg|_{t=0}$$

Relations between cumulants and central moments for some lower orders are as follows

$$
\begin{array}{ll}
\kappa_1 = \mu_1' & \\
\kappa_2 = \mu_2 & \mu_2 = \kappa_2 \\
\kappa_3 = \mu_3 & \mu_3 = \kappa_3 \\
\kappa_4 = \mu_4 - 3\mu_2^2 & \mu_4 = \kappa_4 + 3\kappa_2^2 \\
\kappa_5 = \mu_5 - 10\mu_3\mu_2 & \mu_5 = \kappa_5 + 10\kappa_3\kappa_2 \\
\kappa_6 = \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3 & \mu_6 = \kappa_6 + 15\kappa_4\kappa_2 + 10\kappa_3^2 + 15\kappa_2^3 \\
\kappa_7 = \mu_7 - 21\mu_5\mu_2 - 35\mu_4\mu_3 + 210\mu_3\mu_2^2 & \mu_7 = \kappa_7 + 21\kappa_5\kappa_2 + 35\kappa_4\kappa_3 + 105\kappa_3\kappa_2^2 \\
\kappa_8 = \mu_8 - 28\mu_6\mu_2 - 56\mu_5\mu_3 - 35\mu_4^2 + & \mu_8 = \kappa_8 + 28\kappa_6\kappa_2 + 56\kappa_5\kappa_3 + 35\kappa_4^2 + \\
\qquad +420\mu_4\mu_2^2 + 560\mu_3^2\mu_2 - 630\mu_2^4 & \qquad +210\kappa_4\kappa_2^2 + 280\kappa_3^2\kappa_2 + 105\kappa_2^4
\end{array}
$$

2.6 Random Number Generation

When generating random numbers from different distribution it is assumed that a good generator for uniform pseudorandom numbers between zero and one exist (normally the end-points are excluded).

2.6.1 Cumulative Technique

The most direct technique to obtain random numbers from a continuous probability density function $f(x)$ with a limited range from $x_{\rm min}$ to $x_{\rm max}$ is to solve for x in the equation

$$\xi = \frac{F(x) - F(x_{\rm min})}{F(x_{\rm max}) - F(x_{\rm min})}$$

where ξ is uniformly distributed between zero and one and $F(x)$ is the cumulative distribution (or as statisticians say the distribution function). For a properly normalized probability density function thus

$$x = F^{-1}(\xi)$$

The technique is sometimes also of use in the discrete case if the cumulative sum may be expressed in analytical form as *e.g.* for the geometric distribution.

Also for general cases, discrete or continuous, *e.g.* from an arbitrary histogram the cumulative method is convenient and often faster than more elaborate methods. In this case the task is to construct a cumulative vector and assign a random number according to the value of a uniform random number (interpolating within bins in the continuous case).

2.6.2 Accept-Reject technique

A useful technique is the acceptance-rejection, or hit-miss, method where we choose $f_{\rm max}$ to be greater than or equal to $f(x)$ in the entire interval between $x_{\rm min}$ and $x_{\rm max}$ and proceed as follows

i Generate a pair of uniform pseudorandom numbers ξ_1 and ξ_2.

ii Determine $x = x_{\rm min} + \xi_1 \cdot (x_{\rm max} - x_{\rm min})$.

iii Determine $y = f_{\rm max} \cdot \xi_2$.

iv If $y - f(x) > 0$ reject and go to **i** else accept x as a pseudorandom number from $f(x)$.

The efficiency of this method depends on the average value of $f(x)/f_{\max}$ over the interval. If this value is close to one the method is efficient. On the other hand, if this average is close to zero, the method is extremely inefficient. If α is the fraction of the area $f_{\max} \cdot (x_{\max} - x_{\min})$ covered by the function the average number of rejects in step **iv** is $\frac{1}{\alpha} - 1$ and $\frac{2}{\alpha}$ uniform pseudorandom numbers are required on average.

The efficiency of this method can be increased if we are able to choose a function $h(x)$, from which random numbers are more easily obtained, such that $f(x) \leq \alpha h(x) = g(x)$ over the entire interval under consideration (where α is a constant). A random sample from $f(x)$ is obtained by

i Generate in x a random number from $h(x)$.

ii Generate a uniform random number ξ.

iii If $\xi \geq f(x)/g(x)$ go back to **i** else accept x as a pseudorandom number from $f(x)$.

Yet another situation is when a function $g(x)$, from which fast generation may be obtained, can be inscribed in such a way that a big proportion (f) of the area under the function is covered (as an example see the trapezoidal method for the normal distribution). Then proceed as follows:

i Generate a uniform random number ξ.

ii If $\xi < f$ then generate a random number from $g(x)$.

iii Else use the acceptance/rejection technique for $h(x) = f(x) - g(x)$ (in subintervals if more efficient).

2.6.3 Composition Techniques

If $f(x)$ may be written in the form

$$f(x) = \int_{-\infty}^{\infty} g_z(x) dH(z)$$

where we know how to sample random numbers from the p.d.f. $g(x)$ and the distribution function $H(z)$. A random number from $f(x)$ is then obtained by

i Generate two uniform random numbers ξ_1 and ξ_2.

ii Determine $z = H^{-1}(\xi_1)$.

iii Determine $x = G^{-1}(\xi_2)$ where G is the distribution function corresponding to the p.d.f. $g(x)$.

For more detailed information on the Composition technique see [20] or [21].

A combination of the composition and the rejection method has been proposed by J. C. Butcher [22]. If $f(x)$ can be written

$$f(x) = \sum_{i=0}^{n} \alpha_i f_i(x) g_i(x)$$

where α_i are positive constants, $f_i(x)$ p.d.f.'s for which we know how to sample a random number and $g_i(x)$ are functions taking values between zero and one. The method is then as follows:

i Generate uniform random numbers ξ_1 and ξ_2.

ii Determine an integer k from the discrete distribution $p_i = \alpha_i/(\alpha_1 + \alpha_2 + ... + \alpha_n)$ using ξ_1.

iii Generate a random number x from $f_k(x)$.

iv Determine $g_k(x)$ and if $\xi_2 > g_k(x)$ then go to **i**.

v Accept x as a random number from $f(x)$.

2.7 Multivariate Distributions

Joint probability density functions in several variables are denoted by $f(x_1, x_2, \ldots, x_n)$ and $p(r_1, r_2, \ldots, r_n)$ for continuous and discrete variables, respectively. It is assumed that they are properly normalized *i.e.* integrated (or summed) over all variables the result is unity.

2.7.1 Multivariate Moments

The generalization of algebraic and central moments to multivariate distributions is straightforward. As an example we take a bivariate distribution $f(x, y)$ in two continuous variables x and y and define algebraic and central bivariate moments of order k, ℓ as

$$\mu'_{k\ell} \equiv E(x^k y^\ell) = \iint x^k y^\ell f(x, y) dx dy$$

$$\mu_{k\ell} \equiv E((x - \mu_x)^k (y - \mu_y)^\ell) = \iint (x - \mu_x)^k (y - \mu_y)^\ell f(x, y) dx dy$$

where μ_x and μ_y are the mean values of x and y. The covariance is a central bivariate moment of order $1, 1$ *i.e.* $Cov(x, y) = \mu_{11}$. Similarly one easily defines multivariate moments for distribution in discrete variables.

2.7.2 Errors of Bivariate Moments

Algebraic (m'_{rs}) and central (m_{rs}) *bivariate moments* are defined by:

$$m'_{rs} = \frac{1}{n}\sum_{i=1}^{n} x_i^r y_i^s \quad \text{and} \quad m_{rs} = \frac{1}{n}\sum_{i=1}^{n}(x_i - m'_{10})^r (y_i - m'_{01})^s$$

When there is a risk of ambiguity we write $m_{r,s}$ instead of m_{rs}.

The notations m'_{rs} and m_{rs} are used for the statistics (sample values) while we write μ'_{rs} and μ_{rs} for the population values. The errors of bivariate moments are given by

$$\begin{aligned}
Cov(m'_{rs}, m'_{uv}) &= \frac{1}{n}(\mu'_{r+u,s+v} - \mu'_{rs}\mu'_{uv}) \\
Cov(m_{rs}, m_{uv}) &= \frac{1}{n}(\mu_{r+u,s+v} - \mu_{rs}\mu_{uv} + ru\mu_{20}\mu_{r-1,s}\mu_{u-1,v} + sv\mu_{02}\mu_{r,s-1}\mu_{u,v-1} \\
&\quad + rv\mu_{11}\mu_{r-1,s}\mu_{u,v-1} + su\mu_{11}\mu_{r,s-1}\mu_{u-1,v} - u\mu_{r+1,s}\mu_{u-1,v} \\
&\quad - v\mu_{r,s+1}\mu_{u,v-1} - r\mu_{r-1,s}\mu_{u+1,v} - s\mu_{r,s-1}\mu_{u,v+1})
\end{aligned}$$

especially

$$\begin{aligned}
V(m'_{rs}) &= \frac{1}{n}(\mu'_{2r,2s} - \mu'^2_{rs}) \\
V(m_{rs}) &= \frac{1}{n}(\mu_{2r,2s} - \mu^2_{rs} + r^2\mu_{20}\mu^2_{r-1,s} + s^2\mu_{02}\mu^2_{r,s-1} \\
&\quad + 2rs\mu_{11}\mu_{r-1,s}\mu_{r,s-1} - 2r\mu_{r+1,s}\mu_{r-1,s} - 2s\mu_{r,s+1}\mu_{r,s-1})
\end{aligned}$$

For the *covariance* (m_{11}) we get by error propagation

$$\begin{aligned}
V(m_{11}) &= \frac{1}{n}(\mu_{22} - \mu^2_{11}) \\
Cov(m_{11}, m'_{10}) &= \frac{\mu_{21}}{n} \\
Cov(m_{11}, m_{20}) &= \frac{1}{n}(\mu_{31} - \mu_{20}\mu_{11})
\end{aligned}$$

For the *correlation coefficient* (denoted by $\rho = \mu_{11}/\sqrt{\mu_{20}\mu_{02}}$ for the population value and by r for the sample value) we get

$$V(r) = \frac{\rho^2}{n}\left\{\frac{\mu_{22}}{\mu^2_{11}} + \frac{1}{4}\left[\frac{\mu_{40}}{\mu^2_{20}} + \frac{\mu_{04}}{\mu^2_{02}} + \frac{2\mu_{22}}{\mu_{20}\mu_{02}}\right] - \frac{1}{\mu_{11}}\left[\frac{\mu_{31}}{\mu_{20}} + \frac{\mu_{13}}{\mu_{02}}\right]\right\}$$

Beware, however, that the sampling distribution of r tends to normality very slowly.

2.7.3 Joint Characteristic Function

The joint characteristic function is defined by

$$\begin{aligned}
\phi(t_1, t_2, \ldots, t_n) &= E(e^{\imath t_1 x_1 + \imath t_2 x_2 + \ldots t_n x_n}) = \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} e^{\imath t_1 x_1 + \imath t_2 x_2 + \ldots + \imath t_n x_n} f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n
\end{aligned}$$

From this function multivariate moments may be obtained *e.g.* for a bivariate distribution algebraic bivariate moments are given by

$$\mu'_{rs} = E(x_1^r x_2^s) = \left.\frac{\partial^{r+s}\phi(t_1,t_2)}{\partial(\imath t_1)^r\partial(\imath t_2)^s}\right|_{t_1=t_2=0}$$

2.7.4 Random Number Generation

Random sampling from a many dimensional distribution with a joint probability density function $f(x_1, x_2, ..., x_n)$ can be made by the following method:

- Define the marginal distributions

$$g_m(x_1, x_2, ..., x_m) = \int f(x_1, ..., x_n)dx_{m+1}dx_{m+2}...dx_n = \int g_{m+1}(x_1, ..., x_{m+1})dx_{m+1}$$

- Consider the conditional density function h_m given by

$$h_m(x_m|x_1, x_2, ...x_{m-1}) \equiv g_m(x_1, x_2, ..., x_m)/g_{m-1}(x_1, x_2, ..., x_{m-1})$$

- We see that $g_n = f$ and that

$$\int h_m(x_m|x_1, x_2, ..., x_{m-1})dx_m = 1$$

  from the definitions. Thus h_m is the conditional distribution in x_m given fixed values for $x_1, x_2, ..., x_{m-1}$.

- We can now factorize f as

$$f(x_1, x_2, ..., x_n) = h_1(x_1)h_2(x_2|x_1)\ldots h_n(x_n|x_1, x_2, ..., x_{n-1})$$

- We sample values for $x_1, x_2, ..., x_n$ from the joint probability density function f by:
  - Generate a value for x_1 from $h_1(x_1)$.
  - Use x_1 and sample x_2 from $h_2(x_2|x_1)$.
  - Proceed step by step and use previously sampled values for $x_1, x_2, ..., x_m$ to obtain a value for x_{m+1} from $h_{m+1}(x_{m+1}|x_1, x_2, ..., x_m)$.
  - Continue until all x_i:s have been sampled.

- If all x_i:s are independent the conditional densities will equal the marginal densities and the variables can be sampled in any order.

3 Bernoulli Distribution

3.1 Introduction

The Bernoulli distribution, named after the swiss mathematician Jacques Bernoulli (1654–1705), describes a probabilistic experiment where a trial has two possible outcomes, a success or a failure.

The parameter p is the probability for a success in a single trial, the probability for a failure thus being $1-p$ (often denoted by q). Both p and q is limited to the interval from zero to one. The distribution has the simple form

$$p(r;p) = \begin{cases} 1-p=q & \text{if} \quad r=0 \quad \text{(failure)} \\ p & \text{if} \quad r=1 \quad \text{(success)} \end{cases}$$

and zero elsewhere. The work of J. Bernoulli, which constitutes a foundation of probability theory, was published posthumously in *Ars Conjectandi* (1713) [23].

The probability generating function is $G(z) = q + pz$ and the distribution function given by $P(0) = q$ and $P(1) = 1$. A random numbers are easily obtained by using a uniform random number variate ξ and putting $r = 1$ (success) if $\xi \leq p$ and $r = 0$ else (failure).

3.2 Relation to Other Distributions

From the Bernoulli distribution we may deduce several probability density functions described in this document all of which are based on series of independent Bernoulli trials:

- **Binomial distribution**: expresses the probability for r successes in an experiment with n trials ($0 \leq r \leq n$).

- **Geometric distribution**: expresses the probability of having to wait exactly r trials before the first successful event ($r \geq 1$).

- **Negative Binomial distribution**: expresses the probability of having to wait exactly r trials until k successes have occurred ($r \geq k$). This form is sometimes referred to as the *Pascal distribution*.

  Sometimes this distribution is expressed as the number of failures n occurring while waiting for k successes ($n \geq 0$).

4 Beta distribution

4.1 Introduction

The Beta distribution is given by

$$f(x; p, q) = \frac{1}{B(p, q)} x^{p-1}(1-x)^{q-1}$$

where the parameters p and q are positive real quantities and the variable x satisfies $0 \leq x \leq 1$. The quantity $B(p, q)$ is the Beta function defined in terms of the more common Gamma function as

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

For $p = q = 1$ the Beta distribution simply becomes a uniform distribution between zero and one. For $p = 1$ and $q = 2$ or vise versa we get triangular shaped distributions, $f(x) = 2 - 2x$ and $f(x) = 2x$. For $p = q = 2$ we obtain a distribution of parabolic shape, $f(x) = 6x(1-x)$. More generally, if p and q both are greater than one the distribution has a unique mode at $x = (p-1)/(p+q-2)$ and is zero at the end-points. If p and/or q is less than one $f(0) \to \infty$ and/or $f(1) \to \infty$ and the distribution is said to be J-shaped. In figure 1 below we show the Beta distribution for two cases: $p = q = 2$ and $p = 6, q = 3$.

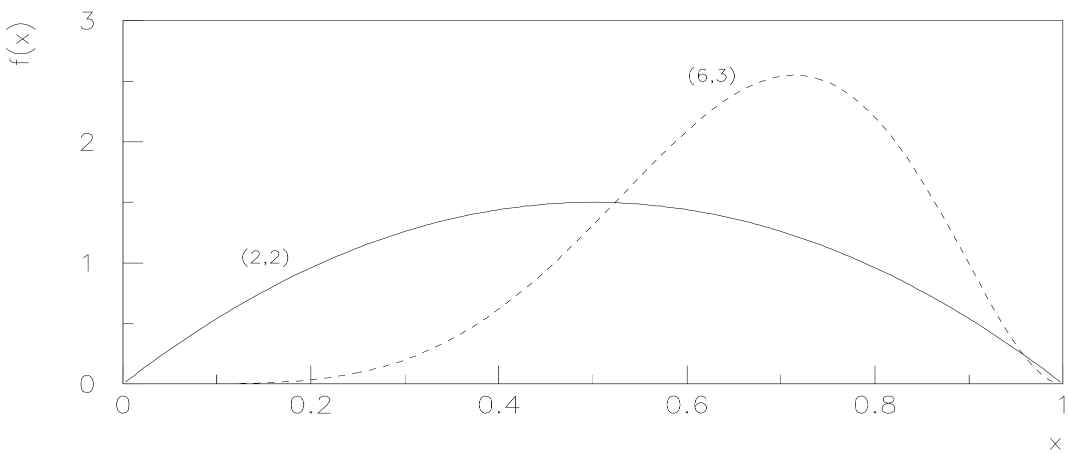


Figure 1: Examples of Beta distributions

4.2 Derivation of the Beta Distribution

If y_m and y_n are two independent variables distributed according to the chi-squared distribution with m and n degrees of freedom, respectively, then the ratio $y_m/(y_m + y_n)$ follows a Beta distribution with parameters $p = \frac{m}{2}$ and $q = \frac{n}{2}$.

To show this we make a change of variables to $x = y_m/(y_m + y_n)$ and $y = y_m + y_n$ which implies that $y_m = xy$ and $y_n = y(1-x)$. We obtain

$$
\begin{aligned}
f(x,y) &= \left\| \begin{array}{cc} \frac{\partial y_m}{\partial x} & \frac{\partial y_m}{\partial y} \\ \frac{\partial y_n}{\partial x} & \frac{\partial y_n}{\partial y} \end{array} \right\| f(y_m, y_n) = \\
&= \left| \begin{array}{cc} y & x \\ -y & 1-x \end{array} \right| \left\{ \frac{\left(\frac{y_m}{2}\right)^{\frac{m}{2}-1} e^{-\frac{y_m}{2}}}{2\Gamma\left(\frac{m}{2}\right)} \right\} \left\{ \frac{\left(\frac{y_n}{2}\right)^{\frac{n}{2}-1} e^{-\frac{y_n}{2}}}{2\Gamma\left(\frac{n}{2}\right)} \right\} = \\
&= \left\{ \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} x^{\frac{m}{2}-1}(1-x)^{\frac{n}{2}-1} \right\} \left\{ \frac{\left(\frac{y}{2}\right)^{\frac{m}{2}+\frac{n}{2}-1} e^{-\frac{y}{2}}}{2\Gamma\left(\frac{m+n}{2}\right)} \right\}
\end{aligned}
$$

which we recognize as a product of a Beta distribution in the variable x and a chi-squared distribution with $m+n$ degrees of freedom in the variable y (as expected for the sum of two independent chi-square variables).

4.3 Characteristic Function

The characteristic function of the Beta distribution may be expressed in terms of the confluent hypergeometric function (see section 43.3) as

$$\phi(t) = M(p, p+q; \imath t)$$

4.4 Moments

The expectation value, variance, third and fourth central moment are given by

$$
\begin{aligned}
E(x) &= \frac{p}{p+q} \\
V(x) &= \frac{pq}{(p+q)^2(p+q+1)} \\
\mu_3 &= \frac{2pq(q-p)}{(p+q)^3(p+q+1)(p+q+2)} \\
\mu_4 &= \frac{3pq(2(p+q)^2 + pq(p+q-6))}{(p+q)^4(p+q+1)(p+q+2)(p+q+3)}
\end{aligned}
$$

More generally algebraic moments are given in terms of the Beta function by

$$\mu'_k = \frac{B(p+k,q)}{B(p,q)}$$

4.5 Probability Content

In order to find the probability content for a Beta distribution we form the cumulative distribution

$$F(x) = \frac{1}{B(p,q)} \int_0^x t^{p-1}(1-t)^{q-1} dt = \frac{B_x(p,q)}{B(p,q)} = I_x(p,q)$$

where both B_x and I_x seems to be called the incomplete Beta function in the literature.

The incomplete Beta function I_x is connected to the binomial distribution for integer values of a by

$$1 - I_x(a,b) = I_{1-x}(b,a) = (1-x)^{a+b-1} \sum_{i=0}^{a-1} \binom{a+b-1}{i} \left(\frac{x}{1-x}\right)^i$$

or expressed in the opposite direction

$$\sum_{s=a}^{n} \binom{n}{s} p^s (1-p)^{n-s} = I_p(a, n-a+1)$$

Also to the negative binomial distribution there is a connection by the relation

$$\sum_{s=a}^{n} \binom{n+s-1}{s} p^n q^s = I_q(a,n)$$

The incomplete Beta function is also connected to the probability content of Student's t-distribution and the F-distribution. See further section 42.7 for more information on I_x.

4.6 Random Number Generation

In order to obtain random numbers from a Beta distribution we first single out a few special cases.

For $p = 1$ and/or $q = 1$ we may easily solve the equation $F(x) = \xi$ where $F(x)$ is the cumulative function and ξ a uniform random number between zero and one. In these cases

$$\begin{aligned} p = 1 \quad &\Rightarrow \quad x = 1 - \xi^{1/q} \\ q = 1 \quad &\Rightarrow \quad x = \xi^{1/p} \end{aligned}$$

For p and q half-integers we may use the relation to the chi-square distribution by forming the ratio

$$\frac{y_m}{y_m + y_n}$$

with y_m and y_n two independent random numbers from chi-square distributions with $m = 2p$ and $n = 2q$ degrees of freedom, respectively.

Yet another way of obtaining random numbers from a Beta distribution valid when p and q are both integers is to take the ℓ:th out of k ($1 \leq \ell \leq k$) independent uniform random numbers between zero and one. Doing this we obtain a Beta distribution with parameters $p = \ell$ and $q = k + 1 - \ell$. Conversely, if we want to generate random numbers from a Beta distribution with integer parameters p and q we could use this technique with $\ell = p$ and $k = p + q - 1$. This last technique implies that for low integer values of p and q simple code may be used, *e.g.* for $p = 2$ and $q = 1$ we may simply take $\max(\xi_1, \xi_2)$ *i.e.* the maximum of two uniform random numbers.

5 Binomial Distribution

5.1 Introduction

The Binomial distribution is given by

$$p(r;N,p) = \binom{N}{r} p^r (1-p)^{N-r}$$

where the variable r with $0 \le r \le N$ and the parameter N ($N > 0$) are integers and the parameter p ($0 \le p \le 1$) is a real quantity.

The distribution describes the probability of exactly r successes in N trials if the probability of a success in a single trial is p (we sometimes also use $q = 1 - p$, the probability for a failure, for convenience). It was first presented by Jacques Bernoulli in a work which was posthumously published [23].

5.2 Moments

The expectation value, variance, third and fourth moment are given by

$$\begin{aligned}
E(r) &= Np \\
V(r) &= Np(1-p) = Npq \\
\mu_3 &= Np(1-p)(1-2p) = Npq(q-p) \\
\mu_4 &= Np(1-p)\left[1+3p(1-p)(N-2)\right] = Npq\left[1+3pq(N-2)\right]
\end{aligned}$$

Central moments of higher orders may be obtained by the recursive formula

$$\mu_{r+1} = pq\left\{Nr\mu_{r-1} + \frac{\partial \mu_r}{\partial p}\right\}$$

starting with $\mu_0 = 1$ and $\mu_1 = 0$.

The coefficients of skewness and kurtosis are given by

$$\gamma_1 = \frac{q-p}{\sqrt{Npq}} \quad \text{and} \quad \gamma_2 = \frac{1-6pq}{Npq}$$

5.3 Probability Generating Function

The probability generating function is given by

$$G(z) = E(z^r) = \sum_{r=0}^{N} z^r \binom{N}{r} p^r (1-p)^{N-r} = (pz+q)^N$$

and the characteristic function thus by

$$\phi(t) = G(e^{\imath t}) = \left(q + pe^{\imath t}\right)^N$$

5.4 Cumulative Function

For fixed N and p one may easily construct the cumulative function $P(r)$ by a recursive formula, see section on random numbers below.

However, an interesting and useful relation exist between $P(r)$ and the incomplete Beta function I_x namely

$$P(k) = \sum_{r=0}^{k} p(r; N, p) = I_{1-p}(N-k, k+1)$$

For further information on I_x see section 42.7.

5.5 Random Number Generation

In order to achieve random numbers from a binomial distribution we may either

- Generate N uniform random numbers and accumulate the number of such that are less or equal to p, or
- Use the cumulative technique, *i.e.* construct the cumulative, distribution, function and by use of this and one uniform random number obtain the required random number, or
- for larger values of N, say $N > 100$, use an approximation to the normal distribution with mean Np and variance Npq.

Except for very small values of N and very high values of p the cumulative technique is the fastest for numerical calculations. This is especially true if we proceed by constructing the cumulative vector once for all[2] (as opposed to making this at each call) using the recursive formula

$$p(i) = p(i-1)\ \frac{p}{q}\ \frac{N+1-i}{i}$$

for $i = 1, 2, \ldots, N$ starting with $p(0) = q^N$.

However, using the relation given in the previous section with a well optimized code for the incomplete Beta function (see [9] or section 42.7) turns out to be a numerically more stable way of creating the cumulative distribution than a simple loop adding up the individual probabilities.

5.6 Estimation of Parameters

Experimentally the quantity $\frac{r}{N}$, the relative number of successes in N trials, often is of more interest than r itself. This variable has expectation $E(\frac{r}{N}) = p$ and variance $V(\frac{r}{N}) = \frac{pq}{N}$.

The estimated value for p in an experiment giving r successes in N trials is $\hat{p} = \frac{r}{N}$.

[2]This is possible only if we require random numbers from one and the same binomial distribution with fixed values of N and p.

If p is unknown a unbiased estimate of the variance of a binomial distribution is given by

$$V(r) = \frac{N}{N-1} N \left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right) = \frac{N}{N-1} N \hat{p}(1-\hat{p})$$

To find lower and upper confidence levels for p we proceed as follows.

- For lower limits find a p_{low} such that

$$\sum_{r=k}^{N} \binom{N}{r} p_{low}^{r} (1-p_{low})^{N-r} = 1 - \alpha$$

  or expressed in terms of the incomplete Beta function $1 - I_{1-p}(N-k+1, k) = 1 - \alpha$

- for upper limits find a p_{up} such that

$$\sum_{r=0}^{k} \binom{N}{r} p_{up}^{r} (1-p_{up})^{N-r} = 1 - \alpha$$

  which is equivalent to $I_{1-p}(N-k, k+1) = 1 - \alpha$ *i.e.* $I_p(k+1, N-k) = \alpha$.

As an example we take an experiment with $N = 10$ where a certain number of successes $0 \le k \le N$ have been observed. The confidence levels corresponding to 90%, 95%, 99% as well as the levels corresponding to one, two and three standard deviations for a normal distribution (84.13%, 97.72% and 99.87% probability content) are given below.

	Lower confidence levels							Upper confidence levels					
k	-3σ	99%	-2σ	95%	90%	$-\sigma$	$\hat{p}$	$-\sigma$	90%	95%	-2σ	99%	-3σ
0							0.00	0.17	0.21	0.26	0.31	0.37	0.48
1	0.00	0.00	0.00	0.01	0.01	0.02	0.10	0.29	0.34	0.39	0.45	0.50	0.61
2	0.01	0.02	0.02	0.04	0.05	0.07	0.20	0.41	0.45	0.51	0.56	0.61	0.71
3	0.02	0.05	0.06	0.09	0.12	0.14	0.30	0.51	0.55	0.61	0.66	0.70	0.79
4	0.05	0.09	0.12	0.15	0.19	0.22	0.40	0.60	0.65	0.70	0.74	0.78	0.85
5	0.10	0.15	0.18	0.22	0.27	0.30	0.50	0.70	0.73	0.78	0.82	0.85	0.90
6	0.15	0.22	0.26	0.30	0.35	0.40	0.60	0.78	0.81	0.85	0.88	0.91	0.95
7	0.21	0.30	0.34	0.39	0.45	0.49	0.70	0.86	0.88	0.91	0.94	0.95	0.98
8	0.29	0.39	0.44	0.49	0.55	0.59	0.80	0.93	0.95	0.96	0.98	0.98	0.99
9	0.39	0.50	0.55	0.61	0.66	0.71	0.90	0.98	0.99	0.99	1.00	1.00	1.00
10	0.52	0.63	0.69	0.74	0.79	0.83	1.00						

5.7 Probability Content

It is sometimes of interest to judge the significance level of a certain outcome given the hypothesis that $p = \frac{1}{2}$. If N trials are made and we find k successes (let's say $k < N/2$ else use $N - k$ instead pf k) we want to estimate the probability to have k or fewer successes plus the probability for $N - k$ or more successes. Since the assumption is that $p = \frac{1}{2}$ we want the two-tailed probability content.

To calculate this either sum the individual probabilities or use the relation to the incomplete beta function. The former may seem more straightforward but the latter may be computationally easier given a routine for the incomplete beta function. If $k = N/2$ we watch up not to add the central term twice (in this case the requested probability is 100% anyway). In the table below we show such confidence levels in % for values of N ranging from 1 to 20. *E.g.* the probability to observe 3 successes (or failures) or less and 12 failures (or successes) or more for $n = 15$ is 3.52%.

	k										
N	0	1	2	3	4	5	6	7	8	9	10
1	100.00										
2	50.00	100.00									
3	25.00	100.00									
4	12.50	62.50	100.00								
5	6.25	37.50	100.00								
6	3.13	21.88	68.75	100.00							
7	1.56	12.50	45.31	100.00							
8	0.78	7.03	28.91	72.66	100.00						
9	0.39	3.91	17.97	50.78	100.00						
10	0.20	2.15	10.94	34.38	75.39	100.00					
11	0.10	1.17	6.54	22.66	54.88	100.00					
12	0.05	0.63	3.86	14.60	38.77	77.44	100.00				
13	0.02	0.34	2.25	9.23	26.68	58.11	100.00				
14	0.01	0.18	1.29	5.74	17.96	42.40	79.05	100.00			
15	0.01	0.10	0.74	3.52	11.85	30.18	60.72	100.00			
16	0.00	0.05	0.42	2.13	7.68	21.01	45.45	80.36	100.00		
17	0.00	0.03	0.23	1.27	4.90	14.35	33.23	62.91	100.00		
18	0.00	0.01	0.13	0.75	3.09	9.63	23.79	48.07	81.45	100.00	
19	0.00	0.01	0.07	0.44	1.92	6.36	16.71	35.93	64.76	100.00	
20	0.00	0.00	0.04	0.26	1.18	4.14	11.53	26.32	50.34	82.38	100.00

6 Binormal Distribution

6.1 Introduction

As a generalization of the normal or Gauss distribution to two dimensions we define the binormal distribution as

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2-2\rho\cdot\frac{x_1-\mu_1}{\sigma_1}\cdot\frac{x_2-\mu_2}{\sigma_2}\right)}$$

where μ_1 and μ_2 are the expectation values of x_1 and x_2, σ_1 and σ_2 their standard deviations and ρ the correlation coefficient between them. Putting $\rho=0$ we see that the distribution becomes the product of two one-dimensional Gauss distributions.

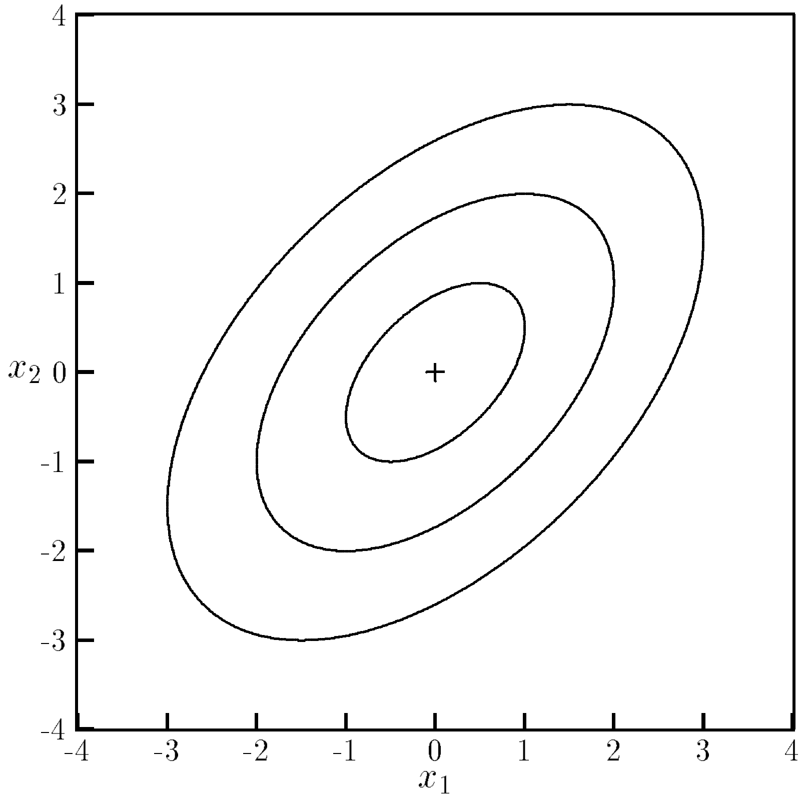


Figure 2: Binormal distribution

In figure 2 we show contours for a standardized Binormal distribution *i.e* putting $\mu_1=\mu_2=0$ and $\sigma_1=\sigma_2=1$ (these parameters are anyway shift- and scale-parameters only). In the example shown $\rho=0.5$. Using standardized variables the contours range from a perfect circle for $\rho=0$ to gradually thinner ellipses in the $\pm 45°$ direction as $\rho\rightarrow\pm 1$. The contours shown correspond to the one, two, and three standard deviation levels. See section on probability content below for details.

6.2 Conditional Probability Density

The conditional density of the binormal distribution is given by

$$
\begin{aligned}
f(x|y) &= f(x,y)/f(y) = \\
&= \frac{1}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}} exp\left\{-\frac{1}{2\sigma_x^2(1-\rho^2)}\left[x - \left(\mu_x + \frac{\rho\sigma_x}{\sigma_y}(y-\mu_y)\right)\right]^2\right\} = \\
&= N\left(\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y), \sigma_x^2(1-\rho^2)\right)
\end{aligned}
$$

which is seen to be a normal distribution which for $\rho = 0$ is, as expected, given by $N(\mu_x, \sigma_x^2)$ but generally has a mean shifted from μ_x and a variance which is smaller than σ_x^2.

6.3 Characteristic Function

The characteristic function of the binormal distribution is given by

$$
\begin{aligned}
\phi(t_1, t_2) &= E(e^{\imath t_1 x_1 + \imath t_2 x_2}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\imath t_1 x_1 + \imath t_2 x_2} f(x_1, x_2) dx_1 dx_2 = \\
&= \exp\left\{\imath t_1 \mu_1 + \imath t_2 \mu_2 + \tfrac{1}{2}\left[(\imath t_1)^2\sigma_1^2 + (\imath t_2)^2\sigma_2^2 + 2(\imath t_1)(\imath t_2)\rho\sigma_1\sigma_2\right]\right\}
\end{aligned}
$$

which shows that if the correlation coefficient ρ is zero then the characteristic function factorizes *i.e.* the variables are independent. This is a unique property of the normal distribution since in general $\rho = 0$ does not imply independence.

6.4 Moments

To find bivariate moments of the binormal distribution the simplest, but still quite tedious, way is to use the characteristic function given above (see section 2.7.3).

Algebraic bivariate moments for the binormal distribution becomes somewhat complicated but normally they are of less interest than the central ones. Algebraic moments of the type μ'_{0k} and μ'_{k0} are, of course, equal to moments of the marginal one-dimensional normal distribution *e.g.* $\mu'_{10} = \mu_2$, $\mu'_{20} = \mu_1^2 + \sigma_1^2$, and $\mu'_{30} = \mu_1(2\sigma_1^2 + \mu_1^2)$ (for μ'_{0k} simply exchange the subscripts on μ and σ). Some other lower order algebraic bivariate moments are given by

$$
\begin{aligned}
\mu'_{11} &= \mu_1\mu_2 + \rho\sigma_1\sigma_2 \\
\mu'_{12} &= 2\rho\sigma_1\sigma_2\mu_2 + \sigma_2^2\mu_1 + \mu_2^2\mu_1 \\
\mu'_{22} &= \sigma_1^2\sigma_2^2 + \sigma_1^2\mu_2^2 + \sigma_2^2\mu_1^2 + \mu_1^2\mu_2^2 + 2\rho^2\sigma_1^2\sigma_2^2 + 4\rho\sigma_1\sigma_2\mu_1\mu_2
\end{aligned}
$$

Beware of the somewhat confusing notation where μ with two subscripts denotes bivariate moments while μ with one subscript denotes expectation values.

Lower order central bivariate moments $\mu_{k\ell}$, arranged in matrix form, are given by

	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
$k=0$	1	0	σ_2^2	0	$3\sigma_2^4$
$k=1$	0	$\rho\sigma_1\sigma_2$	0	$3\rho\sigma_1\sigma_2^3$	0
$k=2$	σ_1^2	0	$\sigma_1^2\sigma_2^2(2\rho^2+1)$	0	$3\sigma_1^2\sigma_2^4(4\rho^2+1)$
$k=3$	0	$3\rho\sigma_1^3\sigma_2$	0	$3\rho\sigma_1^3\sigma_2^3(2\rho^2+3)$	0
$k=4$	$3\sigma_1^4$	0	$3\sigma_1^4\sigma_2^2(4\rho^2+1)$	0	$3\sigma_1^4\sigma_2^4(8\rho^4+24\rho^2+3)$

6.5 Box-Muller Transformation

Recall that if we have a distribution in one set of variables $\{x_1, x_2, ..., x_n\}$ and want to change variables to another set $\{y_1, y_2, ..., y_n\}$ the distribution in the new variables are given by

$$f(y_1, y_2, ..., y_n) = \left\| \begin{matrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{matrix} \right\| f(x_1, x_2, ..., x_n)$$

where the symbol $||J||$ denotes the absolute value of the determinant of the Jacobian J.

Let x_1 and x_2 be two independent stochastic variables from a uniform distribution between zero and one and define

$$\begin{aligned} y_1 &= \sqrt{-2\ln x_1}\sin 2\pi x_2 \\ y_2 &= \sqrt{-2\ln x_1}\cos 2\pi x_2 \end{aligned}$$

Note that with the definition above $-\infty < y_1 < \infty$ and $-\infty < y_2 < \infty$. In order to obtain the joint probability density function in y_1 and y_2 we need to calculate the Jacobian matrix

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{pmatrix}$$

In order to obtain these partial derivatives we express x_1 and x_2 in y_1 and y_2 by rewriting the original equations.

$$\begin{aligned} y_1^2 + y_2^2 &= -2\ln x_1 \\ \frac{y_1}{y_2} &= \tan 2\pi x_2 \end{aligned}$$

which implies

$$\begin{aligned} x_1 &= e^{-\frac{1}{2}(y_1^2+y_2^2)} \\ x_2 &= \frac{1}{2\pi}\arctan\left(\frac{y_1}{y_2}\right) \end{aligned}$$

Then the Jacobian matrix becomes

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{pmatrix} -y_1 e^{-\frac{1}{2}(y_1^2+y_2^2)} & -y_2 e^{-\frac{1}{2}(y_1^2+y_2^2)} \\ \frac{1}{2\pi y_2}\cos^2\arctan\left(\frac{y_1}{y_2}\right) & -\frac{y_1}{2\pi y_2^2}\cos^2\arctan\left(\frac{y_1}{y_2}\right) \end{pmatrix}$$

The distribution $f(y_1, y_2)$ is given by

$$f(y_1, y_2) = \left\|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right\| f(x_1, x_2)$$

where $f(x_1, x_2)$ is the uniform distribution in x_1 and x_2. Now $f(x_1, x_2) = 1$ in the interval $0 \le x_1 \le 1$ and $0 \le x_2 \le 1$ and zero outside this region. and the absolute value of the determinant of the Jacobian is

$$\left\|\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right\| = \frac{1}{2\pi} e^{-\frac{1}{2}(y_1^2+y_2^2)} \left(\frac{y_1^2}{y_2^2} + 1\right) \cos^2 \arctan\left(\frac{y_1}{y_2}\right)$$

but

$$\left(\frac{y_1^2}{y_2^2} + 1\right) \cos^2 \arctan\left(\frac{y_1}{y_2}\right) = (\tan^2 2\pi x_2 + 1) \cos^2 2\pi x_2 = 1$$

and thus

$$f(y_1, y_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(y_1^2+y_2^2)} = \frac{1}{\sqrt{2\pi}} e^{-\frac{y_1^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y_2^2}{2}}$$

i.e. the product of two standard normal distributions.

Thus the result is that y_1 and y_2 are distributed as two independent standard normal variables. This is a well known method, often called the Box-Muller transformation, used in order to achieve pseudorandom numbers from the standard normal distribution given a uniform pseudorandom number generator (see below). The method was introduced by G. E. P. Box and M. E. Muller [24].

6.6 Probability Content

In figure 2 contours corresponding to one, two, and three standard deviations were shown. The projection on each axis for *e.g.* the one standard deviation contour covers the range $-1 \le x_i \le 1$ and contain the probability content of 68.3% which is well known from the one-dimensional case.

More generally, for a contour corresponding to z standard deviations the contour has the equation

$$\frac{(x_1 + x_2)^2}{1 + \rho} + \frac{(x - y)^2}{1 - \rho} = 2z^2$$

i.e. the major and minor semi-axes are $z\sqrt{1+\rho}$ and $z\sqrt{1-\rho}$, respectively. The function value at the contour is given by

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{z^2}{2}\right\}$$

Expressed in polar coordinates (r, ϕ) the contour is described by

$$r^2 = \frac{z^2(1-\rho^2)}{1 - 2\rho \sin\phi \cos\phi}$$

While the projected probability contents follow the usual figures for one-dimensional normal distributions the joint probability content within each ellipse is smaller. For the one, two, and three standard deviation contours the probability content, regardless of the correlation coefficient ρ, inside the ellipse is approximately 39.3%, 86.5%, and 98.9%. If we would like to find the ellipse with a joint probability content of 68.3% we must chose $z \approx 1.5$ (for a content of 95.5% use $z \approx 2.5$ and for 99.7% use $z \approx 3.4$). Se further discussion on probability content for a multinormal distribution in section 28.3.

6.7 Random Number Generation

The joint distribution of y_1 and y_2 in section 6.5 above is a binormal distribution having $\rho = 0$. For arbitrary correlation coefficients ρ the binormal distribution is given by

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\cdot\frac{x_1-\mu_1}{\sigma_1}\cdot\frac{x_2-\mu_2}{\sigma_2}\right)}$$

where μ_1 and μ_2 are the expectation values of x_1 and x_2, σ_1 and σ_2 their standard deviations and ρ the correlation coefficient between them.

Variables distributed according to the binormal distribution may be obtained by transforming the two independent numbers y_1 and y_2 found in the section 6.5 either as

$$\begin{aligned} z_1 &= \mu_1 + \sigma_1\left(y_1\sqrt{1-\rho^2} + y_2\rho\right) \\ z_2 &= \mu_2 + \sigma_2 y_2 \end{aligned}$$

or as

$$\begin{aligned} z_1 &= \mu_1 + \frac{\sigma_1}{\sqrt{2}}\left(y_1\sqrt{1+\rho} + y_2\sqrt{1-\rho}\right) \\ z_2 &= \mu_2 + \frac{\sigma_2}{\sqrt{2}}\left(y_1\sqrt{1+\rho} - y_2\sqrt{1-\rho}\right) \end{aligned}$$

which can be proved by expressing y_1 and y_2 as functions of z_1 and z_2 and evaluate

$$f(z_1, z_2) = \left\|\frac{\partial(y_1, y_2)}{\partial(z_1, z_2)}\right\| f(y_1, y_2) = \left\|\begin{matrix} \frac{\partial y_1}{\partial z_1} & \frac{\partial y_1}{\partial z_2} \\ \frac{\partial y_2}{\partial z_1} & \frac{\partial y_2}{\partial z_2} \end{matrix}\right\| f(y_1, y_2)$$

In the first case

$$\begin{aligned} y_1 &= \frac{1}{\sqrt{1-\rho^2}}\left(\frac{z_1-\mu_1}{\sigma_1} - \rho\frac{z_2-\mu_2}{\sigma_2}\right) \\ y_2 &= \frac{z_2-\mu_2}{\sigma_2} \end{aligned}$$

and in the second case

$$\begin{aligned} y_1 &= \frac{\sqrt{2}}{2\sqrt{1+\rho}}\left(\frac{z_1-\mu_1}{\sigma_1} + \frac{z_2-\mu_2}{\sigma_2}\right) \\ y_2 &= \frac{\sqrt{2}}{2\sqrt{1-\rho}}\left(\frac{z_1-\mu_1}{\sigma_1} - \frac{z_2-\mu_2}{\sigma_2}\right) \end{aligned}$$

In both cases the absolute value of the determinant of the Jacobian is $1/\sigma_1\sigma_2\sqrt{1-\rho^2}$ and we get

$$f(z_1,z_2)=\frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{y_1^2}{2}}\cdot\frac{1}{\sqrt{2\pi}}e^{-\frac{y_2^2}{2}}=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot e^{-\frac{1}{2}(y_1^2+y_2^2)}$$

Inserting the relations expressing y_1 and y_2 in z_1 and z_2 in the exponent we finally obtain the binormal distribution in both cases.

Thus we have found methods which given two independent uniform pseudorandom numbers between zero and one supplies us with a pair of numbers from a binormal distribution with arbitrary means, standard deviations and correlation coefficient.

7 Cauchy Distribution

7.1 Introduction

The Cauchy distribution is given by

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$$

and is defined for $-\infty < x < \infty$. It is a symmetric unimodal distribution as is shown in figure 3.

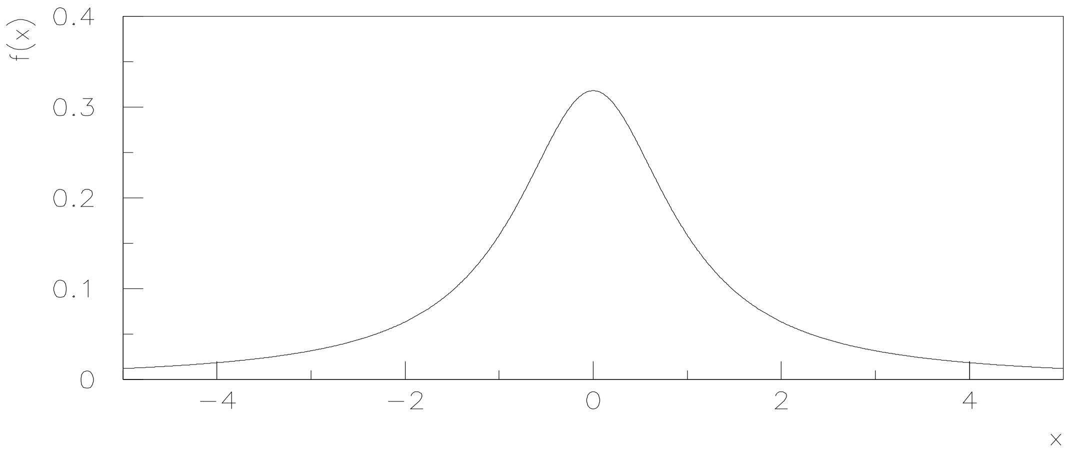


Figure 3: Graph of the Cauchy distribution

The distribution is named after the famous french mathematician Augustin Louis Cauchy (1789-1857) who was a professor at École Polytechnique in Paris from 1816. He was one of the most productive mathematicians which have ever existed.

7.2 Moments

This probability density function is peculiar inasmuch as it has undefined expectation value and all higher moments diverge. For the expectation value the integral

$$E(x) = \frac{1}{\pi} \int\limits_{-\infty}^{\infty} \frac{x}{1+x^2} dx$$

is not completely convergent, *i.e.*

$$\lim_{a\to\infty, b\to\infty} \frac{1}{\pi} \int\limits_{-a}^{b} \frac{x}{1+x^2} dx$$

does not exist. However, the principal value

$$\lim_{a\to\infty} \frac{1}{\pi} \int_{-a}^{a} \frac{x}{1+x^2} dx$$

does exist and is equal to zero. Anyway the convention is to regard the expectation value of the Cauchy distribution as undefined.

Other measures of location and dispersion which are useful in the case of the Cauchy distribution is the *median* and the *mode* which are at $x = 0$ and the *half-width at half-maximum* which is 1 (half-maxima at $x = \pm 1$).

7.3 Normalization

In spite of the somewhat awkward property of not having any moments the distribution at least fulfil the normalization requirement for a proper probability density function *i.e.*

$$\mathcal{N} = \int_{-\infty}^{\infty} f(x)dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \frac{1}{1+\tan^2\phi} \cdot \frac{d\phi}{\cos^2\phi} = 1$$

where we have made the substitution $\tan\phi = x$ in order to simplify the integration.

7.4 Characteristic Function

The characteristic function for the Cauchy distribution is given by

$$\begin{aligned}
\phi(t) &= \int_{-\infty}^{\infty} e^{\imath tx} f(x)dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\cos tx + \imath \sin tx}{1+x^2} dx = \\
&= \frac{1}{\pi} \left(\int_{0}^{\infty} \frac{\cos tx}{1+x^2} dx + \int_{-\infty}^{0} \frac{\cos tx}{1+x^2} dx + \int_{0}^{\infty} \frac{\imath \sin tx}{1+x^2} dx + \int_{-\infty}^{0} \frac{\imath \sin tx}{1+x^2} dx \right) = \\
&= \frac{2}{\pi} \int_{0}^{\infty} \frac{\cos tx}{1+x^2} dx = e^{-|t|}
\end{aligned}$$

where we have used that the two sine integrals are equal but with opposite sign whereas the two cosine integrals are equal. The final integral we have taken from standard integral tables. Note that the characteristic function has no derivatives at $t = 0$ once again telling us that the distribution has no moments.

7.5 Location and Scale Parameters

In the form given above the Cauchy distribution has no parameters. It is useful, however, to introduce location (x_0) and scale ($\Gamma > 0$) parameters writing

$$f(x; x_0, \Gamma) = \frac{1}{\pi} \cdot \frac{\Gamma}{\Gamma^2 + (x - x_0)^2}$$

where x_0 is the *mode* of the distribution and Γ the *half-width at half-maximum (HWHM)*. Including these two parameters the characteristic function is modified to

$$\phi(t) = e^{itx_0 - \Gamma|t|}$$

7.6 Breit-Wigner Distribution

In this last form we recognize the Breit-Wigner formula, named after the two physicists Gregory Breit and Eugene Wigner, which arises in physics *e.g.* in the description of the cross section dependence on energy (mass) for two-body resonance scattering. Resonances like *e.g.* the Δ^{++} in π^+p scattering or the ρ in $\pi\pi$ scattering can be quite well described in terms of the Cauchy distribution. This is the reason why the Cauchy distribution in physics often is referred to as the Breit-Wigner distribution. However, in more elaborate physics calculations the width may be energy-dependent in which case things become more complicated.

7.7 Comparison to Other Distributions

The Cauchy distribution is often compared to the normal (or Gaussian) distribution with mean μ and standard deviation $\sigma > 0$

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

and the double-exponential distribution with mean μ and slope parameter $\lambda > 0$

$$f(x;\mu,\lambda) = \frac{\lambda}{2} e^{-\lambda|x-\mu|}$$

These are also examples of symmetric unimodal distributions. The Cauchy distribution has longer tails than the double-exponential distribution which in turn has longer tails than the normal distribution. In figure 4 we compare the Cauchy distribution with the standard normal ($\mu = 0$ and $\sigma = 1$) and the double-exponential distributions ($\lambda = 1$) for $x > 0$.

The normal and double-exponential distributions have well defined moments. Since they are symmetric all central moments of odd order vanish while central moments of even order are given by $\mu_{2n} = (2n)!\sigma^{2n}/2^n n!$ (for $n \geq 0$) for the normal and by $\mu_n = n!/\lambda^n$ (for even n) for the double-exponential distribution. *E.g.* the variances are σ^2 and $2/\lambda^2$ and the fourth central moments $3\sigma^4$ and $24/\lambda^4$, respectively.

The Cauchy distribution is related to Student's t-distribution with n degrees of freedom (with n a positive integer)

$$f(t;n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} = \frac{\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n}\,B\left(\frac{1}{2},\frac{n}{2}\right)}$$

where $\Gamma(x)$ is the Euler gamma-function not no be mixed up with the width parameter for the Cauchy distribution used elsewhere in this section. B is the beta-function defined in

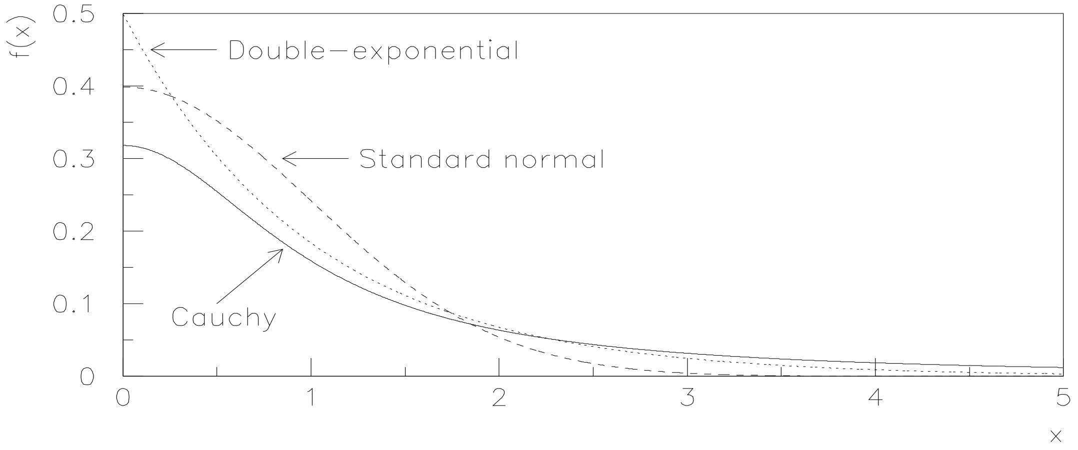


Figure 4: Comparison between the Cauchy distribution, the standard normal distribution, and the double-exponential distribution

terms of the Γ-function as $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$. As can be seen the Cauchy distribution arises as the special case where $n = 1$. If we change variable to $x = t/\sqrt{n}$ and put $m = \frac{n+1}{2}$ the Student's t-distribution becomes

$$f(x;m) = \frac{k}{(1+x^2)^m} \quad \text{with} \quad k = \frac{\Gamma(m)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(m-\frac{1}{2}\right)} = \frac{1}{B\left(m-\frac{1}{2},\frac{1}{2}\right)}$$

where k is simply a normalization constant. Here it is easier to see the more general form of this distribution which for $m = 1$ gives the Cauchy distribution. The requirement $n \geq 1$ corresponds to m being a half-integer ≥ 1 but we could even allow for m being a real number.

As for the Cauchy distribution the Student's t-distribution have problems with divergent moments and moments of order $\geq n$ does not exist. Below this limit odd central moments are zero (the distribution is symmetric) and even central moments are given by

$$\mu_{2r} = n^r \frac{\Gamma\left(r+\frac{1}{2}\right)\Gamma\left(\frac{n}{2}-r\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} = n^r \frac{B\left(r+\frac{1}{2},\frac{n}{2}-r\right)}{B\left(\frac{1}{2},\frac{n}{2}\right)}$$

for r a positive integer ($2r < n$). More specifically the expectation value is $E(t) = 0$, the variance $V(t) = \frac{n}{n-2}$ and the fourth central moment is given by $\mu_4 = \frac{3n^2}{(n-2)(n-4)}$ when they exist. As $n \to \infty$ the Student's t-distribution approaches a standard normal distribution.

7.8 Truncation

In order to avoid the long tails of the distribution one sometimes introduces a truncation. This, of course, also cures the problem with the undefined mean and divergent higher moments. For a symmetric truncation $-X \leq x \leq X$ we obtain the renormalized probability density function

$$f(x) = \frac{1}{2\arctan X} \cdot \frac{1}{1+x^2}$$

which has expectation value $E(x) = 0$, variance $V(x) = \frac{X}{\arctan X} - 1$, third central moment $\mu_3 = 0$ and fourth central moment $\mu_4 = \frac{X}{\arctan X}\left(\frac{X^2}{3} - 1\right) + 1$. The fraction of the original Cauchy distribution within the symmetric interval is $f = \frac{2}{\pi}\arctan X$. We will, however, not make any truncation of the Cauchy distribution in the considerations made in this note.

7.9 Sum and Average of Cauchy Variables

In most cases one would expect the sum and average of many variables drawn from the same population to approach a normal distribution. This follows from the famous Central Limit Theorem. However, due to the divergent variance of the Cauchy distribution the requirements for this theorem to hold is not fulfilled and thus this is not the case here. We define

$$S_n = \sum_{i=1}^{n} x_i \quad \text{and} \quad \overline{S}_n = \frac{1}{n} S_n$$

with x_i independent variables from a Cauchy distribution.

The characteristic function of a sum of independent random variables is equal to the product of the individual characteristic functions and hence

$$\Phi(t) = \phi(t)^n = e^{-n|t|}$$

for S_n. Turning this into a probability density function we get (putting $x = S_n$ for convenience)

$$\begin{aligned}
f(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi(t) e^{-\imath xt} dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-(\imath xt + n|t|)} dt = \frac{1}{2\pi}\left(\int_{-\infty}^{0} e^{nt - \imath xt} dt + \int_{0}^{\infty} e^{-\imath xt - nt} dt\right) = \\
&= \frac{1}{2\pi}\left(\left[\frac{e^{t(n-\imath x)}}{n - \imath x}\right]_{-\infty}^{0} + \left[\frac{e^{-t(\imath x + n)}}{-n - \imath x}\right]_{0}^{\infty}\right) = \frac{1}{2\pi}\left(\frac{1}{n - \imath x} + \frac{1}{n + \imath x}\right) = \frac{1}{\pi} \cdot \frac{n}{n^2 + x^2}
\end{aligned}$$

This we recognize as a Cauchy distribution with scale parameter $\Gamma = n$ and thus for each additional Cauchy variable the HWHM increases by one unit.

Moreover, the probability density function of $\overline{S}_n$ is given by

$$f(\overline{S}_n) = \frac{dS_n}{d\overline{S}_n} f(S_n) = \frac{1}{\pi} \cdot \frac{1}{1 + \overline{S}_n^2}$$

i.e. the somewhat amazing result is that the average of any number of independent random variables from a Cauchy distribution is also distributed according to the Cauchy distribution.

7.10 Estimation of the Median

For the Cauchy distribution the sample mean is not a consistent estimator of the median of the distribution. In fact, as we saw in the previous section, the sample mean is itself distributed according to the Cauchy distribution and therefore has divergent variance. However, the sample median for a sample of n independent observations from a Cauchy distribution is a consistent estimator of the true median.

In the table below we give the expectations and variances of the sample mean and sample median estimators for the normal, double-exponential and Cauchy distributions (see above for definitions of distributions). Sorting all the observations the median is taken as the value for the central observation for odd n and as the average of the two central values for even n. The variance of the sample mean is simply the variance of the distribution divided by the sample size n. For large n the variance of the sample median m is given by $V(m) = 1/4nf^2$ where f is the function value at the median.

Distribution	$E(\overline{x})$	$V(\overline{x})$	$E(m)$	$V(m)$
Normal	μ	$\frac{\sigma^2}{n}$	μ	$\frac{\pi\sigma^2}{2n}$
Double-exponential	μ	$\frac{2}{n\lambda^2}$	μ	$\frac{1}{n\lambda^2}$
Cauchy	undef.	∞	x_0	$\frac{\pi^2\Gamma^2}{4n}$

For a normal distribution the sample mean is superior to the median as an estimator of the mean (*i.e.* it has the smaller variance). However, the double-exponential distribution is an example of a distribution where the sample median is the best estimator of the mean of the distribution. In the case of the Cauchy distribution only the median works of the above alternatives but even better is a proper Maximum Likelihood estimator. In the case of the normal and double-exponential the mean and median, respectively, are identical to the maximum likelihood estimators but for the Cauchy distribution such an estimator may not be expressed in a simple way.

The large n approximation for the variance of the sample median gives conservative estimates for lower values of n in the case of the normal distribution. Beware, however, that for the Cauchy and the double-exponential distributions it is *not* conservative but gives too small values. Calculating the standard deviation this is within 10% of the true value already at $n = 5$ for the normal distribution whereas for the Cauchy distribution this is true at about $n = 20$ and for the double-exponential distribution only at about $n = 60$.

7.11 Estimation of the HWHM

To find an estimator for the half-width at half-maximum is not trivial. It implies binning the data, finding the maximum and then locating the positions where the curve is at half-maximum. Often it is preferable to fit the probability density function to the observations in such a case.

However, it turns out that another measure of dispersion the so called *semi-interquartile range* can be used as an estimator. The semi-interquartile range is defined as half the difference between the upper and the lower quartiles. The quartiles are the values which divide the probability density function into four parts with equal probability content, *i.e.* 25% each. The second quartile is thus identical to the median. The definition is thus

$$\mathcal{S} = \frac{1}{2}(Q_3 - Q_1) = \Gamma$$

where Q_1 is the lower and Q_3 the upper quartile which for the Cauchy distribution is equal to $x_0 - \Gamma$ and $x_0 + \Gamma$, respectively. As is seen $\mathcal{S} = HWHM = \Gamma$ and thus this estimator may be used in order to estimate Γ.

We gave above the large n approximation for the variance of the median. The median and the quartiles are examples of the more general concept *quantiles*. Generally the large n approximation for the variance of a quantile $\mathcal{Q}$ is given by $V(\mathcal{Q}) = pq/nf^2$ where f is the ordinate at the quantile and p and $q = 1 - p$ are the probability contents above and below the quantile, respectively. The covariance between two quantiles $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is, with similar notations, given by $Cov(\mathcal{Q}_1, \mathcal{Q}_2) = p_2 q_1 / n f_1 f_2$ where $\mathcal{Q}_1$ should be the leftmost quantile.

For large n the variance of the semi-interquartile range for a sample of size n is thus found by error propagation inserting the formulæ above

$$V(\mathcal{S}) = \frac{1}{4}\left(V(Q_1) + V(Q_3) - 2Cov(Q_1, Q_3)\right) = \frac{1}{64n}\left(\frac{3}{f_1^2} + \frac{3}{f_3^2} - \frac{2}{f_1 f_3}\right) = \frac{1}{16nf_1^2} = \frac{\pi^2\Gamma^2}{4n}$$

where f_1 and f_3 are the function values at the lower and upper quartile which are both equal to $1/2\pi\Gamma$. This turns out to be exactly the same as the variance we found for the median in the previous section.

After sorting the sample the quartiles are determined by extrapolation between the two observations closest to the quartile. In the case where $n + 2$ is a multiple of 4 *i.e.* the series $n = 2, 6, 10...$ the lower quartile is exactly at the $\frac{n+2}{4}$:th observation and the upper quartile at the $\frac{3n+2}{4}$:th observation. In the table below we give the expectations and variances of the estimator of $\mathcal{S}$ as well as the variance estimator s^2 for the normal, double-exponential and Cauchy distributions. The variance estimator s^2 and its variance are given by

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 \quad \text{and} \quad V(s^2) = \frac{\mu_4 - \mu_2^2}{n} + \frac{2\mu_2^2}{n(n-1)}$$

with μ_2 and μ_4 the second and fourth central moments. The expectation value of s^2 is equal to the variance and thus it is a unbiased estimator.

Distribution	HWHM	$E(s^2)$	$V(s^2)$	$E(\mathcal{S})$	$V(\mathcal{S})$
Normal	$\sigma\sqrt{2\ln 2}$	σ^2	$\frac{2\sigma^4}{n-1}$	0.6745σ	$\frac{1}{16nf(Q_1)^2}$
Double-exponential	$\frac{\ln 2}{\lambda}$	$\frac{2}{\lambda^2}$	$\frac{20}{n\lambda^4}\alpha$	$\frac{\ln 2}{\lambda}$	$\frac{1}{n\lambda^2}$
Cauchy	Γ	∞	∞	Γ	$\frac{\pi^2\Gamma^2}{4n}$

In this table $\alpha = 1 + \frac{0.4}{n-1}$ if we include the second term in the expression of $V(s^2)$ above and $\alpha = 1$ otherwise. It can be seen that the double-exponential distribution also has $HWHM = \mathcal{S}$ but for the normal distribution $HWHM \approx 1.1774\sigma$ as compared to $\mathcal{S} \approx 0.6745\sigma$.

For the three distributions tested the semi-interquartile range estimator is biased. In the case of the normal distribution the values are approaching the true value from below while for the Cauchy and double-exponential distributions from above. The large n approximation for $V(\mathcal{S})$ is conservative for the normal and double-exponential distribution but *not* conservative for the Cauchy distribution. In the latter case the standard deviation of $\mathcal{S}$ is within 10% of the true value for $n > 50$ but for small values of n it is substantially larger than given by the formula. The estimated value for Γ is less than 10% too big for $n > 25$.

7.12 Random Number Generation

In order to generate pseudorandom numbers from a Cauchy distribution we may solve the equation $F(x) = \xi$ where $F(x)$ is the cumulative distribution function and ξ is a uniform pseudorandom number between 0 and 1. This means solving for x in the equation

$$F(x) = \frac{\Gamma}{\pi} \int_{-\infty}^{x} \frac{1}{\Gamma^2 + (t - x_0)^2} dt = \xi$$

If we make the substitution $\tan\phi = (t - x_0)/\Gamma$ using that $d\phi/\cos^2\phi = dt/\Gamma$ we obtain

$$\frac{1}{\pi} \arctan\left(\frac{x - x_0}{\Gamma}\right) + \tfrac{1}{2} = \xi$$

which finally gives

$$x = x_0 + \Gamma \tan\left(\pi\left(\xi - \tfrac{1}{2}\right)\right)$$

as a pseudorandom number from a Cauchy distribution. One may easily see that it is equivalent to use

$$x = x_0 + \Gamma \tan(2\pi\xi)$$

which is a somewhat simpler expression.

An alternative method (see also below) to achieve random numbers from a Cauchy distribution would be to use

$$x = x_0 + \Gamma \frac{z_1}{z_2}$$

where z_1 and z_2 are two independent random numbers from a standard normal distribution. However, if the standard normal random numbers are achieved through the Box-Muller transformation then $z_1/z_2 = \tan 2\pi\xi$ and we are back to the previous method.

In generating pseudorandom numbers one may, if profitable, avoid the tangent by

a Generate in u and v two random numbers from a uniform distribution between -1 and 1.

b If $u^2 + v^2 > 1$ (outside circle with radius one in uv-plane) go back to **a**.

c Obtain $x = x_0 + \Gamma \frac{u}{v}$ as a random number from a Cauchy distribution.

7.13 Physical Picture

A physical picture giving rise to the Cauchy distribution is as follows: Regard a plane in which there is a point source which emits particles isotropically in the plane (either in the full 2π region or in one hemisphere π radians wide). The source is at the x-coordinate x_0 and the particles are detected in a detector extending along a line Γ length units from the source. This scenario is depicted in figure 5

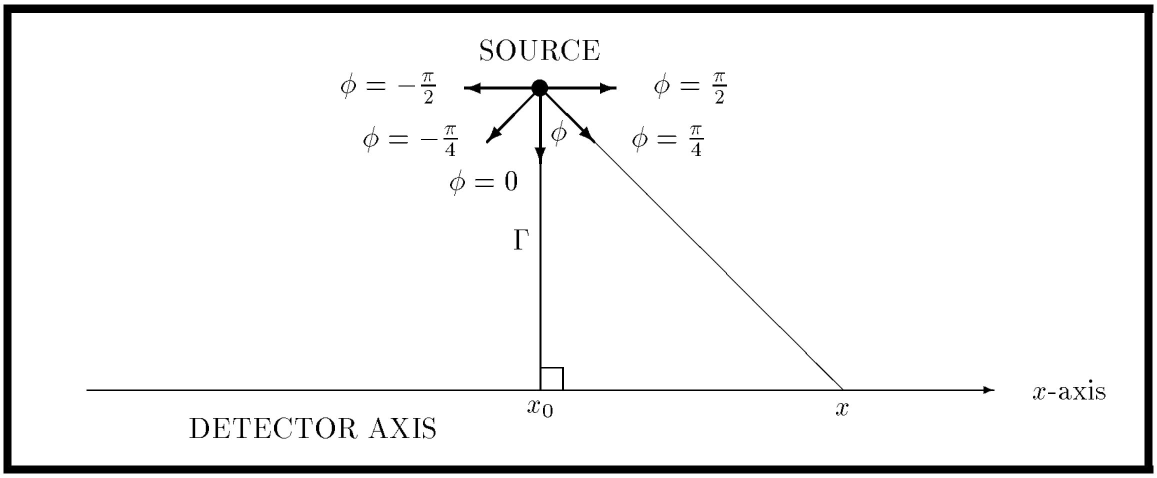


Figure 5: Physical scenario leading to a Cauchy distribution

The distribution in the variable x along the detector will then follow the Cauchy distribution. As can be seen by pure geometrical considerations this is in accordance with the result above where pseudorandom numbers from a Cauchy distribution could be obtained by $x = x_0 + \Gamma \tan\phi$, *i.e.* $\tan\phi = \frac{x - x_0}{\Gamma}$, with ϕ uniformly distributed between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$.

To prove this let us start with the distribution in ϕ

$$f(\phi) = \frac{1}{\pi} \quad \text{for} \quad -\frac{\pi}{2} \le \phi \le \frac{\pi}{2}$$

To change variables from ϕ to x requires the derivative $d\phi/dx$ which is given by

$$\frac{d\phi}{dx} = \frac{\cos^2\phi}{\Gamma} = \frac{1}{\Gamma}\cos^2\arctan\left(\frac{x-x_0}{\Gamma}\right)$$

Note that the interval from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ in ϕ maps onto the interval $-\infty < x < \infty$. We get

$$\begin{aligned} f(x) &= \left|\frac{d\phi}{dx}\right| f(\phi) = \frac{1}{\pi\Gamma}\cos^2\arctan\left(\frac{x-x_0}{\Gamma}\right) = \frac{1}{\pi\Gamma}\cos^2\phi = \\ &= \frac{1}{\pi\Gamma}\cdot\frac{\Gamma^2}{\Gamma^2+(x-x_0)^2} = \frac{1}{\pi}\cdot\frac{\Gamma}{\Gamma^2+(x-x_0)^2} \end{aligned}$$

i.e. the Cauchy distribution.

It is just as easy to make the proof in the reversed direction, *i.e.* given a Cauchy distribution in x one may show that the ϕ-distribution is uniform between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$.

7.14 Ratio Between Two Standard Normal Variables

As mentioned above the Cauchy distribution also arises if we take the ratio between two standard normal variables z_1 and z_2, *viz.*

$$x = x_0 + \Gamma\frac{z_1}{z_2}\ .$$

In order to deduce the distribution in x we first introduce a dummy variable y which we simply take as z_2 itself. We then make a change of variables from z_1 and z_2 to x and y. The transformation is given by

$$\begin{aligned} x &= x_0 + \Gamma\frac{z_1}{z_2} \\ y &= z_2 \end{aligned}$$

or if we express z_1 and z_2 in x and y

$$\begin{aligned} z_1 &= y(x-x_0)/\Gamma \\ z_2 &= y \end{aligned}$$

The distribution in x and y is given by

$$f(x,y) = \left|\left|\frac{\partial(z_1,z_2)}{\partial(x,y)}\right|\right| f(z_1,z_2)$$

where the absolute value of the determinant of the Jacobian is equal to y/Γ and $f(z_1,z_2)$ is the product of two independent standard normal distributions. We get

$$f(x,y) = \frac{y}{\Gamma}\cdot\frac{1}{2\pi}e^{-\frac{1}{2}(z_1^2+z_2^2)} = \frac{y}{2\pi\Gamma}e^{-\frac{1}{2}\left(\frac{y^2(x-x_0)^2}{\Gamma^2}+y^2\right)}$$

In order to obtain the marginal distribution in x we integrate over y

$$f(x) = \int_{-\infty}^{\infty} f(x,y)dy = \frac{1}{2\pi\Gamma} \int_{-\infty}^{\infty} y e^{-\alpha y^2} dy$$

where we have put

$$\alpha = \frac{1}{2}\left(\left(\frac{x-x_0}{\Gamma}\right)^2 + 1\right)$$

for convenience. If we make the substitution $z = y^2$ we get

$$f(x) = 2\frac{1}{2\pi\Gamma} \int_{0}^{\infty} e^{-\alpha z} \frac{dz}{2} = \frac{1}{2\pi\Gamma\alpha}$$

Note that the first factor of 2 comes from the fact that the region $-\infty < y < \infty$ maps twice onto the region $0 < z < \infty$. Finally

$$f(x) = \frac{1}{2\pi\Gamma\alpha} = \frac{1}{2\pi\Gamma} \cdot \frac{2}{\left(\frac{x-x_0}{\Gamma}\right)^2 + 1} = \frac{1}{\pi} \cdot \frac{\Gamma}{(x-x_0)^2 + \Gamma^2}$$

i.e. a Cauchy distribution.

8 Chi-square Distribution

8.1 Introduction

The chi-square distribution is given by

$$f(x;n) = \frac{\left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2\Gamma\left(\frac{n}{2}\right)}$$

where the variable $x \geq 0$ and the parameter n, the number of degrees of freedom, is a positive integer. In figure 6 the distribution is shown for n-values of 1, 2, 5 and 10. For $n \geq 2$ the distribution has a maximum at $n-2$.

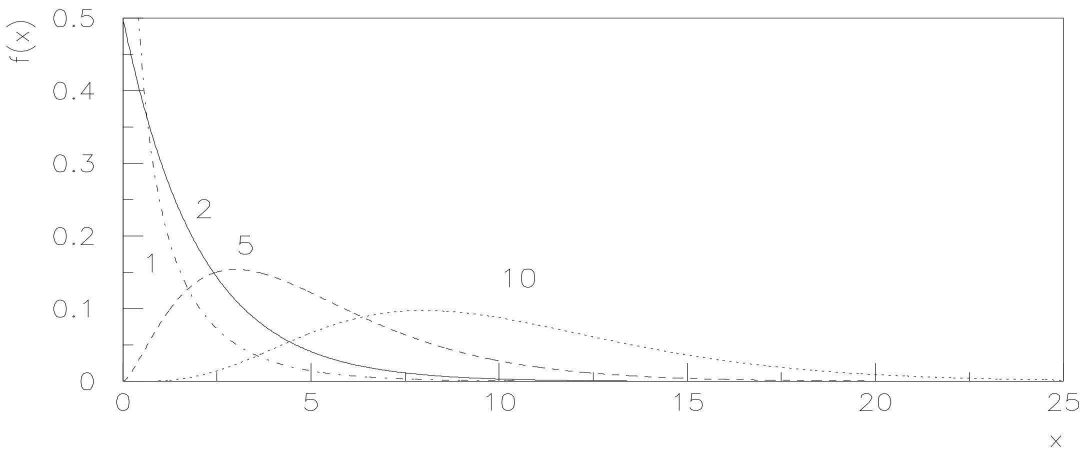


Figure 6: Graph of chi-square distribution for some values of n

8.2 Moments

Algebraic moments of order k are given by

$$\begin{aligned}\mu'_k &= E(x^k) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)} \int_0^\infty x^k \left(\frac{x}{n}\right)^{\frac{n}{2}-1} e^{-\frac{x^2}{2}} dx = \frac{2^k}{\Gamma\left(\frac{n}{2}\right)} \int_0^\infty y^{\frac{n}{2}-1+k} e^{-y} dy = \frac{2^k \Gamma\left(\frac{n}{2}+k\right)}{\Gamma\left(\frac{n}{2}\right)} = \\ &= 2^k \cdot \frac{n}{2}(\frac{n}{2}+1)\cdots(\frac{n}{2}+k-2)(\frac{n}{2}+k-1) = n(n+2)(n+4)\cdots(n+2k-2)\end{aligned}$$

e.g. the first algebraic moment which is the expectation value is equal to n. A recursive formula to calculate algebraic moments is thus given by

$$\mu'_k = \mu'_{k-1} \cdot (n+2k-2)$$

where we may start with $\mu'_0 = 1$ to find the expectation value $\mu'_1 = n$, $\mu'_2 = n(n+1)$ etc.

From this we may calculate the central moments which for the lowest orders become $\mu_2 = 2n$, $\mu_3 = 8n$, $\mu_4 = 12n(n+4)$, $\mu_5 = 32n(5n+12)$ and $\mu_6 = 40n(3n^2+52n+96)$. The coefficients of skewness and kurtosis thus becomes $\gamma_1 = 2\sqrt{2/n}$ and $\gamma_2 = 12/n$.

The fact that the expectation value of a chi-square distribution equals the number of degrees of freedom has led to a bad habit to give the ratio between a found chi-square value and the number of degrees of freedom. This is, however, not a very good variable and it may be misleading. We strongly recommend that one always should give both the chi-square value and degrees of freedom *e.g.* as χ^2/n.d.f.=9.7/5.

To judge the quality of the fit we want a better measure. Since the exact sampling distribution is known one should stick to the chi-square probability as calculated from an integral of the tail *i.e.* given a specific chi-square value for a certain number of degrees of freedom we integrate from this value to infinity (see below).

As an illustration we show in figure 7 the chi-square probability for constant ratios of χ^2/n.d.f.

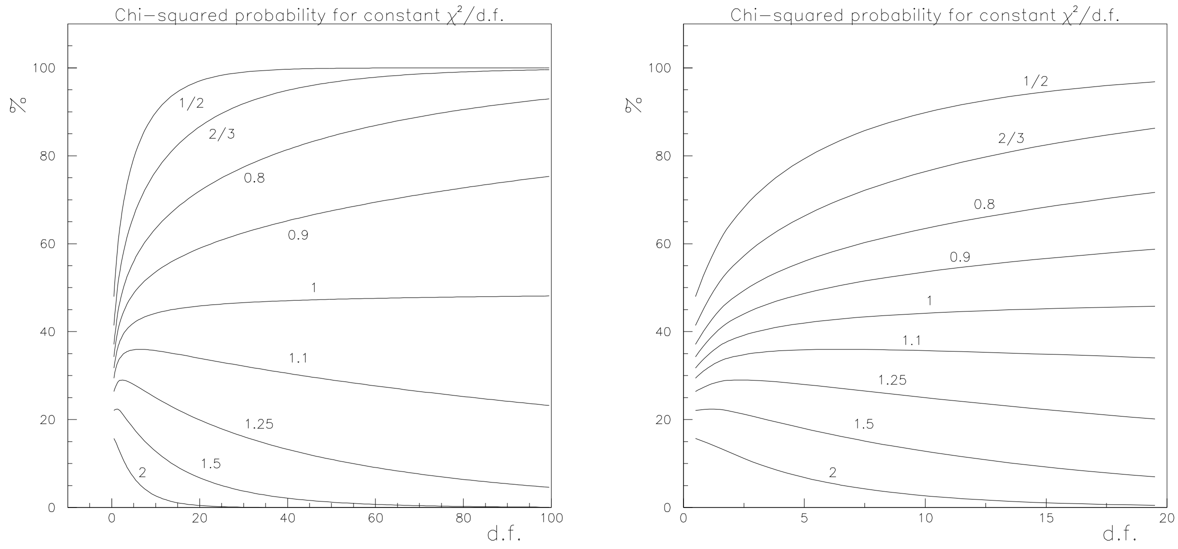


Figure 7: Chi-square probability for constant ratios of χ^2/n.d.f.

Note *e.g.* that for few degrees of freedom we may have an acceptable chi-square value even for larger ratios.

8.3 Characteristic Function

The characteristic function for a chi-square distribution with n degrees of freedom is given by

$$\begin{aligned}\phi(t) &= E(e^{\imath tx}) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)}\int\limits_0^\infty \left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-(\frac{1}{2}-\imath t)x}dx = \frac{1}{2\Gamma\left(\frac{n}{2}\right)}\int\limits_0^\infty \left(\frac{y}{1-2\imath t}\right)^{\frac{n}{2}-1} e^{-y}\frac{dy}{\frac{1}{2}-\imath t} = \\ &= \frac{1}{\Gamma\left(\frac{n}{2}\right)(1-2\imath t)^{\frac{n}{2}}}\int\limits_0^\infty y^{\frac{n}{2}-1}e^{-y}dy = (1-2\imath t)^{-\frac{n}{2}}\end{aligned}$$

8.4 Cumulative Function

The cumulative, or distribution, function for a chi-square distribution with n degrees of freedom is given by

$$\begin{aligned}F(x) &= \frac{1}{2\Gamma\left(\frac{n}{2}\right)}\int\limits_0^x \left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}}dx = \frac{1}{2\Gamma\left(\frac{n}{2}\right)}\int\limits_0^{\frac{x}{2}} y^{\frac{n}{2}-1}e^{-y}2dy = \\ &= \frac{\gamma\left(\frac{n}{2},\frac{x}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = P\left(\frac{n}{2},\frac{x}{2}\right)\end{aligned}$$

where $P\left(\frac{n}{2},\frac{x}{2}\right)$ is the incomplete Gamma function (see section 42.5). In this calculation we have made the simple substitution $y = x/2$ in simplifying the integral.

8.5 Origin of the Chi-square Distribution

If $z_1, z_2, ..., z_n$ are n independent standard normal random variables then $\sum\limits_{i=1}^{n} z_i^2$ is distributed as a chi-square variable with n degrees of freedom.

In order to prove this first regard the characteristic function for the square of a standard normal variable

$$E(e^{\imath tz^2}) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{z^2}{2}(1-2\imath t)}dz = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\frac{dy}{\sqrt{1-2\imath t}} = \frac{1}{\sqrt{1-2\imath t}}$$

where we made the substitution $y = z\sqrt{1-2\imath t}$.

For a sum of n such independent variables the characteristic function is then given by

$$\phi(t) = (1-2it)^{-\frac{n}{2}}$$

which we recognize as the characteristic function for a chi-square distribution with n degrees of freedom.

This property implies that if x and y are independently distributed according to the chi-square distribution with n and m degrees of freedom, respectively, then $x+y$ is distributed as a chi-square variable with $m+n$ degrees of freedom.

Indeed the requirement that all z's come from a standard normal distribution is more than what is needed. The result is the same if all observations x_i come from different normal populations with means μ_i and variance σ_i^2 if we in each case calculate a standardized variable by subtracting the mean and dividing with the standard deviation *i.e.* taking $z_i = (x_i - \mu_i)/\sigma_i$.

8.6 Approximations

For large number of degrees of freedom n the chi-square distribution may be approximated by a normal distribution. There are at least three different approximations. Firstly we may naïvely construct a standardized variable

$$z_1 = \frac{x - E(x)}{\sqrt{V(x)}} = \frac{x-n}{\sqrt{2n}}$$

which would tend to normality as n increases. Secondly an approximation, due to R. A. Fisher, is that the quantity

$$z_2 = \sqrt{2x} - \sqrt{2n-1}$$

approaches a standard normal distribution faster than the standardized variable. Thirdly a transformation, due to E. B. Wilson and M. M. Hilferty, is that the cubic root of x/n is closely distributed as a standard normal distribution using

$$z_3 = \frac{\left(\frac{x}{n}\right)^{\frac{1}{3}} - \left(1 - \frac{2}{9n}\right)}{\sqrt{\frac{2}{9n}}}$$

The second approximation is probably the most well known but the latter is approaching normality even faster. In fact there are even correction factors which may be applied to z_3 to give an even more accurate approximation (see *e.g.* [25])

$$z_4 = z_3 + h_n = z_3 + \frac{60}{n} h_{60}$$

with h_{60} given for values of z_2 from −3.5 to 3.5 in steps of 0.5 (in this order the values of h_{60} are −0.0118, −0.0067, −0.0033, −0.0010, 0.0001, 0.0006, 0.0006, 0.0002, −0.0003, −0.0006, −0.0005, 0.0002, 0.0017, 0.0043, and 0.0082).

To compare the quality of all these approximations we calculate the maximum deviation between the cumulative function for the true chi-square distribution and each of these approximations for $n = 30$ and $n = 100$. The results are shown in the table below. Normally one accepts z_2 for $n > 100$ while z_3, and certainly z_4, are even better already for $n > 30$.

Approximation	$n = 30$	$n = 100$
z_1	0.034	0.019
z_2	0.0085	0.0047
z_3	0.00039	0.00011
z_4	0.000044	0.000035

8.7 Random Number Generation

As we saw above the sum of n independent standard normal random variables gave a chi-square distribution with n degrees of freedom. This may be used as a technique to produce pseudorandom numbers from a chi-square distribution. This required a generator for standard normal random numbers and may be quite slow. However, if we make use of the Box-Muller transformation in order to obtain the standard normal random numbers we may simplify the calculations.

First we recall the Box-Muller transformation which given two pseudorandom numbers uniformly distributed between zero and one through the transformation

$$\begin{aligned} z_1 &= \sqrt{-2\ln\xi_1}\cos 2\pi\xi_2 \\ z_2 &= \sqrt{-2\ln\xi_1}\sin 2\pi\xi_2 \end{aligned}$$

gives, in z_1 and z_2, two independent pseudorandom numbers from a standard normal distribution.

Adding n such squared random numbers implies that

$$\begin{aligned} y_{2k} &= -2\ln(\xi_1 \cdot \xi_2 \cdots \xi_k) \\ y_{2k+1} &= -2\ln(\xi_1 \cdot \xi_2 \cdots \xi_k) - 2\ln\xi_{k+1}\cos^2 2\pi\xi_{k+2} \end{aligned}$$

for k a positive integer will be distributed as chi-square variable with even or odd number of degrees of freedom. In this manner a lot of unnecessary operations are avoided.

Since the chi-square distribution is a special case of the Gamma distribution we may also use a generator for this distribution.

8.8 Confidence Intervals for the Variance

If $x_1, x_2, ..., x_n$ are independent normal random variables from a $N(\mu, \sigma^2)$ distribution then $\frac{(n-1)s^2}{\sigma^2}$ is distributed according to the chi-square distribution with $n-1$ degrees of freedom.

A $1-\alpha$ confidence interval for the variance is then given by

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}$$

where $\chi_{\alpha,n}$ is the chi-square value for a distribution with n degrees of freedom for which the probability to be greater or equal to this value is given by α. See also below for calculations of the probability content of the chi-square distribution.

8.9 Hypothesis Testing

Let $x_1, x_2, ..., x_n$ be n independent normal random variables distributed according to a $N(\mu, \sigma^2)$ distribution. To test the null hypothesis H_0: $\sigma^2 = \sigma_0^2$ versus H_1: $\sigma^2 \neq \sigma_0^2$ at the α level of significance, we would reject the null hypothesis if $(n-1)s^2/\sigma_0^2$ is less than $\chi^2_{\alpha/2,n-1}$ or greater than $\chi^2_{1-\alpha/2,n-1}$.

8.10 Probability Content

In testing hypotheses using the chi-square distribution we define $x_\alpha = \chi^2_{\alpha,n}$ from

$$F(x_\alpha) = \int\limits_0^{x_\alpha} f(x;n)dx = 1 - \alpha$$

i.e. α is the probability that a variable distributed according to the chi-square distribution with n degrees of freedom exceeds x_α.

This formula can be used in order to determine confidence levels for certain values of α. This is what is done in producing the tables which is common in all statistics text-books. However, more often the equation is used in order to calculate the confidence level α given an experimentally determined chi-square value x_α.

In calculating the probability content of a chi-square distribution we differ between the case with even and odd number of degrees of freedom. This is described in the two following subsections.

Note that one may argue that it is as unlikely to obtain a very small chi-square value as a very big one. It is customary, however, to use only the upper tail in calculation of significance levels. A too small chi-square value is regarded as not a big problem. However, in such a case one should be somewhat critical since it indicates that one either is cheating, are using selected (biased) data or has (undeliberately) overestimated measurement errors (*e.g.* included systematic errors).

To proceed in calculating the cumulative function we write

$$1 - \alpha = F(x_\alpha) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)} \int\limits_0^{x_\alpha} \left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{\Gamma\left(\frac{n}{2}\right)} \int\limits_0^{x_\alpha/2} z^{\frac{n}{2}-1} e^{-z} dz = P\left(\frac{n}{2}, \frac{x_\alpha}{2}\right)$$

where we have made the substitution $z = x/2$. From this we see that we may use the incomplete Gamma function P (see section 42.5) in evaluating probability contents but for historical reasons we have solved the problem by considering the cases with even and odd degrees of freedom separately as is shown in the next two subsections.

Although we prefer exact routines to calculate the probability in each specific case a classical table may sometimes be useful. In table 1 on page 173 we show percentage points, *i.e.* points where the cumulative probability is $1-\alpha$, for different degrees of freedom.

It is sometimes of interest *e.g.* when rejecting a hypothesis using a chi-square test to scrutinize extremely small confidence levels. In table 2 on page 174 we show this for confidence levels down to 10^{-12} as chi-square values. In table 3 on page 175 we show the same thing in terms of chi-square over degrees of freedom ratios (reluctantly since we do not like such ratios). As discussed in section 8.2 we see, perhaps even more clearly, that for few degrees of freedom the ratios may be very high while for large number of degrees of freedom this is not the case for the same confidence level.

8.11 Even Number of Degrees of Freedom

With even n the power of z in the last integral in the formula for $F(x_\alpha)$ above is an integer. From standard integral tables we find

$$\int x^m e^{ax} dx = e^{ax} \sum_{r=0}^{m} (-1)^r \frac{m! x^{m-r}}{(m-r)! a^{r+1}}$$

where, in our case, $a = -1$. Putting $m = \frac{n}{2} - 1$ and using this integral we obtain

$$\begin{aligned} 1 - \alpha &= \frac{1}{\Gamma\left(\frac{n}{2}\right)} \int_0^{x_\alpha/2} z^{\frac{n}{2}-1} e^{-z} dz = \left[\frac{1}{m!} e^{-z} \sum_{r=0}^{m} (-1)^r \frac{m! z^{m-r}}{(m-r)!(-1)^{r+1}} \right]_0^{\frac{x_\alpha}{2}} \\ &= 1 - e^{-\frac{x_\alpha}{2}} \sum_{r=0}^{m} \frac{x_\alpha^{m-r}}{2^{m-r}(m-r)!} = 1 - e^{-\frac{x_\alpha}{2}} \sum_{r=0}^{\frac{n}{2}-1} \frac{x_\alpha^r}{2^r r!} \end{aligned}$$

a result which indeed is identical to the formula for $P(n, x)$ for integer n given on page 162.

8.12 Odd Number of Degrees of Freedom

In the case of odd number of degrees of freedom we make the substitution $z^2 = x$ yielding

$$\begin{aligned} 1 - \alpha &= F(x_\alpha) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)} \int_0^{x_\alpha} \left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \int_0^{\sqrt{x_\alpha}} \left(z^2\right)^{\frac{n}{2}-1} e^{-\frac{z^2}{2}} 2z dz = \\ &= \frac{1}{2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)} \int_0^{\sqrt{x_\alpha}} \left(z^2\right)^{\frac{n-1}{2}} e^{-\frac{z^2}{2}} dz = \frac{1}{2^{m-\frac{1}{2}}\Gamma\left(m + \frac{1}{2}\right)} \int_0^{\sqrt{x_\alpha}} z^{2m} e^{-\frac{z^2}{2}} dz \end{aligned}$$

where we have put $m = \frac{n-1}{2}$ which for odd n is an integer. By partial integration in m steps

$$\begin{aligned} \int z^{2m} e^{-\frac{z^2}{2}} dz &= \int z^{2m-1} z e^{-\frac{z^2}{2}} dz = -z^{2m-1} e^{-\frac{z^2}{2}} + (2m-1) \int z^{2m-2} e^{-\frac{z^2}{2}} dz \\ \int z^{2m-2} e^{-\frac{z^2}{2}} dz &= -z^{2m-3} e^{-\frac{z^2}{2}} + (2m-3) \int z^{2m-4} e^{-\frac{z^2}{2}} dz \\ &\vdots \\ \int z^4 e^{-\frac{z^2}{2}} dz &= -z^3 e^{-\frac{z^2}{2}} + 3 \int z^2 e^{-\frac{z^2}{2}} dz \\ \int z^2 e^{-\frac{z^2}{2}} dz &= -z e^{-\frac{z^2}{2}} + \int e^{-\frac{z^2}{2}} dz \end{aligned}$$

we obtain

$$\int z^{2m} e^{-\frac{z^2}{2}} = (2m-1)!! \int e^{-\frac{z^2}{2}} dz - \sum_{r=0}^{m-1} \frac{(2m-1)!!}{(2r+1)!!} z^{2r+1} e^{-\frac{z^2}{2}}$$

Applying this to our case gives

$$
\begin{aligned}
1-\alpha &= \frac{1}{2^{m-\frac{1}{2}}\Gamma\left(m+\frac{1}{2}\right)}\left((2m-1)!!\int\limits_0^{\sqrt{x_\alpha}} e^{-\frac{z^2}{2}}dz - \left[\sum_{r=0}^{m-1}\frac{(2m-1)!!}{(2r+1)!!}z^{2r+1}e^{-\frac{z^2}{2}}\right]_0^{\sqrt{x_\alpha}}\right) = \\
&= \sqrt{\frac{2}{\pi}}\left(\int\limits_0^{\sqrt{x_\alpha}} e^{-\frac{z^2}{2}}dz - \left[\sum_{r=0}^{m-1}\frac{1}{(2r+1)!!}z^{2r+1}e^{-\frac{z^2}{2}}\right]_0^{\sqrt{x_\alpha}}\right) = \\
&= 2G(\sqrt{x_\alpha}) - 1 - \sqrt{\frac{2x_\alpha}{\pi}}e^{-\frac{x_\alpha}{2}}\sum_{r=0}^{m-1}\frac{x_\alpha^r}{(2r+1)!!}
\end{aligned}
$$

where $G(z)$ is the integral of the standard normal distribution from $-\infty$ to z. Here we have used $\Gamma\left(m+\frac{1}{2}\right) = \frac{(2m-1)!!}{2^m}\sqrt{\pi}$ in order to simplify the coefficient. This result may be compared to the formula given on page 163 for the incomplete Gamma function when the first argument is a half-integer.

8.13 Final Algorithm

The final algorithm to evaluate the probability content from $-\infty$ to x for a chi-square distribution with n degrees of freedom is

- For n even:
  - Put $m = \frac{n}{2} - 1$.
  - Set $u_0 = 1, s = 0$ and $i = 0$.
  - For $i = 0, 1, ..., m$ set $s = s + u_i, i = i + 1$ and $u_i = u_{i-1} \cdot \frac{x}{2i}$.
  - $\alpha = s \cdot e^{-\frac{x}{2}}$.
- For n odd:
  - Put $m = \frac{n-1}{2}$.
  - Set $u_0 = 1, s = 0$ and $i = 0$.
  - For $i = 0, 1, ..., m-1$ set $s = s + u_i, i = i + 1$ and $u_i = u_{i-1} \cdot \frac{x}{2i+1}$.
  - $\alpha = 2 - 2G(\sqrt{x}) + \sqrt{\frac{2x}{\pi}}e^{-\frac{x}{2}} \cdot s$.

8.14 Chi Distribution

Sometimes, but less often, the *chi distribution i.e.* the distribution of $y = \sqrt{x}$ is used. By a simple change of variables this distribution is given by

$$
f(y) = \left|\frac{dx}{dy}\right| f(y^2) = 2y \cdot \tfrac{1}{2}\left(\frac{y^2}{2}\right)^{\frac{n}{2}-1}\frac{e^{-\frac{y^2}{2}}}{\Gamma\left(\frac{n}{2}\right)} = y^{n-1}\left(\tfrac{1}{2}\right)^{\frac{n}{2}-1}\frac{e^{-\frac{y^2}{2}}}{\Gamma\left(\frac{n}{2}\right)}
$$

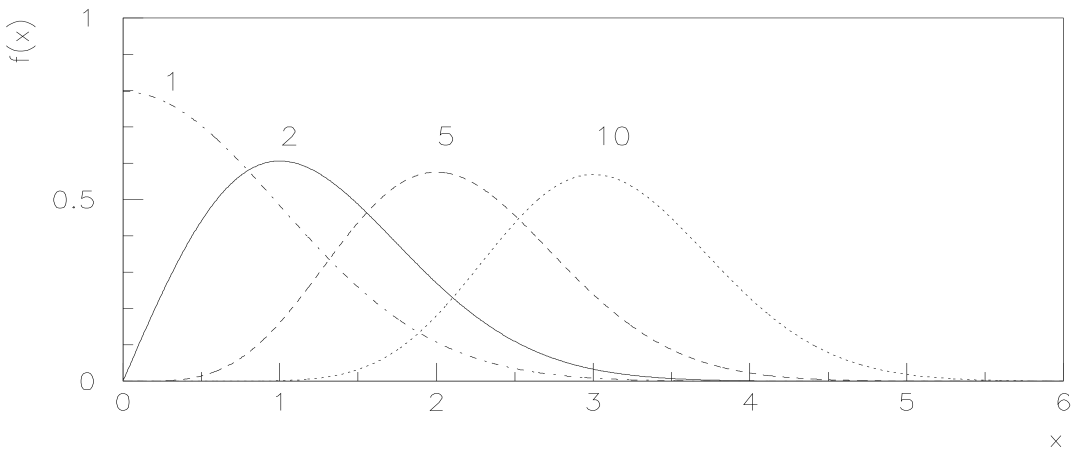


Figure 8: Graph of chi distribution for some values of n

In figure 8 the chi distribution is shown for n-values of 1, 2, 5, and 10. The mode of the distribution is at $\sqrt{n-1}$.

The cumulative function for the chi distribution becomes

$$F(y) = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right)} \int_0^y x^{n-1} e^{-\frac{x^2}{2}} dx = P\left(\frac{n}{2}, \frac{y^2}{2}\right)$$

and algebraic moments are given by

$$\mu'_k = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right)} \int_0^\infty y^k y^{n-1} e^{-\frac{y^2}{2}} dy = \frac{2^{\frac{k}{2}} \Gamma\left(\frac{n}{2} + \frac{k}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}$$

9 Compound Poisson Distribution

9.1 Introduction

The compound Poisson distribution describes the branching process for Poisson variables and is given by

$$p(r;\mu,\lambda) = \sum_{n=0}^{\infty} \frac{(n\mu)^r e^{-n\mu}}{r!} \frac{\lambda^n e^{-\lambda}}{n!}$$

where the integer variable $r \geq 0$ and the parameters μ and λ are positive real quantities.

9.2 Branching Process

The distribution describes the branching of n Poisson variables n_i all with mean μ where n is also distributed according to the Poisson distribution with mean λ *i.e.*

$$r = \sum_{i=1}^{n} n_i \quad \text{with} \quad p(n_i) = \frac{\mu^{n_i} e^{-\mu}}{n_i!} \quad \text{and} \quad p(n) = \frac{\lambda^n e^{-\lambda}}{n!}$$

and thus

$$p(r) = \sum_{n=0}^{\infty} p(r|n)p(n)$$

Due to the so called addition theorem (see page 122) for Poisson variables with mean μ the sum of n such variables are distributed as a Poisson variable with mean $n\mu$ and thus the distribution given above results.

9.3 Moments

The expectation value and variance of the Compound Poisson distribution are given by

$$E(r) = \lambda\mu \quad \text{and} \quad V(r) = \lambda\mu(1+\mu)$$

while higher moments gets slightly more complicated:

$$\begin{aligned}
\mu_3 &= \mu\lambda\left\{\mu + (\mu+1)^2\right\} \\
\mu_4 &= \mu\lambda\left\{\mu^3 + 6\mu^2 + 7\mu + 1 + 3\mu\lambda(1+\mu)^2\right\} \\
\mu_5 &= \mu\lambda\left\{\mu^4 + 10\mu^3 + 25\mu^2 + 15\mu + 1 + 10\mu\lambda(\mu+1)(\mu + (1+\mu)^2)\right\} \\
\mu_6 &= \mu\lambda\Big\{\mu^5 + 15\mu^4 + 65\mu^3 + 90\mu^2 + 31\mu + 1 \\
&\quad + 5\mu\lambda\left(5\mu^4 + 33\mu^3 + 61\mu^2 + 36\mu + 5\right) + 15\mu^2\lambda^2(\mu+1)^3\Big\}
\end{aligned}$$

9.4 Probability Generating Function

The probability generating function of the compound Poisson distribution is given by

$$G(z) = \exp\left\{-\lambda + \lambda e^{-\mu+\mu z}\right\}$$

This is easily found by using the rules for branching processes where the probability generating function (p.g.f.) is given by

$$G(z) = G_P(G_P(z))$$

where $G_P(z)$ is the p.g.f. for the Poisson distribution.

9.5 Random Number Generation

Using the basic definition we may proceed by first generate a random number n from a Poisson distribution with mean λ and then another one with mean $n\mu$.

For fixed μ and λ it is, however, normally much faster to prepare a cumulative vector for values ranging from zero up to the point where computer precision gives unity and then use this vector for random number generation. Using a binary search technique one may allow for quite long vectors giving good precision without much loss in efficiency.

10 Double-Exponential Distribution

10.1 Introduction

The Double-exponential distribution is given by

$$f(x;\mu,\lambda) = \frac{\lambda}{2}e^{-\lambda|x-\mu|}$$

where the variable x is a real number as is the location parameter μ while the parameter λ is a real positive number.

The distribution is sometimes called the *Laplace distribution* after the french astronomer, mathematician and physicist marquis Pierre Simon de Laplace (1749–1827). It is a symmetric distribution whose tails fall off less sharply than the Gaussian distribution but faster than the Cauchy distribution. It has a cusp, discontinuous first derivative, at $x = \mu$.

The distribution has an interesting feature inasmuch as the best estimator for the mean μ is the median and not the sample mean. See further the discussion in section 7 on the Cauchy distribution where the Double-exponential distribution is discussed in some detail.

10.2 Moments

For the Double-exponential distribution central moments are more easy to determine than algebraic moments (the mean is $\mu_1' = \mu$). They are given by

$$\begin{aligned}
\mu_n &= \int\limits_{-\infty}^{\infty}(x-\mu)^n f(x)dx = \frac{\lambda}{2}\left\{\int\limits_{-\infty}^{\mu}(x-\mu)^n e^{-\lambda(\mu-x)} + \int\limits_{\mu}^{\infty}(x-\mu)^n e^{-\lambda(x-\mu)}\right\} = \\
&= \frac{1}{2}\left\{\int\limits_{-\infty}^{0}\left(-\frac{y}{\lambda}\right)^n e^{-y}dy + \int\limits_{0}^{\infty}\left(\frac{y}{\lambda}\right)^n e^{-y}dy\right\} = \frac{n!}{2\lambda^n} + (-1)^n\frac{n!}{2\lambda^n}
\end{aligned}$$

i.e. odd moments vanish as they should due to the symmetry of the distribution and even moments are given by the simple relation $\mu_n = n!/\lambda^n$. From this one easily finds that the coefficient of skewness is zero and the coefficient of kurtosis 3.

If required algebraic moments may be calculated from the central moments especially the lowest order algebraic moments become

$$\mu_1' = \mu, \quad \mu_2' = \frac{2}{\lambda^2} + \mu^2, \quad \mu_3' = \frac{6\mu}{\lambda^2} + \mu^3, \quad \text{and} \quad \mu_4' = \frac{24}{\lambda^4} + \frac{12\mu^2}{\lambda^2} + \mu^4$$

but more generally

$$\mu_n' = \sum_{r=0}^{n}\binom{n}{r}\mu^r\mu_{n-r}$$

10.3 Characteristic Function

The characteristic function which generates central moments is given by

$$\phi_{x-\mu}(t) = \frac{\lambda^2}{\lambda^2 + t^2}$$

from which we may find the characteristic function which generates algebraic moments

$$\phi_x(t) = E(e^{\imath tx}) = e^{\imath t\mu}E(e^{\imath t(x-\mu)}) = e^{\imath t\mu}\phi_{x-\mu}(t) = e^{\imath t\mu}\frac{\lambda^2}{\lambda^2+t^2}$$

Sometimes an alternative which generates the sequence $\mu'_1, \mu_2, \mu_3, \ldots$ is given as

$$\phi(t) = \imath t\mu + \frac{\lambda^2}{\lambda^2+t^2}$$

10.4 Cumulative Function

The cumulative function, or distribution function, for the Double-exponential distribution is given by

$$F(x) = \begin{cases} \frac{1}{2}e^{-\lambda(\mu-x)} & \text{if} \quad x \le \mu \\ 1 - \frac{1}{2}e^{-\lambda(x-\mu)} & \text{if} \quad x > \mu \end{cases}$$

From this we see not only the obvious that the median is at $x = \mu$ but also that the lower and upper quartile is located at $\mu \mp \ln 2/\lambda$.

10.5 Random Number Generation

Given a uniform random number between zero and one in ξ a random number from a Double-exponential distribution is given by solving the equation $F(x) = \xi$ for x giving

$$\begin{array}{lll} \text{For} & \xi \le \frac{1}{2} & x = \mu + \ln(2\xi)/\lambda \\ \text{for} & \xi > \frac{1}{2} & x = \mu - \ln(2-2\xi)/\lambda \end{array}$$

11 Doubly Non-Central F-Distribution

11.1 Introduction

If x_1 and x_2 are independently distributed according to two non-central chi-square distributions with n_1 and n_2 degrees of freedom and non-central parameters λ_1 and λ_2, respectively, then the variable

$$F' = \frac{x_1/n_1}{x_2/n_2}$$

is said to have a *doubly non-central F-distribution* with n_1, n_2 degrees of freedom (positive integers) and non-centrality parameters λ_1, λ_2 (both ≥ 0).

This distribution may be written as

$$f(x; n_1, n_2, \lambda_1, \lambda_2) = \frac{n_1}{n_2} e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \frac{\left(\frac{\lambda_1}{2}\right)^r}{r!} \frac{\left(\frac{\lambda_2}{2}\right)^s}{s!} \frac{\left(\frac{n_1 x}{n_2}\right)^{\frac{n_1}{2}+r-1}}{\left(1 + \frac{n_1 x}{n_2}\right)^{\frac{n}{2}+r+s}} \frac{1}{B\left(\frac{n_1}{2}+r, \frac{n_2}{2}+s\right)}$$

where we have put $n = n_1 + n_2$ and $\lambda = \lambda_1 + \lambda_2$. For $\lambda_2 = 0$ we obtain the (singly) non-central F-distribution (see section 32) and if also $\lambda_1 = 0$ we are back to the ordinary variance ratio, or F-, distribution (see section 16).

With four parameters a variety of shapes are possible. As an example figure 9 shows the doubly non-central F-distribution for the case with $n_1 = 10, n_2 = 5$ and $\lambda_1 = 10$ varying λ_2 from zero (an ordinary non-central F-distribution) to five.

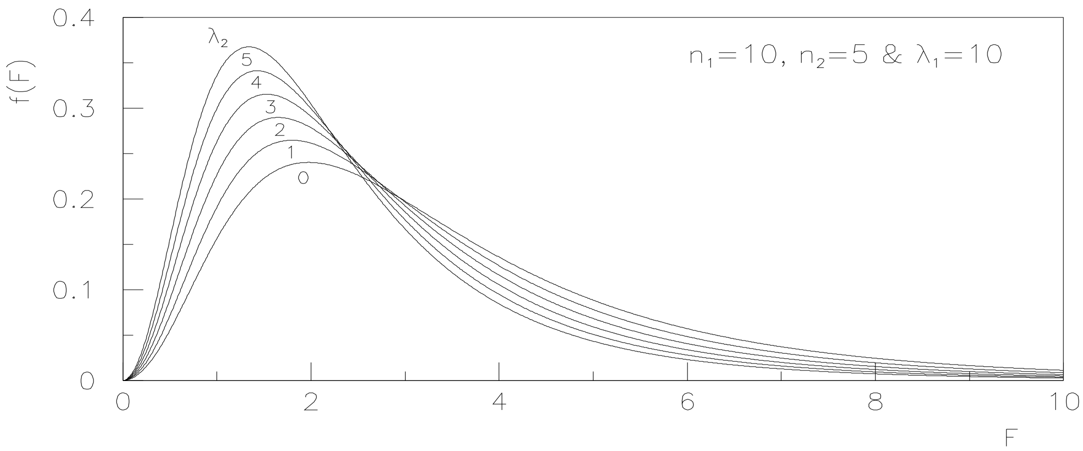


Figure 9: Examples of doubly non-central F-distributions

11.2 Moments

Algebraic moments of this distributions become

$$E(x^k) \;=\; \left(\frac{n_2}{n_1}\right)^k e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty} \frac{\left(\frac{\lambda_1}{2}\right)^r}{r!} \frac{\Gamma\left(\frac{n_1}{2}+r+k\right)}{\Gamma\left(\frac{n_1}{2}+r\right)} \sum_{s=0}^{\infty} \frac{\left(\frac{\lambda_2}{2}\right)^s}{s!} \frac{\Gamma\left(\frac{n_2}{2}+s-k\right)}{\Gamma\left(\frac{n_2}{2}+s\right)} =$$

$$
\begin{aligned}
&= \left(\frac{n_2}{n_1}\right)^k e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty} \frac{\left(\frac{\lambda_1}{2}\right)^r}{r!} \left(\tfrac{n_1}{2}+r+k-1\right)\cdots\left(\tfrac{n_1}{2}+r\right) \cdot \\
&\quad \cdot \sum_{s=0}^{\infty} \frac{\left(\frac{\lambda_2}{2}\right)^s}{s!} \left(\tfrac{n_2}{2}+s-1\right)^{-1}\cdots\left(\tfrac{n_2}{2}+s-k\right)^{-1}
\end{aligned}
$$

The r-sum involved, with a polynomial in the numerator, is quite easily solvable giving similar expressions as for the (singly) non-central F-distribution. The s-sums, however, with polynomials in the denominator give rise to confluent hypergeometric functions $M(a,b;x)$ (see appendix B). Lower order algebraic moments are given by

$$
\begin{aligned}
E(x) &= e^{-\frac{\lambda_2}{2}} \frac{n}{m} \cdot \frac{m+\lambda}{n-2} M\left(\tfrac{n-2}{2}, \tfrac{n}{2}; \tfrac{\lambda_2}{2}\right) \\
E(x^2) &= e^{-\frac{\lambda_2}{2}} \left(\frac{n}{m}\right)^2 \frac{\lambda^2+(2\lambda+m)(m+2)}{(n-2)(n-4)} M\left(\tfrac{n-4}{2}, \tfrac{n}{2}; \tfrac{\lambda_2}{2}\right) \\
E(x^3) &= e^{-\frac{\lambda_2}{2}} \left(\frac{n}{m}\right)^3 \frac{\lambda^3+3(m+4)\lambda^2+(3\lambda+m)(m+2)(m+4)}{(n-2)(n-4)(n-6)} M\left(\tfrac{n-6}{2}, \tfrac{n}{2}; \tfrac{\lambda_2}{2}\right) \\
E(x^4) &= e^{-\frac{\lambda_2}{2}} \left(\frac{n}{m}\right)^4 \frac{\lambda^4+(m+6)\left\{4\lambda^3+(m+4)\left[6\lambda^2+(4\lambda+m)(m+2)\right]\right\}}{(n-2)(n-4)(n-6)(n-8)} \cdot \\
&\quad \cdot M\left(\tfrac{n-8}{2}, \tfrac{n}{2}; \tfrac{\lambda_2}{2}\right)
\end{aligned}
$$

11.3 Cumulative Distribution

The cumulative, or distribution, function may be deduced by simple integration

$$
\begin{aligned}
F(x) &= \frac{n_1}{n_2} e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \frac{\left(\frac{\lambda_1}{2}\right)^r}{r!} \frac{\left(\frac{\lambda_2}{2}\right)^s}{s!} \frac{1}{B\left(\frac{n_1}{2}+r, \frac{n_2}{2}+s\right)} \int_0^x \frac{\left(\frac{n_1 u}{n_2}\right)^{\frac{n_1}{2}+r-1}}{\left(1+\frac{n_1 u}{n_2}\right)^{\frac{n}{2}+r+s}} du = \\
&= e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \frac{\left(\frac{\lambda_1}{2}\right)^r}{r!} \frac{\left(\frac{\lambda_2}{2}\right)^s}{s!} \frac{B_q\left(\frac{n_1}{2}+r, \frac{n_2}{2}+s\right)}{B\left(\frac{n_1}{2}+r, \frac{n_2}{2}+s\right)} = \\
&= e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \frac{\left(\frac{\lambda_1}{2}\right)^r}{r!} \frac{\left(\frac{\lambda_2}{2}\right)^s}{s!} I_q\left(\tfrac{n_1}{2}+r, \tfrac{n_2}{2}+s\right)
\end{aligned}
$$

with

$$
q = \frac{\frac{n_1 x}{n_2}}{1+\frac{n_1 x}{n_2}}
$$

11.4 Random Number Generation

Random numbers from a doubly non-central F-distribution is easily obtained using the definition in terms of the ratio between two independent random numbers from non-central chi-square distributions. This ought to be sufficient for most applications but if needed more efficient techniques may easily be developed *e.g.* using more general techniques.

12 Doubly Non-Central t-Distribution

12.1 Introduction

If x and y are independent and x is normally distributed with mean δ and unit variance while y is distributed according a non-central chi-square distribution with n degrees of freedom and non-centrality parameter λ then the variable

$$t = x/\sqrt{y/n}$$

is said to have a *doubly non-central t-distribution* with n degrees of freedom (positive integer) and non-centrality parameters δ and λ (with $\lambda \geq 0$).

This distribution may be expressed as

$$f(t;n,\delta,\lambda) = \frac{e^{-\frac{\delta^2}{2}}e^{-\frac{\lambda}{2}}}{\sqrt{n\pi}}\sum_{r=0}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{1}{\Gamma\left(\frac{n}{2}+r\right)}\sum_{s=0}^{\infty}\frac{(t\delta)^s}{s!\left(\frac{n}{2}\right)^{\frac{s}{2}}}\left(1+\frac{t^2}{n}\right)^{-\left(\frac{n+s+1}{2}+r\right)}\Gamma\left(\frac{n+s+1}{2}+r\right)$$

For $\lambda = 0$ we obtain the (singly) non-central t-distribution (see section 33) and if also $\delta = 0$ we are back to the ordinary t-distribution (see section 38).

Examples of doubly non-central t-distributions are shown in figure 9 for the case with $n = 10$ and $\delta^2 = 5$ varying λ from zero (an ordinary non-central t-distribution) to ten.

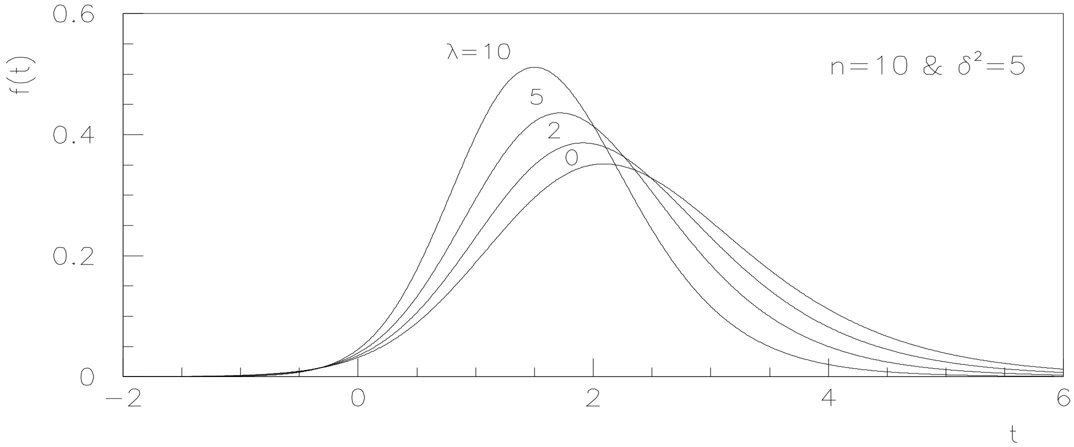


Figure 10: Examples of doubly non-central t-distributions

12.2 Moments

Algebraic moments may be deduced from the expression

$$E(t^k) = \frac{e^{-\frac{\delta^2}{2}}e^{-\frac{\lambda}{2}}}{\sqrt{n\pi}}\sum_{r=0}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{1}{\Gamma\left(\frac{n}{2}+r\right)}\sum_{s=0}^{\infty}\frac{\delta^s}{s!n^{\frac{s}{2}}}2^{\frac{s}{2}}\Gamma\left(\frac{n+s+1}{2}+r\right)\int_{-\infty}^{\infty}\frac{t^{s+k}}{\left(1+\frac{t^2}{n}\right)^{\frac{n+s+1}{2}+r}}dt =$$

$$= \frac{e^{-\frac{\delta^2}{2}}e^{-\frac{\lambda}{2}}}{\sqrt{\pi}}\sum_{r=0}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{1}{\Gamma\left(\frac{n}{2}+r\right)}\sum_{s=0}^{\infty}\frac{\delta^s}{s!}2^{\frac{s}{2}}n^{\frac{k}{2}}\Gamma\left(\frac{s+k+1}{2}\right)\Gamma\left(\frac{n-k}{2}+r\right)$$

where the sum should be taken for even values of $s+k$ *i.e.* for even (odd) orders sum only over even (odd) s-values.

Differing between moments of even and odd order the following expressions for lower order algebraic moments of the doubly non-central t-distribution may be expressed in terms of the confluent hypergeometric function $M(a,b;x)$ (see appendix B for details) as

$$
\begin{aligned}
E(t) &= \delta\sqrt{\frac{n}{2}}e^{-\frac{\lambda}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}M\left(\tfrac{n-1}{2},\tfrac{n}{2};\tfrac{\lambda}{2}\right) \\
E(t^2) &= \frac{ne^{-\frac{\lambda}{2}}\left(\delta^2+1\right)}{n-2}M\left(\tfrac{n-2}{2},\tfrac{n}{2};\tfrac{\lambda}{2}\right) \\
E(t^3) &= \delta(\delta^2+3)\sqrt{\frac{n^3}{8}}e^{-\frac{\lambda}{2}}\frac{\Gamma\left(\frac{n-3}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}M\left(\tfrac{n-3}{2},\tfrac{n}{2};\tfrac{\lambda}{2}\right) \\
E(t^4) &= \frac{n^2}{(n-2)(n-4)}\left(\delta^4+6\delta^2+3\right)e^{-\frac{\lambda}{2}}M\left(\tfrac{n-4}{2},\tfrac{n}{2};\tfrac{\lambda}{2}\right)
\end{aligned}
$$

12.3 Cumulative Distribution

The cumulative, distribution, function is given by

$$
\begin{aligned}
F(t) &= \frac{e^{-\frac{\delta^2}{2}}e^{-\frac{\lambda}{2}}}{\sqrt{n\pi}}\sum_{r=0}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{1}{\Gamma\left(\frac{n}{2}+r\right)}\sum_{s=0}^{\infty}\frac{\delta^s}{s!\left(\frac{n}{2}\right)^{\frac{s}{2}}}\Gamma\left(\tfrac{n+s+1}{2}+r\right)\int_{-\infty}^{t}\frac{u^s}{\left(1+\frac{u^2}{n}\right)^{\frac{n+s+1}{2}+r}}du = \\
&= \frac{e^{-\frac{\delta^2}{2}}e^{-\frac{\lambda}{2}}}{\sqrt{\pi}}\sum_{r=0}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\sum_{s=0}^{\infty}\frac{\delta^s}{s!}2^{\frac{s}{2}-1}\Gamma\left(\tfrac{s+1}{2}\right)\left\{s_1+s_2 I_q\left(\tfrac{s+1}{2},\tfrac{n}{2}+r\right)\right\}
\end{aligned}
$$

where $q=(t^2/n)/(1+t^2/n)$ and s_1, s_2 are signs differing between cases with positive or negative t as well as odd or even s in the summation. More specific, the sign s_1 is -1 if s is odd and $+1$ if it is even while s_2 is $+1$ unless $t<0$ and s is even in which case it is -1.

12.4 Random Number Generation

Random numbers from a doubly non-central t-distribution is easily obtained with the definition given above using random numbers from a normal distribution and a non-central chi-square distribution. This ought to be sufficient for most applications but if needed more efficient techniques may easily be developed *e.g.* using more general techniques.

13 Error Function

13.1 Introduction

A function, related to the probability content of the normal distribution, which often is referred to is the *error function*

$$\text{erf } z = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

and its complement

$$\text{erfc } z = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt = 1 - \text{erf } z$$

These functions may be defined for complex arguments, for many relations concerning the error function see [26], but here we are mainly interested in the function for real positive values of z. However, sometimes one may still want to define the function values for negative real values of z using symmetry relations

$$\begin{aligned} \text{erf}(-z) &= -\text{erf}(z) \\ \text{erfc}(-z) &= 1 - \text{erf}(-z) = 1 + \text{erf}(z) \end{aligned}$$

13.2 Probability Density Function

As is seen the error function erf is a distribution (or cumulative) function and the corresponding probability density function is given by

$$f(z) = \frac{2}{\sqrt{\pi}} e^{-z^2}$$

If we make the transformation $z = (x - \mu)/\sigma\sqrt{2}$ we obtain a folded normal distribution

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

where the function is defined for $x > \mu$ corresponding to $z > 0$, μ may be any real number while $\sigma > 0$.

This implies that $\text{erf}(z/\sqrt{2})$ is equal to the symmetric integral of a standard normal distribution between $-z$ and z.

The error function may also be expressed in terms of the incomplete Gamma function

$$\text{erf } x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = P\left(\tfrac{1}{2}, x^2\right)$$

defined for $x \geq 0$.

14 Exponential Distribution

14.1 Introduction

The exponential distribution is given by

$$f(x;\alpha) = \frac{1}{\alpha} e^{-\frac{x}{\alpha}}$$

where the variable x as well as the parameter α is positive real quantities.

The exponential distribution occur in many different connections such as the radioactive or particle decays or the time between events in a Poisson process where events happen at a constant rate.

14.2 Cumulative Function

The cumulative (distribution) function is

$$F(x) = \int_0^x f(x)dx = 1 - e^{-\frac{x}{\alpha}}$$

and it is thus straightforward to calculate the probability content in any given situation. *E.g.* we find that the median and the lower and upper quartiles are at

$$\mathcal{M} = \alpha \ln 2 \approx 0.693\alpha, \quad \mathcal{Q}_1 = -\alpha \ln \tfrac{3}{4} \approx 0.288\alpha, \quad \text{and} \quad \mathcal{Q}_3 = \alpha \ln 4 \approx 1.386\alpha$$

14.3 Moments

The expectation value, variance, and lowest order central moments are given by

$$E(x) = \alpha, \quad V(x) = \alpha^2, \quad \mu_3 = 2\alpha^3, \quad \mu_4 = 9\alpha^4,$$

$$\mu_5 = 44\alpha^5, \quad \mu_6 = 265\alpha^6, \quad \mu_7 = 1854\alpha^7, \quad \text{and} \quad \mu_8 = 14833\alpha^8$$

More generally algebraic moments are given by

$$\mu'_n = \alpha^n n!$$

Central moments thereby becomes

$$\mu_n = \alpha^n n! \sum_{m=0}^{n} \frac{(-1)^m}{m!} \to \frac{\alpha^n n!}{e} = \frac{\mu'_n}{e} \quad \text{when} \quad n \to \infty$$

the approximation is, in fact, quite good already for $n = 5$ where the absolute error is $0.146\alpha^5$ and the relative error 0.3%.

14.4 Characteristic Function

The characteristic function of the exponential distribution is given by

$$\phi(t) = E(e^{\imath tx}) = \frac{1}{\alpha}\int_0^\infty e^{(\imath t - \frac{1}{\alpha})x}dx = \frac{1}{1 - \imath t\alpha}$$

14.5 Random Number Generation

The most common way to achieve random numbers from an exponential distribution is to use the inverse to the cumulative distribution such that

$$x = F^{-1}(\xi) = -\alpha \ln(1 - \xi) = -\alpha \ln \xi'$$

where ξ is a uniform random number between zero and one (aware not to include exactly zero in the range) and so is, of course, also $\xi' = 1 - \xi$.

There are, however, alternatives some of which may be of some interest and useful if the penalty of using the logarithm would be big on any system [28].

14.5.1 Method by von Neumann

The first of these is due to J. von Neumann [29] and is as follows (with different ξ's denoting independent uniform random numbers between zero and one)

i Set $a = 0$.

ii Generate ξ and put $\xi_0 = \xi$.

iii Generate ξ^* and if $\xi^* < \xi$ then go to **vi**.

iv Generate ξ and if $\xi < \xi^*$ then go to **iii**.

v Put $a = a + 1$ and go to **ii**.

vi Put $x = \alpha(a + \xi_0)$ as a random number from an exponential distribution.

14.5.2 Method by Marsaglia

The second technique is attributed to G. Marsaglia [30].

- Prepare

$$p_n = 1 - e^{-n} \quad \text{and} \quad q_n = \frac{1}{e-1}\left(\frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!}\right)$$

for $n = 1, 2, \ldots$ until the largest representable fraction below one is exceeded in both vectors.

i Put $i = 0$ and generate ξ.

ii If $\xi > p_{i+1}$ put $i = i + 1$ and perform this step again.

iii Put $k = 1$, generate ξ and ξ^*, and set $\xi_{\min} = \xi^*$.

iv If $\xi \le q_k$ then go to **vi** else set $k = k + 1$.

v Generate a new ξ^* and if $\xi^* < \xi_{\min}$ set $\xi_{\min} = \xi^*$ and go to **iv**.

vi Put $x = \alpha(i + \xi_{\min})$ as an exponentially distributed random number.

14.5.3 Method by Ahrens

The third method is due to J. H. Ahrens [28]

- Prepare
$$q_n = \frac{\ln 2}{1!} + \frac{(\ln 2)^2}{2!} + \cdots + \frac{(\ln 2)^n}{n!}$$
for $n = 1, 2, \ldots$ until the largest representable fraction less than one is exceeded.

i Put $a = 0$ and generate ξ.

ii If $\xi < \frac{1}{2}$ set $a = a + \ln 2 = a + q_1$, $\xi = 2\xi$ and perform this step again.

iii Set $\xi = 2\xi - 1$ and if $\xi \le \ln 2 = q_1$ then exit with $x = \alpha(a + \xi)$ else put $i = 2$ and generate $\xi_{\min}$.

iv Generate ξ and put $\xi_{\min} = \xi$ if $\xi < \xi_{\min}$ then if $\xi > q_i$ put $i = i + 1$ and perform this step again else exit with $x = \alpha(a + q_1\xi_{\min})$.

Of these three methods the method by Ahrens is the fastest. This is much due to the fact that the average number of uniform random numbers consumed by the three methods is 1.69 for Ahrens, 3.58 for Marsaglia, and 4.31 for von Neumann. The method by Ahrens is often as fast as the direct logarithm method on many computers.

15 Extreme Value Distribution

15.1 Introduction

The extreme value distribution is given by

$$f(x;\mu,\sigma) = \frac{1}{\sigma}\exp\left\{\mp\frac{x-\mu}{\sigma} - e^{\mp\frac{x-\mu}{\sigma}}\right\}$$

where the upper sign is for the maximum and the lower sign for the minimum (often only the maximum is considered). The variable x and the parameter μ (the mode) are real numbers while σ is a positive real number. The distribution is sometimes referred to as the *Gumbel distribution* after E. J. Gumbel (1891–1966).

The extreme value distribution gives the limiting distribution for the largest or smallest elements of a set of independent observations from a distribution of exponential type (normal, gamma, exponential, *etc.*).

A normalized form, useful to simplify calculations, is obtained by making the substitution to the variable $z = \pm\frac{x-\mu}{\sigma}$ which has the distribution

$$g(z) = e^{-z-e^{-z}}$$

In figure 11 we show the distribution in this normalized form. The shape corresponds to the case for the maximum value while the distribution for the minimum value would be mirrored in $z = 0$.

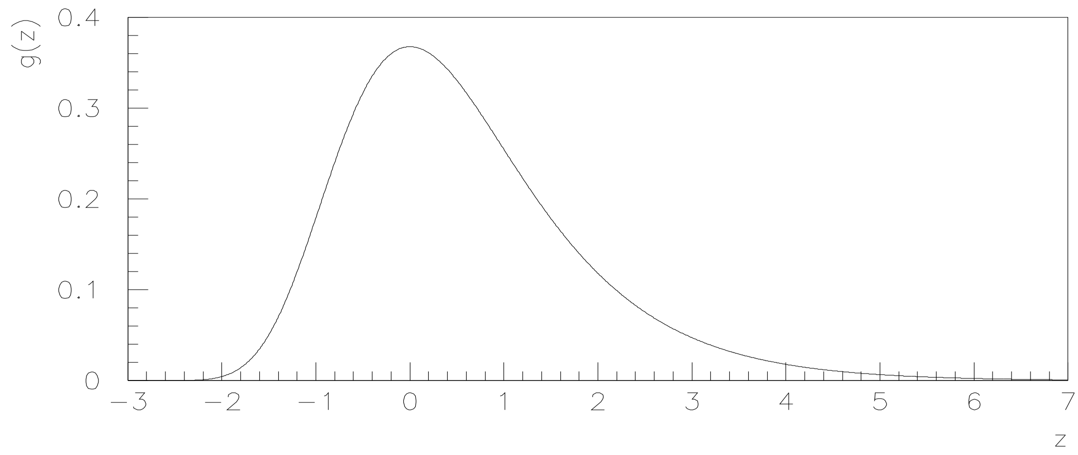


Figure 11: The normalized Extreme Value Distribution

15.2 Cumulative Distribution

The cumulative distribution for the extreme value distribution is given by

$$F(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{\pm\frac{x-\mu}{\sigma}} g(z)dz = G(\pm\frac{x-\mu}{\sigma})$$

where $G(z)$ is the cumulative function of $g(z)$ which is given by

$$G(z) = \int_{-\infty}^{z} e^{-u-e^{-u}}du = \int_{e^{-z}}^{\infty} e^{-y}dy = e^{-e^{-z}}$$

where we have made the substitution $y = e^{-u}$ in simplifying the integral. From this, and using $x = \mu \pm \sigma z$, we find the position of the median and the lower and upper quartile as

$$\begin{aligned}
\mathcal{M} &= \mu \mp \sigma \ln\ln 2 \approx \mu \pm 0.367\sigma, \\
\mathcal{Q}_1 &= \mu \mp \sigma \ln\ln 4 \approx \mu \mp 0.327\sigma, \quad \text{and} \\
\mathcal{Q}_3 &= \mu \mp \sigma \ln\ln \tfrac{4}{3} \approx \mu \pm 1.246\sigma
\end{aligned}$$

15.3 Characteristic Function

The characteristic function of the extreme value distribution is given by

$$\begin{aligned}
\phi(t) &= E\left(e^{\imath tx}\right) = \int_{-\infty}^{\infty} e^{\imath tx}\frac{1}{\sigma}\exp\left\{\mp\frac{x-\mu}{\sigma} - e^{\mp\frac{x-\mu}{\sigma}}\right\}dx = \\
&= \mp\int_{0}^{\infty} e^{\imath t(\mu\mp\sigma\ln z)}\ ze^{-z}\left(\mp\frac{\sigma dz}{z}\right) = e^{\imath t\mu}\int_{0}^{\infty} z^{\mp\imath t\sigma}e^{-z}dz = e^{\imath t\mu}\Gamma(1\mp\imath t\sigma)
\end{aligned}$$

where we have made the substitution $z = \exp\left(\mp(x-\mu)/\sigma\right)$ *i.e.* $x = \mu \mp \sigma\ln z$ and thus $dx = \mp\sigma dz/z$ to achieve an integral which could be expressed in terms of the Gamma function (see section 42.2). As a check we may calculate the first algebraic moment, the mean, by

$$\mu_1' = \frac{1}{\imath}\left.\frac{d\phi(t)}{dt}\right|_{t=0} = \frac{1}{\imath}\left[\imath\mu\Gamma(1) + \Gamma(1)\psi(1)(\mp\imath\sigma)\right] = \mu \pm \sigma\gamma$$

Here $\psi(1) = -\gamma$ is the digamma function, see section 42.3, and γ is Euler's constant. Similarly higher moments could be obtained from the characteristic function or, perhaps even easier, we may find cumulants from the cumulant generating function $\ln\phi(t)$. In the section below, however, moments are determined by a more direct technique.

15.4 Moments

Algebraic moments for $f(x)$ are given by

$$E(x^n) = \int_{-\infty}^{\infty} x^n f(x)dx = \int_{-\infty}^{\infty} (\mu\pm\sigma z)^n g(z)dz$$

which are related to moments of $g(z)$

$$E(z^n) = \int_{-\infty}^{\infty} z^n e^{-z-e^{-z}} dz = \int_0^{\infty} (-\ln y)^n e^{-y} dy$$

The first six such integrals, for n values from 1 to 6, are given by

$$\begin{aligned}
\int_0^{\infty} (-\ln x) e^{-x} dx &= \gamma \\
\int_0^{\infty} (-\ln x)^2 e^{-x} dx &= \gamma^2 + \frac{\pi^2}{6} \\
\int_0^{\infty} (-\ln x)^3 e^{-x} dx &= \gamma^3 + \frac{\gamma\pi^2}{2} + 2\zeta_3 \\
\int_0^{\infty} (-\ln x)^4 e^{-x} dx &= \gamma^4 + \gamma^2\pi^2 + \frac{3\pi^4}{20} + 8\gamma\zeta_3 \\
\int_0^{\infty} (-\ln x)^5 e^{-x} dx &= \gamma^5 + \frac{5\gamma^3\pi^2}{3} + \frac{3\gamma\pi^4}{4} + 20\gamma^2\zeta_3 + \frac{10\pi^2\zeta_3}{3} + 24\zeta_5 \\
\int_0^{\infty} (-\ln x)^6 e^{-x} dx &= \gamma^6 + \frac{5\gamma^4\pi^2}{2} + \frac{9\gamma^2\pi^4}{4} + \frac{61\pi^6}{168} + 40\gamma^3\zeta_3 + \\
&\quad + 20\gamma\pi^2\zeta_3 + 40\zeta_3^2 + 144\gamma\zeta_5
\end{aligned}$$

corresponding to the six first algebraic moments of $g(z)$. Here γ is Euler's (or Euler-Mascheroni) constant

$$\gamma = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \ln n\right) = 0.57721\ 56649\ 01532\ 86060\ 65120\ ...$$

and ζ_n is a short hand notation for Riemann's zeta-function $\zeta(n)$ given by

$$\zeta(z) = \sum_{k=1}^{\infty}\frac{1}{k^z} = \frac{1}{\Gamma(z)}\int_0^{\infty}\frac{x^{z-1}}{e^x - 1}dx \quad \text{for} \quad z > 1$$

(see also [27]). For z an even integer we may use

$$\zeta(2n) = \frac{2^{2n-1}\pi^{2n}|B_{2n}|}{(2n)!} \quad \text{for} \quad n = 1, 2, ...$$

where B_{2n} are the Bernoulli numbers given by $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}$, $B_8 = -\frac{1}{30}$ etc (see table 4 on page 176 for an extensive table of the Bernoulli numbers). This implies $\zeta_2 = \frac{\pi^2}{6}$, $\zeta_4 = \frac{\pi^4}{90}$, $\zeta_6 = \frac{\pi^6}{945}$ etc.

For odd integer arguments no similar relation as for even integers exists but evaluating the sum of reciprocal powers the two numbers needed in the calculations above are given by $\zeta_3 =$1.20205 69031 59594 28540 ... and $\zeta_5 =$1.03692 77551 43369 92633 ....

Using the algebraic moments of $g(z)$ as given above we may find the low order central moments of $g(z)$ as

$$\begin{aligned}
\mu_2 &= \frac{\pi^2}{6} = \zeta_2 \\
\mu_3 &= 2\zeta_3 \\
\mu_4 &= 3\pi^4/20 \\
\mu_5 &= \frac{10\pi^2\zeta_3}{3} + 24\zeta_5 \\
\mu_6 &= \frac{61\pi^6}{168} + 40\zeta_3^2
\end{aligned}$$

and thus the coefficients of skewness γ_1 and kurtosis γ_2 are given by

$$\begin{aligned}
\gamma_1 &= \mu_3/\mu_2^{\frac{3}{2}} = 12\sqrt{6}\zeta_3/\pi^3 \approx 1.13955 \\
\gamma_2 &= \mu^4/\mu_2^2 - 3 = 2.4
\end{aligned}$$

Algebraic moments of $f(x)$ may be found from this with some effort. Central moments are simpler being connected to those for $g(z)$ through the relation $\mu_n(x) = (\pm 1)^n \sigma^n \mu_n(z)$.

In particular the expectation value and the variance of $f(x)$ are given by

$$\begin{aligned}
E(x) &= \mu \pm \sigma E(z) = \mu \pm \sigma\gamma \\
V(x) &= \sigma^2 V(z) = \frac{\sigma^2\pi^2}{6}
\end{aligned}$$

The coefficients of skewness (except for a sign ± 1) and kurtosis are the same as for $g(z)$.

15.5 Random Number Generation

Using the expression for the cumulative distribution we may use a random number ξ, uniformly distributed between zero and one, to obtain a random number from the extreme value distribution by

$$G(z) = e^{-e^{-z}} = \xi \quad \Rightarrow \quad z = -\ln(-\ln \xi)$$

which gives a random number from the normalized function $g(z)$. A random number from $f(x)$ is then easily obtained by $x = \mu \pm \sigma z$.

16 F-distribution

16.1 Introduction

The F-distribution is given by

$$f(F;m,n) = \frac{m^{\frac{m}{2}} n^{\frac{n}{2}} \Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \cdot \frac{F^{\frac{m}{2}-1}}{(mF+n)^{\frac{m+n}{2}}} = \frac{m^{\frac{m}{2}} n^{\frac{n}{2}}}{B\left(\frac{m}{2},\frac{n}{2}\right)} \cdot \frac{F^{\frac{m}{2}-1}}{(mF+n)^{\frac{m+n}{2}}}$$

where the parameters m and n are positive integers, degrees of freedom and the variable F is a positive real number. The functions Γ and B are the usual Gamma and Beta functions. The distribution is often called the Fisher F-distribution, after the famous british statistician Sir Ronald Aylmer Fisher (1890-1962), sometimes the Snedecor F-distribution and sometimes the Fisher-Snedecor F-distribution. In figure 12 we show the F-distribution for low values of m and n.

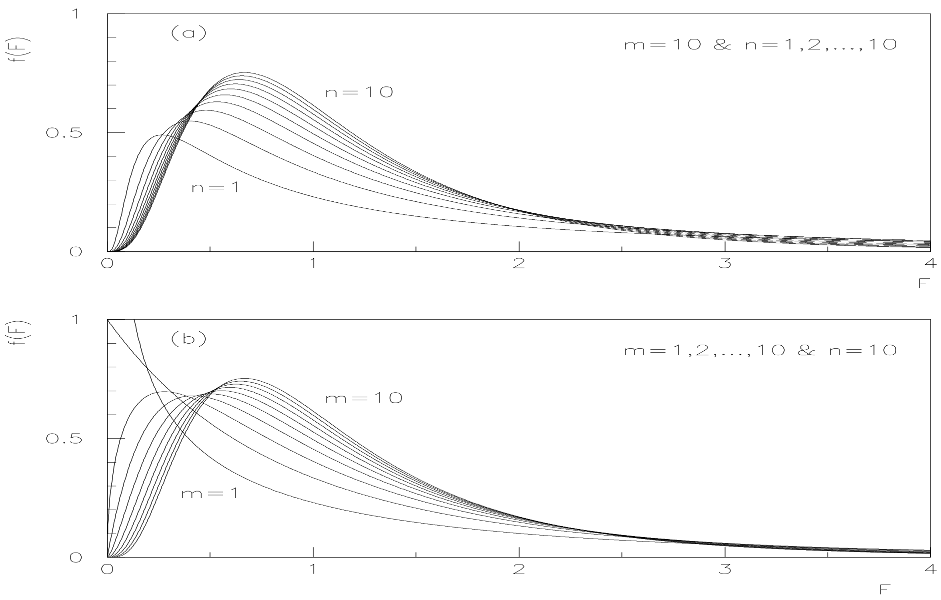


Figure 12: The F-distribution (a) for $m = 10$ and $n = 1, 2, \ldots, 10$ and (b) for $m = 1, 2, \ldots, 10$ and $n = 10$

For $m \leq 2$ the distribution has its maximum at $F = 0$ and is monotonically decreasing. Otherwise the distribution has the *mode* at

$$F_{mode} = \frac{m-2}{m} \cdot \frac{n}{n+2}$$

This distribution is also known as the *variance-ratio distribution* since it, as will be shown below, describes the distribution of the ratio of the estimated variances from two independent samples from normal distributions with equal variance.

16.2 Relations to Other Distributions

For $m = 1$ we obtain a t^2-distribution, the distribution of the square of a variable distributed according to Student's t-distribution. As $n \to \infty$ the quantity mF approaches a chi-square distribution with m degrees of freedom.

For large values of m and n the F-distribution tends to a normal distribution. There are several approximations found in the literature all of which are better than a simpleminded standardized variable. One is

$$z_1 = \frac{\sqrt{2n-1}\,\frac{mF}{n} - \sqrt{2m-1}}{\sqrt{1+\frac{mF}{n}}}$$

and an even better choice is

$$z_2 = \frac{F^{\frac{1}{3}}\left(1-\frac{2}{9n}\right) - \left(1-\frac{2}{9m}\right)}{\sqrt{\frac{2}{9m} + F^{\frac{2}{3}}\cdot\frac{2}{9n}}}$$

For large values of m and n also the distribution in the variable $z = \frac{\ln F}{2}$, the distribution of which is known as the Fisher z-distribution, is approximately normal with mean $\frac{1}{2}\left(\frac{1}{n}-\frac{1}{m}\right)$ and variance $\frac{1}{2}\left(\frac{1}{m}+\frac{1}{n}\right)$. This approximation is, however, not as good as z_2 above.

16.3 1/F

If F is distributed according to the F-distribution with m and n degrees of freedom then $\frac{1}{F}$ has the F-distribution with n and m degrees of freedom. This is easily verified by a change of variables. Putting $G = \frac{1}{F}$ we have

$$f(G) = \left|\frac{dF}{dG}\right| f(F) = \frac{1}{G^2}\cdot\frac{m^{\frac{m}{2}}n^{\frac{n}{2}}}{B(\frac{m}{2},\frac{n}{2})}\cdot\frac{\left(\frac{1}{G}\right)^{\frac{m}{2}-1}}{\left(\frac{m}{G}+n\right)^{\frac{m+n}{2}}} = \frac{m^{\frac{m}{2}}n^{\frac{n}{2}}}{B(\frac{m}{2},\frac{n}{2})}\cdot\frac{G^{\frac{n}{2}-1}}{(m+nG)^{\frac{m+n}{2}}}$$

which is seen to be identical to a F-distribution with n and m degrees of freedom for $G = \frac{1}{F}$.

16.4 Characteristic Function

The characteristic function for the F-distribution may be expressed in terms of the confluent hypergeometric function M (see section 43.3) as

$$\phi(t) = E(e^{\imath Ft}) = M\left(\tfrac{m}{2}, -\tfrac{n}{2}; -\tfrac{n}{m}\imath t\right)$$

16.5 Moments

Algebraic moments are given by

$$\mu_r' = \frac{m^{\frac{m}{2}}n^{\frac{n}{2}}}{B(\frac{m}{2},\frac{n}{2})}\int_0^\infty \frac{F^{\frac{m}{2}-1+r}}{(mF+n)^{\frac{m+n}{2}}}dF = \left(\frac{m}{n}\right)^{\frac{m}{2}}\frac{1}{B(\frac{m}{2},\frac{n}{2})}\int_0^\infty \frac{F^{\frac{m}{2}-1+r}}{\left(\frac{mF}{n}+1\right)^{\frac{m+n}{2}}}dF =$$

$$= \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{1}{B(\frac{m}{2},\frac{n}{2})} \int\limits_0^\infty \frac{\left(\frac{un}{m}\right)^{\frac{m}{2}-1+r}}{(u+1)^{\frac{m+n}{2}}} \frac{n}{m} du = \left(\frac{n}{m}\right)^r \frac{B(\frac{m}{2}+r,\frac{n}{2}-r)}{B(\frac{m}{2},\frac{n}{2})} =$$

$$= \left(\frac{n}{m}\right)^r \cdot \frac{\Gamma\left(\frac{m}{2}+r\right)\Gamma\left(\frac{n}{2}-r\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}$$

and are defined for $r < \frac{n}{2}$. This may be written

$$\mu'_r = \left(\frac{n}{m}\right)^r \cdot \frac{\frac{m}{2}(\frac{m}{2}+1)\cdots(\frac{m}{2}+r-1)}{(\frac{n}{2}-r)(\frac{n}{2}-r+1)\cdots(\frac{n}{2}-1)}$$

a form which may be more convenient for computations especially when m or n are large. A recursive formula to obtain the algebraic moments would thus be

$$\mu'_r = \mu'_{r-1} \cdot \left(\frac{n}{m}\right) \cdot \frac{\frac{m}{2}+r-1}{\frac{n}{2}-r}$$

starting with $\mu'_0 = 1$.

The first algebraic moment, the *mean*, becomes

$$E(F) = \frac{n}{n-2} \quad \text{for} \quad n > 2$$

and the variance is given by

$$V(F) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad \text{for} \quad n > 4$$

16.6 F-ratio

Regard $F = \frac{u/m}{v/n}$ where u and v are two independent variables distributed according to the chi-square distribution with m and n degrees of freedom, respectively.

The independence implies that the joint probability function in u and v is given by the product of the two chi-square distributions

$$f(u,v;m,n) = \left(\frac{\left(\frac{u}{2}\right)^{\frac{m}{2}-1} e^{-\frac{u}{2}}}{2\Gamma\left(\frac{m}{2}\right)}\right)\left(\frac{\left(\frac{v}{2}\right)^{\frac{n}{2}-1} e^{-\frac{v}{2}}}{2\Gamma\left(\frac{n}{2}\right)}\right)$$

If we change variables to $x = \frac{u/m}{v/n}$ and $y = v$ the distribution in x and y becomes

$$f(x,y;m,n) = \left\|\frac{\partial(u,v)}{\partial(x,y)}\right\| f(u,v;m,n)$$

The determinant of the Jacobian of the transformation is $\frac{ym}{n}$ and thus we have

$$f(x,y;m,n) = \frac{ym}{n}\left(\frac{\left(\frac{xym}{n}\right)^{\frac{m}{2}-1} e^{-\frac{xym}{2n}}}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)}\right)\left(\frac{y^{\frac{n}{2}-1}e^{-\frac{y}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\right)$$

Finally, since we are interested in the marginal distribution in x we integrate over y

$$
\begin{aligned}
f(x;m,n) &= \int_0^\infty f(x,y;m,n)dy = \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \int_0^\infty y^{\frac{m+n}{2}-1} e^{-\frac{y}{2}\left(\frac{xm}{n}+1\right)} dy = \\
&= \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \cdot \frac{2^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)}{\left(\frac{xm}{n}+1\right)^{\frac{m+n}{2}}} = \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}}}{B\left(\frac{m}{2},\frac{n}{2}\right)} \cdot \frac{x^{\frac{m}{2}-1}}{\left(\frac{xm}{n}+1\right)^{\frac{m+n}{2}}}
\end{aligned}
$$

which with $x = F$ is the F-distribution with m and n degrees of freedom. Here we used the integral

$$\int_0^\infty t^{z-1}e^{-\alpha t}dt = \frac{\Gamma(z)}{\alpha^z}$$

in simplifying the expression.

16.7 Variance Ratio

A practical example where the F-distribution is applicable is when estimates of the variance for two independent samples from normal distributions

$$s_1^2 = \sum_{i=1}^{m} \frac{(x_i - \overline{x})^2}{m-1} \quad \text{and} \quad s_2^2 = \sum_{i=1}^{n} \frac{(y_i - \overline{y})^2}{n-1}$$

have been made. In this case s_1^2 and s_2^2 are so called normal theory estimates of σ_1^2 and σ_2^2 *i.e.* $(m-1)s_1^2/\sigma_1^2$ and $(n-1)s_2^2/\sigma_2^2$ are distributed according to the chi-square distribution with $m-1$ and $n-1$ degrees of freedom, respectively.

In this case the quantity

$$F = \frac{s_1^2}{\sigma_1^2} \cdot \frac{\sigma_2^2}{s_2^2}$$

is distributed according to the F-distribution with $m-1$ and $n-1$ degrees of freedom. If the true variances of the two populations are indeed the same then the variance ratio s_1^2/s_2^2 have the F-distribution. We may thus use this ratio to test the null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ versus the alternative $H_1 : \sigma_1^2 \neq \sigma_2^2$ using the F-distribution. We would reject the null hypotheses at the α confidence level if the F-ratio is less than $F_{1-\alpha/2,m-1,n-1}$ or greater than $F_{\alpha/2,m-1,n-1}$ where $F_{\alpha,m,n}$ is defined by

$$\int_0^{F_{\alpha,m,n}} f(F;m,n)dF = 1-\alpha$$

i.e. α is the probability content of the distribution above the value $F_{\alpha,m-1,n-1}$. Note that the following relation between F-values corresponding to the same upper and lower confidence levels is valid

$$F_{1-\alpha,m,n} = \frac{1}{F_{\alpha,n,m}}$$

16.8 Analysis of Variance

As a simple example, which is often called *analysis of variance*, we regard n observations of a dependent variable x with overall mean $\overline{x}$ divided into k classes on an independent variable. The mean in each class is denoted $\overline{x}_j$ for $j = 1, 2, ..., k$. In each of the k classes there are n_j observations together adding up to n, the total number of observations. Below we denote by x_{ji} the i:th observation in class j.

Rewrite the total sum of squares of the deviations from the mean

$$\begin{aligned}
SS_x &= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ji}-\overline{x})^2 = \sum_{j=1}^{k}\sum_{i=1}^{n_j}((x_{ji}-\overline{x}_j)+(\overline{x}_j-\overline{x}))^2 = \\
&= \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left[(x_{ji}-\overline{x}_j)^2+(\overline{x}_j-\overline{x})^2+2(x_{ji}-\overline{x}_j)(\overline{x}_j-\overline{x})\right] = \\
&= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ji}-\overline{x}_j)^2+\sum_{j=1}^{k}\sum_{i=1}^{n_j}(\overline{x}_j-\overline{x})^2+2\sum_{j=1}^{k}(\overline{x}_j-\overline{x})\sum_{i=1}^{n_j}(x_{ji}-\overline{x}_j) = \\
&= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ji}-\overline{x}_j)^2+\sum_{j=1}^{k}n_j(\overline{x}_j-\overline{x})^2 = SS_{within}+SS_{between}
\end{aligned}$$

i.e. the total sum of squares is the sum of the sum of squares within classes and the sum of squares between classes. Expressed in terms of variances

$$nV(x) = \sum_{j=1}^{k} n_j V_j(x) + \sum_{j=1}^{k} n_j(\overline{x}_j-\overline{x})^2$$

If the variable x is independent on the classification then the variance within groups and the variance between groups are both estimates of the same true variance. The quantity

$$F = \frac{SS_{between}/(k-1)}{SS_{within}/(n-k)}$$

is then distributed according to the F-distribution with $k-1$ and $n-k$ degrees of freedom. This may then be used in order to test the hypothesis of no dependence. A too high F-value would be unlikely and thus we can choose a confidence level at which we would reject the hypothesis of no dependence of x on the classification.

Sometimes one also defines $\eta^2 = SS_{between}/SS_x$, the proportion of variance explained, as a measure of the strength of the effects of classes on the variable x.

16.9 Calculation of Probability Content

In order to set confidence levels for the F-distribution we need to evaluate the cumulative function *i.e.* the integral

$$1-\alpha = \int_0^{F_\alpha} f(F; m, n)dF$$

where we have used the notation F_α instead of $F_{\alpha,m,n}$ for convenience.

$$
\begin{aligned}
1-\alpha &= \frac{m^{\frac{m}{2}}n^{\frac{n}{2}}}{B(\frac{m}{2},\frac{n}{2})}\int\limits_0^{F_\alpha}\frac{F^{\frac{m}{2}-1}}{(mF+n)^{\frac{m+n}{2}}}dF = \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}}}{B(\frac{m}{2},\frac{n}{2})}\int\limits_0^{F_\alpha}\frac{F^{\frac{m}{2}-1}}{\left(\frac{mF}{n}+1\right)^{\frac{m+n}{2}}}dF = \\
&= \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}}}{B(\frac{m}{2},\frac{n}{2})}\int\limits_0^{\frac{mF_\alpha}{n}}\frac{\left(\frac{un}{m}\right)^{\frac{m}{2}-1}}{(u+1)^{\frac{m+n}{2}}}\frac{n}{m}du = \frac{1}{B(\frac{m}{2},\frac{n}{2})}\int\limits_0^{\frac{mF_\alpha}{n}}\frac{u^{\frac{m}{2}-1}}{(1+u)^{\frac{m+n}{2}}}du
\end{aligned}
$$

where we made the substitution $u = \frac{mF}{n}$. The last integral we recognize as the incomplete Beta function B_x defined for $0 \le x \le 1$ as

$$B_x(p,q) = \int\limits_0^x t^{p-1}(1-t)^{q-1}dt = \int\limits_0^{\frac{x}{1-x}}\frac{u^{p-1}}{(1+u)^{p+q}}du$$

where we made the substitution $u = \frac{t}{1-t}$ *i.e.* $t = \frac{u}{1+u}$. We thus obtain

$$1-\alpha = \frac{B_x(\frac{m}{2},\frac{n}{2})}{B(\frac{m}{2},\frac{n}{2})} = I_x(\frac{m}{2},\frac{n}{2})$$

with $\frac{x}{1-x} = \frac{mF_\alpha}{n}$ *i.e.* $x = \frac{mF_\alpha}{n+mF\alpha}$. The variable x thus has a Beta distribution. Note that also $I_x(a,b)$ is called the incomplete Beta function (for historical reasons discussed below but see also section 42.7).

16.9.1 The Incomplete Beta function

In order to evaluate the incomplete Beta function we may use the serial expansion

$$B_x(p,q) = x^p\left[\frac{1}{p}+\frac{1-q}{p+1}x+\frac{(1-q)(2-q)}{2!(p+2)}x^2+\ldots+\frac{(1-q)(2-q)\cdots(n-q)}{n!(p+n)}x^n+\ldots\right]$$

For integer values of q corresponding to even values of n the sum may be stopped at $n = q-1$ since all remaining terms will be identical to zero in this case.

We may express the sum with successive terms expressed recursively in the previous term

$$B_x(p,q) = x^p\sum_{r=0}^{\infty}t_r \quad \text{with} \quad t_r = t_{r-1}\cdot\frac{x(r-q)(p+r-1)}{r(p+r)} \quad \text{starting with} \quad t_0 = \frac{1}{p}$$

The sum normally converges quite fast but beware that *e.g.* for $p = q = \frac{1}{2}$ ($m = n = 1$) the convergence is very slow. Also some cases with q very big but p small seem pathological since in these cases big terms with alternate signs cancel each other causing roundoff problems. It seems preferable to keep $q < p$ to assure faster convergence. This may be done by using the relation

$$B_x(p,q) = B_1(q,p) - B_{1-x}(q,p)$$

which if inserted in the formula for $1-\alpha$ gives

$$1-\alpha = \frac{B_1(\frac{n}{2},\frac{m}{2}) - B_{1-x}(\frac{n}{2},\frac{m}{2})}{B(\frac{m}{2},\frac{n}{2})} \quad \Rightarrow \quad \alpha = \frac{B_{1-x}(\frac{n}{2},\frac{m}{2})}{B(\frac{m}{2},\frac{n}{2})} = I_{1-x}(\frac{n}{2},\frac{m}{2})$$

since $B_1(p,q) = B(p,q) = B(q,p)$.

A numerically better way to evaluate the incomplete Beta function $I_x(a,b)$ is by the continued fraction formula [9]

$$I_x(a,b) = \frac{x^a(1-x)^b}{aB(a,b)}\left[\frac{1}{1+}\frac{d_1}{1+}\frac{d_2}{1+}\cdots\right]$$

Here

$$d_{2m+1} = -\frac{(a+m)(a+b+m)x}{(a+2m)(a+2m+1)} \quad \text{and} \quad d_{2m} = \frac{m(b-m)x}{(a+2m-1)(a+2m)}$$

and the formula converges rapidly for $x < (a+1)/(a+b+1)$. For other x-values the same formula may be used after applying the symmetry relation

$$I_x(a,b) = 1 - I_{1-x}(b,a)$$

16.9.2 Final Formulæ

Using the serial expression for B_x given in the previous subsection the probability content of the F-distribution may be calculated. The numerical situation is, however, not ideal. For integer a- or b-values[3] the following relation to the binomial distribution valid for integer values of a is useful

$$1 - I_x(a,b) = I_{1-x}(b,a) = \sum_{i=0}^{a-1}\binom{a+b-1}{i}x^i(1-x)^{a+b-1-i}$$

Our final formulæ are taken from [25], using $x = \frac{n}{n+mF}$ (note that this is one minus our previous definition of x),

- Even m:

$$\begin{aligned} 1-\alpha &= x^{\frac{n}{2}}\cdot\left[1 + \frac{n}{2}(1-x) + \frac{n(n+1)}{2\cdot 4}(1-x)^2 + \ldots\right. \\ &\qquad \left.\ldots + \frac{n(n+2)\ldots(m+n-4)}{2\cdot 4\ldots(m-2)}(1-x)^{\frac{m-2}{2}}\right] \end{aligned}$$

- Even n:

$$\begin{aligned} 1-\alpha &= 1-(1-x)^{\frac{m}{2}}\left[1 + \frac{m}{2}x + \frac{m(m+2)}{2\cdot 4}x^2 + \ldots\right. \\ &\qquad \left.\ldots + \frac{m(m+2)\ldots(m+n-4)}{2\cdot 4\ldots(n-2)}x^{\frac{n-2}{2}}\right] \end{aligned}$$

[3] If only b is an integer use the relation $I_x(a,b) = 1 - I_{1-x}(b,a)$.

- Odd m and n:

$$\begin{aligned}
1-\alpha &= 1-A+\beta \quad \text{with} \\
A &= \frac{2}{\pi}\left[\theta+\sin\theta\left(\cos\theta+\frac{2}{3}\cos^3\theta+\ldots\right.\right. \\
&\left.\left.\ldots+\frac{2\cdot 4\ldots(n-3)}{1\cdot 3\ldots(n-2)}\cos^{n-2}\theta\right)\right] \quad \text{for} \quad n>1 \quad \text{and} \\
\beta &= \frac{2}{\sqrt{\pi}}\cdot\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\cdot\sin\theta\cdot\cos^n\theta\cdot\left[1+\frac{n+1}{3}\sin^2\theta+\ldots\right. \\
&\left.\ldots+\frac{(n+1)(n+3)\ldots(m+n-4)}{3\cdot 5\ldots(n-2)}\sin^{m-3}\theta\right] \quad \text{for} \quad m>1 \quad \text{where} \\
\theta &= \arctan\sqrt{\frac{nF}{m}} \quad \text{and} \quad \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}=\frac{(n-1)!!}{(n-2)!!}\cdot\frac{1}{\sqrt{\pi}}
\end{aligned}$$

  If $n=1$ then $A=2\theta/\pi$ and if $m=1$ then $\beta=0$.

- For large values of m and n we use an approximation using the standard normal distribution where

$$z=\frac{F^{\frac{1}{3}}\left(1-\frac{2}{9n}\right)-\left(1-\frac{2}{9m}\right)}{\sqrt{\frac{2}{9m}+F^{\frac{2}{3}}\cdot\frac{2}{9n}}}$$

  is approximately distributed according to the standard normal distribution. Confidence levels are obtained by

$$1-\alpha=\frac{1}{\sqrt{2\pi}}\int_z^{\infty}e^{-\frac{x^2}{2}}dx$$

In table 5 on page 177 we show some percentage points for the F-distribution. Here n is the degrees of freedom of the greater mean square and m the degrees of freedom for the lesser mean square. The values express the values of F which would be exceeded by pure chance in 10%, 5% and 1% of the cases, respectively.

16.10 Random Number Generation

Following the definition the quantity

$$F=\frac{y_m/m}{y_n/n}$$

where y_n and y_m are two variables distributed according to the chi-square distribution with n and mx degrees of freedom respectively follows the F-distribution. We may thus use this relation inserting random numbers from chi-square distributions (see section 8.7).

17 Gamma Distribution

17.1 Introduction

The Gamma distribution is given by

$$f(x; a, b) = a(ax)^{b-1}e^{-ax}/\Gamma(b)$$

where the parameters a and b are positive real quantities as is the variable x. Note that the parameter a is simply a scale factor.

For $b \leq 1$ the distribution is J-shaped and for $b > 1$ it is unimodal with its maximum at $x = \frac{b-1}{a}$.

In the special case where b is a positive integer this distribution is often referred to as the *Erlangian* distribution.

For $b = 1$ we obtain the exponential distribution and with $a = \frac{1}{2}$ and $b = \frac{n}{2}$ with n an integer we obtain the chi-squared distribution with n degrees of freedom.

In figure 13 we show the Gamma distribution for b-values of 2 and 5.

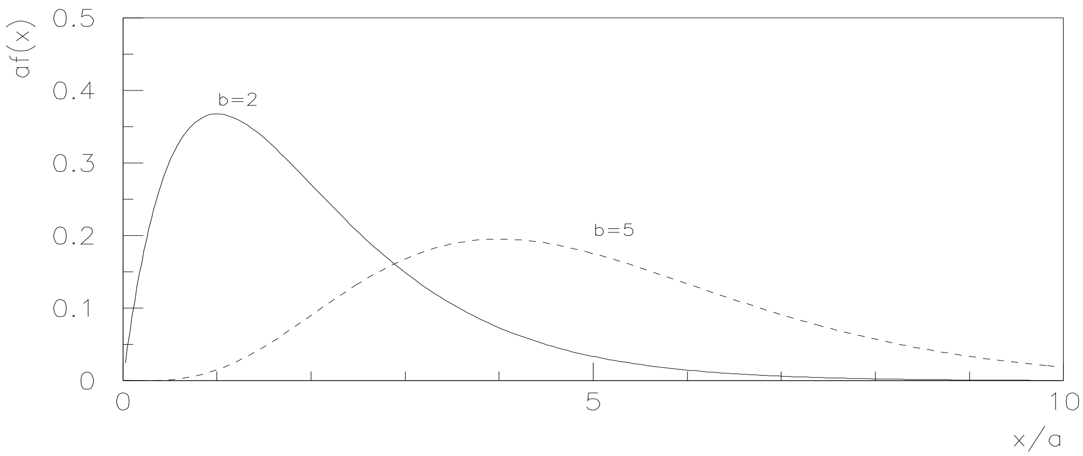


Figure 13: Examples of Gamma distributions

17.2 Derivation of the Gamma Distribution

For integer values of b, *i.e.* for Erlangian distributions, we may derive the Gamma distribution from the Poisson assumptions. For a Poisson process where events happen at a rate of λ the number of events in a time interval t is given by Poisson distribution

$$P(r) = \frac{(\lambda t)^r e^{-\lambda t}}{r!}$$

The probability that the k:th event occur at time t is then given by

$$\sum_{r=0}^{k-1} P(r) = \sum_{r=0}^{k-1} \frac{(\lambda t)^r e^{-\lambda t}}{r!}$$

i.e. the probability that there are at least k events in the time t is given by

$$F(t) = \sum_{r=k}^{\infty} P(r) = 1 - \sum_{r=0}^{k-1} \frac{(\lambda t)^r e^{-\lambda t}}{r!} = \int_0^{\lambda t} \frac{z^{k-1}e^{-z}}{(k-1)!}dz = \int_0^{t} \frac{\lambda^k z^{k-1}e^{-\lambda z}}{(k-1)!}dz$$

where the sum has been replaced by an integral (no proof given here) and the substitution $z = \lambda z$ made at the end. This is the cumulative Gamma distribution with $a = \lambda$ and $b = k$, *i.e.* the time distribution for the k:th event follows a Gamma distribution. In particular we may note that the time distribution for the occurrence of the first event follows an exponential distribution.

The Erlangian distribution thus describes the time distribution for exponentially distributed events occurring in a series. For exponential processes in parallel the appropriate distribution is the hyperexponential distribution.

17.3 Moments

The distribution has expectation value, variance, third and fourth central moments given by

$$E(x) = \frac{b}{a}, \quad V(x) = \frac{b}{a^2}, \quad \mu_3 = \frac{2b}{a^3}, \quad \text{and} \quad \mu_4 = \frac{3b(2+b)}{a^4}$$

The coefficients of skewness and kurtosis is given by

$$\gamma_1 = \frac{2}{\sqrt{b}} \quad \text{and} \quad \gamma_2 = \frac{6}{b}$$

More generally algebraic moments are given by

$$\begin{aligned}
\mu'_n &= \int_0^\infty x^n f(x)dx = \frac{a^b}{\Gamma(b)} \int_0^\infty x^{n+b-1}e^{-ax}dx = \\
&= \frac{a^b}{\Gamma(b)} \int_0^\infty \left(\frac{y}{a}\right)^{n+b-1} e^{-y}\frac{dy}{a} = \frac{\Gamma(n+b)}{a^n\Gamma(b)} = \\
&= \frac{b(b+1)\cdots(b+n-1)}{a^n}
\end{aligned}$$

where we have made the substitution $y = ax$ in simplifying the integral.

17.4 Characteristic Function

The characteristic function is

$$\phi(t) \;=\; E(e^{\imath tx}) = \frac{a^b}{\Gamma(b)} \int_0^\infty x^{b-1}e^{-x(a-\imath t)}dx =$$

$$= \frac{a^b}{\Gamma(b)} \cdot \frac{1}{(a-\imath t)^b} \int\limits_0^\infty y^{b-1} e^{-y} dy = \left(1 - \frac{\imath t}{a}\right)^{-b}$$

where we made the transformation $y = x(a - \imath t)$ in evaluating the integral.

17.5 Probability Content

In order to calculate the probability content for a Gamma distribution we need the cumulative (or distribution) function

$$\begin{aligned} F(x) &= \int\limits_0^x f(x)dx = \frac{a^b}{\Gamma(b)} \int\limits_0^x u^{b-1} e^{-au} du = \\ &= \frac{a^b}{\Gamma(b)} \int\limits_0^{ax} \left(\frac{v}{a}\right)^{b-1} e^{-v} \frac{dv}{a} = \frac{1}{\Gamma(b)} \int\limits_0^{ax} v^{b-1} e^{-v} dv = \frac{\gamma(b, ax)}{\Gamma(b)} \end{aligned}$$

where $\gamma(b, ax)$ denotes the incomplete gamma function[4].

17.6 Random Number Generation

17.6.1 Erlangian distribution

In the case of an Erlangian distribution (b a positive integer) we obtain a random number by adding b independent random numbers from an exponential distribution *i.e.*

$$x = -\ln(\xi_1 \cdot \xi_2 \cdot \ldots \cdot \xi_b)/a$$

where all the ξ_i are uniform random numbers in the interval from zero to one. Note that care must be taken if b is large in which case the product of uniform random numbers may become zero due to machine precision. In such cases simply divide the product in pieces and add the logarithms afterwards.

17.6.2 General case

In a more general case we use the so called Johnk's algorithm

i Denote the integer part of b with i and the fractional part with f and put $r = 0$. Let ξ denote uniform random numbers in the interval from zero to one.

ii If $i > 0$ then put $r = -\ln(\xi_1 \cdot \xi_2 \cdot \ldots \cdot \xi_i)$.

iii If $f = 0$ then go to **vii**.

iv Calculate $w_1 = \xi_{i+1}^{1/f}$ and $w_2 = \xi_{i+2}^{1/(1-f)}$.

[4] When integrated from zero to x the incomplete gamma function is often denoted by $\gamma(a, x)$ while for the complement, integrated from x to infinity, it is denoted $\Gamma(a, x)$. Sometimes the ratio $P(a, x) = \gamma(a, x)/\Gamma(a)$ is called the incomplete Gamma function.

v If $w_1 + w_2 > 1$ then go back to **iv**.

vi Put $r = r - \ln(\xi_{i+3}) \cdot \frac{w_1}{w_1+w_2}$.

vii Quit with $r = r/a$.

17.6.3 Asymptotic Approximation

For b big, say $b > 15$, we may use the Wilson-Hilferty approximation:

i Calculate $q = 1 + \frac{1}{9b} + \frac{z}{3\sqrt{b}}$ where z is a random number from a standard normal distribution.

ii Calculate $r = b \cdot q^3$.

iii If $r < 0$ then go back to **i**.

iv Quit with $r = r/a$.

18 Generalized Gamma Distribution

18.1 Introduction

The Gamma distribution is often used to describe variables bounded on one side. An even more flexible version of this distribution is obtained by adding a third parameter giving the so called generalized Gamma distribution

$$f(x; a, b, c) = ac(ax)^{bc-1}e^{-(ax)^c}/\Gamma(b)$$

where a (a scale parameter) and b are the same real positive parameters as is used for the Gamma distribution but a third parameter c has been added ($c = 1$ for the ordinary Gamma distribution). This new parameter may in principle take any real value but normally we consider the case where $c > 0$ or even $c \geq 1$. Put $|c|$ in the normalization for $f(x)$ if $c < 0$.

According to Hegyi [31] this density function first appeared in 1925 when L. Amoroso used it in analyzing the distribution of economic income. Later it has been used to describe the sizes of grains produced in comminution and drop size distributions in sprays *etc.*

In figure 14 we show the generalized Gamma distribution for different values of c for the case $a = 1$ and $b = 2$.

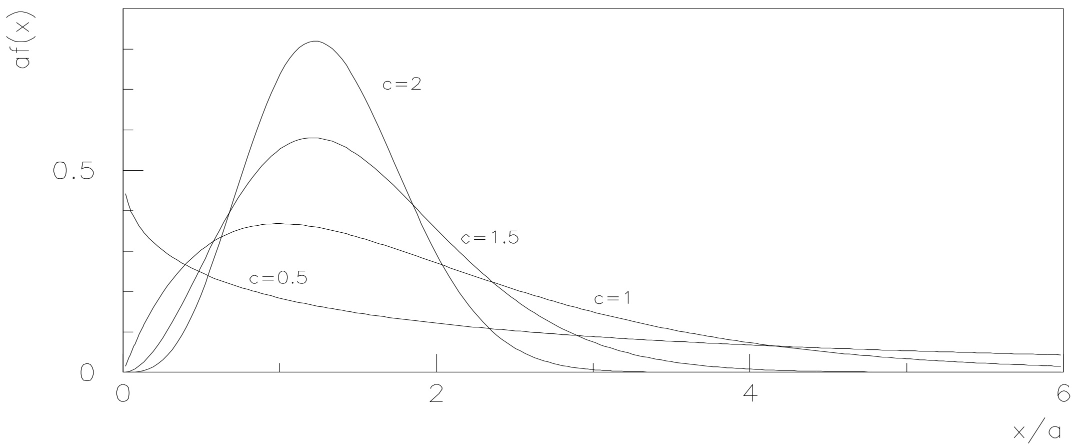


Figure 14: Examples of generalized Gamma distributions

18.2 Cumulative Function

The cumulative function is given by

$$F(x) = \gamma\left(b, (ax)^c\right)/\Gamma(b) = P\left(b, (ax)^c\right)$$

where P is the incomplete Gamma function.

18.3 Moments

Algebraic moments are given by

$$\mu'_n = \frac{1}{a^n} \cdot \frac{\Gamma\left(b + \frac{n}{c}\right)}{\Gamma(b)}$$

For negative values of c the moments are finite for ranks n satisfying $n/c > -b$ (or even just avoiding the singularities $\frac{1}{a} + \frac{n}{c} \neq 0, -1, -2 \ldots$).

18.4 Relation to Other Distributions

The generalized Gamma distribution is a general form which for certain parameter combinations gives many other distributions as special cases. In the table below we indicate some such relations. For notations see the corresponding section.

Distribution	a	b	c	Section
Generalized gamma	a	b	c	18
Gamma	a	b	1	17
Chi-squared	$\frac{1}{2}$	$\frac{n}{2}$	1	8
Exponential	$\frac{1}{\alpha}$	1	1	14
Weibull	$\frac{1}{\sigma}$	1	η	41
Rayleigh	$\frac{1}{\alpha\sqrt{2}}$	1	2	37
Maxwell	$\frac{1}{\alpha\sqrt{2}}$	$\frac{3}{2}$	2	25
Standard normal (folded)	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	2	34

In reference [31], where this distribution is used in the description of multiplicity distributions in high energy particle collisions, more examples on special cases as well as more details regarding the distribution are given.

19 Geometric Distribution

19.1 Introduction

The geometric distribution is given by

$$p(r; p) = p(1-p)^{r-1}$$

where the integer variable $r \geq 1$ and the parameter $0 < p < 1$ (no need to include limits since this give trivial special cases). It expresses the probability of having to wait exactly r trials before the first successful event if the probability of a success in a single trial is p (probability of failure $q = 1 - p$). It is a special case of the negative binomial distribution (with $k = 1$).

19.2 Moments

The expectation value, variance, third and fourth moment are given by

$$E(r) = \frac{1}{p} \quad V(r) = \frac{1-p}{p^2} \quad \mu_3 = \frac{(1-p)(2-p)}{p^3} \quad \mu_4 = \frac{(1-p)(p^2-9p+9)}{p^4}$$

The coefficients of skewness and kurtosis is thus

$$\gamma_1 = \frac{2-p}{\sqrt{1-p}} \quad \text{and} \quad \gamma_2 = \frac{p^2-6p+6}{1-p}$$

19.3 Probability Generating Function

The probability generating function is

$$G(z) = E(z^r) = \sum_{r=1}^{\infty} z^r p(1-p)^{r-1} = \frac{pz}{1-qz}$$

19.4 Random Number Generation

The cumulative distribution may be written

$$P(k) = \sum_{r=1}^{k} p(r) = 1 - q^k \quad \text{with} \quad q = 1 - p$$

which can be used in order to obtain a random number from a geometric distribution by generating uniform random numbers between zero and one until such a number (the k:th) is above q^k.

A more straightforward technique is to generate uniform random numbers ξ_i until we find a success where $\xi_k \leq p$.

These two methods are both very inefficient for low values of p. However, the first technique may be solved explicitly

$$\sum_{r=1}^{k} P(r) = \xi \quad \Rightarrow \quad k = \frac{\ln \xi}{\ln q}$$

which implies taking the largest integer less than $k+1$ as a random number from a geometric distribution. This method is quite independent of the value of p and we found [13] that a reasonable breakpoint below which to use this technique is $p = 0.07$ and use the first method mentioned above this limit. With such a method we do not gain by creating a cumulative vector for the random number generation as we do for many other discrete distributions.

20 Hyperexponential Distribution

20.1 Introduction

The hyperexponential distribution describes exponential processes in parallel and is given by

$$f(x; p, \lambda_1, \lambda_2) = p\lambda_1 e^{-\lambda_1 x} + q\lambda_2 e^{-\lambda_2 x}$$

where the variable x and the parameters λ_1 and λ_2 are positive real quantities and $0 \leq p \leq 1$ is the proportion for the first process and $q = 1 - p$ the proportion of the second.

The distribution describes the time between events in a process where the events are generated from two independent exponential distributions. For exponential processes in *series* we obtain the Erlangian distribution (a special case of the Gamma distribution).

The hyperexponential distribution is easily generalized to the case with k exponential processes in parallel

$$f(x) = \sum_{i=1}^{k} p_i \lambda_i e^{-\lambda_i x}$$

where λ_i is the slope and p_i the proportion for each process (with the constraint that $\sum p_i = 1$).

The cumulative (distribution) function is

$$F(x) = p\left(1 - e^{-\lambda_1 x}\right) + q\left(1 - e^{-\lambda_2 x}\right)$$

and it is thus straightforward to calculate the probability content in any given situation.

20.2 Moments

Algebraic moments are given by

$$\mu'_n = n!\left(\frac{p}{\lambda_1^n} + \frac{q}{\lambda_2^n}\right)$$

Central moments becomes somewhat complicated but the second central moment, the variance of the distribution, is given by

$$\mu_2 = V(x) = \frac{p}{\lambda_1^2} + \frac{q}{\lambda_2^2} + pq\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right)^2$$

20.3 Characteristic Function

The characteristic function of the hyperexponential distribution is given by

$$\phi(t) = \frac{p}{1 - \frac{\imath t}{\lambda_1}} + \frac{q}{1 - \frac{\imath t}{\lambda_2}}$$

20.4 Random Number Generation

Generating two uniform random numbers between zero and one, ξ_1 and ξ_2, we obtain a random number from a hyperexponential distribution by

- If $\xi_1 \leq p$ then put $x = -\frac{\ln \xi_2}{\lambda_1}$.
- If $\xi_1 > p$ then put $x = -\frac{\ln \xi_2}{\lambda_2}$.

i.e. using ξ_1 we choose which of the two processes to use and with ξ_2 we generate an exponential random number for this process. The same technique is easily generalized to the case with k processes.

21 Hypergeometric Distribution

21.1 Introduction

The Hypergeometric distribution is given by

$$p(r; n, N, M) = \frac{\binom{M}{r}\binom{N-M}{n-r}}{\binom{N}{n}}$$

where the discrete variable r has limits from $\max(0, n - N + M)$ to $\min(n, M)$ (inclusive). The parameters n ($1 \leq n \leq N$), N ($N \geq 1$) and M ($M \geq 1$) are all integers.

This distribution describes the experiment where elements are picked at random *without replacement*. More precisely, suppose that we have N elements out of which M has a certain attribute (and $N - M$ has not). If we pick n elements at random without replacement $p(r)$ is the probability that exactly r of the selected elements come from the group with the attribute.

If $N \gg n$ this distribution approaches a binomial distribution with $p = \frac{M}{N}$.

If instead of two groups there are k groups with different attributes the generalized hypergeometric distribution

$$p(\underline{r}; n, N, \underline{M}) = \frac{\prod\limits_{i=1}^{k}\binom{M_i}{r_i}}{\binom{N}{n}}$$

where, as before, N is the total number of elements, n the number of elements picked and $\underline{M}$ a vector with the number of elements of each attribute (whose sum should equal N). Here $n = \sum r_i$ and the limits for each r_k is given by $\max(0, n-N+M_k) \leq r_k \leq \min(n, M_k)$.

21.2 Probability Generating Function

The Hypergeometric distribution is closely related to the hypergeometric function, see appendix B on page 169, and the probability generating function is given by

$$G(z) = \frac{\binom{N-M}{n}}{\binom{N}{n}} \, {}_2F_1(-n, -M; N-M-n+1; z)$$

21.3 Moments

With the notation $p = \frac{M}{N}$ and $q = 1 - p$, *i.e.* the proportions of elements with and without the attribute, the expectation value, variance, third and fourth central moments are given by

$$\begin{aligned}
E(r) &= np \\
V(r) &= npq\frac{N-n}{N-1} \\
\mu_3 &= npq(q-p)\frac{(N-n)(N-2n)}{(N-1)(N-2)}
\end{aligned}$$

$$\mu_4 = npq(N-n)\frac{N(N+1)-6n(N-n)+3pq(N^2(n-2)-Nn^2+6n(N-n))}{(N-1)(N-2)(N-3)}$$

For the generalized hypergeometric distribution using $p_i = M_i/N$ and $q_i = 1 - p_i$ we find moments of r_i using the formulæ above regarding the group i as having an attribute and all other groups as not having the attribute. the covariances are given by

$$Cov(r_i, r_j) = np_ip_j\frac{N-n}{N-1}$$

21.4 Random Number Generation

To generate random numbers from a hypergeometric distribution one may construct a routine which follow the recipe above by picking elements at random. The same technique may be applied for the generalized hypergeometric distribution. Such techniques may be sufficient for many purposes but become quite slow.

For the hypergeometric distribution a better choice is to construct the cumulative function by adding up the individual probabilities using the recursive formula

$$p(r) = \frac{(M-r+1)(n-r+1)}{r(N-M-n+r)}p(r-1)$$

for the appropriate r-range (see above) starting with $p(r_{min})$. With the cumulative vector and one single uniform random number one may easily make a fast algorithm in order to obtain the required random number.

22 Logarithmic Distribution

22.1 Introduction

The logarithmic distribution is given by

$$p(r;p) = -\frac{(1-p)^r}{r \ln p}$$

where the variable $r \geq 1$ is an integer and the parameter $0 < p < 1$ is a real quantity.

It is a limiting form of the negative binomial distribution when the zero class has been omitted and the parameter $k \to 0$ (see section 29.4.3).

22.2 Moments

The expectation value and variance are given by

$$E(r) = -\frac{\alpha q}{p} \quad \text{and} \quad V(r) = -\frac{\alpha q(1+\alpha q)}{p^2}$$

where we have introduced $q = 1 - p$ and $\alpha = 1/\ln p$ for convenience. The third and fourth central moments are given by

$$\begin{aligned}
\mu_3 &= -\frac{\alpha q}{p^3}\left(1 + q + 3\alpha q + 2\alpha^2 q^2\right) \\
\mu_4 &= -\frac{\alpha q}{p^4}\left(1 + 4q + q^2 + 4\alpha q(1+q) + 6\alpha^2 q^2 + 3\alpha^3 q^3\right)
\end{aligned}$$

More generally factorial moments are easily found using the probability generating function

$$E(r(r-1)\cdots(r-k+1)) = \left.\frac{d^k}{dz^k}G(z)\right|_{z=1} = -(n-1)!\alpha\frac{q^k}{p^k}$$

From these moments ordinary algebraic and central moments may be found by straightforward but somewhat tedious algebra.

22.3 Probability Generating Function

The probability generating function is given by

$$G(z) = E(z^r) = \sum_{r=0}^{\infty} -\frac{z^r(1-p)^r}{r \ln p} = -\frac{1}{\ln p}\sum_{r=0}^{\infty}\frac{(zq)^r}{r} = \frac{\ln(1-zq)}{\ln(1-q)}$$

where $q = 1 - p$ and since

$$\ln(1-x) = -\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \ldots\right) \quad \text{for} \quad -1 \leq x < 1$$

22.4 Random Number Generation

The most straightforward way to obtain random numbers from a logarithmic distribution is to use the cumulative technique. If p is fixed the most efficient way is to prepare a cumulative vector starting with $p(1) = -\alpha q$ and subsequent elements by the recursive formula $p(i) = p(i-1)q/i$. The cumulative vector may, however, become very long for small values of p. Ideally it should extend until the cumulative vector element is exactly one due to computer precision. It p is not fixed the same procedure has to be made at each generation.

23 Logistic Distribution

23.1 Introduction

The Logistic distribution is given by

$$f(x; a, k) = \frac{e^z}{k(1+e^z)^2} \quad \text{with} \quad z = \frac{x-a}{k}$$

where the variable x is a real quantity, the parameter a a real location parameter (the mode, median, and mean) and k a positive real scale parameter (related to the standard deviation). In figure 15 the logistic distribution with parameters $a=0$ and $k=1$ (*i.e.* $z=x$) is shown.

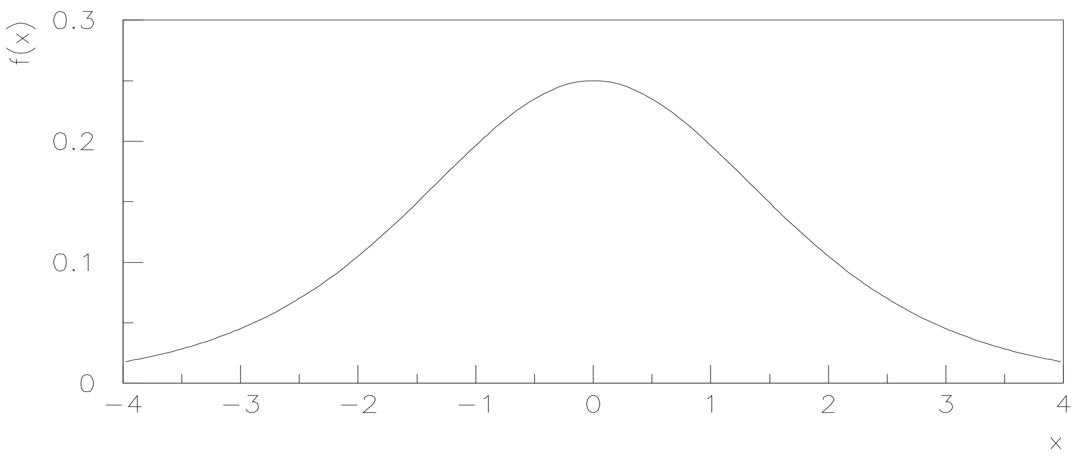


Figure 15: Graph of logistic distribution for $a = 0$ and $k = 1$

23.2 Cumulative Distribution

The distribution function is given by

$$F(x) = 1 - \frac{1}{1+e^z} = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-\frac{x-a}{k}}}$$

The inverse function is found by solving $F(x)=\alpha$ giving

$$G(\alpha) = F^{-1}(\alpha) = a - k \ln\left(\frac{1-\alpha}{\alpha}\right)$$

from which we may find *e.g.* the median as $\mathcal{M}=a$. Similarly the lower and upper quartiles are given by $\mathcal{Q}_{1,2}=a \mp k\ln 3$.

23.3 Characteristic Function

The characteristic function is given by

$$\begin{aligned}
\phi(t) &= E(e^{\imath tx}) = \int\limits_{-\infty}^{\infty} e^{\imath tx} \frac{e^{\frac{x-a}{k}}}{k\left(1+e^{\frac{x-a}{k}}\right)^2} dx = e^{\imath ta} \int\limits_{-\infty}^{\infty} \frac{e^{\imath tzk} e^z}{k(1+e^z)^2} k dz = \\
&= e^{\imath ta} \int\limits_{0}^{\infty} \frac{y^{\imath tk} y}{(1+y)^2} \cdot \frac{dy}{y} = e^{\imath ta} B(1+\imath tk, 1-\imath tk) = \\
&= e^{\imath ta} \frac{\Gamma(1+\imath tk)\Gamma(1-\imath tk)}{\Gamma(2)} = e^{\imath ta} \imath tk \Gamma(\imath tk)\Gamma(1-\imath tk) = e^{\imath ta} \frac{\imath tk\pi}{\sin \pi \imath tk}
\end{aligned}$$

where we have used the transformations $z = (x-a)/k$ and $y = e^z$ in simplifying the integral, at the end identifying the beta function, and using relation of this in terms of Gamma functions and their properties (see appendix A in section 42).

23.4 Moments

The characteristic function is slightly awkward to use in determining the algebraic moments by taking partial derivatives in t. However, using

$$\ln \phi(t) = \imath ta + \ln \Gamma(1+\imath tk) + \ln \Gamma(1-\imath tk)$$

we may determine the cumulants of the distributions. In the process we take derivatives of $\ln \phi(t)$ which involves polygamma functions (see section 42.4) but all of them with argument 1 when inserting $t=0$ a case which may be explicitly written in terms of Riemann's zeta-functions with even real argument (see page 60). It is quite easily found that all cumulants of odd order except $\kappa_1 = a$ vanish and that for even orders

$$\kappa_{2n} = 2k^{2n}\psi^{(2n-1)}(1) = 2(2n-1)!k^{2n}\zeta(2n) = 2(2n-1)!k^{2n}\frac{2^{2n-1}\pi^{2n}|B_{2n}|}{(2n)!}$$

for $n = 1, 2, \ldots$ and where B_{2n} are the Bernoulli numbers (see table 4 on page 176).

Using this formula lower order moments and the coefficients of skewness and kurtosis is found to be

$$\begin{aligned}
\mu_1' &= E(x) = \kappa_1 = a \\
\mu_2 &= V(x) = \kappa_2 = k^2\pi^2/3 \\
\mu_3 &= 0 \\
\mu_4 &= \kappa_4 + 3\kappa_2^2 = \frac{2k^4\pi^4}{15} + \frac{k^4\pi^4}{3} = \frac{7k^4\pi^4}{15} \\
\mu_5 &= 0 \\
\mu_6 &= \kappa_6 + 15\kappa_4\kappa_2 + 10\kappa_3^2 + 15\kappa_2^3 = \\
&= \frac{16k^6\pi^6}{63} + \frac{2k^6\pi^6}{3} + \frac{15k^6\pi^6}{27} = \frac{31k^6\pi^6}{11} \\
\gamma_1 &= 0 \\
\gamma_2 &= 1.2 \quad \text{(exact)}
\end{aligned}$$

23.5 Random numbers

Using the inverse cumulative function one easily obtains a random number from a logistic distribution by

$$x = a + k \ln\left(\frac{\xi}{1-\xi}\right)$$

with ξ a uniform random number between zero and one (limits not included).

24 Log-normal Distribution

24.1 Introduction

The log-normal distribution or is given by

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

where the variable $x > 0$ and the parameters μ and $\sigma > 0$ all are real numbers. It is sometimes denoted $\Lambda(\mu,\sigma^2)$ in the same spirit as we often denote a normally distributed variable by $N(\mu,\sigma^2)$.

If u is distributed as $N(\mu,\sigma^2)$ and $u = \ln x$ then x is distributed according to the log-normal distribution.

Note also that if x has the distribution $\Lambda(\mu,\sigma^2)$ then $y = e^a x^b$ is distributed as $\Lambda(a + b\mu, b^2\sigma^2)$.

In figure 16 we show the log-normal distribution for the basic form, with $\mu = 0$ and $\sigma = 1$.

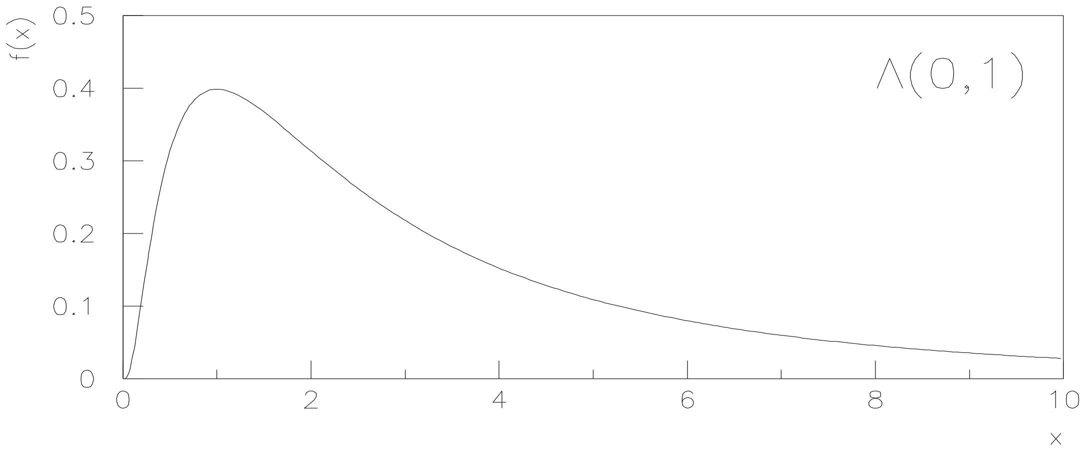


Figure 16: Log-normal distribution

The log-normal distribution is sometimes used as a first approximation to the Landau distribution describing the energy loss by ionization of a heavy charged particle (*cf* also the Moyal distribution in section 26).

24.2 Moments

The expectation value and the variance of the distribution are given by

$$E(x) = e^{\mu+\frac{\sigma^2}{2}} \quad \text{and} \quad V(x) = e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$$

and the coefficients of skewness and kurtosis becomes

$$\gamma_1 = \sqrt{e^{\sigma^2}-1}\left(e^{\sigma^2}+2\right) \quad \text{and} \quad \gamma_2 = \left(e^{\sigma^2}-1\right)\left(e^{3\sigma^2}+3e^{2\sigma^2}+6e^{\sigma^2}+6\right)$$

More generally algebraic moments of the log-normal distribution are given by

$$\begin{aligned} \mu'_k = E(x^k) &= \frac{1}{\sigma\sqrt{2\pi}}\int\limits_0^\infty x^{k-1}e^{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2}dx = \frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{yk}e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy = \\ &= \frac{1}{\sigma\sqrt{2\pi}}e^{k\mu+\frac{k^2\sigma^2}{2}}\int\limits_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{y-\mu-k\sigma^2}{\sigma}\right)^2}dy = e^{k\mu+\frac{k^2\sigma^2}{2}} \end{aligned}$$

where we have used the transformation $y = \ln x$ in simplifying the integral.

24.3 Cumulative Distribution

The cumulative distribution, or distribution function, for the log-normal distribution is given by

$$\begin{aligned} F(x) &= \frac{1}{\sigma\sqrt{2\pi}}\int\limits_0^x \frac{1}{t}e^{-\frac{1}{2}\left(\frac{\ln t-\mu}{\sigma}\right)^2}dt = \frac{1}{\sigma\sqrt{2\pi}}\int\limits_0^{\ln x} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}dy = \\ &= \frac{1}{2} \pm \frac{1}{2}P\left(\frac{1}{2}, \frac{z^2}{2}\right) \end{aligned}$$

where we have put $z = (\ln x - \mu)/\sigma$ and the positive sign is valid for $z \geq 0$ and the negative sign for $z < 0$.

24.4 Random Number Generation

The most straightforward way of achieving random numbers from a log-normal distribution is to generate a random number u from a normal distribution with mean μ and standard deviation σ and construct $r = e^u$.

25 Maxwell Distribution

25.1 Introduction

The Maxwell distribution is given by

$$f(x;\alpha) = \frac{1}{\alpha^3}\sqrt{\frac{2}{\pi}}x^2 e^{-\frac{x^2}{2\alpha^2}}$$

where the variable x with $x \geq 0$ and the parameter α with $\alpha > 0$ are real quantities. It is named after the famous scottish physicist James Clerk Maxwell (1831–1879).

The parameter α is simply a scale factor and the variable $y = x/\alpha$ has the simplified distribution

$$g(y) = \sqrt{\frac{2}{\pi}}y^2 e^{-\frac{y^2}{2}}$$

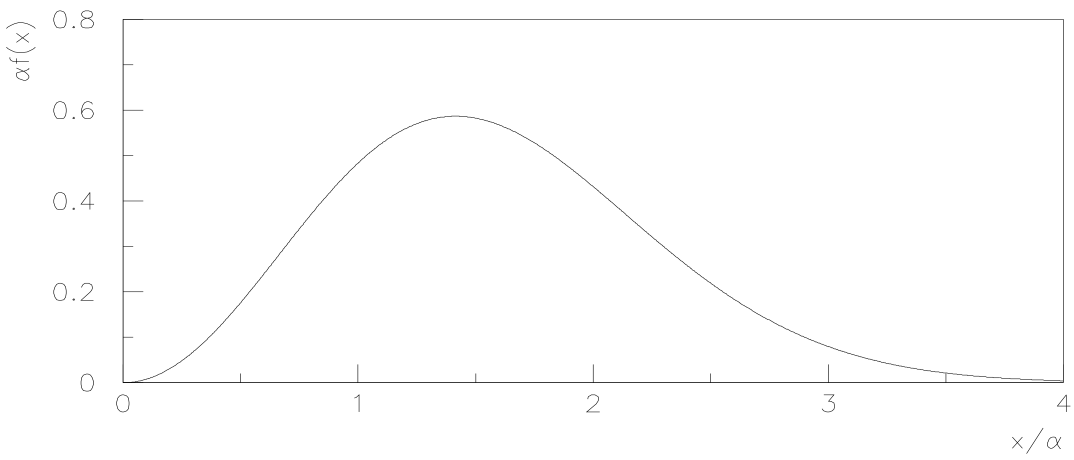


Figure 17: The Maxwell distribution

The distribution, shown in figure 17, has a mode at $x = \alpha$ and is positively skewed.

25.2 Moments

Algebraic moments are given by

$$E(x^n) = \int\limits_0^\infty x^n f(x)dx = \frac{1}{2\alpha^3}\sqrt{\frac{2}{\pi}}\int\limits_{-\infty}^{\infty} |x|^{n+2} e^{-x^2/2\alpha^2}$$

i.e. we have a connection to the absolute moments of the Gauss distribution. Using these (see section on the normal distribution) the result is

$$E(x^n) = \begin{cases} \sqrt{\frac{2}{\pi}}2^k k! \alpha^{2k-1} & \text{for } n = 2k-1 \\ (n+1)!!\alpha^n & \text{for } n \text{ even} \end{cases}$$

Specifically we note that the expectation value, variance, and the third and fourth central moments are given by

$$E(x) = 2\alpha\sqrt{\frac{2}{\pi}}, \quad V(x) = \alpha^2\left(3 - \frac{8}{\pi}\right), \quad \mu_3 = 2\alpha^3\left(\frac{16}{\pi} - 5\right)\sqrt{\frac{2}{\pi}}, \quad \text{and} \quad \mu_4 = \alpha^4\left(15 - \frac{8}{\pi}\right)$$

The coefficients of skewness and kurtosis is thus

$$\gamma_1 = \frac{2\left(\frac{16}{\pi} - 5\right)\sqrt{\frac{2}{\pi}}}{\left(3 - \frac{8}{\pi}\right)^{\frac{3}{2}}} \approx 0.48569 \quad \text{and} \quad \gamma_2 = \frac{15 - \frac{8}{\pi}}{\left(3 - \frac{8}{\pi}\right)^2} - 3 \approx 0.10818$$

25.3 Cumulative Distribution

The cumulative distribution, or the distribution function, is given by

$$F(x) = \int_0^x f(y)dy = \frac{1}{a^3}\sqrt{\frac{2}{\pi}}\int_0^x y^2 e^{-\frac{y^2}{2\alpha^2}}dy = \frac{2}{\sqrt{\pi}}\int_0^{\frac{x^2}{2\alpha^2}}\sqrt{z}e^{-z}dz = \frac{\gamma\left(\frac{3}{2}, \frac{x^2}{2\alpha^2}\right)}{\Gamma\left(\frac{3}{2}\right)} = P\left(\frac{3}{2}, \frac{x^2}{2\alpha^2}\right)$$

where we have made the substitution $z = \frac{y^2}{2\alpha^2}$ in order to simplify the integration. Here $P(a,x)$ is the incomplete Gamma function.

Using the above relation we may estimate the median $\mathcal{M}$ and the lower and upper quartile, $\mathcal{Q}_1$ and $\mathcal{Q}_3$, as

$$\begin{aligned}
\mathcal{Q}_1 &= \alpha\sqrt{P^{-1}(\tfrac{3}{2}, \tfrac{1}{2})} \approx 1.10115\ \alpha \\
\mathcal{M} &= \alpha\sqrt{P^{-1}(\tfrac{3}{2}, \tfrac{1}{2})} \approx 1.53817\ \alpha \\
\mathcal{Q}_3 &= \alpha\sqrt{P^{-1}(\tfrac{3}{2}, \tfrac{1}{2})} \approx 2.02691\ \alpha
\end{aligned}$$

where $P^{-1}(a,p)$ denotes the inverse of the incomplete Gamma function *i.e.* the value x for which $P(a,x) = p$.

25.4 Kinetic Theory

The following is taken from kinetic theory, see *e.g.* [32]. Let $v = (v_x, v_y, v_z)$ be the velocity vector of a particle where each component is distributed independently according to normal distributions with zero mean and the same variance σ^2.

First construct

$$w = \frac{v^2}{\sigma^2} = \frac{v_x^2}{\sigma^2} + \frac{v_y^2}{\sigma^2} + \frac{v_z^2}{\sigma^2}$$

Since v_x/σ, v_y/σ, and v_z/σ are distributed as standard normal variables the sum of their squares has the chi-squared distribution with 3 degrees of freedom *i.e.* $g(w) = \sqrt{\frac{w}{2\pi}}e^{-w/2}$ which leads to

$$f(v) = g(w)\left|\frac{dw}{dv}\right| = g\left(\frac{v^2}{\sigma^2}\right)\frac{2v}{\sigma^2} = \frac{1}{\sigma^3}\sqrt{\frac{2}{\pi}}v^2 e^{-\frac{v^2}{2\sigma^2}}$$

which we recognize as a Maxwell distribution with $\alpha = \sigma$.

In kinetic theory $\sigma = kT/m$, where k is Boltzmann's constant, T the temperature, and m the mass of the particles, and we thus have

$$f(v) = \sqrt{\frac{2m^3}{\pi k^3 T^3}}\; v^2 e^{-\frac{mv^2}{2kT}}$$

The distribution in kinetic energy $E = mv^2/2$ becomes

$$g(E) = \sqrt{\frac{4E}{\pi k^3 T^3}} e^{-\frac{E}{kT}}$$

which is a Gamma distribution with parameters $a = 1/kT$ and $b = \frac{3}{2}$.

25.5 Random Number Generation

To obtain random numbers from the Maxwell distribution we first make the transformation $y = x^2/2\alpha^2$ a variable which follow the Gamma distribution $g(y) = \sqrt{y}e^{-y}/\Gamma\left(\frac{3}{2}\right)$.

A random number from this distribution may be obtained using the so called Johnk's algorithm which in this particular case becomes (denoting independent pseudorandom numbers from a uniform distribution from zero to one by ξ_i)

i Put $r = -\ln \xi_1$ *i.e.* a random number from an exponential distribution.

ii Calculate $w_1 = \xi_2^2$ and $w_2 = \xi_3^2$ (with new uniform random numbers ξ_2 and ξ_3 each iteration, of course).

iii If $w = w_1 + w_2 > 1$ then go back to **ii** above.

iv Put $r = r - \frac{w_1}{w} \ln \xi_4$

v Finally construct $a\sqrt{2r}$ as a random number from the Maxwell distribution with parameter r.

Following the examples given above we may also use three independent random numbers from a standard normal distribution, z_1, z_2, and z_3, and construct

$$r = \frac{1}{\alpha}\sqrt{z_1^2 + z_2^2 + z_3^2}$$

However, this technique is not as efficient as the one outlined above.

As a third alternative we could also use the cumulative distribution putting

$$F(x) = \xi \quad \Rightarrow \quad P\left(\tfrac{3}{2}, \tfrac{x^2}{2\alpha^2}\right) = \xi \quad \Rightarrow \quad x = \alpha\sqrt{2P^{-1}\left(\tfrac{3}{2}, \xi\right)}$$

where $P^{-1}(a, p)$, as above, denotes the value x where $P(a, x) = p$. This technique is, however, much slower than the alternatives given above.

The first technique described above is not very fast but still the best alternative presented here. Also it is less dependent on numerical algorithms (such as those to find the inverse of the incomplete Gamma function) which may affect the precision of the method.

26 Moyal Distribution

26.1 Introduction

The Moyal distribution is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(z + e^{-z}\right)\right\}$$

for real values of z. A scale shift and a scale factor is introduced by making the standardized variable $z = (x - \mu)/\sigma$ and hence the distribution in the variable x is given by

$$g(x) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right)$$

Without loss of generality we treat the Moyal distribution in its simpler form, $f(z)$, in this document. Properties for $g(x)$ are easily obtained from these results which is sometimes indicated.

The Moyal distribution is a universal form for

(a) the energy loss by ionization for a fast charged particle and

(b) the number of ion pairs produced in this process.

It was proposed by J. E. Moyal [33] as a good approximation to the Landau distribution. It was also shown that it remains valid taking into account quantum resonance effects and details of atomic structure of the absorber.

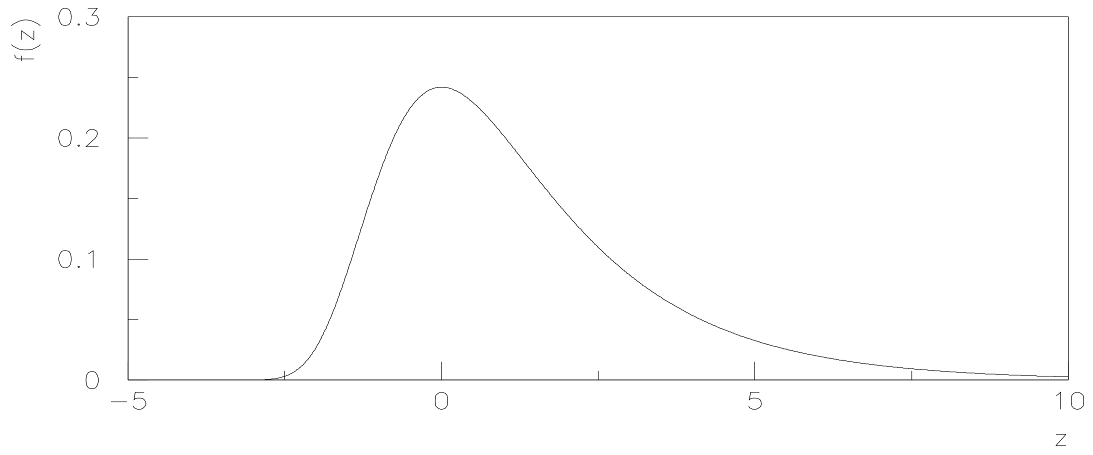


Figure 18: The Moyal distribution

The distribution, shown in figure 18, has a mode at $z = 0$ and is positively skewed. This implies that the mode of the x−distribution, $g(x)$, is equal to the parameter μ.

26.2 Normalization

Making the transformation $x = e^{-z}$ we find that

$$\begin{aligned}\int_{-\infty}^{\infty} f(z)dz &= \int_0^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2}\left(-\ln x + x\right)\right\} \frac{dx}{x} = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{e^{-\frac{x}{2}}}{\sqrt{x}} dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{e^{-y}}{\sqrt{2y}} 2dy = \frac{1}{\sqrt{\pi}} \int_0^{\infty} \frac{e^{-y}}{\sqrt{y}} dy = \frac{1}{\sqrt{\pi}} \Gamma\left(\tfrac{1}{2}\right) = 1\end{aligned}$$

where we have made the simple substitution $y = x/2$ in order to clearly recognize the Gamma function at the end. The distribution is thus properly normalized.

26.3 Characteristic Function

The characteristic function for the Moyal distribution becomes

$$\begin{aligned}\phi(t) &= E(e^{\imath tz}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\imath tz} e^{-\frac{1}{2}\left(z+e^{-z}\right)} dz = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} (2x)^{\frac{1}{2}(1-2\imath t)} e^{-x} \frac{dx}{x} = \\ &= \frac{2^{\frac{1}{2}(1-2\imath t)}}{\sqrt{2\pi}} \int_0^{\infty} x^{-\frac{1}{2}(1+2\imath t)} e^{-x} dx = \frac{2^{-\imath t}}{\sqrt{\pi}} \Gamma\left(\tfrac{1}{2} - \imath t\right)\end{aligned}$$

where we made the substitution $x = e^{-z}/2$ in simplifying the integral. The last relation to the Gamma function with complex argument is valid when the real part of the argument is positive which indeed is true in the case at hand.

26.4 Moments

As in some other cases the most convenient way to find the moments of the distribution is via its cumulants (see section 2.5). We find that

$$\begin{aligned}\kappa_1 &= -\ln 2 - \psi(\tfrac{1}{2}) = \ln 2 + \gamma \\ \kappa_n &= (-1)^n \psi^{(n-1)}(\tfrac{1}{2}) = (n-1)!(2^n - 1)\zeta_n \quad \text{for} \quad n \geq 2\end{aligned}$$

with $\gamma \approx 0.5772156649$ Euler's constant, $\psi^{(n)}$ polygamma functions (see section 42.4) and ζ Riemann's zeta-function (see page 60). Using the cumulants we find the lower order moments and the coefficients of skewness and kurtosis to be

$$\begin{aligned}\mu_1' &= E(z) = \kappa_1 = \ln 2 + \gamma \approx 1.27036 \\ \mu_2 &= V(z) = \kappa_2 = \psi^{(1)}(\tfrac{1}{2}) = \frac{\pi^2}{2} \approx 4.93480 \\ \mu_3 &= \kappa_3 = -\psi^{(2)}(\tfrac{1}{2}) = 14\zeta_3 \\ \mu_4 &= \kappa_4 + 3\kappa_2^2 = \psi^{(3)}(\tfrac{1}{2}) + 3\psi^{(1)}(\tfrac{1}{2})^2 = \frac{7\pi^4}{4} \\ \gamma_1 &= \frac{28\sqrt{2}\zeta_3}{\pi^3} \approx 1.53514 \\ \gamma_2 &= 4\end{aligned}$$

For the distribution $g(x)$ we have $E(x) = \sigma E(z) + \mu$, $V(x) = \sigma^2 V(z)$ or more generally central moments are obtained by $\mu_n(x) = \sigma^n \mu_n(z)$ for $n \geq 2$ while γ_1 and γ_2 are identical.

26.5 Cumulative Distribution

Using the same transformations as was used above in evaluating the normalization of the distribution we write the cumulative (or distribution) function as

$$\begin{aligned} F(Z) &= \int_{-\infty}^{Z} f(z)dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z} \exp\left\{-\tfrac{1}{2}\left(z + e^{-z}\right)\right\} = \frac{1}{\sqrt{2\pi}} \int_{e^{-Z}}^{\infty} \frac{e^{-\frac{x}{2}}}{\sqrt{x}} dx = \\ &= \frac{1}{\sqrt{\pi}} \int_{e^{-Z}/2}^{\infty} \frac{e^{-y}}{\sqrt{y}} dy = \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}, \frac{e^{-Z}}{2}\right) = \frac{\Gamma\left(\frac{1}{2}, \frac{e^{-Z}}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} = 1 - P\left(\frac{1}{2}, \frac{e^{-Z}}{2}\right) \end{aligned}$$

where P is the incomplete Gamma function.

Using the inverse of the cumulative function we find the median $\mathcal{M} \approx 0.78760$ and the lower and upper quartiles $\mathcal{Q}_1 \approx$ -0.28013 and $\mathcal{Q}_3 \approx 2.28739$.

26.6 Random Number Generation

To obtain random numbers from the Moyal distribution we may either make use of the inverse to the incomplete Gamma function such that given a pseudorandom number ξ we get a random number by solving the equation

$$1 - P\left(\frac{1}{2}, \frac{e^{-z}}{2}\right) = \xi$$

for z. If $P^{-1}(a, p)$ denotes the value x where $P(a, x) = p$ then

$$z = -\ln\left\{2P^{-1}\left(\tfrac{1}{2}, 1 - \xi\right)\right\}$$

is a random number from a Moyal distribution.

This is, however, a very slow method and one may instead use a straightforward reject-accept (or hit-miss) method. To do this we prefer to transform the distribution to get it into a finite interval. For this purpose we make the transformation $\tan y = x$ giving

$$h(y) = f(\tan y)\frac{1}{\cos^2 y} = \frac{1}{\sqrt{2\pi}} \frac{1}{\cos^2 y} \exp\left\{-\frac{1}{2}\left(\tan y + e^{-\tan y}\right)\right\}$$

This distribution, shown in figure 19, has a maximum of about 0.911 and is limited to the interval $-\frac{\pi}{2} \leq y \leq \frac{\pi}{2}$.

A simple algorithm to get random numbers from a Moyal distribution, either $f(z)$ or $g(x)$, using the reject-accept technique is as follows:

a Get into ξ_1 and ξ_2 two uniform random numbers uniformly distributed between zero and one using a good basic pseudorandom number generator.

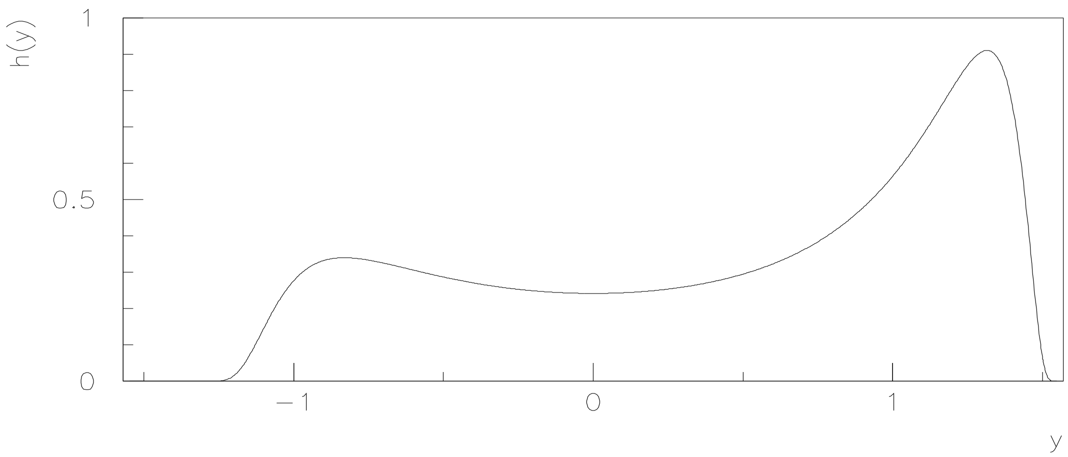


Figure 19: Transformed Moyal distribution

b Calculate uniformly distributed variables along the horizontal and vertical direction by $y = \pi\xi_1 - \frac{\pi}{2}$ and $h = \xi_2 h_{\max}$ where $h_{max} = 0.912$ is chosen slightly larger than the maximum value of the function.

c Calculate $z = \tan y$ and the function value $h(y)$.

d If $h \leq h(y)$ then accept z as a random number from the Moyal distribution $f(z)$ else go back to point **a** above.

e If required then scale and shift the result by $x = z\sigma + \mu$ in order to obtain a random number from $g(x)$.

This method is easily improved *e.g.* by making a more tight envelope to the distribution than a uniform distribution. The efficiency of the reject-accept technique outlined here is only $1/0.912\pi \approx 0.35$ (the ratio between the area of the curve and the uniform distribution). The method seems, however, fast enough for most applications.

27 Multinomial Distribution

27.1 Introduction

The Multinomial distribution is given by

$$p(\underline{r}; N, k, \underline{p}) = \frac{N!}{r_1! r_2! \cdots r_k!} p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k} = N! \prod_{i=1}^{k} \frac{p_i^{r_i}}{r_i!}$$

where the variable $\overline{r}$ is a vector with k integer elements for which $0 \leq r_i \leq N$ and $\sum r_i = N$. The parameters $N > 0$ and $k > 2$ are integers and $\overline{p}$ is a vector with elements $0 \leq p_i \leq 1$ with the constraint that $\sum p_i = 1$.

The distribution is a generalization of the Binomial distribution ($k = 2$) to many dimensions where, instead of two groups, the N elements are divided into k groups each with a probability p_i with i ranging from 1 to k. A common example is a histogram with N entries in k bins.

27.2 Histogram

The histogram example is valid when the total number of events N is regarded as a fixed number. The variance in each bin then becomes, see also below, $V(r_i) = Np_i(1 - p_1) \approx r_i$ if $p_i \ll 1$ which normally is the case for a histogram with many bins.

If, however, we may regard the total number of events N as a random variable distributed according to the Poisson distribution we find: Given a multinomial distribution, here denoted $M(\underline{r}; N, \underline{p})$, for the distribution of events into bins for fixed N and a Poisson distribution, denoted $P(N; \nu)$, for the distribution of N we write the joint distribution

$$\begin{aligned}
\mathcal{P}(\underline{r}, N) &= M(\underline{r}; N, \underline{p}) P(N; \nu) = \left(\frac{N!}{r_1! r_2! \ldots r_k!} p_1^{r_1} p_2^{r_2} \ldots p_k^{r_k} \right) \left(\frac{\nu^N e^{-\nu}}{N!} \right) = \\
&= \left(\frac{1}{r_1!} (\nu p_1)^{r_1} e^{-\nu p_1} \right) \left(\frac{1}{r_2!} (\nu p_2)^{r_2} e^{-\nu p_2} \right) \ldots \left(\frac{1}{r_k!} (\nu p_k)^{r_k} e^{-\nu p_k} \right)
\end{aligned}$$

where we have used that

$$\sum_{i=1}^{k} p_i = 1 \quad \text{and} \quad \sum_{i=1}^{k} r_i = N$$

i.e. we get a product of independent Poisson distributions with means νp_i for each individual bin.

As seen, in both cases, we find justification for the normal rule of thumb to assign the square root of the bin contents as the error in a certain bin. Note, however, that in principle we should insert the *true* value of r_i for this error. Since this normally is unknown we use the observed number of events in accordance with the law of large numbers. This means that caution must be taken in bins with few entries.

27.3 Moments

For each specific r_i we may obtain moments using the Binomial distribution with $q_i = 1 - p_i$

$$E(r_i) = Np_i \quad \text{and} \quad V(r_i) = Np_i(1 - p_i) = Np_i q_i$$

The covariance between two groups are given by

$$Cov(r_i, r_j) = -Np_i p_j \quad \text{for} \quad i \neq j$$

27.4 Probability Generating Function

The probability generating function for the multinomial distribution is given by

$$G(\underline{z}) = \left(\sum_{i=1}^{k} p_i z_i \right)^N$$

27.5 Random Number Generation

The straightforward but time consuming way to generate random numbers from a multinomial distribution is to follow the definition and generate N uniform random numbers which are assigned to specific bins according to the cumulative value of the p-vector.

27.6 Significance Levels

To determine a significance level for a certain outcome from a multinomial distribution one may add all outcomes which are *as likely or less likely* than the probability of the observed outcome. This may be a non-trivial calculation for large values of N since the number of possible outcomes grows very fast. An alternative, although quite clumsy, is to generate a number of multinomial random numbers and evaluate how often these outcomes are as likely or less likely than the observed one.

If we as an example observe the outcome $r = (4, 1, 0, 0, 0, 0)$ for a case with 5 observations in 6 groups ($N = 5$ and $k = 6$) and the probability for all groups are the same $p_i = 1/k = 1/6$ we obtain a probability of $p \approx 0.02$. This includes all orderings of the same outcome since these are all equally probable but also all less likely outcomes of the type $p = (5, 0, 0, 0, 0, 0)$.

If a probability calculated in this manner is too small one may conclude that the null hypothesis that all probabilities are equal is wrong. Thus if our confidence level is preset to 95% this conclusion would be drawn in the above example. Of course, the conclusion would be wrong in 2% of all cases.

27.7 Equal Group Probabilities

A common case or null hypothesis for a multinomial distribution is that the probability of the k groups is the same *i.e.* $p = 1/k$. In this case the multinomial distribution is simplified and since ordering become insignificant much fewer unique outcomes are possible.

Take as an example a game where five dices are thrown. The probabilities for different outcomes may quite readily be evaluated from basic probability theory properly accounting for the $6^5 = 7776$ possible outcomes. But one may also use the multinomial distribution with $k = 6$ and $N = 5$ to find probabilities for different outcomes. If we properly take care of combinatorial coefficients for each outcome we obtain (with zeros for empty groups suppressed)

name	outcome	# combinations	probability
one doublet	2,1,1,1	3600	0.46296
two doublets	2,2,1	1800	0.23148
triplets	3,1,1	1200	0.15432
nothing	1,1,1,1,1	720	0.09259
full house	3,2	300	0.03858
quadruplets	4,1	150	0.01929
quintuplets	5	6	0.00077
total		7776	1.00000

The experienced dice player may note that the "nothing" group includes 240 combinations giving straights (1 to 5 or 2 to 6). From this table we may verify the statement from the previous subsection that the probability to get an outcome with quadruplets or less likely outcomes is given by 0.02006.

Generally we have for $N < k$ that the two extremes of either all observations in separate groups p_{sep} or all observations in one group p_{all}

$$
\begin{aligned}
p_{sep} &= \frac{k!}{k^N(k-N)!} = \frac{k}{k} \cdot \frac{k-1}{k} \cdots \frac{k-N+1}{k} \\
p_{all} &= \frac{1}{k^{N-1}}
\end{aligned}
$$

which we could have concluded directly from a quite simple probability calculation.

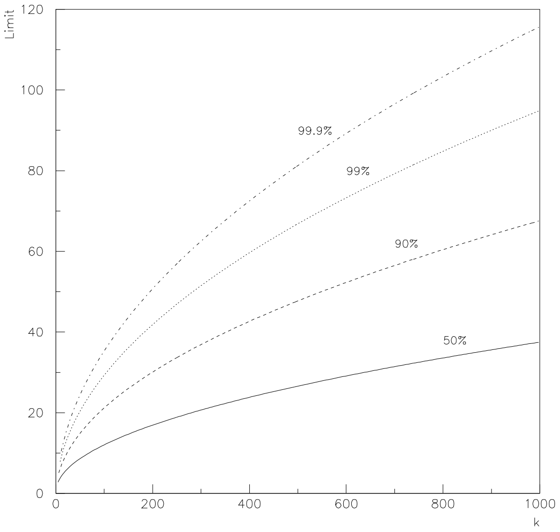
The first case is the formula which shows the quite well known fact that if 23 people or more are gathered the probability that at least two have the same birthday, *i.e.* $1 - p_{sep}$, is greater than 50% (using $N = 23$ and $k = 365$ and not bothering about leap-years or possible deviations from the hypothesis of equal probabilities for each day). This somewhat non-intuitive result becomes even more pronounced for higher values of k and the level above which $p_{sep} < 0.5$ is approximately given by

$$N \approx 1.2\sqrt{k}$$

For higher significance levels we may note that in the case with $k = 365$ the probability $1 - p_{sep}$ becomes greater than 90% at $N = 41$, greater than 99% at $N = 57$ and greater than 99.9% at $N = 70$ *i.e.* already for $N << k$ a bet would be almost certain.

In Fig.20 we show, in linear scale to the left and logarithmic scale to the right, the lower limit on N for which the probability to have $1 - p_{sep}$ above 50%, 90%, 99% and 99.9.% for k-values ranging up to 1000. By use of the gamma function the problem has been generalized to real numbers. Note that the curves start at certain values where $k = N$ since for $N > k$ it is impossible to have all events in separate groups[5].

[5] This limit is at $N = k = 2$ for the 50%-curve, 3.92659 for 90%, 6.47061 for 99% and 8.93077 for 99.9%



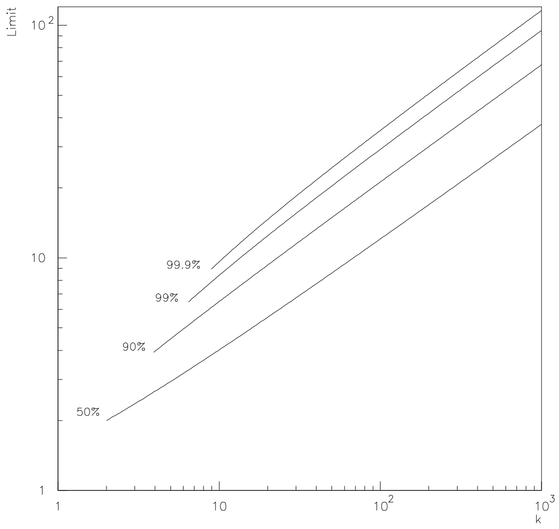


Figure 20: Limits for N at several confidence levels as a function of k (linear scale to the left and logarithmic scale to the right).

28 Multinormal Distribution

28.1 Introduction

As a generalization of the normal or Gauss distribution to many dimensions we define the multinormal distribution.

A multinormal distribution in $x = \{x_1, x_2, \ldots, x_n\}$ with parameters μ (mean vector) and V (variance matrix) is given by

$$f(x|\mu, V) = \frac{e^{-\frac{1}{2}(x-\mu)V^{-1}(x-\mu)^T}}{(2\pi)^{\frac{n}{2}}\sqrt{|V|}}$$

The variance matrix V has to be a positive semi-definite matrix in order for f to be a proper probability density function (necessary in order that the normalization integral $\int f(x)dx$ should converge).

If x is normal and V non-singular then $(x-\mu)V^{-1}(x-\mu)^T$ is called the *covariance form* of x and has a χ^2-distribution with n degrees of freedom. Note that the distribution has constant probability density for constant values of the covariance form.

The *characteristic function* is given by

$$\phi(t) = e^{\imath t\mu - \frac{1}{2}t^T V t}$$

where t is a vector of length n.

28.2 Conditional Probability Density

The conditional density for a fixed value of any x_i is given by a multinormal density with $n-1$ dimensions where the new variance matrix is obtained by deleting the i:th row and column of V^{-1} and inverting the resulting matrix.

This may be compared to the case where we instead just want to neglect one of the variables x_i. In this case the remaining variables has a multinormal distribution with $n-1$ dimensions with a variance matrix obtained by deleting the i:th row and column of V.

28.3 Probability Content

As discussed in section 6.6 on the binormal distribution the joint probability content of a multidimensional normal distribution is different, and smaller, than the corresponding well known figures for the one-dimensional normal distribution. In the case of the binormal distribution the ellipse (see figure 2 on page 20) corresponding to one standard deviation has a joint probability content of 39.3%.

The same is even more true for the probability content within the hyperellipsoid in the case of a multinormal distribution. In the table below we show, for different dimensions n, the probability content for the one (denoted $z = 1$), two and three standard deviation contours. We also give z-values z_1, z_2, and z_3 adjusted to give a probability content within the hyperellipsoid corresponding to the one-dimensional one, two, and three standard deviation contents (68.3%, 95.5%, and 99.7%). Finally z-value corresponding to joint probability contents of 90%, 95% and 99% in z_{90}, z_{95}, and z_{99}, respectively, are given. Note that these probability contents are independent of the variance matrix which only has the effect to change the shape of the hyperellipsoid from a perfect hypersphere with radius z when all variables are uncorrelated to *e.g.* cigar shapes when correlations are large.

Note that this has implications on errors estimated from a chi-square or a maximum likelihood fit. If a multiparameter confidence limit is requested and the chi-square minimum is at $\chi^2_{\min}$ or the logarithmic likelihood maximum at $\ln \mathcal{L}_{\max}$, one should look for the error contour at $\chi^2_{\min} + z^2$ or $\ln \mathcal{L}_{\max} - z^2/2$ using a z-value from the right-hand side of the table below.

	Probability content in %			Adjusted z-values					
n	$z=1$	$z=2$	$z=3$	z_1	z_2	z_3	z_{90}	z_{95}	z_{99}
1	68.3	95.5	99.7	1.00	2.00	3.00	1.64	1.96	2.58
2	39.3	86.5	98.9	1.51	2.49	3.44	2.15	2.45	3.03
3	19.9	73.9	97.1	1.88	2.83	3.76	2.50	2.80	3.37
4	9.02	59.4	93.9	2.17	3.12	4.03	2.79	3.08	3.64
5	3.74	45.1	89.1	2.43	3.36	4.27	3.04	3.33	3.88
6	1.44	32.3	82.6	2.65	3.58	4.48	3.26	3.55	4.10
7	0.52	22.0	74.7	2.84	3.79	4.67	3.47	3.75	4.30
8	0.175	14.3	65.8	3.05	3.97	4.86	3.66	3.94	4.48
9	0.056	8.86	56.3	3.23	4.15	5.03	3.83	4.11	4.66
10	0.017	5.27	46.8	3.40	4.31	5.19	4.00	4.28	4.82
11	0.005	3.01	37.8	3.56	4.47	5.34	4.16	4.44	4.97
12	0.0015	1.66	29.7	3.71	4.62	5.49	4.31	4.59	5.12
13	0.0003	0.88	22.7	3.85	4.76	5.63	4.45	4.73	5.26
14	0.0001	0.45	16.9	3.99	4.90	5.76	4.59	4.87	5.40
15	≈ 0	0.23	12.2	4.13	5.03	5.89	4.72	5.00	5.53
16	≈ 0	0.11	8.66	4.26	5.16	6.02	4.85	5.13	5.66
17	≈ 0	0.052	5.97	4.38	5.29	6.14	4.98	5.25	5.78
18	≈ 0	0.024	4.03	4.50	5.41	6.26	5.10	5.37	5.90
19	≈ 0	0.011	2.65	4.62	5.52	6.37	5.22	5.49	6.02
20	≈ 0	0.005	1.71	4.74	5.60	6.13	5.33	5.60	6.13
25	≈ 0	0.0001	0.14	5.27	6.17	7.01	5.86	6.14	6.66
30	≈ 0	≈ 0	0.007	5.75	6.65	7.49	6.34	6.62	7.13

28.4 Random Number Generation

In order to obtain random numbers from a multinormal distribution we proceed as follows:

- If $x = \{x_1, x_2, \ldots, x_n\}$ is distributed multinormally with mean 0 (zero vector) and variance matrix I (unity matrix) then each x_i ($i = 1, 2, \ldots, n$) can be found independently from a standard normal distribution.

- If x is multinormally distributed with mean μ and variance matrix V then any linear combination $y = Sx$ is also multinormally distributed with mean $S\mu$ and variance matrix SVS^T,

- If we want to generate vectors, y, from a multinormal distribution with mean μ and variance matrix V we may make a so called *Cholesky decomposition* of V, *i.e.* we find a triangular matrix S such that $V = SS^T$. We then calculate $y = Sx + \mu$ with the components of x generated independently from a standard normal distribution.

Thus we have found a quite nice way of generating multinormally distributed random numbers which is important in many simulations where correlations between variables may not be ignored. If many random numbers are to be generated for multinormal variables from the same distribution it is beneficial to make the Cholesky decomposition once and store the matrix S for further usage.

29 Negative Binomial Distribution

29.1 Introduction

The Negative Binomial distribution is given by

$$p(r;k,p) = \binom{r-1}{k-1} p^k (1-p)^{r-k}$$

where the variable $r \geq k$ and the parameter $k > 0$ are integers and the parameter p ($0 \leq p \leq 1$) is a real number.

The distribution expresses the probability of having to wait exactly r trials until k successes have occurred if the probability of a success in a single trial is p (probability of failure $q = 1 - p$).

The above form of the Negative Binomial distribution is often referred to as the *Pascal distribution* after the french mathematician, physicist and philosopher Blaise Pascal (1623–1662).

The distribution is sometimes expressed in terms of the number of failures occurring while waiting for k successes, $n = r - k$, in which case we write

$$p(n;k,p) = \binom{n+k-1}{n} p^k (1-p)^n$$

where the new variable $n \geq 0$.

Changing variables, for this last form, to $\overline{n}$ and k instead of p and k we sometimes use

$$p(n;\overline{n},k) = \binom{n+k-1}{n} \frac{\overline{n}^n k^k}{(\overline{n}+k)^{n+k}} = \binom{n+k-1}{n} \left(\frac{\overline{n}}{\overline{n}+k}\right)^n \left(\frac{k}{\overline{n}+k}\right)^k$$

The distribution may also be generalized to real values of k, although this may seem obscure from the above probability view-point ("fractional success"), writing the binomial coefficient as $(n+k-1)(n+k-2)\cdots(k+1)k/n!$.

29.2 Moments

In the first form given above the expectation value, variance, third and fourth central moments of the distribution are

$$E(r) = \frac{k}{p}, \quad V(r) = \frac{kq}{p^2}, \quad \mu_3 = \frac{kq(2-p)}{p^3}, \quad \text{and} \quad \mu_4 = \frac{kq(p^2-6p+6+3kq)}{p^4}$$

The coefficients of skewness and kurtosis are

$$\gamma_1 = \frac{2-p}{\sqrt{kq}} \quad \text{and} \quad \gamma_2 = \frac{p^2-6p+6}{kq}$$

In the second formulation above, $p(n)$, the only difference is that the expectation value becomes

$$E(n) = E(r) - k = \frac{k(1-p)}{p} = \frac{kq}{p}$$

while higher moments remain unchanged as they should since we have only shifted the scale by a fixed amount.

In the last form given, using the parameters $\overline{n}$ and k, the expectation value and the variance are

$$E(n) = \overline{n} \quad \text{and} \quad V(n) = \overline{n} + \frac{\overline{n}^2}{k}$$

29.3 Probability Generating Function

The probability generating function is given by

$$G(z) = \left(\frac{pz}{1 - zq}\right)^k$$

in the first case ($p(r)$) and

$$G(z) = \left(\frac{p}{1 - zq}\right)^k = \left(\frac{1}{1 + (1 - z)\frac{\overline{n}}{k}}\right)^k$$

in the second case ($p(n)$) for the two different parameterizations.

29.4 Relations to Other Distributions

There are several interesting connections between the Negative Binomial distribution and other standard statistical distributions. In the following subsections we briefly address some of these connections.

29.4.1 Poisson Distribution

Regard the negative binomial distribution in the form

$$p(n; \overline{n}, k) = \binom{n + k - 1}{n} \left(\frac{1}{1 + \overline{n}/k}\right)^k \left(\frac{\overline{n}/k}{1 + \overline{n}/k}\right)^n$$

where $n \geq 0$, $k > 0$ and $\overline{n} > 0$.

As $k \to \infty$ the three terms become

$$\binom{n + k - 1}{n} = \frac{(n + k - 1)(n + k - 2) \ldots k}{n!} \to \frac{k^n}{n!},$$

$$\left(\frac{1}{1 + \overline{n}/k}\right)^k = 1 - k\frac{\overline{n}}{k} + \frac{k(k + 1)}{2}\left(\frac{\overline{n}}{k}\right)^2 - \frac{k(k + 1)(k + 2)}{6}\left(\frac{\overline{n}}{k}\right)^3 + \ldots \to e^{-\overline{n}} \quad \text{and}$$

$$k^n \left(\frac{\overline{n}/k}{1 + \overline{n}/k}\right)^n \to \overline{n}^n$$

where, for the last term we have incorporated the factor k^n from the first term.

Thus we have shown that

$$\lim_{k\to\infty} p(n;\overline{n},k) = \frac{\overline{n}^n e^{-\overline{n}}}{n!}$$

i.e. a Poisson distribution.

This "proof" could perhaps better be made using the probability generating function of the negative binomial distribution

$$G(z) = \left(\frac{p}{1-zq}\right)^k = \left(\frac{1}{1-(z-1)\overline{n}/k}\right)^k$$

Making a Taylor expansion of this for $(z-1)\overline{n}/k \ll 1$ we get

$$G(z) = 1 + (z-1)\overline{n} + \frac{k+1}{k}\frac{(z-1)^2\overline{n}^2}{2} + \frac{(k+1)(k+2)}{k^2}\frac{(z-1)^3\overline{n}^3}{6} + \ldots \to e^{(z-1)\overline{n}}$$

as $k \to \infty$. This result we recognize as the probability generating function of the Poisson distribution.

29.4.2 Gamma Distribution

Regard the negative binomial distribution in the form

$$p(n;k,p) = \binom{n+k-1}{n} p^k q^n$$

where $n \geq 0$, $k > 0$ and $0 \leq p \leq 1$ and where we have introduced $q = 1-p$. If we change parameters from k and p to k and $\overline{n} = kq/p$ this may be written

$$p(n;\overline{n},k) = \binom{n+k-1}{n} \left(\frac{1}{1+\overline{n}/k}\right)^k \left(\frac{\overline{n}/k}{1+\overline{n}/k}\right)^n$$

Changing variable from n to $z = n/\overline{n}$ we get $(dn/dz = \overline{n})$

$$\begin{aligned}
p(z;\overline{n},k) &= p(n;\overline{n},k)\frac{dn}{dz} = \overline{n}\binom{z\overline{n}+k-1}{z\overline{n}} \left(\frac{1}{1+\overline{n}/k}\right)^k \left(\frac{\overline{n}/k}{1+\overline{n}/k}\right)^{z\overline{n}} = \\
&= \overline{n}\,\frac{(z\overline{n}+k-1)(z\overline{n}+k-2)\ldots(z\overline{n}+1)}{\Gamma(k)} k^k \left(\frac{1}{k+\overline{n}}\right)^k \left(\frac{1}{k/\overline{n}+1}\right)^{z\overline{n}} \to \\
&\to \overline{n}\,k^k \frac{(z\overline{n})^{k-1}}{\Gamma(k)} \left(\frac{1}{k+\overline{n}}\right)^k \left(\frac{1}{k/\overline{n}+1}\right)^{z\overline{n}} = \\
&= \frac{z^{k-1}k^k}{\Gamma(k)} \left(\frac{\overline{n}}{k+\overline{n}}\right)^k \left(\frac{1}{k/\overline{n}+1}\right)^{z\overline{n}} \to \frac{z^{k-1}k^k e^{-kz}}{\Gamma(k)}
\end{aligned}$$

where we have used that for $k \ll \overline{n} \to \infty$

$$\left(\frac{\overline{n}}{k+\overline{n}}\right)^k \to 1 \quad \text{and}$$

$$
\begin{aligned}
\left(\frac{1}{k/\overline{n}+1}\right)^{z\overline{n}} &= 1 - z\overline{n}\frac{k}{\overline{n}} + \frac{z\overline{n}(z\overline{n}+1)}{2}\left(\frac{k}{\overline{n}}\right)^2 - \frac{z\overline{n}(z\overline{n}+1)(z\overline{n}+2)}{6}\left(\frac{k}{\overline{n}}\right)^3 + \ldots \to \\
&\to 1 - zk + \frac{z^2k^2}{2} - \frac{z^3k^3}{6} + \ldots = e^{-kz}
\end{aligned}
$$

as $\overline{n} \to \infty$.

Thus we have "shown" that as $\overline{n} \to \infty$ and $\overline{n} \gg k$ we obtain a gamma distribution in the variable $z = n/\overline{n}$.

29.4.3 Logarithmic Distribution

Regard the negative binomial distribution in the form

$$
p(n; k, p) = \binom{n+k-1}{n} p^k q^n
$$

where $n \geq 0$, $k > 0$ and $0 \leq p \leq 1$ and where we have introduced $q = 1 - p$.

The probabilities for $n = 0, 1, 2, 3...$ are given by

$$
\{\, p(0),\ p(1),\ p(2),\ p(3),\ \ldots\} = p^k \left\{ 1,\ kq,\ \frac{k(k+1)}{2!}q^2,\ \frac{k(k+1)(k+2)}{3!}q^3,\ \ldots \right\}
$$

if we omit the zero class (n=0) and renormalize we get

$$
\frac{kp^k}{1-p^k}\left\{ 0,\ q,\ \frac{k+1}{2!}q^2,\ \frac{(k+1)(k+2)}{3!}q^3,\ \ldots \right\}
$$

and if we let $k \to 0$ we finally obtain

$$
-\frac{1}{\ln p}\left\{ 0,\ q,\ \frac{q^2}{2},\ \frac{q^3}{3},\ \ldots \right\}
$$

where we have used that

$$
\lim_{k \to 0} \frac{k}{p^{-k} - 1} = -\frac{1}{\ln p}
$$

which is easily realized expanding $p^{-k} = e^{-k \ln p}$ into a power series.

This we recognize as the logarithmic distribution

$$
p(n; p) = -\frac{1}{\ln p}\frac{(1-p)^n}{n}
$$

thus we have shown that omitting the zero class and letting $k \to 0$ the negative binomial distribution becomes the logarithmic distribution.

29.4.4 Branching Process

In a process where a branching occurs from a Poisson to a logarithmic distribution the most elegant way to determine the resulting distribution is by use of the probability generating function. The probability generating functions for a Poisson distribution with parameter (mean) μ and for a logarithmic distribution with parameter p ($q = 1 - p$) are given by

$$G_P(z) = e^{\mu(z-1)} \quad \text{and} \quad G_L(z) = \ln(1 - zq)/\ln(1 - q) = \alpha \ln(1 - zq)$$

where $\mu > 0$, $0 \le q \le 1$ and $\alpha = 1/\ln p$.

For a branching process in n steps

$$G(z) = G_1(G_2(\ldots G_{n-1}(G_n(z))\ldots))$$

where $G_k(z)$ is the probability generating function in the k:th step. In the above case this gives

$$\begin{aligned} G(z) &= G_P(G_L(z)) = \exp\{\mu(\alpha \ln(1 - zq) - 1)\} = \\ &= \exp\{\alpha\mu \ln(1 - zq) - \mu\} = (1 - zq)^{\alpha\mu} e^{-\mu} = \\ &= (1 - zq)^{-k}(1 - q)^k = p^k/(1 - zq)^k \end{aligned}$$

where we have put $k = -\alpha\mu$. This we recognize as the probability generating function of a negative binomial distribution with parameters k and p.

We have thus shown that a Poisson distribution with mean μ branching into a logarithmic distribution with parameter p gives rise to a negative binomial distribution with parameters $k = -\alpha\mu = -\mu/\ln p$ and p (or $\overline{n} = kq/p$).

Conversely a negative binomial distribution with parameters k and p or $\overline{n}$ *could* arise from the combination of a Poisson distribution with parameter $\mu = -k \ln p = k \ln(1 + \frac{\overline{n}}{k})$ and a logarithmic distribution with parameter p and mean $\overline{n}/\mu$.

A particle physics example would be a charged multiplicity distribution arising from the production of independent clusters subsequently decaying into charged particles according to a logarithmic distribution. The UA5 experiment [34] found on the $S\overline{p}pS$ collider at CERN that at a centre of mass energy of 540 GeV a negative binomial distribution with $\overline{n} = 28.3$ and $k = 3.69$ fitted the data well. With the above scenario this would correspond to ≈ 8 clusters being independently produced (Poisson distribution with $\mu = 7.97$) each one decaying, according to a logarithmic distribution, into 3.55 charged particles on average.

29.4.5 Poisson and Gamma Distributions

If a Poisson distribution with mean $\mu > 0$

$$p(n; \mu) = \frac{e^{-\mu}\mu^n}{n!} \quad \text{for} \quad n \ge 0$$

is weighted by a gamma distribution with parameters $a > 0$ and $b > 0$

$$f(x; a, b) = \frac{a(ax)^{b-1}e^{-ax}}{\Gamma(b)} \quad \text{for} \quad x > 0$$

we obtain

$$\begin{aligned}
\mathcal{P}(n) &= \int_0^\infty p(n;\mu) f(\mu;a,b) d\mu = \int_0^\infty \frac{e^{-\mu}\mu^n}{n!} \frac{a(a\mu)^{b-1}e^{-a\mu}}{\Gamma(b)} d\mu = \\
&= \frac{a^b}{n!\Gamma(b)} \int_0^\infty \mu^{n+b-1} e^{-\mu(a+1)} d\mu = \frac{a^b}{n!(b-1)!}(n+b-1)!(a+1)^{-(n+b)} = \\
&= \binom{n+b-1}{n} \left(\frac{a}{a+1}\right)^b \left(\frac{1}{a+1}\right)^n
\end{aligned}$$

which is a negative binomial distribution with parameters $p = \frac{a}{a+1}$, *i.e.* $q = 1 - p = \frac{1}{a+1}$, and $k = b$. If we aim at a negative binomial distribution with parameters $\overline{n}$ and k we should thus weight a Poisson distribution with a gamma distribution with parameters $a = k/\overline{n}$ and $b = k$. This is the same as superimposing Poisson distributions with means coming from a gamma distribution with mean $\overline{n}$.

In the calculation above we have made use of integral tables for the integral

$$\int_0^\infty x^n e^{-\alpha x} dx = n! \alpha^{-(n+1)}$$

29.5 Random Number Generation

In order to obtain random numbers from a Negative Binomial distribution we may use the recursive formula

$$p(r+1) = p(r) \frac{qr}{r+1-k} \quad \text{or} \quad p(n+1) = p(n) \frac{q(k+n)}{n+1}$$

for $r = k, k+1, \ldots$ and $n = 0, 1, \ldots$ in the two cases starting with the first term ($p(k)$ or $p(0)$) being equal to p^k. This technique may be speeded up considerably, if p and k are constants, by preparing a cumulative vector once for all.

One may also use some of the relations described above such as the branching of a Poisson to a Logarithmic distribution[6] or a Poisson distribution weighted by a Gamma distribution[7]. This, however, will always be less efficient than the straightforward cumulative technique.

[6] Generating random numbers from a Poisson distribution with mean $\mu = -k \ln p$ branching to a Logarithmic distribution with parameter p will give a Negative Binomial distribution with parameters k and p.

[7] Taking a Poisson distribution with a mean distributed according to a Gamma distribution with parameters $a = k/\overline{n}$ and $b = k$.

30 Non-central Beta-distribution

30.1 Introduction

The non-central Beta-distribution is given by

$$f(x;p,q) = \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^r}{r!} \frac{x^{p+r-1}(1-x)^{q-1}}{B\left(p+r,q\right)}$$

where p and q are positive real quantities and the non-centrality parameter $\lambda \geq 0$.

In figure 21 we show examples of a non-central Beta distribution with $p = \frac{3}{2}$ and $q = 3$ varying the non-central parameter λ from zero (an ordinary Beta distribution) to ten in steps of two.

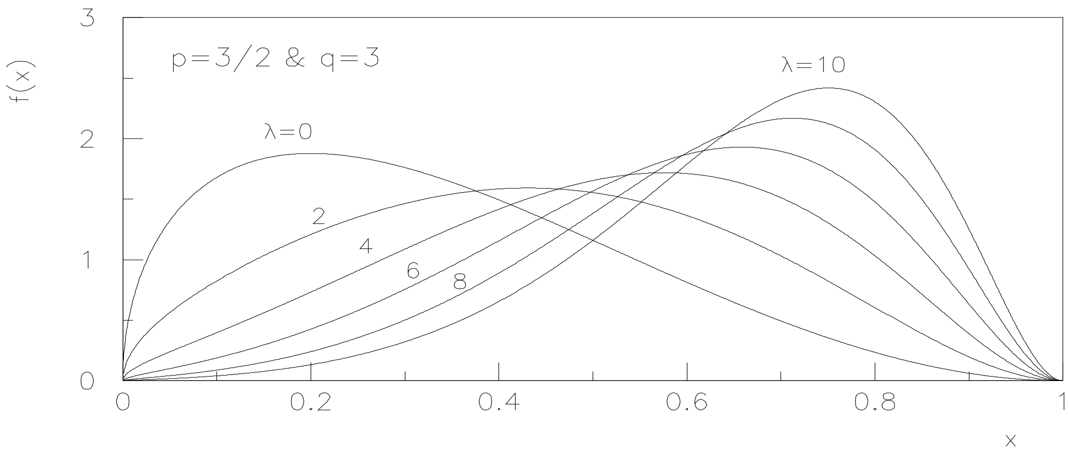


Figure 21: Graph of non-central Beta-distribution for $p = \frac{3}{2}$, $q = 3$ and some values of λ

30.2 Derivation of distribution

If y_m and y_n are two independent variables distributed according to the chi-squared distribution with m and n degrees of freedom, respectively, then the ratio $y_m/(y_m + y_n)$ follows a Beta distribution with parameters $p = \frac{m}{2}$ and $q = \frac{n}{2}$. If instead y_m follows a non-central chi-square distribution we may proceed in a similar way as was done for the derivation of the Beta-distribution (see section 4.2).

We make a change of variables to $x = y_m/(y_m + y_n)$ and $y = y_m + y_n$ which implies that $y_m = xy$ and $y_n = y(1-x)$ obtaining

$$\begin{aligned} f(x,y) &= \left\| \begin{matrix} \frac{\partial y_m}{\partial x} & \frac{\partial y_m}{\partial y} \\ \frac{\partial y_n}{\partial x} & \frac{\partial y_n}{\partial y} \end{matrix} \right\| f(y_m, y_n) = \\ &= \left| \begin{matrix} y & x \\ -y & 1-x \end{matrix} \right| \left\{ \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^r}{r!} \frac{\left(\frac{y_m}{2}\right)^{\frac{m}{2}+r-1} e^{-\frac{y_m}{2}}}{2\Gamma\left(\frac{m}{2}+r\right)} \right\} \left\{ \frac{\left(\frac{y_n}{2}\right)^{\frac{n}{2}-1} e^{-\frac{y_n}{2}}}{2\Gamma\left(\frac{n}{2}\right)} \right\} = \end{aligned}$$

$$
\begin{aligned}
&= y\left\{\sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{\left(\frac{xy}{2}\right)^{\frac{m}{2}+r-1}e^{-\frac{xy}{2}}}{2\Gamma\left(\frac{m}{2}+r\right)}\right\}\left\{\frac{\left(\frac{y(1-x)}{2}\right)^{\frac{n}{2}-1}e^{-\frac{y(1-x)}{2}}}{2\Gamma\left(\frac{n}{2}\right)}\right\} = \\
&= \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{x^{\frac{m}{2}+r-1}(1-x)^{\frac{n}{2}-1}}{B\left(\frac{m}{2}+r,\frac{n}{2}\right)}\left\{\frac{\left(\frac{y}{2}\right)^{\frac{m+n}{2}+r-1}e^{-\frac{y}{2}}}{2\Gamma\left(\frac{m+n}{2}+r\right)}\right\}
\end{aligned}
$$

In the last braces we see a chi-square distribution in y with $m+n+2r$ degrees of freedom and integrating $f(x,y)$ over y in order to get the marginal distribution in x gives us the non-central Beta-distribution as given above with $p = m/2$ and $q = n/2$.

If instead y_n were distributed as a non-central chi-square distribution we would get a very similar expression (not amazing since $y_m/(y_m+y_n) = 1 - y_n/(y_m+y_n)$) but it's the form obtained when y_m is non-central, that is normally referred to as the non-central Beta-distribution.

30.3 Moments

Algebraic moments of the non-central Beta-distribution are given in terms of the hypergeometric function ${}_2F_2$ as

$$
\begin{aligned}
E(x^k) &= \int_0^1 x^k f(x;p,q)dx = \int_0^1 \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{x^{p+r+k-1}(1-x)^{q-1}}{B(p+r,q)}dx = \\
&= \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{B(p+r+k,q)}{B(p+r,q)} = \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{\Gamma(p+r+k)}{\Gamma(p+r)}\frac{\Gamma(p+r+q)}{\Gamma(p+r+q+k)} = \\
&= \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{(p+r+k-1)\cdots(p+r+1)(p+r)}{(p+q+r+k-1)\cdots(p+q+r+1)(p+q+r)} = \\
&= e^{-\frac{\lambda}{2}}\cdot\frac{\Gamma(p+k)}{\Gamma(p)}\cdot\frac{\Gamma(p+q)}{\Gamma(p+q+k)}\cdot {}_2F_2\left(p+q,p+k;p,p+q+k;\tfrac{\lambda}{2}\right)
\end{aligned}
$$

However, to evaluate the hypergeometric function involves a summation so it is more efficient to directly use the penultimate expression above.

30.4 Cumulative distribution

The cumulative distribution is found by straightforward integration

$$
F(x) = \int_0^x \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}\frac{u^{p+r-1}(1-u)^{q-1}}{B(p+r,q)}du = \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}I_x(p+r,q)
$$

30.5 Random Number Generation

Random numbers from a non-central Beta-distribution with integer or half-integer $p-$ and $q-$values is easily obtained using the definition above *i.e.* by using a random number from a non-central chi-square distribution and another from a (central) chi-square distribution.

31 Non-central Chi-square Distribution

31.1 Introduction

If we instead of adding squares of n independent standard normal, $N(0,1)$, variables, giving rise to the chi-square distribution with n degrees of freedom, add squares of $N(\mu_i, 1)$ variables we obtain the *non-central chi-square distribution*

$$f(x;n,\lambda) = \sum_{r=0}^{\infty} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^r}{r!} f(x;n+2r) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{1}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{1}{2}(x+\lambda)} \sum_{r=0}^{\infty} \frac{(\lambda x)^r}{(2r)!} \frac{\Gamma\left(\frac{1}{2}+r\right)}{\Gamma\left(\frac{n}{2}+r\right)}$$

where $\lambda = \sum \mu_i^2$ is the non-central parameter and $f(x;n)$ the ordinary chi-square distribution. As for the latter the variable $x \geq 0$ and the parameter n a positive integer. The additional parameter $\lambda \geq 0$ and in the limit $\lambda = 0$ we retain the ordinary chi-square distribution. According to [2] pp 227–229 the non-central chi-square distribution was first introduced by R. A. Fisher in 1928. In figure 22 we show the distribution for $n = 5$ and non-central parameter $\lambda = 0, 1, 2, 3, 4, 5$ (zero corresponding to the ordinary chi-squared distribution).

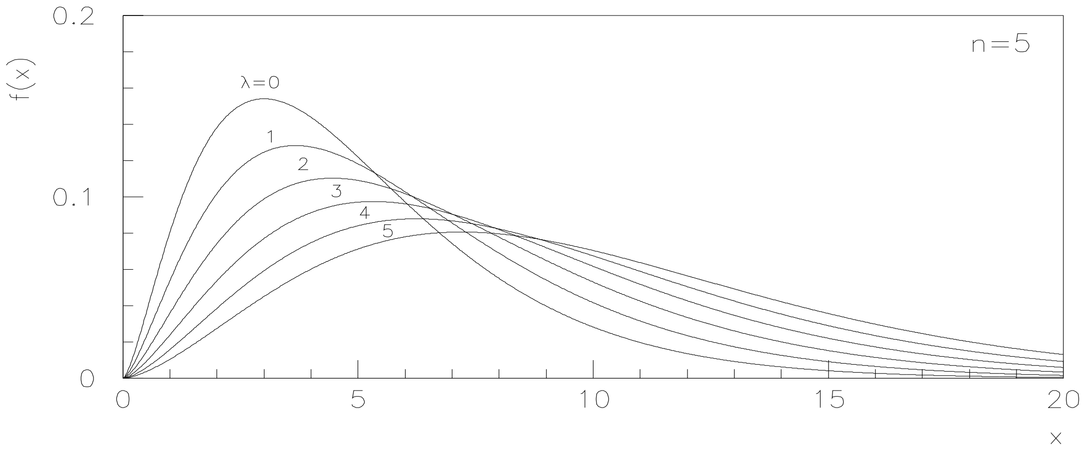


Figure 22: Graph of non-central chi-square distribution for $n = 5$ and some values of λ

31.2 Characteristic Function

The characteristic function for the non-central chi-square distribution is given by

$$\phi(t) = \frac{\exp\left(\frac{\imath t\lambda}{1-2\imath t}\right)}{(1-2\imath t)^{\frac{n}{2}}}$$

but even more useful in determining moments is

$$\ln \phi(t) = \frac{\imath t\lambda}{1-2\imath t} - \frac{n}{2}\ln(1-2\imath t)$$

from which cumulants may be determined in a similar manner as we normally obtain algebraic moments from $\phi(t)$ (see below).

By looking at the characteristic function one sees that the sum of two non-central chi-square variates has the same distribution with degrees of freedoms as well as non-central parameters being the sum of the corresponding parameters for the individual distributions.

31.3 Moments

To use the characteristic function to obtain algebraic moments is not trivial but the cumulants (see section 2.5) are easily found to be given by the formula

$$\kappa_r = 2^{r-1}(r-1)!(n+r\lambda) \quad \text{for} \quad r \geq 1$$

from which we may find the lower order algebraic and central moments (with $a = n + \lambda$ and $b = \lambda/a$) as

$$\begin{aligned}
\mu_1' &= \kappa_1 = a = n + \lambda \\
\mu_2 &= \kappa_2 = 2a(1+b) = 2(n+2\lambda) \\
\mu_3 &= \kappa_3 = 8a(1+2b) = 8(n+3\lambda) \\
\mu_4 &= \kappa_4 + 3\kappa_2^2 = 48(n+4\lambda) + 12(n+2\lambda)^2 \\
\mu_5 &= \kappa_5 + 10\kappa_3\kappa_2 = 384(n+5\lambda) + 160(n+2\lambda)(n+3\lambda) \\
\mu_6 &= \kappa_6 + 15\kappa_4\kappa_2 + 10\kappa_3^2 + 15\kappa_2^3 = \\
&= 3840(n+6\lambda) + 1440(n+2\lambda)(n+4\lambda) + 640(n+3\lambda)^2 + 120(n+2\lambda)^3 \\
\gamma_1 &= \left(\frac{2}{1+b}\right)^{\frac{3}{2}} \cdot \frac{1+2b}{\sqrt{a}} = \frac{8(n+3\lambda)}{[2(n+2\lambda)]^{\frac{3}{2}}} \\
\gamma_2 &= \frac{12}{a} \cdot \frac{1+3b}{(1+b)^2} = \frac{12(n+4\lambda)}{(n+2\lambda)^2}
\end{aligned}$$

31.4 Cumulative Distribution

The cumulative, or distribution, function may be found by

$$\begin{aligned}
F(x) &= \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{1}{2}\right)} e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty} \frac{\lambda^r}{(2r)!} \frac{\Gamma\left(\frac{1}{2}+r\right)}{\Gamma\left(\frac{n}{2}+r\right)} \int_0^x u^{\frac{n}{2}+r-1} e^{-\frac{u}{2}} du = \\
&= \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{1}{2}\right)} e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty} \frac{\lambda^r}{(2r)!} \frac{\Gamma\left(\frac{1}{2}+r\right)}{\Gamma\left(\frac{n}{2}+r\right)} 2^{\frac{n}{2}+r} \gamma\left(\tfrac{n}{2}+r, \tfrac{x}{2}\right) = \\
&= e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty} \frac{\left(\frac{\lambda}{2}\right)^r}{r!} P\left(\tfrac{n}{2}+r, \tfrac{x}{2}\right)
\end{aligned}$$

31.5 Approximations

An approximation to a chi-square distribution is found by equating the first two cumulants of a non-central chi-square distribution with those of ρ times a chi-square distribution.

Here ρ is a constant to be determined. The result is that with

$$\rho = \frac{n+2\lambda}{n+\lambda} = 1 + \frac{\lambda}{n+\lambda} \quad \text{and} \quad n^* = \frac{(n+\lambda)^2}{n+2\lambda} = n + \frac{\lambda^2}{n+2\lambda}$$

we may approximate a non-central chi-square distribution $f(x;n,\lambda)$ with a (central) chi-square distribution in x/ρ with n^* degrees of freedom (n^* in general being fractional).

Approximations to the standard normal distribution are given using

$$z = \sqrt{\frac{2x}{1+b}} - \sqrt{\frac{2a}{1+b} - 1} \quad \text{or} \quad z = \frac{\left(\frac{x}{a}\right)^{\frac{1}{3}} - \left[1 - \frac{2}{9}\cdot\frac{1+b}{a}\right]}{\sqrt{\frac{2}{9}\cdot\frac{1+b}{a}}}$$

31.6 Random Number Generation

Random numbers from a non-central chi-square distribution is easily obtained using the definition above by *e.g.*

- Put $\mu = \sqrt{\lambda/n}$
- Sum n random numbers from a normal distribution with mean μ and variance unity. Note that this is not a unique choice. The only requirement is that $\lambda = \sum \mu_i^2$.
- Return the sum as a random number from a non-central chi-square distribution with n degrees of freedom and non-central parameter λ.

This ought to be sufficient for most applications but if needed more efficient techniques may easily be developed *e.g.* using more general techniques.

32 Non-central F-Distribution

32.1 Introduction

If x_1 is distributed according to a non-central chi-square distribution with m degrees of freedom and non-central parameter λ and x_2 according to a (central) chi-square distribution with n degrees of freedom then, provided x_1 and x_2 are independent, the variable

$$F' = \frac{x_1/m}{x_2/n}$$

is said to have a *non-central F-distribution* with m, n degrees of freedom (positive integers) and non-central parameter $\lambda \geq 0$. As the non-central chi-square distribution it was first discussed by R. A. Fisher in 1928.

This distribution in F' may be written

$$f(F'; m, n, \lambda) = e^{-\frac{\lambda}{2}} \sum_{r=0}^{\infty} \frac{1}{r!} \left(\frac{\lambda}{2}\right)^r \frac{\Gamma\left(\frac{m+n}{2}+r\right)}{\Gamma\left(\frac{m}{2}+r\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}+r} \frac{(F')^{\frac{m}{2}-1+r}}{\left(1+\frac{mF'}{n}\right)^{\frac{1}{2}(m+n)+r}}$$

In figure 23 we show the non-central F-distribution for the case with $m = 10$ and $n = 5$ varying λ from zero (an ordinary, central, F-distribution) to five.

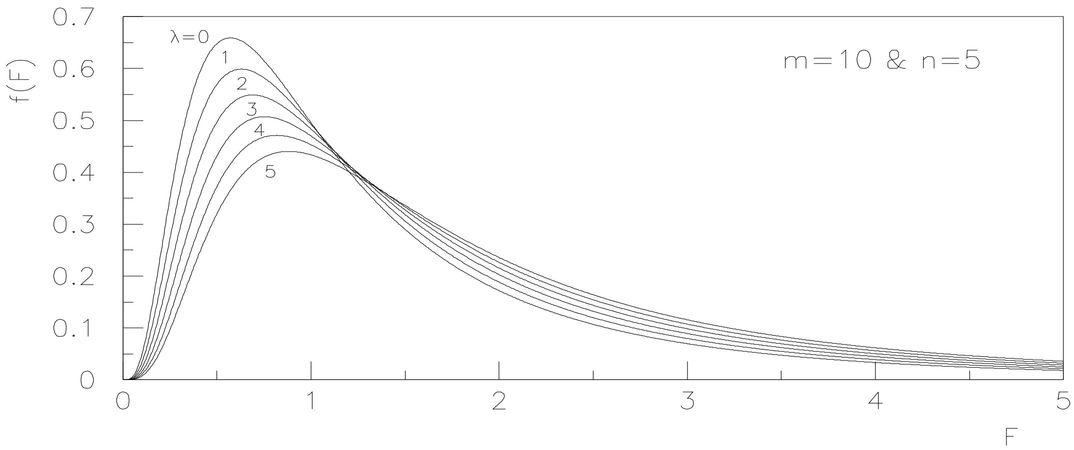


Figure 23: Graph of non-central F-distribution for $m = 10$, $n = 5$ and some values of λ

When $m = 1$ the non-central F-distribution reduces to a non-central t^2-distribution with $\delta^2 = \lambda$. As $n \to \infty$ then nF' approaches a non-central chi-square distribution with m degrees of freedom and non-central parameter λ.

32.2 Moments

Algebraic moments of the non-central F-distribution may be achieved by straightforward, but somewhat tedious, algebra as

$$
\begin{aligned}
E(F'^k) &= \int_0^\infty x^k f(x;m,n,\lambda)dx = \\
&= e^{-\frac{\lambda}{2}}\left(\frac{n}{m}\right)^k \frac{\Gamma\left(\frac{n}{2}-k\right)}{\Gamma\left(\frac{n}{2}\right)} \sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{\lambda}{2}\right)^r \frac{\Gamma\left(\frac{m}{2}+r+k\right)}{\Gamma\left(\frac{m}{2}+r\right)}
\end{aligned}
$$

an expression which may be used to find lower order moments (defined for $n > 2k$)

$$
\begin{aligned}
E(F') &= \frac{n}{m}\cdot\frac{m+\lambda}{n-2} \\
E(F'^2) &= \left(\frac{n}{m}\right)^2 \frac{1}{(n-2)(n-4)}\left\{\lambda^2+(2\lambda+m)(m+2)\right\} \\
E(F'^3) &= \left(\frac{n}{m}\right)^3 \frac{1}{(n-2)(n-4)(n-6)}\cdot\left\{\lambda^3+3(m+4)\lambda^2+(3\lambda+m)(m+4)(m+2)\right\} \\
E(F'^4) &= \left(\frac{n}{m}\right)^4 \frac{1}{(n-2)(n-4)(n-6)(n-8)}\cdot\left\{\lambda^4+4(m+6)\lambda^3+6(m+6)(m+4)\lambda^2+\right. \\
&\qquad \left.+(4\lambda+m)(m+6)(m+4)(m+2)\right\} \\
V(F') &= \left(\frac{n}{m}\right)^2 \frac{2}{(n-2)(n-4)}\left\{\frac{(\lambda+m)^2}{n-2}+2\lambda+m\right\}
\end{aligned}
$$

32.3 Cumulative Distribution

The cumulative, or distribution, function may be found by

$$
\begin{aligned}
F(x) &= \int_0^x u^k f(u;m,n,\lambda)du = \\
&= e^{-\frac{\lambda}{2}}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\lambda}{2}\right)^r \frac{\Gamma\left(\frac{m+n}{2}+r\right)}{\Gamma\left(\frac{m}{2}+r\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{\frac{m}{2}+r}\int_0^x \frac{u^{\frac{m}{2}-1+r+k}}{\left(1+\frac{mu}{n}\right)^{\frac{m+n}{2}+r}}du = \\
&= e^{-\frac{\lambda}{2}}\sum_{r=0}^{\infty}\frac{1}{r!}\left(\frac{\lambda}{2}\right)^r \frac{B_q\left(\frac{m}{2}+r,\frac{n}{2}\right)}{B\left(\frac{m}{2}+r,\frac{n}{2}\right)} = e^{-\frac{\lambda}{2}}\sum_{r=0}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^r}{r!}I_q\left(\frac{m}{2}+r,\frac{n}{2}\right)
\end{aligned}
$$

with

$$q = \frac{\frac{mx}{n}}{1+\frac{mx}{n}}$$

32.4 Approximations

Using the approximation of a non-central chi-square distribution to a (central) chi-square distribution given in the previous section we see that

$$\frac{m}{m+\lambda}F'$$

is approximately distributed according to a (central) F-distribution with $m^* = m + \frac{\lambda^2}{m+2\lambda}$ and n degrees of freedom.

Approximations to the standard normal distribution is achieved with

$$\begin{aligned}
z_1 &= \frac{F' - E(F')}{\sqrt{V(F')}} = \frac{F' - \frac{n(m+\lambda)}{m(n-2)}}{\frac{n}{m}\left[\frac{2}{(n-2)(n-4)}\left\{\frac{(m+\lambda)^2}{n-2} + m + 2\lambda\right\}\right]^{\frac{1}{2}}} \\
\text{or} \quad z_2 &= \frac{\left(\frac{mF'}{m+\lambda}\right)^{\frac{1}{3}}\left(1 - \frac{2}{9n}\right) - \left(1 - \frac{2}{9}\cdot\frac{m+2\lambda}{(m+\lambda)^2}\right)}{\left[\frac{2}{9}\cdot\frac{m+2\lambda}{(m+\lambda)^2} + \frac{2}{9n}\cdot\left(\frac{mF'}{m+\lambda}\right)^{\frac{2}{3}}\right]^{\frac{1}{2}}}
\end{aligned}$$

32.5 Random Number Generation

Random numbers from a non-central chi-square distribution is easily obtained using the definition above *i.e.* by using a random number from a non-central chi-square distribution and another from a (central) chi-square distribution.

33 Non-central t-Distribution

33.1 Introduction

If x is distributed according to a normal distribution with mean δ and variance 1 and y according to a chi-square distribution with n degrees of freedom (independent of x) then

$$t' = \frac{x}{\sqrt{y/n}}$$

has a *non-central t-distribution* with n degrees of freedom (positive integer) and non-central parameter δ (real).

We may also write

$$t' = \frac{z+\delta}{\sqrt{w/n}}$$

where z is a standard normal variate and w is distributed as a chi-square variable with n degrees of freedom.

The distribution is given by (see comments on derivation in section below)

$$f(t'; n, \delta) = \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \sum_{r=0}^{\infty} \frac{(t'\delta)^r}{r! n^{\frac{r}{2}}} \left(1 + \frac{t'^2}{n}\right)^{-\frac{n+r+1}{2}} 2^{\frac{r}{2}} \Gamma\left(\frac{n+r+1}{2}\right)$$

In figure 24 we show the non-central t-distribution for the case with $n = 10$ varying δ from zero (an ordinary t-distribution) to five.

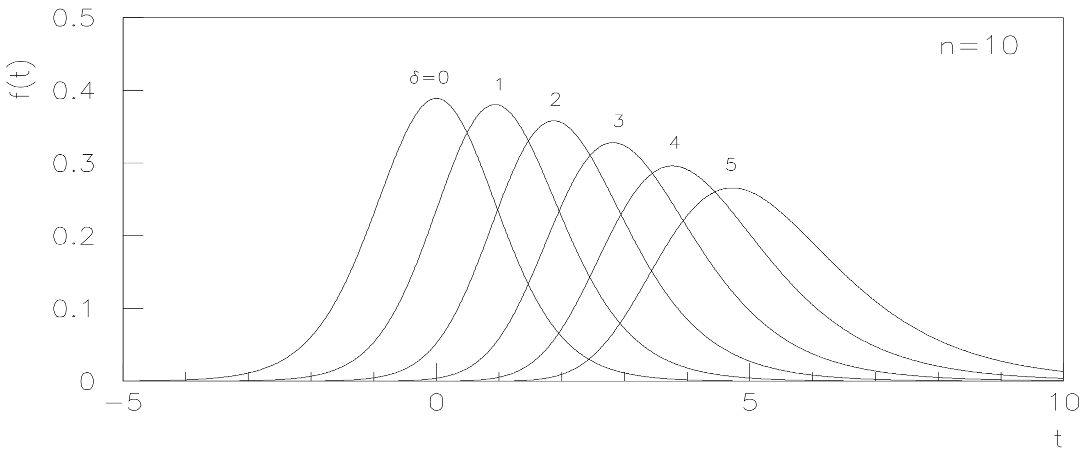


Figure 24: Graph of non-central t-distribution for $n = 10$ and some values of δ

This distribution is of importance in hypotheses testing if we are interested in the probability of committing a Type II error implying that we would accept an hypothesis although it was wrong, see discussion in section 38.11 on page 147.

33.2 Derivation of distribution

Not many text-books include a formula for the non-central t-distribution and some turns out to give erroneous expressions. A non-central F-distribution with $m = 1$ becomes a non-central t^2-distribution which then may be transformed to a non-central t-distribution. However, with this approach one easily gets into trouble for $t' < 0$. Instead we adopt a technique very similar to what is used in section 38.6 to obtain the normal (central) t-distribution from a t-ratio.

The difference in the non-central case is the presence of the δ-parameter which introduces two new exponential terms in the equations to be solved. One is simply $\exp(-\delta^2/2)$ but another factor we treat by a serial expansion leading to the p.d.f. above. This may not be the 'best' possible expression but empirically it works quite well.

33.3 Moments

With some effort the p.d.f. above may be used to calculate algebraic moments of the distribution yielding

$$E(t'^k) = \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)}\Gamma\left(\tfrac{n-k}{2}\right) n^{\frac{k}{2}} \sum_{r=0}^{\infty} \frac{\delta^r 2^{\frac{r}{2}}}{r!}\Gamma\left(\tfrac{r+k+1}{2}\right)$$

where the sum should be made for odd (even) values of r if k is odd (even). This gives for low orders

$$\begin{aligned}
\mu'_1 &= \sqrt{\frac{n}{2}}\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\delta \\
\mu'_2 &= \frac{n\Gamma\left(\frac{n-2}{2}\right)}{2\Gamma\left(\frac{n}{2}\right)}\left(1+\delta^2\right) = \frac{n}{n-2}\left(1+\delta^2\right) \\
\mu'_3 &= \frac{n^{\frac{3}{2}}\sqrt{2}\Gamma\left(\frac{n-3}{2}\right)}{4\Gamma\left(\frac{n}{2}\right)}\delta\left(3+\delta^2\right) \\
\mu'_4 &= \frac{n^2\Gamma\left(\frac{n-4}{2}\right)}{4\Gamma\left(\frac{n}{2}\right)}\left(\delta^4+6\delta^2+3\right) = \frac{n^2}{(n-2)(n-4)}\left(\delta^4+6\delta^2+3\right)
\end{aligned}$$

from which expressions for central moments may be found *e.g.* the variance

$$\mu_2 = V(t') = \frac{n}{2}\left\{\left(1+\delta^2\right)\frac{2}{n-2} - \delta^2\left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2\right\}$$

33.4 Cumulative Distribution

The cumulative, or distribution, function may be found by

$$F(t) = \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\sum_{r=0}^{\infty}\frac{\delta^r}{r!n^{\frac{r}{2}}}2^{\frac{r}{2}}\Gamma\left(\tfrac{n+r+1}{2}\right)\int_{-\infty}^{t}\frac{u^r}{\left(1+\frac{u^2}{n}\right)^{\frac{n+r+1}{2}}}du =$$

$$= \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)} \sum_{r=0}^{\infty} \frac{\delta^r}{r!} 2^{\frac{r}{2}-1} \Gamma\left(\tfrac{n+r+1}{2}\right) \left\{ s_1 B\left(\tfrac{r+1}{2}, \tfrac{n}{2}\right) + s_2 B_q\left(\tfrac{r+1}{2}, \tfrac{n}{2}\right) \right\} =$$

$$= \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{\pi}} \sum_{r=0}^{\infty} \frac{\delta^r}{r!} 2^{\frac{r}{2}-1} \Gamma\left(\tfrac{r+1}{2}\right) \left\{ s_1 + s_2 I_q\left(\tfrac{r+1}{2}, \tfrac{n}{2}\right) \right\}$$

where s_1 are and s_2 are signs differing between cases with positive or negative t as well as odd or even r in the summation. The sign s_1 is -1 if r is odd and $+1$ if it is even while s_2 is $+1$ unless $t < 0$ and r is even in which case it is -1.

33.5 Approximation

An approximation is given by

$$z = \frac{t'\left(1 - \frac{1}{4n}\right) - \delta}{\sqrt{1 + \frac{t'^2}{2n}}}$$

which is asymptotically distributed as a standard normal variable.

33.6 Random Number Generation

Random numbers from a non-central t-distribution is easily obtained using the definition above *i.e.* by using a random number from a normal distribution and another from a chi-square distribution. This ought to be sufficient for most applications but if needed more efficient techniques may easily be developed *e.g.* using more general techniques.

34 Normal Distribution

34.1 Introduction

The normal distribution or, as it is often called, the Gauss distribution is the most important distribution in statistics. The distribution is given by

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

where μ is a location parameter, equal to the mean, and σ the standard deviation. For $\mu = 0$ and $\sigma = 1$ we refer to this distribution as the *standard normal distribution*. In many connections it is sufficient to use this simpler form since μ and σ simply may be regarded as a shift and scale parameter, respectively. In figure 25 we show the standard normal distribution.

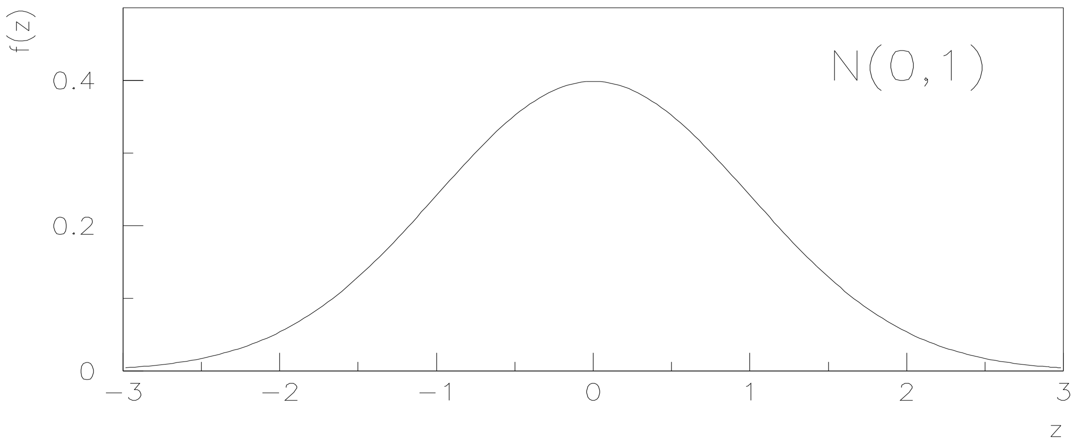


Figure 25: Standard normal distribution

Below we give some useful information in connection with the normal distribution. Note, however, that this is only a minor collection since there is no limit on important and interesting statistical connections to this distribution.

34.2 Moments

The expectation value of the distribution is $E(x) = \mu$ and the variance $V(x) = \sigma^2$.

Generally odd central moments vanish due to the symmetry of the distribution and even central moments are given by

$$\mu_{2r} = \frac{(2r)!}{2^r r!}\sigma^{2r} = (2r-1)!!\sigma^{2r}$$

for $r \geq 1$.

It is sometimes also useful to evaluate absolute moments $E(|x|^n)$ for the normal distribution. To do this we make use of the integral

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

which if differentiated k times with respect to a yields

$$\int_{-\infty}^{\infty} x^{2k} e^{-ax^2} dx = \frac{(2k-1)!!}{2^k} \sqrt{\frac{\pi}{a^{2k+1}}}$$

In our case $a = 1/2\sigma^2$ and since even absolute moments are identical to the algebraic moments it is enough to evaluate odd absolute moments for which we get

$$E(|x|^{2k+1}) = \frac{2}{\sigma\sqrt{2\pi}} \int_0^{\infty} x^{2k+1} e^{-\frac{x^2}{2\sigma^2}} dx = \sqrt{\frac{2}{\pi}} \frac{(2\sigma^2)^{k+1}}{2\sigma} \int_0^{\infty} y^k e^{-y} dy$$

The last integral we recognize as being equal to $k!$ and we finally obtain the absolute moments of the normal distribution as

$$E(|x|^n) = \begin{cases} (n-1)!!\sigma^n & \text{for } n = 2k \\ \sqrt{\frac{2}{\pi}} 2^k k! \sigma^{2k+1} & \text{for } n = 2k+1 \end{cases}$$

The half-width at half-height of the normal distribution is given by $\sqrt{2\ln 2}\sigma \approx 1.177\sigma$ which may be useful to remember when estimating σ using a ruler.

34.3 Cumulative Function

The distribution function, or cumulative function, may be expressed in term of the incomplete gamma function P as

$$F(z) = \begin{cases} \frac{1}{2} + \frac{1}{2} P\left(\frac{1}{2}, \frac{z^2}{2}\right) & \text{if } z \geq 0 \\ \frac{1}{2} - \frac{1}{2} P\left(\frac{1}{2}, \frac{z^2}{2}\right) & \text{if } z < 0 \end{cases}$$

or we may use the *error function* $\mathrm{erf}(z/\sqrt{2})$ in place of the incomplete gamma function.

34.4 Characteristic Function

The characteristic function for the normal distribution is easily found from the general definition

$$\phi(t) = E\left(e^{\imath t x}\right) = \exp\left\{\mu \imath t - \tfrac{1}{2}\sigma^2 t^2\right\}$$

34.5 Addition Theorem

The so called *Addition theorem* for normally distributed variables states that any linear combination of independent normally distributed random variables x_i ($i = 1, 2, \ldots, n$) is also distributed according to the normal distribution.

If each x_i is drawn from a normal distribution with mean μ_i and variance σ_i^2 then regard the linear combination

$$\mathcal{S} = \sum_{i=1}^{n} a_i x_i$$

where a_i are real coefficients. Each term $a_i x_i$ has characteristic function

$$\phi_{a_i x_i}(t) = \exp\left\{(a_i\mu_i)\imath t - \tfrac{1}{2}(a_i^2\sigma_i^2)t^2\right\}$$

and thus $\mathcal{S}$ has characteristic function

$$\phi_{\mathcal{S}}(t) = \prod_{i=1}^{n} \phi_{a_i x_i}(t) = \exp\left\{\left(\sum_{i=1}^{n} a_i\mu_i\right)\imath t - \tfrac{1}{2}\left(\sum_{i=1}^{n} a_i^2\sigma_i^2\right)t^2\right\}$$

which is seen to be a normal distribution with mean $\sum a_i\mu_i$ and variance $\sum a_i^2\sigma_i^2$.

34.6 Independence of $\overline{x}$ and s^2

A unique property of the normal distribution is the independence of the sample statistics $\overline{x}$ and s^2, estimates of the mean and variance of the distribution. Recall that the definition of these quantities are

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad s^2 = \frac{1}{n-1}\sum_{i=0}^{n}(x_i - \overline{x})^2$$

where $\overline{x}$ is an estimator of the true mean μ and s^2 is the usual unbiased estimator for the true variance σ^2.

For a population of n events from a normal distribution $\overline{x}$ has the distribution $N(\mu, \sigma^2/n)$ and $(n-1)s^2/\sigma^2$ is distributed according to a chi-square distribution with $n-1$ degrees of freedom. Using the relation

$$\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^2 = \frac{(n-1)s^2}{\sigma^2} + \left(\frac{\overline{x} - \mu}{\sigma/\sqrt{n}}\right)^2$$

and creating the joint characteristic function for the variables $(n-1)s^2/\sigma^2$ and $(\sqrt{n}(\overline{x} - \mu)/\sigma^2)^2$ one may show that this function factorizes thus implying independence of these quantities and thus also of $\overline{x}$ and s^2.

In summary the *"independence theorem"* states that given n independent random variables with identical normal distributions the two statistics $\overline{x}$ and s^2 are independent. Also conversely it holds that if the mean $\overline{x}$ and the variance s^2 of a random sample are independent then the population is normal.

34.7 Probability Content

The probability content of the normal distribution is often referred to in statistics. When the term *one standard deviation* is mentioned one immediately thinks in terms of a probability content of 68.3% within the symmetric interval from the value given.

Without loss of generality we may treat the standard normal distribution only since the transformation from a more general case is straightforward putting $z = (x-\mu)/\sigma$. In different situation one may want to find

- the probability content, two-side or one-sided, to exceed a certain number of standard deviations, or
- the number of standard deviations corresponding to a certain probability content.

In calculating this we need to evaluate integrals like

$$\alpha = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

There are no explicit solution to this integral but it is related to the *error function* (see section 13) as well as the *incomplete gamma function* (see section 42).

$$\frac{1}{\sqrt{2\pi}} \int\limits_{-z}^{z} e^{-\frac{z^2}{2}} = \text{erf}\left(\frac{z}{\sqrt{2}}\right) = P\left(\frac{1}{2}, \frac{z^2}{2}\right)$$

These relations may be used to calculate the probability content. Especially the error function is often available as a system function on different computers. Beware, however, that it seems to be implemented such that erf(z) is the symmetric integral from $-z$ to z and thus the $\sqrt{2}$ factor should not be supplied. Besides from the above relations there are also excellent approximations to the integral which may be used.

In the tables below we give the probability content for exact z-values (left-hand table) as well as z-values for exact probability contents (right-hand table).

z	$\int\limits_{-\infty}^{z}$	$\int\limits_{-z}^{z}$	$\int\limits_{z}^{\infty}$
0.0	0.50000	0.00000	0.50000
0.5	0.69146	0.38292	0.30854
1.0	0.84134	0.68269	0.15866
1.5	0.93319	0.86639	0.06681
2.0	0.97725	0.95450	0.02275
2.5	0.99379	0.98758	$6.210 \cdot 10^{-3}$
3.0	0.99865	0.99730	$1.350 \cdot 10^{-3}$
3.5	0.99977	0.99953	$2.326 \cdot 10^{-4}$
4.0	0.99997	0.99994	$3.167 \cdot 10^{-5}$
4.5	1.00000	0.99999	$3.398 \cdot 10^{-6}$
5.0	1.00000	1.00000	$2.867 \cdot 10^{-7}$
6.0	1.00000	1.00000	$9.866 \cdot 10^{-10}$
7.0	1.00000	1.00000	$1.280 \cdot 10^{-12}$
8.0	1.00000	1.00000	$6.221 \cdot 10^{-16}$

z	$\int\limits_{-\infty}^{z}$	$\int\limits_{-z}^{z}$	$\int\limits_{z}^{\infty}$
0.00000	0.5	0.0	0.5
0.25335	0.6	0.2	0.4
0.67449	0.75	0.5	0.25
0.84162	0.8	0.6	0.2
1.28155	0.9	0.8	0.1
1.64485	0.95	0.9	0.05
1.95996	0.975	0.95	0.025
2.32635	0.99	0.98	0.01
2.57583	0.995	0.99	0.005
3.09023	0.999	0.998	0.001
3.29053	0.9995	0.999	0.0005
3.71902	0.9999	0.9998	0.0001
3.89059	0.99995	0.9999	0.00005
4.26489	0.99999	0.99998	0.00001

It is sometimes of interest to scrutinize extreme significance levels which implies integrating the far tails of a normal distribution. In the table below we give the number of standard deviations, z, required in order to achieve a one-tailed probability content of 10^{-n}.

z-values for which $\frac{1}{\sqrt{2\pi}}\int\limits_z^\infty e^{-z^2/2}dz = 10^{-n}$ for $n = 1, 2, \ldots, 23$

n	z	n	z	n	z	n	z	n	z
1	1.28155	6	4.75342	11	6.70602	16	8.22208	21	9.50502
2	2.32635	7	5.19934	12	7.03448	17	8.49379	22	9.74179
3	3.09023	8	5.61200	13	7.34880	18	8.75729	23	9.97305
4	3.71902	9	5.99781	14	7.65063	19	9.01327		
5	4.26489	10	6.36134	15	7.94135	20	9.26234		

Below are also given the one-tailed probability content for a standard normal distribution in the region from z to ∞ (or $-\infty$ to $-z$). The information in the previous as well as this table is taken from [25].

Probability content $Q(z) = \frac{1}{\sqrt{2\pi}}\int\limits_z^\infty e^{-z^2/2}dz$ for $z = 1, 2, \ldots, 50, 60, \ldots, 100, 150, \ldots, 500$

z	$-\log Q(z)$	z	$-\log Q(z)$	z	$-\log Q(z)$	z	$-\log Q(z)$	z	$-\log Q(z)$
1	0.79955	14	44.10827	27	160.13139	40	349.43701	80	1392.04459
2	1.64302	15	50.43522	28	172.09024	41	367.03664	90	1761.24604
3	2.86970	16	57.19458	29	184.48283	42	385.07032	100	2173.87154
4	4.49934	17	64.38658	30	197.30921	43	403.53804	150	4888.38812
5	6.54265	18	72.01140	31	210.56940	44	422.43983	200	8688.58977
6	9.00586	19	80.06919	32	224.26344	45	441.77568	250	13574.49960
7	11.89285	20	88.56010	33	238.39135	46	461.54561	300	19546.12790
8	15.20614	21	97.48422	34	252.95315	47	481.74964	350	26603.48018
9	18.94746	22	106.84167	35	267.94888	48	502.38776	400	34746.55970
10	23.11805	23	116.63253	36	283.37855	49	523.45999	450	43975.36860
11	27.71882	24	126.85686	37	299.24218	50	544.96634	500	54289.90830
12	32.75044	25	137.51475	38	315.53979	60	783.90743		
13	38.21345	26	148.60624	39	332.27139	70	1066.26576		

Beware, however, that extreme significance levels are purely theoretical and that one seldom or never should trust experimental limits at these levels. In an experimental situations one rarely fulfills the statistical laws to such detail and any bias or background may heavily affect statements on extremely small probabilities.

Although one normally would use a routine to find the probability content for a normal distribution it is sometimes convenient to have a "classical" table available. In table 6 on page 178 we give probability contents for a symmetric region from $-z$ to z for z-values ranging from 0.00 to 3.99 in steps of 0.01. Conversely we give in table 7 on page 179 the z-values corresponding to specific probability contents from 0.000 to 0.998 in steps of 0.002.

34.8 Random Number Generation

There are many different methods to obtain random numbers from a normal distribution some of which are reviewed below. It is enough to consider the case of a standard normal distribution since given such a random number z we may easily obtain one from a general normal distribution by making the transformation $x = \mu + \sigma z$.

Below $f(x)$ denotes the standard normal distribution and if not explicitly stated all variables denoted by ξ are uniform random numbers in the range from zero to one.

34.8.1 Central Limit Theory Approach

The sum of n independent random numbers from a uniform distribution between zero and one, R_n, has expectation value $E(R_n) = n/2$ and variance $V(R_n) = n/12$. By the central limit theorem the quantity

$$z_n = \frac{R_n - E(R_n)}{\sqrt{V(n)}} = \frac{R_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$$

approaches the standard normal distribution as $n \to \infty$. A practical choice is $n = 12$ since this expression simplifies to $z_{12} = R_{12} - 6$ which could be taken as a random number from a standard normal distribution. Note, however, that this method is neither accurate nor fast.

34.8.2 Exact Transformation

The Box-Muller transformation used to find random numbers from the binormal distribution (see section 6.5 on page 22), using two uniform random numbers between zero and one in ξ_1 and ξ_2,

$$\begin{aligned} z_1 &= \sqrt{-2\ln x_1}\sin 2\pi x_2 \\ z_2 &= \sqrt{-2\ln x_1}\cos 2\pi x_2 \end{aligned}$$

may be used to obtain two independent random numbers from a standard normal distribution.

34.8.3 Polar Method

The above method may be altered in order to avoid the cosine and sine by

i Generate u and v as two uniformly distributed random numbers in the range from -1 to 1 by $u = 2\xi_1 - 1$ and $v = 2\xi_2 - 1$.

ii Calculate $w = u^2 + v^2$ and if $w > 1$ then go back to **i**.

iii Return $x = uz$ and $y = vz$ with $z = \sqrt{-2\ln w/w}$ as two independent random numbers from a standard normal distribution.

This method is often faster than the previous since it eliminates the sine and cosine at the slight expense of $1 - \pi/4 \approx 21\%$ rejection in step **iii** and a few more arithmetic operations. As is easily seen $u/\sqrt{w}$ and $v/\sqrt{w}$ plays the role of the cosine and the sine in the previous method.

34.8.4 Trapezoidal Method

The maximum trapezoid that may be inscribed under the standard normal curve covers an area of 91.95% of the total area. Random numbers from a trapezoid is easily obtained by a linear combination of two uniform random numbers. In the remaining cases a tail-technique and accept-reject techniques, as described in figure 26, are used.

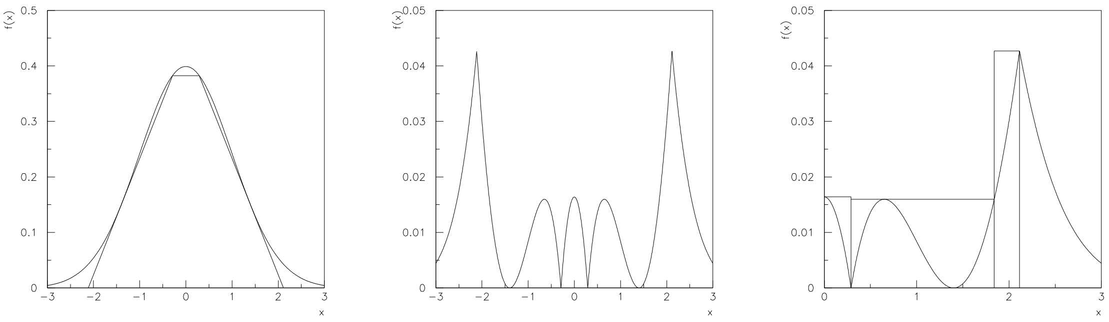


Figure 26: Trapezoidal method

Below we describe, in some detail, a slightly modified version of what is presented in [28]. For more exact values of the constants used see this reference.

i Generate two uniform random numbers between zero and one ξ and ξ_0

ii If $\xi < 0.9195$ generate a random number from the trapezoid by $x = 2.404\xi_0 + 1.984\xi - 2.114$ and exit

iii Else if $\xi < 0.9541$ (3.45% of all cases) generate a random number from the tail $x > 2.114$

- **a** Generate two uniform random numbers ξ_1 and ξ_2
- **b** Put $x = 2.114^2 - 2\ln\xi_1$ and if $x\xi_2^2 > 2.114^2$ then go back to **a**
- **c** Put $x = \sqrt{x}$ and go to **vii**

iv Else if $\xi < 0.9782$ (2.41% of all cases) generate a random number from the region $0.290 < x < 1.840$ between the normal curve and the trapezoid

- **a** Generate two uniform random numbers ξ_1 and ξ_2
- **b** Put $x = 0.290 + 1.551\xi_1$ and if $f(x) - 0.443 + 0.210x < 0.016\xi_2$ then go to **a**

c Go to **vii**

v Else if $\xi < 0.9937$ (1.55% of all cases) generate a random number from the region $1.840 < \xi < 2.114$ between the normal curve and the trapezoid

a Generate two uniform random numbers ξ_1 and ξ_2

b Put $x = 1.840 + 0.274\xi_1$ and if $f(x) - 0.443 + 0.210x < 0.043\xi_2$ then go to **a**

c Go to **vii**

vi Else, in 0.63% of all cases, generate a random number from the region $0 < x < 0.290$ between the normal curve and the trapezoid by

a Generate two uniform random numbers ξ_1 and ξ_2

b Put $x = 0.290\xi_1$ and if $f(x) - 0.383 < 0.016\xi_2$ then go back to **a**

vii Assign a minus sign to x if $\xi_0 \geq \frac{1}{2}$

34.8.5 Center-tail method

Ahrens and Dieter [28] also proposes a so called center-tail method. In their article they treat the tails outside $|z| > \sqrt{2}$ with a special tail method which avoids the logarithm. However, it turns out that using the same tail method as in the previous method is even faster. The method is as follows:

i Generate a uniform random number ξ and use the first bit after the decimal point as a sign bit s *i.e.* for $\xi \leq \frac{1}{2}$ put $\xi = 2\xi$ and $s = -1$ and for $\xi > \frac{1}{2}$ put $\xi = 2\xi - 1$ and $s = 1$

ii If $\xi > 0.842700792949715$ (the area for $-\sqrt{2} < z < \sqrt{2}$) go to **vi**.

iii Center method: Generate ξ_0 and set $\nu = \xi + 0$

iv Generate ξ_1 and ξ_2 and set $\nu^* = max(\xi_1, \xi_2)$.
If $\nu < \nu^*$ calculate $y = \xi_0\sqrt{2}$ and go to **viii**

v Generate ξ_1 and ξ_2 and set $\nu = max(\xi_1, \xi_2)$
If $\nu < \nu^*$ go to **iv** else go to **iii**

vi Tail method: Generate ξ_1 and set $y = 1 - \ln \xi_1$

vii Generate ξ_2 and if $y\xi_2^2 > 1$ go to **vi** else put $y = \sqrt{y}$

viii Set $x = sy\sqrt{2}$.

34.8.6 Composition-rejection Methods

In reference [20] two methods using the composition-rejection method is proposed. The first one, attributed to Butcher [22] and Kahn, uses only one term in the sum and has $\alpha = \sqrt{2e/\pi}$, $f(x) = \exp\{-x\}$ and $g(x) = \exp\{-(x-1)^2/2\}$. The algorithm is as follows:

i Generate ξ_1 and ξ_2

ii Determine $x = -\ln \xi_1$, *i.e.* a random number from $f(x)$

iii Determine $g(x) = \exp\{-(x-1)^2/2\}$

iv If $\xi_2 > g(x)$ then go to **i**

v Decide a sign either by generating a new random number, or by using ξ_2 for which $0 < \xi_2 \leq g(x)$ here, and exit with x with this sign.

The second method is originally proposed by J. C. Butcher [22] and uses two terms

$$
\begin{array}{llll}
\alpha_1 = \sqrt{\frac{2}{\pi}} & f_1(x) = 1 & g_1(x) = e^{-\frac{x^2}{2}} & \text{for } 0 \leq x \leq 1 \\
\alpha_2 = 1/\sqrt{2\pi} & f_2(x) = 2e^{-2(x-1)} & g_2(x) = e^{-\frac{(x-2)^2}{2}} & \text{for } x > 1
\end{array}
$$

i Generate ξ_1 and ξ_2

ii If $\xi_1 - \frac{2}{3} > 0$ then determine $x = 1 - \frac{1}{2}\ln(3\xi_1 - 2)$ and $z = \frac{1}{2}(x-2)^2$ else determine $x = 3\xi_1/2$ and $z = x^2/2$

iii Determine $g = e^{-z}$

iv If $\xi_2 > g$ the go to **i**

v Determine the sign of $\xi_2 - g/2$ and exit with x with this sign.

34.8.7 Method by Marsaglia

A nice method proposed by G. Marsaglia is based on inscribing a spline function beneath the standard normal curve and subsequently a triangular distribution beneath the remaining difference. See figure 27 for a graphical presentation of the method. The algorithm used is described below.

- The sum of three uniform random numbers ξ_1, ξ_2, and ξ_3 follow a parabolic spline function. Using $x = 2(\xi_1 + \xi_2 + \xi_3 - \frac{3}{2})$ we obtain a distribution

$$
f_1(x) = \begin{cases} (3-x^2)/8 & \text{if} \quad |x| \leq 1 \\ (3-|x|)^2/16 & \text{if} \quad 1 < |x| \leq 3 \\ 0 & \text{if} \quad |x| > 3 \end{cases}
$$

  Maximizing α_1 with the constraint $f(x) - \alpha_1 f_1(x) \geq 0$ in the full interval $|x| \leq 3$ gives $\alpha_1 = 16e^{-2}/\sqrt{2\pi} \approx 0.8638554$ *i.e.* in about 86% of all cases such a combination is made.

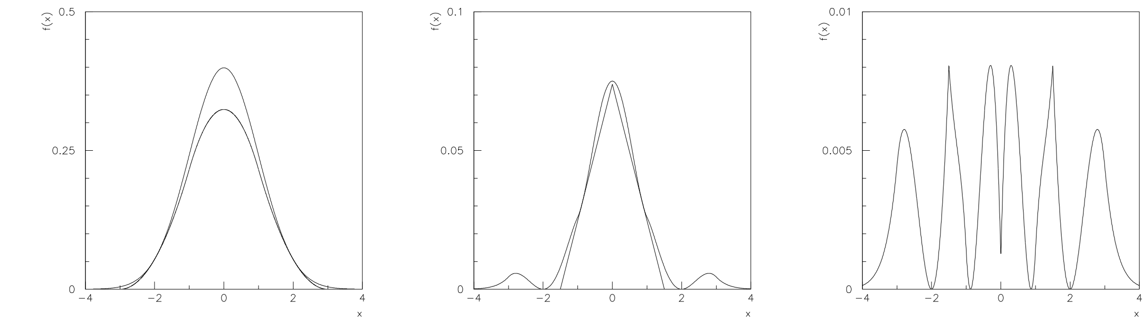


Figure 27: Marsaglia method

- Moreover, a triangular distribution given by making the combination $x = \frac{3}{2}(\xi_1+\xi_2-1)$ leading to a function

$$f_2(x) = \frac{4}{9}\left(\frac{3}{2} - |x|\right) \quad \text{for} \quad |x| < \frac{3}{2}$$

  and zero elsewhere. This function may be inscribed under the remaining curve $f(x) - f_1(x)$ maximizing α_2 such that $f_3(x) = f(x) - \alpha_1 f_1(x) - \alpha_2 f_2(x) \geq 0$ in the interval $|x| \leq \frac{3}{2}$. This leads to a value $\alpha_2 \approx 0.1108$ *i.e.* in about 11% of all cases this combination is used

- The maximum value of $f_3(x)$ in the region $|x| \leq 3$ is 0.0081 and here we use a straightforward reject-accept technique. This is done in about 2.26% of all cases.

- Finally, the tails outside $|x| > 3$, covering about 0.27% of the total area is dealt with with a standard tail-method where

  a Put $x = 9 - 2\ln\xi_1$

  b If $x\xi_2^2 > 9$ then go to **a**

  c Else generate a sign $s = +1$ or $s = -1$ with equal probability and exit with $x = s\sqrt{x}$

34.8.8 Histogram Technique

Yet another method due to G. Marsaglia and collaborators [35] is one where a histogram with k bins and bin-width c is inscribed under the (folded) normal curve. The difference between the normal curve and the histogram is treated with a combination of triangular distributions and accept-reject techniques as well as the usual technique for the tails. Trying to optimize fast generation we found $k = 9$ and $c = \frac{1}{3}$ to be a fair choice. This may, however, not be true on all computers. See figure 28 for a graphical presentation of the method.

The algorithm is as follows:

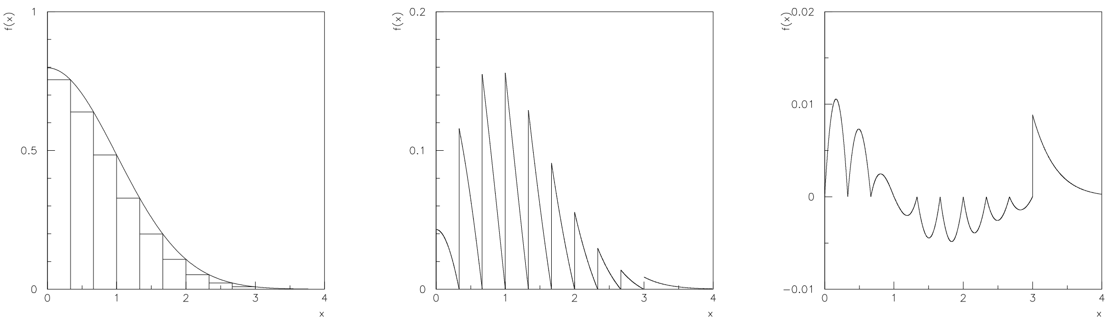


Figure 28: Histogram method

i Generate ξ_1 and chose which region to generate from. This is done *e.g.* with a sequential search in a cumulative vector where the areas of the regions have been sorted in descending order. The number of elements in this vector is $2k + \frac{1}{c} + 1$ which for the parameters mentioned above becomes 22.

ii If a histogram bin i ($i = 1, 2, \ldots, k$) is selected then determine $x = (\xi_2 + i - 1)c$ and go to **vii**.

iii If an inscribed triangle i ($i = 1, 2, \ldots, \frac{1}{c}$) then determine $x = (min(\xi_2, \xi_3) + 1 - i)c$ and go to **vii**.

iv If subscribed triangle i ($i = \frac{1}{c} + 1, \ldots, k$) then determine $x = (min(\xi_2, \xi_3) + i - 1)c$ and accept this value with a probability equal to the ratio between the normal curve and the triangle at this x-value (histogram subtracted in both cases) else iterate. When a value is accepted then go to **vii**.

v For the remaining $\frac{1}{c}$ regions between the inscribed triangles and the normal curve for $x < 1$ use a standard reject accept method in each bin and then go to **vii**.

vi If the tail region is selected then use a standard technique *e.g.* **(a)** $x = (kc)^2 - 2\ln\xi_2$, **(b)** if $x\xi_3^2 > (kc)^2$ then go to **a** else use $x = \sqrt{x}$.

vii Attach a random sign to x and exit. This is done by either generating a new uniform random number or by saving the first bit of ξ_1 in step **i**. The latter is faster and the degradation in precision is negligible.

34.8.9 Ratio of Uniform Deviates

A technique using the ratio of two uniform deviates was propose by A. J. Kinderman and J. F. Monahan in 1977 [36]. It is based on selecting an acceptance region such that the ratio of two uniform pseudorandom numbers follow the standard normal distribution. With u

and v uniform random numbers, u between 0 and 1 and v between $-\sqrt{2/e}$ and $\sqrt{2/e}$, such a region is defined by

$$v^2 < -4\; u^2\; \ln u$$

as is shown in the left-hand side of figure 29.

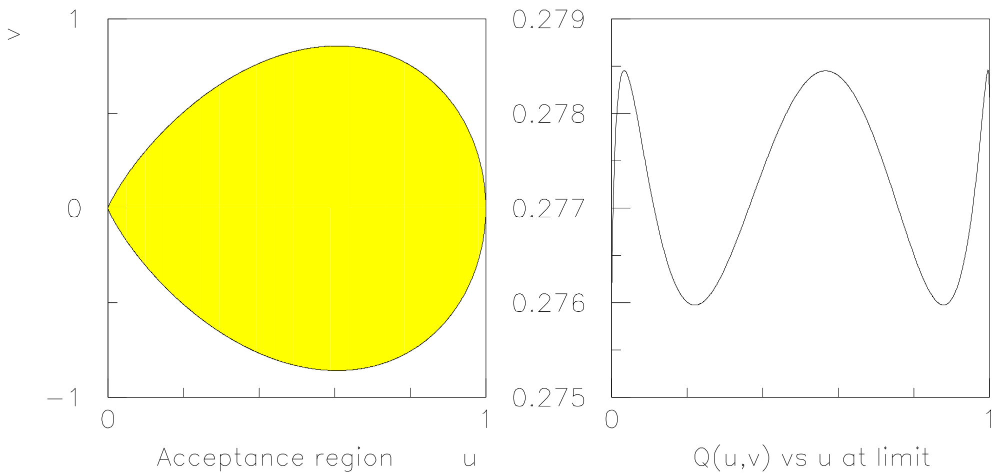


Figure 29: Method using ratio between two uniform deviates

Note that it is enough to consider the upper part ($v > 0$) of the acceptance limit due to the symmetry of the problem. In order to avoid taking the logarithm, which may slow the algorithm down, simpler boundary curves were designed. An improvement to the original proposal was made by Joseph L. Leva in 1992 [37,38] choosing the same quadratic form for both the lower and the upper boundary namely

$$Q(u,v) = (u-s)^2 - b(u-s)(v-t) + (a-v)^2$$

Here $(s,t) = (0.449871, \text{-}0.386595)$ is the center of the ellipses and $a = 0.196$ and $b = 0.25472$ are suitable constants to obtain tight boundaries. In the right-hand side of figure 29 we show the value of the quadratic form at the acceptance limit $Q(u, 2u\sqrt{-\ln u})$ as a function of u. It may be deduced that only in the interval $r_1 < Q < r_2$ with $r_1 = 0.27597$ and $r_2 = 0.27846$ we still have to evaluate the logarithm.

The algorithm is as follows:

i Generate uniform random numbers $u = \xi_1$ and $v = 2\sqrt{2/e}(\xi_2 - \frac{1}{2})$.

ii Evaluate the quadratic form $Q = x^2 + y(ay - bx)$ with $x = u - s$ and $y = |v| - t$.

iii Accept if inside inner boundary, *i.e.* if $Q < r_1$, then go to **vi**.

iv Reject if outside upper boundary, *i.e.* if $Q > r_2$, then go to **i**.

v Reject if outside acceptance region, *i.e.* if $v^2 > -4u^2 \ln u$, then go to **i**.

vi Return the ratio v/u as a pseudorandom number from a standard normal distribution.

On average 2.738 uniform random numbers are consumed and 0.012 logarithms are computed per each standard normal random number obtained by this algorithm. As a comparison the number of logarithmic evaluations without cutting on the boundaries, skipping steps **ii** through **iv** above, would be 1.369. The penalty when using logarithms on modern computers is not as severe as it used to be but still some efficiency is gained by using the proposed algorithm.

34.8.10 Comparison of random number generators

Above we described several methods to achieve pseudorandom numbers from a standard normal distribution. Which one is the most efficient may vary depending on the actual implementation and the computer it is used at. To give a rough idea we found the following times per random number[8] (in the table are also given the average number of uniform pseudorandom numbers consumed per random number in our implementations)

Method	section	μs/r.n.	N_ξ/r.n.	comment
Trapezoidal method	34.8.4	0.39	2.246	
Polar method	34.8.3	0.41	1.273	pair
Histogram method	34.8.8	0.42	2.121	
Box-Muller transformation	34.8.2	0.44	1.000	pair
Spline functions	34.8.7	0.46	3.055	
Ratio of two uniform deviates	34.8.9	0.55	2.738	
Composition-rejection, two terms	34.8.6	0.68	2.394	
Center-tail method	34.8.5	0.88	5.844	
Composition-rejection, one term	34.8.6	0.90	2.631	
Central limit method approach	34.8.1	1.16	12.000	inaccurate

The trapezoidal method is thus fastest but the difference is not great as compared to some of the others. The central limit theorem method is slow as well as inaccurate although it might be the easiest to remember. The other methods are all exact except for possible numerical problems. "Pair" indicates that these generators give two random numbers at a time which may implies that either one is not used or one is left pending for the next call (as is the case in our implementations).

[8] The timing was done on a Digital Personal Workstation 433au workstation running Unix version 4.0D and all methods were programmed in standard Fortran as functions giving one random number at each call.

34.9 Tests on Parameters of a Normal Distribution

For observations from a normal sample different statistical distributions are applicable in different situations when it comes to estimating one or both of the parameters μ and σ. In the table below we try to summarize this in a condensed form.

TESTS OF MEAN AND VARIANCE OF NORMAL DISTRIBUTION			
H_0	Condition	Statistic	Distribution
$\mu = \mu_0$	σ^2 known	$\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$	$N(0,1)$
	σ^2 unknown	$\frac{\overline{x}-\mu_0}{s/\sqrt{n}}$	t_{n-1}
$\sigma^2 = \sigma_0^2$	μ known	$\frac{(n-1)s^2}{\sigma_0^2} = \sum\limits_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma_0^2}$	χ_n^2
	μ unknown	$\frac{(n-1)s^2}{\sigma_0^2} = \sum\limits_{i=1}^{n} \frac{(x_i-\overline{x})^2}{\sigma_0^2}$	χ_{n-1}^2
$\mu_1 = \mu_2 = \mu$	$\sigma_1^2 = \sigma_2^2 = \sigma^2$ known	$\frac{\overline{x}-\overline{y}}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}$	$N(0,1)$
	$\sigma_1^2 \neq \sigma_2^2$ known	$\frac{\overline{x}-\overline{y}}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}$	$N(0,1)$
	$\sigma_1^2 = \sigma_2^2 = \sigma^2$ unknown	$\frac{\overline{x}-\overline{y}}{s\sqrt{\frac{1}{n}+\frac{1}{m}}}$	t_{n+m-2}
		$s = \frac{(n-1)s_1^2+(m-1)s_2^2}{n+m-2}$	
	$\sigma_1^2 \neq \sigma_2^2$ unknown	$\frac{\overline{x}-\overline{y}}{\sqrt{\frac{s_1^2}{n}+\frac{s_2^2}{m}}}$	$\approx N(0,1)$
$\sigma_1^2 = \sigma_2^2$	$\mu_1 \neq \mu_2$ known	$\frac{s_1^2}{s_2^2} = \frac{\frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\mu_1)^2}{\frac{1}{m-1}\sum\limits_{i=1}^{m}(y_i-\mu_2)^2}$	$F_{n,m}$
	$\mu_1 \neq \mu_2$ unknown	$\frac{s_1^2}{s_2^2} = \frac{\frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\overline{x})^2}{\frac{1}{m-1}\sum\limits_{i=1}^{m}(y_i-\overline{y})^2}$	$F_{n-1,m-1}$

35 Pareto Distribution

35.1 Introduction

The Pareto distribution is given by

$$f(x;\alpha,k) = \alpha k^\alpha / x^{\alpha+1}$$

where the variable $x \geq k$ and the parameter $\alpha > 0$ are real numbers. As is seen k is only a scale factor.

The distribution has its name after its inventor the italian Vilfredo Pareto (1848–1923) who worked in the fields of national economy and sociology (professor in Lausanne, Switzerland). It was introduced in order to explain the distribution of wages in society.

35.2 Cumulative Distribution

The cumulative distribution is given by

$$F(x) = \int_k^x f(u)du = 1 - \left(\frac{k}{x}\right)^\alpha$$

35.3 Moments

Algebraic moments are given by

$$E(x^n) = \int_k^\infty x^n f(x) = \int_k^\infty x^n \frac{\alpha k^\alpha}{x^{\alpha+1}} = \left[-\frac{\alpha k^\alpha}{x^{\alpha-n+1}}\right]_k^\infty = \frac{\alpha k^\alpha}{\alpha - n}$$

which is defined for $\alpha > n$.

Especially the expectation value and variance are given by

$$\begin{aligned} E(x) &= \frac{\alpha k}{\alpha - 1} \quad \text{for} \quad \alpha > 1 \\ V(x) &= \frac{\alpha k^2}{(\alpha-2)(\alpha-1)^2} \quad \text{for} \quad \alpha > 2 \end{aligned}$$

35.4 Random Numbers

To obtain a random number from a Pareto distribution we use the straightforward way of solving the equation $F(x) = \xi$ with ξ a random number uniformly distributed between zero and one. This gives

$$F(x) = 1 - \left(\frac{k}{x}\right)^\alpha = \xi \quad \Rightarrow \quad x = \frac{k}{(1-\xi)^{\frac{1}{\alpha}}}$$

36 Poisson Distribution

36.1 Introduction

The Poisson distribution is given by

$$p(r; \mu) = \frac{\mu^r e^{-\mu}}{r!}$$

where the variable r is an integer ($r \geq 0$) and the parameter μ is a real positive quantity. It is named after the french mathematician Siméon Denis Poisson (1781–1840) who was the first to present this distribution in 1837 (implicitly the distribution was known already in the beginning of the 18th century).

As is easily seen by comparing two subsequent r-values the distribution increases up to $r+1 < \mu$ and then declines to zero. For low values of μ it is very skewed (for $\mu < 1$ it is J-shaped).

The Poisson distribution describes the probability to find exactly r events in a given length of time if the events occur independently at a constant rate μ. An unbiased and efficient estimator of the Poisson parameter μ for a sample with n observations x_i is $\hat{\mu} = \bar{x}$, the sample mean, with variance $V(\hat{\mu}) = \mu/n$.

For $\mu \to \infty$ the distribution tends to a normal distribution with mean μ and variance μ.

The Poisson distribution is one of the most important distributions in statistics with many applications. Along with the properties of the distribution we give a few examples here but for a more thorough description we refer to standard text-books.

36.2 Moments

The expectation value, variance, third and fourth central moments of the Poisson distribution are

$$\begin{aligned} E(r) &= \mu \\ V(r) &= \mu \\ \mu_3 &= \mu \\ \mu_4 &= \mu(1+3\mu) \end{aligned}$$

The coefficients of skewness and kurtosis are $\gamma_1 = 1/\sqrt{\mu}$ and $\gamma_2 = 1/\mu$ respectively, *i.e.* they tend to zero as $\mu \to \infty$ in accordance with the distribution becoming approximately normally distributed for large values of μ.

Algebraic moments may be found by the recursive formula

$$\mu'_{k+1} = \mu \left\{ \mu'_k + \frac{d\mu'_k}{d\mu} \right\}$$

and central moments by a similar formula

$$\mu_{k+1} = \mu \left\{ k\mu_{k-1} + \frac{d\mu_k}{d\mu} \right\}$$

For a Poisson distribution one may note that factorial moments g_k (cf page 6) and cumulants κ_k (see section 2.5) become especially simple

$$\begin{aligned} g_k &= E(r(r-1)\cdots(r-k+1)) = \mu^k \\ \kappa_r &= \mu \quad \text{for all } r \geq 1 \end{aligned}$$

36.3 Probability Generating Function

The probability generating function is given by

$$G(z) = E(z^r) = \sum_{r=0}^{\infty} z^r \frac{\mu^r e^{-\mu}}{r!} = e^{-\mu} \sum_{r=0}^{\infty} \frac{(\mu z)}{r!} = e^{\mu(z-1)}$$

Although we mostly use the probability generating function in the case of a discrete distribution we may also define the characteristic function

$$\phi(t) = E(e^{\imath tr}) = e^{-\mu} \sum_{r=0}^{\infty} e^{\imath tr} \frac{\mu^r}{r!} = \exp\left\{\mu\left(e^{\imath t} - 1\right)\right\}$$

a result which could have been given directly since $\phi(t) = G(e^{\imath t})$.

36.4 Cumulative Distribution

When calculating the probability content of a Poisson distribution we need the cumulative, or distribution, function. This is easily obtained by finding the individual probabilities *e.g.* by the recursive formula $p(r) = p(r-1)\frac{\mu}{r}$ starting with $p(0) = e^{-\mu}$.

There is, however, also an interesting connection to the incomplete Gamma function [9]

$$P(r) = \sum_{k=0}^{r} \frac{\mu^k e^{-\mu}}{k!} = 1 - P(r+1, \mu)$$

with $P(a, x)$ the incomplete Gamma function not to be confused with $P(r)$.

Since the cumulative chi-square distribution also has a relation to the incomplete Gamma function one may obtain a relation between these cumulative distributions namely

$$P(r) = \sum_{k=0}^{r} \frac{\mu^k e^{-\mu}}{k!} = 1 - \int_0^{2\mu} f(x; \nu = 2r+2) dx$$

where $f(x; \nu = 2r+2)$ denotes the chi-square distribution with ν degrees of freedom.

36.5 Addition Theorem

The so called addition theorem states that the sum of any number of independent Poisson-distributed variables is also distributed according to a Poisson distribution.

For n variables each distributed according to the Poisson distribution with parameters (means) μ_i we find characteristic function

$$\phi(r_1 + r_2 + \ldots + r_n) = \prod_{i=1}^{n} \exp\left\{\mu_i\left(e^{\imath t} - 1\right)\right\} = \exp\left\{\sum_{i=1}^{n} \mu_i \left(e^{\imath t} - 1\right)\right\}$$

which is the characteristic function for a Poisson variable with parameter $\mu = \sum \mu_i$.

36.6 Derivation of the Poisson Distribution

For a binomial distribution the rate of "success" p may be very small but in a long series of trials the total number of successes may still be a considerable number. In the limit $p \to 0$ and $N \to \infty$ but with $Np = \mu$ a finite constant we find

$$\begin{aligned} p(r) &= \binom{N}{r} p^r (1-p)^{N-r} \approx \frac{1}{r!} \frac{\sqrt{2\pi N} N^N e^{-N}}{\sqrt{2\pi(N-r)}(N-r)^{N-r} e^{-(N-r)}} \left(\frac{\mu}{N}\right)^r \left(1-\frac{\mu}{N}\right)^{N-r} = \\ &= \frac{1}{r!} \sqrt{\frac{N}{N-r}} \frac{1}{\left(1-\frac{r}{N}\right)^N} e^{-r} \mu^r \left(1-\frac{\mu}{N}\right)^{N-r} \to \frac{\mu^r e^{-\mu}}{r!} \end{aligned}$$

as $N \to \infty$ and where we have used that $\lim_{n\to\infty}(1-\frac{x}{n})^n = e^{-x}$ and Stirling's formula (se section 42.2) for the factorial of a large number $n! \approx \sqrt{2\pi n}\; n^n\; e^{-n}$.

It was this approximation to the binomial distribution which S. D. Poisson presented in his book in 1837.

36.7 Histogram

In a histogram of events we would regard the distribution of the bin contents as multinomially distributed if the total number of events N were regarded as a fixed number. If, however, we would regard the total number of events not as fixed but as distributed according to a Poisson distribution with mean ν we obtain (with k bins in the histogram and the multinomial probabilities for each bin in the vector $\underline{p}$)

Given a multinomial distribution, denoted $M(\underline{r}; N, \underline{p})$, for the distribution of events into bins for fixed N and a Poisson distribution, denoted $P(N;\nu)$, for the distribution of N we write the joint distribution

$$\begin{aligned} \mathcal{P}(\underline{r}, N) &= M(\underline{r}; N, \underline{p}) P(N;\nu) = \left(\frac{N!}{r_1! r_2! \ldots r_k!} p_1^{r_1} p_2^{r_2} \ldots p_k^{r_k}\right) \left(\frac{\nu^N e^{-\nu}}{N!}\right) = \\ &= \left(\frac{1}{r_1!}(\nu p_1)^{r_1} e^{-\nu p_1}\right) \left(\frac{1}{r_2!}(\nu p_2)^{r_2} e^{-\nu p_2}\right) \ldots \left(\frac{1}{r_k!}(\nu p_k)^{r_k} e^{-\nu p_k}\right) \end{aligned}$$

where we have used that

$$\sum_{i=1}^{k} p_i = 1 \quad \text{and} \quad \sum_{i=1}^{k} r_i = N$$

i.e. we get a product of independent Poisson distributions with means νp_i for each individual bin. A simpler case leading to the same result would be the classification into only two groups using a binomial and a Poisson distribution.

The assumption of independent Poisson distributions for the number events in each bin is behind the usual rule of using $\sqrt{N}$ as the standard deviation in a bin with N entries and neglecting correlations between bins in a histogram.

36.8 Random Number Generation

By use of the cumulative technique *e.g.* forming the cumulative distribution by starting with $P(0) = e^{-\mu}$ and using the recursive formula

$$P(r) = P(r-1)\frac{\mu}{r}$$

a random number from a Poisson distribution is easily obtained using one uniform random number between zero and one. If μ is a constant the by far fastest generation is obtained if the cumulative vector is prepared once for all.

An alternative is to obtain, in ρ, a random number from a Poisson distribution by multiplying independent uniform random numbers ξ_i until

$$\prod_{i=0}^{\rho} \xi_i \leq e^{-\mu}$$

For large values of μ use the normal approximation but beware of the fact that the Poisson distribution is a function in a discrete variable.

37 Rayleigh Distribution

37.1 Introduction

The Rayleigh distribution is given by

$$f(x;\alpha) = \frac{x}{\alpha^2}\, e^{-\frac{x^2}{2\alpha^2}}$$

for real positive values of the variable x and a real positive parameter α. It is named after the british physicist Lord Rayleigh (1842–1919), also known as Baron John William Strutt Rayleigh of Terling Place and Nobel prize winner in physics 1904.

Note that the parameter α is simply a scale factor and that the variable $y = x/\alpha$ has the simplified distribution $g(y) = ye^{-y^2/2}$.

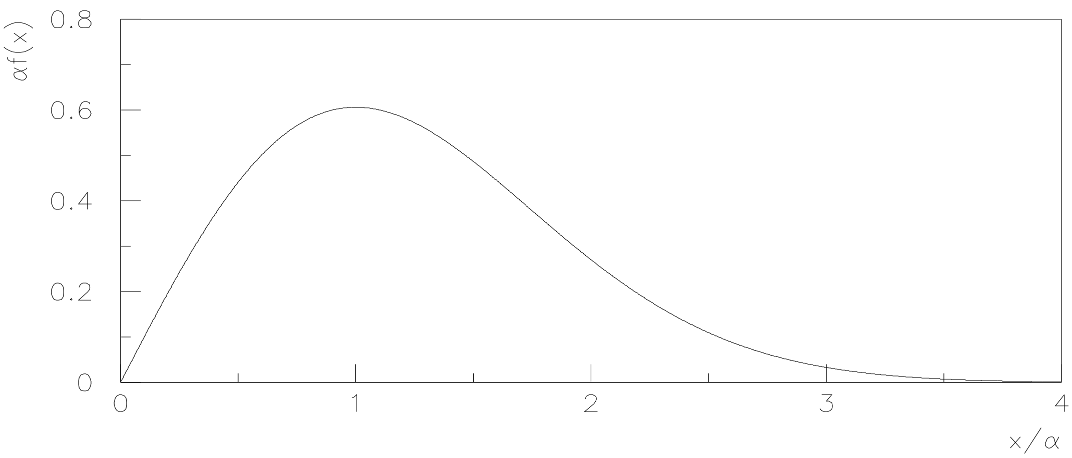


Figure 30: The Rayleigh distribution

The distribution, shown in figure 30, has a mode at $x = \alpha$ and is positively skewed.

37.2 Moments

Algebraic moments are given by

$$E(x^n) = \int_0^\infty x^n f(x)dx = \frac{1}{2\alpha^2}\int_{-\infty}^{\infty} |x|^{n+1} e^{-x^2/2\alpha^2}$$

i.e. we have a connection to the absolute moments of the Gauss distribution. Using these (see section 34 on the normal distribution) the result is

$$E(x^n) = \begin{cases} \sqrt{\frac{\pi}{2}}\, n!!\alpha^n & \text{for } n \text{ odd} \\ 2^k k!\alpha^{2k} & \text{for } n = 2k \end{cases}$$

Specifically we note that the expectation value, variance, and the third and fourth central moments are given by

$$E(x) = \alpha\sqrt{\frac{\pi}{2}}, \quad V(x) = \alpha^2\left(2 - \frac{\pi}{2}\right), \quad \mu_3 = \alpha^3(\pi - 3)\sqrt{\frac{\pi}{2}}, \quad \text{and} \quad \mu_4 = \alpha^4\left(8 - \frac{3\pi^2}{4}\right)$$

The coefficients of skewness and kurtosis is thus

$$\gamma_1 = \frac{(\pi - 3)\sqrt{\frac{\pi}{2}}}{\left(2 - \frac{\pi}{2}\right)^{\frac{3}{2}}} \approx 0.63111 \quad \text{and} \quad \gamma_2 = \frac{8 - \frac{3\pi^2}{4}}{\left(2 - \frac{\pi}{2}\right)^2} - 3 \approx 0.24509$$

37.3 Cumulative Distribution

The cumulative distribution, or the distribution function, is given by

$$F(x) = \int_0^x f(y)dy = \frac{1}{a^2}\int_0^x ye^{-\frac{y^2}{2\alpha^2}}dy = \int_0^{\frac{x^2}{2\alpha^2}} e^{-z}dz = 1 - e^{-\frac{x^2}{2\alpha^2}}$$

where we have made the substitution $z = \frac{y^2}{2\alpha^2}$ in order to simplify the integration. As it should we see that $F(0) = 0$ and $F(\infty) = 1$.

Using this we may estimate the median $\mathcal{M}$ by

$$F(\mathcal{M}) = \frac{1}{2} \quad \Rightarrow \mathcal{M} = \alpha\sqrt{2\ln 2} \approx 1.17741\alpha$$

and the lower and upper quartiles becomes

$$\mathcal{Q}_1 = \alpha\sqrt{-2\ln\tfrac{3}{4}} \approx 0.75853\alpha \quad \text{and} \quad \mathcal{Q}_3 = \alpha\sqrt{2\ln 4} \approx 1.66511\alpha$$

and the same technique is useful when generating random numbers from the Rayleigh distribution as is described below.

37.4 Two-dimensional Kinetic Theory

Given two independent coordinates x and y from normal distributions with zero mean and the same variance σ^2 the distance $z = \sqrt{x^2 + y^2}$ is distributed according to the Rayleigh distribution. The x and y may *e.g.* be regarded as the velocity components of a particle moving in a plane.

To realize this we first write

$$w = \frac{z^2}{\sigma^2} = \frac{x^2}{\sigma^2} + \frac{y^2}{\sigma^2}$$

Since x/σ and y/σ are distributed as standard normal variables the sum of their squares has the chi-squared distribution with 2 degrees of freedom *i.e.* $g(w) = e^{-w/2}/2$ from which we find

$$f(z) = g(w)\left|\frac{dw}{dz}\right| = g\left(\frac{z^2}{\sigma^2}\right)\frac{2z}{\sigma^2} = \frac{z}{\sigma^2}e^{-\frac{z^2}{2\sigma^2}}$$

which we recognize as the Rayleigh distribution. This may be compared to the three-dimensional case where we end up with the Maxwell distribution.

37.5 Random Number Generation

To obtain random numbers from the Rayleigh distribution in an efficient way we make the transformation $y = x^2/2\alpha^2$ a variable which follow the exponential distribution $g(y) = e^{-y}$. A random number from this distribution is easily obtained by taking minus the natural logarithm of a uniform random number. We may thus find a random number r from a Rayleigh distribution by the expression

$$r = \alpha\sqrt{-2\ln\xi}$$

where ξ is a random number uniformly distributed between zero and one.

This could have been found at once using the cumulative distribution putting

$$F(x) = \xi \quad \Rightarrow \quad 1 - e^{-\frac{x^2}{2\alpha^2}} = \xi \quad \Rightarrow \quad x = \alpha\sqrt{-2\ln(1-\xi)}$$

a result which is identical since if ξ is uniformly distributed between zero and one so is $1 - \xi$.

Following the examples given above we may also have used two independent random numbers from a standard normal distribution, z_1 and z_2, and construct

$$r = \frac{1}{\alpha}\sqrt{z_1^2 + z_2^2}$$

However, this technique is not as efficient as the one outlined above.

38 Student's t-distribution

38.1 Introduction

The Student's t-distribution is given by

$$f(t;n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} = \frac{\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)}$$

where the parameter n is a positive integer and the variable t is a real number. The functions Γ and B are the usual Gamma and Beta functions. In figure 31 we show the t-distribution for n values of 1 (lowest maxima), 2, 5 and ∞ (fully drawn and identical to the standard normal distribution).

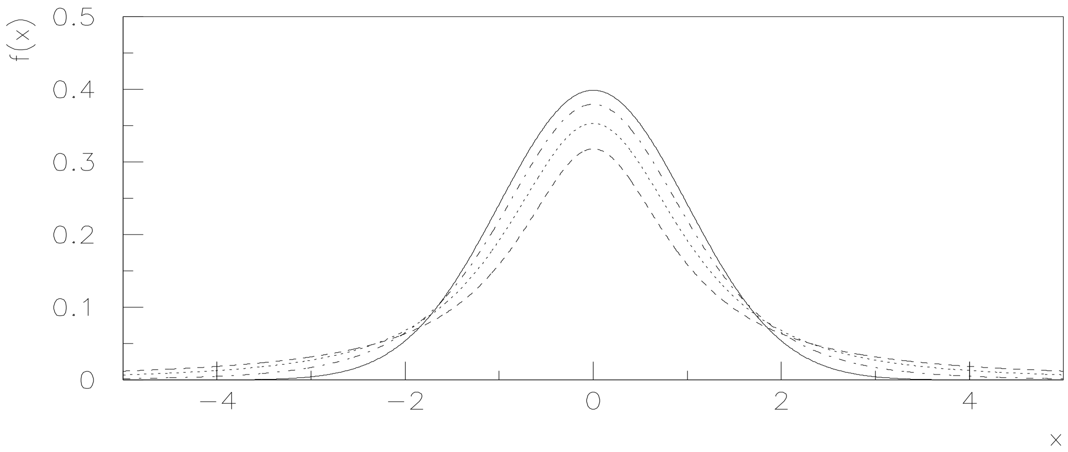


Figure 31: Graph of t-distribution for some values of n

If we change variable to $x = t/\sqrt{n}$ and put $m = \frac{n+1}{2}$ the Student's t-distribution becomes

$$f(x;m) = \frac{k}{(1+x^2)^m} \quad \text{with} \quad k = \frac{\Gamma(m)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(m-\frac{1}{2}\right)} = \frac{1}{B\left(\frac{1}{2},m-\frac{1}{2}\right)}$$

where k is simply a normalization constant and m is a positive half-integer.

38.2 History

A brief history behind this distribution and its name is the following. William Sealy Gosset (1876-1937) had a degree in mathematics and chemistry from Oxford when he in 1899 began working for Messrs. Guinness brewery in Dublin. In his work at the brewery he developed a small-sample theory of statistics which he needed in making small-scale experiments. Due to company policy it was forbidden for employees to publish scientific papers and his work on the t-ratio was published under the pseudonym "Student". It is a very important contribution to statistical theory.

38.3 Moments

The Student's t-distribution is symmetrical around $t = 0$ and thus all odd central moments vanish. In calculating even moments (note that algebraic and central moments are equal) we make use of the somewhat simpler $f(x;m)$ form given above with $x = \frac{t}{\sqrt{n}}$ which implies the following relation between expectation values $E(t^{2r}) = n^r E(x^{2r})$. Central moments of even order are given by, with r an integer ≥ 0,

$$\mu_{2r}(x) = \int_{-\infty}^{\infty} f(x;m)dx = k \int_{-\infty}^{\infty} \frac{x^{2r}}{(1+x^2)^m} dx = 2k \int_0^{\infty} \frac{x^{2r}}{(1+x^2)^m} dx$$

If we make the substitution $y = \frac{x^2}{1+x^2}$ implying $\frac{1}{1+x^2} = 1 - y$ and $x = \sqrt{\frac{y}{1-y}}$ then $dy = \frac{2x}{(1+x^2)^2} dx$ and we obtain

$$\begin{aligned}
\mu_{2r}(x) &= 2k \int_0^1 \frac{x^{2r}}{(1+x^2)^m} \cdot \frac{(1+x^2)^2}{2x} dy = k \int_0^1 \frac{x^{2r-1}}{(1+x^2)^{m-2}} dy = \\
&= k \int_0^1 (1-y)^{m-2} \left(\sqrt{\frac{y}{1-y}} \right)^{2r-1} dy = k \int_0^1 (1-y)^{m-r-\frac{3}{2}} y^{r-\frac{1}{2}} dy = \\
&= kB(r+\frac{1}{2}, m-r-\frac{1}{2}) = \frac{B(r+\frac{1}{2}, m-r-\frac{1}{2})}{B(\frac{1}{2}, m-\frac{1}{2})}
\end{aligned}$$

The normalization constant k was given above and we may now verify this expression by looking at $\mu_0 = 1$ giving $k = 1/B(\frac{1}{2}, m-\frac{1}{2})$ and thus finally, including the n^r factor giving moments in t we have

$$\mu_{2r}(t) = n^r \mu_{2r}(x) = n^r \frac{B(r+\frac{1}{2}, m-r-\frac{1}{2})}{B(\frac{1}{2}, m-\frac{1}{2})} = n^r \frac{B(r+\frac{1}{2}, \frac{n}{2}-r)}{B(\frac{1}{2}, \frac{n}{2})}$$

As can be seen from this expression we get into problems for $2r \geq n$ and indeed those moments are undefined or divergent[9]. The formula is thus valid only for $2r < n$. A recursive formula to obtain even algebraic moments of the t-distribution is

$$\mu'_{2r} = \mu'_{2r-2} \cdot n \cdot \frac{r - \frac{1}{2}}{\frac{n}{2} - r}$$

starting with $\mu'_0 = 1$.

[9]See *e.g.* the discussion in the description of the moments for the Cauchy distribution which is the special case where $m = n = 1$.

Especially we note that, when n is big enough so that these moments are defined, the second central moment (*i.e.* the variance) is $\mu_2 = V(t) = \frac{n}{n-2}$ and the fourth central moment is given by $\mu_4 = \frac{3n^2}{(n-2)(n-4)}$. The coefficients of skewness and kurtosis are given by $\gamma_1 = 0$ and $\gamma_2 = \frac{6}{n-4}$, respectively.

38.4 Cumulative Function

In calculating the cumulative function for the t-distribution it turns out to be simplifying to first estimate the integral for a symmetric region

$$\begin{aligned}
\int\limits_{-t}^{t} f(u)du &= \frac{1}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)} \int\limits_{-t}^{t} \left(1+\frac{u^2}{n}\right)^{-\frac{n+1}{2}} du = \\
&= \frac{2}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)} \int\limits_{0}^{t} \left(1+\frac{u^2}{n}\right)^{-\frac{n+1}{2}} du = \\
&= \frac{-2}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)} \int\limits_{1}^{\frac{n}{n+t^2}} \frac{x^{\frac{n+1}{2}} n\sqrt{x}}{2x^2\sqrt{n}\sqrt{1-x}} dx = \\
&= \frac{1}{B\left(\frac{1}{2},\frac{n}{2}\right)} \int\limits_{\frac{n}{n+t^2}}^{1} (1-x)^{-\frac{1}{2}} x^{\frac{n}{2}-1} dx = \\
&= \frac{1}{B\left(\frac{1}{2},\frac{n}{2}\right)} \left(B\left(\tfrac{n}{2},\tfrac{1}{2}\right) - B_{\frac{n}{n+t^2}}\left(\tfrac{n}{2},\tfrac{1}{2}\right)\right) = \\
&= 1 - I_{\frac{n}{n+t^2}}\left(\tfrac{n}{2},\tfrac{1}{2}\right) = I_{\frac{t^2}{n+t^2}}\left(\tfrac{1}{2},\tfrac{n}{2}\right)
\end{aligned}$$

where we have made the substitution $x = n/(n+u^2)$ in order to simplify the integration. From this we find the cumulative function as

$$F(t) = \begin{cases} \frac{1}{2} - \frac{1}{2} I_{\frac{t^2}{n+t^2}}\left(\frac{1}{2},\frac{n}{2}\right) & \text{for } -\infty < x < 0 \\ \frac{1}{2} + \frac{1}{2} I_{\frac{x^2}{n+x^2}}\left(\frac{1}{2},\frac{n}{2}\right) & \text{for } 0 \le x < \infty \end{cases}$$

38.5 Relations to Other Distributions

The distribution in $F = t^2$ is given by

$$f(F) = \left|\frac{dt}{dF}\right| f(t) = \frac{1}{2\sqrt{F}} \cdot \frac{\left(1+\frac{F}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n}B(\frac{1}{2},\frac{n}{2})} = \frac{n^{\frac{n}{2}} F^{-\frac{1}{2}}}{B(\frac{1}{2},\frac{n}{2})(F+n)^{\frac{n+1}{2}}}$$

which we recognize as a F-distribution with 1 and n degrees of freedom.

As $n \to \infty$ the Student's t-distribution approaches the standard normal distribution. However, a better approximation than to create a simpleminded standardized variable, dividing by the square root of the variance, is to use

$$z = \frac{t\left(1-\frac{1}{4n}\right)}{\sqrt{1+\frac{t^2}{2n}}}$$

which is more closely distributed according to the standard normal distribution.

38.6 t-ratio

Regard $t = x/\sqrt{\frac{y}{n}}$ where x and y are independent variables distributed according to the standard normal and the chi-square distribution with n degrees of freedom, respectively. The independence implies that the joint probability function in x and y is given by

$$f(x,y;n) = \left(\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\right)\left(\frac{y^{\frac{n}{2}-1}e^{-\frac{y}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\right)$$

where $-\infty < x < \infty$ and $y > 0$. If we change variables to $t = x/\sqrt{\frac{y}{n}}$ and $u = y$ the distribution in t and u, with $-\infty < t < \infty$ and $u > 0$, becomes

$$f(t,u;n) = \left\|\frac{\partial(x,y)}{\partial(t,u)}\right\| f(x,y;n)$$

The determinant is $\sqrt{\frac{u}{n}}$ and thus we have

$$f(t,u;n) = \sqrt{\frac{u}{n}}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{ut^2}{2n}}\right)\left(\frac{u^{\frac{n}{2}-1}e^{-\frac{u}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\right) = \frac{u^{\frac{1}{2}(n+1)-1}e^{-\frac{u}{2}\left(1+\frac{t^2}{n}\right)}}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)2^{\frac{n+1}{2}}}$$

Finally, since we are interested in the marginal distribution in t we integrate over u

$$\begin{aligned}
f(t;n) &= \int_0^\infty f(t,u;n)du = \frac{1}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)2^{\frac{n+1}{2}}}\int_0^\infty u^{\frac{n+1}{2}-1}e^{-\frac{u}{2}\left(1+\frac{t^2}{n}\right)}du = \\
&= \frac{1}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)2^{\frac{n+1}{2}}}\int_0^\infty\left(\frac{2v}{1+\frac{t^2}{n}}\right)^{\frac{n+1}{2}-1}e^{-v}\frac{dv}{\frac{1}{2}\left(1+\frac{t^2}{n}\right)} = \\
&= \frac{\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\int_0^\infty v^{\frac{n+1}{2}-1}e^{-v}dv = \frac{\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\Gamma\left(\frac{n+1}{2}\right) = \frac{\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)}
\end{aligned}$$

where we made the substitution $v = \frac{u}{2}\left(1+\frac{t^2}{n}\right)$ in order to simplify the integral which in the last step is recognized as being equal to $\Gamma\left(\frac{n+1}{2}\right)$.

38.7 One Normal Sample

Regard a sample from a normal population $N(\mu,\sigma^2)$ where the mean value $\overline{x}$ is distributed as $N(\mu,\frac{\sigma^2}{n})$ and $\frac{(n-1)s^2}{\sigma^2}$ is distributed according to the chi-square distribution with $n-1$ degrees of freedom. Here s^2 is the usual unbiased variance estimator $s^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\overline{x})^2$ which in the case of a normal distribution is independent of $\overline{x}$. This implies that

$$t = \frac{\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)s^2}{\sigma^2}/(n-1)}} = \frac{\overline{x}-\mu}{s/\sqrt{n}}$$

is distributed according to Student's t-distribution with $n-1$ degrees of freedom. We may thus use Student's t-distribution to test the hypothesis that $\overline{x}=\mu$ (see below).

38.8 Two Normal Samples

Regard two samples $\{x_1,x_2,...,x_m\}$ and $\{y_1,y_2,...,y_n\}$ from normal distributions having the same variance σ^2 but possibly different means μ_x and μ_y, respectively. Then the quantity $(\overline{x}-\overline{y})-(\mu_x-\mu_y)$ has a normal distribution with zero mean and variance equal to $\sigma^2\left(\frac{1}{m}+\frac{1}{n}\right)$. Furthermore the pooled variance estimate

$$s^2 = \frac{(m-1)s_x^2+(n-1)s_y^2}{m+n-2} = \frac{\sum\limits_{i=1}^{m}(x_i-\overline{x})^2+\sum\limits_{i=1}^{n}(y_i-\overline{y})^2}{m+n-2}$$

is a normal theory estimate of σ^2 with $m+n-2$ degrees of freedom[10].

Since s^2 is independent of μ for normal populations the variable

$$t = \frac{(\overline{x}-\overline{y})-(\mu_x-\mu_y)}{s\sqrt{\frac{1}{m}+\frac{1}{n}}}$$

has the t-distribution with $m+n-2$ degrees of freedom. We may thus use Student's t-distribution to test the hypotheses that $\overline{x}-\overline{y}$ consistent with $\delta=\mu_x-\mu_y$. In particular we may test if $\delta=0$ *i.e.* if the two samples originate from population having the same means as well as variances.

38.9 Paired Data

If observations are made in pairs (x_i,y_i) for $i=1,2,...,n$ the appropriate test statistic is

$$t = \frac{\overline{d}}{s_{\overline{d}}} = \frac{\overline{d}}{s_d/\sqrt{n}} = \frac{\overline{d}}{\sqrt{\frac{\sum\limits_{i=1}^{n}(d_i-\overline{d})^2}{n(n-1)}}}$$

[10]If y is a normal theory estimate of σ^2 with k degrees of freedom then ky/σ^2 is distributed according to the chi-square distribution with k degrees of freedom.

where $d_i = x_i - y_i$ and $\overline{d} = \overline{x} - \overline{y}$. This quantity has a t-distribution with $n-1$ degrees of freedom. We may also write this t-ratio as

$$t = \frac{\sqrt{n} \cdot \overline{d}}{\sqrt{s_x^2 + s_y^2 - 2C_{xy}}}$$

where s_x^2 and s_y^2 are the estimated variances of x and y and C_{xy} is the covariance between them. If we would not pair the data the covariance term would be zero but the number of degrees of freedom $2n-2$ *i.e.* twice as large. The smaller number of degrees of freedom in the paired case is, however, often compensated for by the inclusion of the covariance.

38.10 Confidence Levels

In determining confidence levels or testing hypotheses using the t-distribution we define the quantity $t_{\alpha,n}$ from

$$F(t_{\alpha,n}) = \int\limits_{-\infty}^{t_{\alpha,n}} f(t;n)dt = 1 - \alpha$$

i.e. α is the probability that a variable distributed according to the t-distribution with n degrees of freedom exceeds $t_{\alpha,n}$. Note that due to the symmetry about zero of the t-distribution $t_{\alpha,n} = -t_{1-\alpha,n}$.

In the case of one normal sample described above we may set a $1-\alpha$ confidence interval for μ

$$\overline{x} - \frac{s}{\sqrt{n}} t_{\alpha/2,n-1} \leq \mu \leq \overline{x} + \frac{s}{\sqrt{n}} t_{\alpha/2,n-1}$$

Note that in the case where σ^2 is known we would not use the t-distribution. The appropriate distribution to use in order to set confidence levels in this case would be the normal distribution.

38.11 Testing Hypotheses

As indicated above we may use the t-statistics in order to test hypotheses regarding the means of populations from normal distributions.

In the case of one sample the null hypotheses would be H_0: $\mu = \mu_0$ and the alternative hypothesis H_1: $\mu \neq \mu_0$. We would then use $t = \frac{\overline{x}-\mu_0}{s/\sqrt{n}}$ as outlined above and reject H_0 at the α confidence level of significance if $|t| > \frac{s}{\sqrt{n}} t_{\alpha/2,n-1}$. This test is two-tailed since we do not assume any *a priori* knowledge of in which direction an eventual difference would be. If the alternate hypothesis would be *e.g.* $H_1 : \mu > \mu_0$ then a one-tailed test would be appropriate.

The probability to reject the hypothesis H_0 if it is indeed true is thus α. This is a so called Type I error. However, we might also be interested in the probability of committing a Type II error implying that we would accept the hypothesis although it was wrong and the distribution instead had a mean μ_1. In addressing this question the t-distribution could be modified yielding the non-central t-distribution. The probability content β of this distribution in the confidence interval used would then be the probability of wrongly accepting the hypothesis. This calculation would depend on the choice of α as well as on the number of observations n. However, we do not describe details about this here.

In the two sample case we may want to test the null hypothesis H_0: $\mu_x = \mu_y$ as compared to H_1: $\mu_x \neq \mu_y$. Once again we would reject H_0 if the absolute value of the quantity $t = (\overline{x} - \overline{y})/s\sqrt{\frac{1}{m} + \frac{1}{n}}$ would exceed $t_{\alpha/2,n+m-2}$.

38.12 Calculation of Probability Content

In order to find confidence intervals or to test hypotheses we must be able to calculate integrals of the probability density function over certain regions. We recall the formula

$$F(t_{\alpha,n}) = \int_{-\infty}^{t_{\alpha,n}} f(t;n)dt = 1 - \alpha$$

which defines the quantity $t_{\alpha,n}$ for a specified confidence level α. The probability to get a value equal to $t_{\alpha,n}$ or higher is thus α.

Classically all text-books in statistics are equipped with tables giving values of $t_{\alpha,n}$ for specific α-values. This is sometimes useful and in table 8 on page 180 we show such a table giving points where the distribution has a cumulative probability content of $1-\alpha$ for different number of degrees of freedom.

However, it is often preferable to calculate directly the exact probability that one would observe the actual t-value or worse. To calculate the integral on the left-hand side we differ between the case where the number of degrees of freedom is an odd or even integer. The equation above may either be adjusted such that a required α is obtained or we may replace $t_{\alpha,n}$ with the actual t-value found in order to calculate the probability for the present outcome of the experiment.

The algorithm proposed for calculating the probability content of the t-distribution is described in the following subsections.

38.12.1 Even number of degrees of freedom

For even n we have putting $m = \frac{n}{2}$ and making the substitution $x = \frac{t}{\sqrt{n}}$

$$1 - \alpha = \int_{-\infty}^{t_{\alpha,n}} f(t;n)dt = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \int_{-\infty}^{t_{\alpha,n}} \frac{dt}{\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}} = \frac{\Gamma\left(m+\frac{1}{2}\right)}{\sqrt{\pi}\Gamma(m)} \int_{-\infty}^{t_{\alpha,n}/\sqrt{n}} \frac{dx}{(1+x^2)^{m+\frac{1}{2}}}$$

For convenience (or maybe it is rather laziness) we make use of standard integral tables where we find the integral

$$\int \frac{dx}{(ax^2+c)^{m+\frac{1}{2}}} = \frac{x}{\sqrt{ax^2+c}} \sum_{r=0}^{m-1} \frac{2^{2m-2r-1}(m-1)!m!(2r)!}{(2m)!(r!)^2 c^{m-r}\left(ax^2+c\right)^r}$$

where in our case $a = c = 1$. Introducing $x_\alpha = t_{\alpha,n}/\sqrt{n}$ for convenience this gives

$$1 - \alpha = \frac{\Gamma\left(m+\frac{1}{2}\right)(m-1)!m!2^{2m}}{\sqrt{\pi}\Gamma(m)(2m)!} \cdot \left[\frac{x_\alpha}{2\sqrt{1+x_\alpha^2}} \sum_{r=0}^{m-1} \frac{(2r)!}{2^{2r}(r!)^2\left(1+x_\alpha^2\right)^r} + \frac{1}{2}\right]$$

The last term inside the brackets is the value of the integrand at $-\infty$ which is seen to equal $-\frac{1}{2}$. Looking at the factor outside the brackets using that $\Gamma(n) = (n-1)!$ for n a positive integer, $\Gamma\left(m+\frac{1}{2}\right) = \frac{(2m-1)!!}{2^m}\sqrt{\pi}$, and rewriting $(2m)! = (2m)!!(2m-1)!! = 2^m m!(2m-1)!!$ we find that it in fact is equal to one. We thus have

$$1-\alpha = \frac{x_\alpha}{2\sqrt{1+x_\alpha^2}} \sum_{r=0}^{m-1} \frac{(2r)!}{2^{2r}(r!)^2\left(1+x_\alpha^2\right)^r} + \frac{1}{2}$$

In evaluating the sum it is useful to look at the individual terms. Denoting these by u_r we find the recurrence relation

$$u_r = u_{r-1} \cdot \frac{2r(2r-1)}{r^2 2^2 (1+x_\alpha^2)} = u_{r-1} \cdot \frac{1-\frac{1}{2r}}{1+x_\alpha^2}$$

where we start with $u_0 = 1$.

To summarize: in order to determine the probability α to observe a value t or bigger from a t-distribution with an even number of degrees of freedom n we calculate

$$1-\alpha = \frac{\frac{t}{\sqrt{n}}}{2\sqrt{1+\frac{t^2}{n}}} \sum_{r=0}^{m-1} u_r + \frac{1}{2}$$

where $u_0 = 1$ and $u_r = u_{r-1} \cdot \frac{1-\frac{1}{2r}}{1+t^2/n}$.

38.12.2 Odd number of degrees of freedom

For odd n we have putting $m = \frac{n-1}{2}$ and making the substitution $x = \frac{t}{\sqrt{n}}$

$$1-\alpha = \int_{-\infty}^{t_{\alpha,n}} f(t;n)dt = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \int_{-\infty}^{t_{\alpha,n}} \frac{dt}{\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}} = \frac{\Gamma(m+1)}{\sqrt{\pi}\Gamma\left(m+\frac{1}{2}\right)} \int_{-\infty}^{x_\alpha} \frac{dx}{\left(1+x^2\right)^{m+1}}$$

where we again have introduced $x_\alpha = t_{\alpha,n}/\sqrt{n}$. Once again we make use of standard integral tables where we find the integral

$$\int \frac{dx}{\left(a+bx^2\right)^{m+1}} = \frac{(2m)!}{(m!)^2}\left[\frac{x}{2a}\sum_{r=1}^{m} \frac{r!(r-1)!}{(4a)^{m-r}(2r)!\left(a+bx^2\right)^r} + \frac{1}{(4a)^m}\int \frac{dx}{a+bx^2}\right]$$

where in our case $a = b = 1$. We obtain

$$1-\alpha = \frac{\Gamma(m+1)(2m)!}{\sqrt{\pi}\Gamma\left(m+\frac{1}{2}\right)m!^2 4^m}\left[\frac{x_\alpha}{2}\sum_{r=1}^{m} \frac{4^r r!(r-1)!}{(2r)!\left(1+x_\alpha^2\right)^r} + \arctan x_\alpha + \frac{\pi}{2}\right]$$

where the last term inside the brackets is the value of the integrand at $-\infty$. The factor outside the brackets is equal to $\frac{1}{\pi}$ which is found using that $\Gamma(n) = (n-1)!$ for n a positive integer, $\Gamma\left(m+\frac{1}{2}\right) = \frac{(2m-1)!!}{2^m}\sqrt{\pi}$, and $(2m)! = (2m)!!(2m-1)!! = 2^m(m)!(2m-1)!!$. We get

$$\begin{aligned}1-\alpha &= \frac{1}{\pi}\left[\frac{x_\alpha}{2}\sum_{r=1}^{m}\frac{4^r r!(r-1)!}{(2r)!\left(1+x_\alpha^2\right)^r} + \arctan x_\alpha + \frac{\pi}{2}\right] = \\ &= \frac{1}{\pi}\left[\frac{x_\alpha}{1+x_\alpha^2}\sum_{r=1}^{m}\frac{2^{2r-1}r!(r-1)!}{(2r)!(1+x_\alpha^2)^{r-1}} + \arctan x_\alpha + \frac{\pi}{2}\right]\end{aligned}$$

To compute the sum we denote the terms by v_r and find the recurrence relation

$$v_r = v_{r-1}\frac{4r(r-1)}{2r(2r-1)(1+x_\alpha^2)} = v_{r-1}\frac{\left(1-\frac{1}{2r-1}\right)}{(1+x_\alpha^2)}$$

starting with $v_1 = 1$.

To summarize: in order to determine the probability α to observe a value t or bigger from a t-distribution with an odd number of degrees of freedom n we calculate

$$1-\alpha = \frac{1}{\pi}\left[\frac{\frac{t}{\sqrt{n}}}{1+\frac{t^2}{n}}\sum_{r=1}^{\frac{n-1}{2}} v_r + \arctan\frac{t}{\sqrt{n}}\right] + \frac{1}{2}$$

where $v_1 = 1$ and $v_r = v_{r-1}\cdot\frac{1-\frac{1}{2r-1}}{1+t^2/n}$.

38.12.3 Final algorithm

The final algorithm to evaluate the probability content from $-\infty$ to t for a t-distribution with n degrees of freedom is

- Calculate $x = \frac{t}{\sqrt{n}}$
- For n even:
  - Put $m = \frac{n}{2}$
  - Set $u_0 = 1, s = 0$ and $i = 0$.
  - For $i = 0, 1, 2, ..., m-1$ set $s = s + u_i$, $i = i+1$ and $u_i = u_{i-1}\frac{1-\frac{1}{2i}}{1+x^2}$.
  - $\alpha = \frac{1}{2} - \frac{1}{2}\cdot\frac{x}{\sqrt{1+x^2}}s$.
- For n odd:
  - Put $m = \frac{n-1}{2}$.
  - Set $v_1 = 1, s = 0$ and $i = 1$.
  - For $i = 1, 2, ..., m$ set $s = s + v_i$, $i = i+1$ and $v_i = v_{i-1}\cdot\frac{1-\frac{1}{2i-1}}{1+x^2}$.
  - $\alpha = \frac{1}{2} - \frac{1}{\pi}(\frac{x}{1+x^2}\cdot s + \arctan x)$.

38.13 Random Number Generation

Following the definition we may define a random number t from a t-distribution, using random numbers from a normal and a chi-square distribution, as

$$t = \frac{z}{\sqrt{y_n/n}}$$

where z is a standard normal and y_n a chi-squared variable with n degrees of freedom. To obtain random numbers from these distributions see the appropriate sections.

39 Triangular Distribution

39.1 Introduction

The triangular distribution is given by

$$f(x; \mu, \Gamma) = \frac{-|x-\mu|}{\Gamma^2} + \frac{1}{\Gamma}$$

where the variable x is bounded to the interval $\mu - \Gamma \leq x \leq \mu + \Gamma$ and the location and scale parameters μ and Γ ($\Gamma > 0$) all are real numbers.

39.2 Moments

The expectation value of the distribution is $E(x) = \mu$. Due to the symmetry of the distribution odd central moments vanishes while even moments are given by

$$\mu_n = \frac{2\Gamma^n}{(n+1)(n+2)}$$

for even values of n. In particular the variance $V(x) = \mu_2 = \Gamma^2/6$ and the fourth central moment $\mu_4 = \Gamma^4/15$. The coefficient of skewness is zero and the coefficient of kurtosis $\gamma_2 = -0.6$.

39.3 Random Number Generation

The sum of two pseudorandom numbers uniformly distributed between $(\mu - \Gamma)/2$ and $(\mu + \Gamma)/2$ is distributed according to the triangular distribution. If ξ_1 and ξ_2 are uniformly distributed between zero and one then

$$x = \mu + (\xi_1 + \xi_2 - 1)\Gamma \quad \text{or} \quad x = \mu + (\xi_1 - \xi_2)\Gamma$$

follow the triangular distribution.

Note that this is a special case of a combination

$$x = (a+b)\xi_1 + (b-a)\xi_2 - b$$

with $b > a \geq 0$ which gives a random number from a symmetric trapezoidal distribution with vertices at $(\pm b, 0)$ and $(\pm a, \frac{1}{a+b})$.

40 Uniform Distribution

40.1 Introduction

The uniform distribution is, of course, a very simple case with

$$f(x; a, b) = \frac{1}{b-a} \quad \text{for} \quad a \leq x \leq b$$

The cumulative, distribution, function is thus given by

$$F(x; a, b) = \begin{cases} 0 & \text{if} \quad x \leq a \\ \frac{x-a}{b-a} & \text{if} \quad a \leq x \leq b \\ 1 & \text{if} \quad b \leq x \end{cases}$$

40.2 Moments

The uniform distribution has expectation value $E(x) = (a+b)/2$, variance $V(x) = (b-a)^2/12$, $\mu_3 = 0$, $\mu_4 = (b-a)^4/80$, coefficient of skewness $\gamma_1 = 0$ and coefficient of kurtosis $\gamma_2 = -1.2$. More generally all odd central moments vanish and for n an even integer

$$\mu_n = \frac{(b-a)^n}{2^n(n+1)}$$

40.3 Random Number Generation

Since we assume the presence of a pseudorandom number generator giving random numbers ξ between zero and one a random number from the uniform distribution is simply given by

$$x = (b-a)\xi + a$$

41 Weibull Distribution

41.1 Introduction

The Weibull distribution is given by

$$f(x;\eta,\sigma) = \frac{\eta}{\sigma}\left(\frac{x}{\sigma}\right)^{\eta-1} e^{-\left(\frac{x}{\sigma}\right)^{\eta}}$$

where the variable x and the parameters η and σ all are positive real numbers. The distribution is named after the swedish physicist Waloddi Weibull (1887–1979) a professor at the Technical Highschool in Stockholm 1924–1953.

The parameter σ is simply a scale parameter and the variable $y = x/\sigma$ has the distribution

$$g(y) = \eta\ y^{\eta-1}\ e^{-y^{\eta}}$$

In figure 32 we show the distribution for a few values of η. For $\eta < 1$ the distribution has its mode at $y = 0$, at $\eta = 1$ it is identical to the exponential distribution, and for $\eta > 1$ the distribution has a mode at

$$x = \left(\frac{\eta-1}{\eta}\right)^{\frac{1}{\eta}}$$

which approaches $x = 1$ as η increases (at the same time the distribution gets more symmetric and narrow).

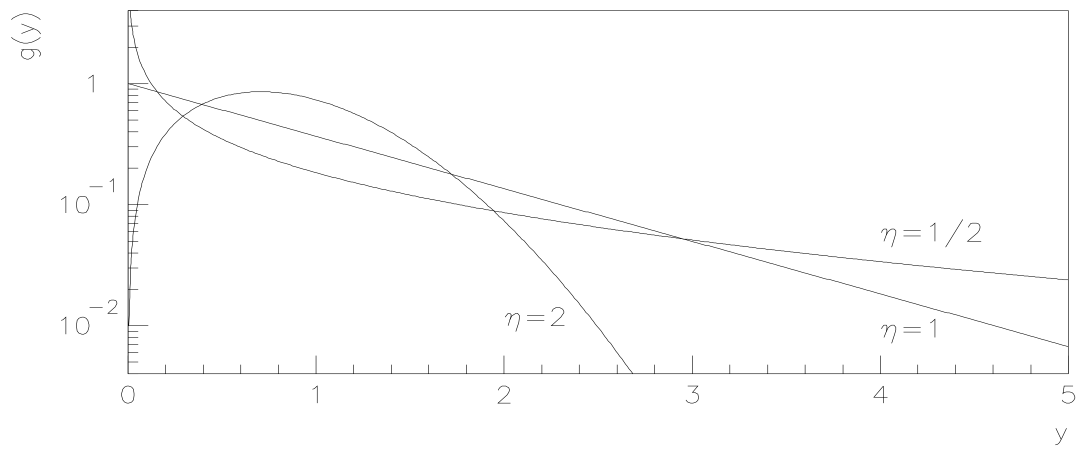


Figure 32: The Weibull distribution

41.2 Cumulative Distribution

The cumulative distribution is given ,by

$$F(x) = \int_0^x f(u)du = \int_0^x \frac{\eta}{\sigma}\left(\frac{u}{\sigma}\right)^{\eta-1} e^{-\left(\frac{u}{\sigma}\right)^\eta} du = \int_0^{(x/\sigma)^\eta} e^{-y}dy = 1 - e^{-\left(\frac{x}{\sigma}\right)^\eta}$$

where we have made the substitution $y = (u/\sigma)^\eta$ in order to simplify the integration.

41.3 Moments

Algebraic moments are given by

$$E(x^k) = \int_0^\infty x^k f(x)dx = \sigma^k \int_0^\infty y^{\frac{k}{\eta}} e^{-y}dy = \sigma^k \Gamma\left(\frac{k}{\eta} + 1\right)$$

where we have made the same substitution as was used when evaluating the cumulative distribution above.

Especially the expectation value and the variance are given by

$$E(x) = \sigma\Gamma\left(\frac{1}{\eta} + 1\right) \quad \text{and} \quad V(x) = \sigma^2 \left\{\Gamma\left(\frac{2}{\eta} + 1\right) - \Gamma\left(\frac{1}{\eta} + 1\right)^2\right\}$$

41.4 Random Number Generation

To obtain random numbers from Weibull's distribution using ξ, a random number uniformly distributed from zero to one, we may solve the equation $F(x) = \xi$ to obtain a random number in x.

$$F(x) = 1 - e^{-\left(\frac{x}{\sigma}\right)^\eta} = \xi \quad \Rightarrow \quad x = \sigma(-\ln\xi)^{\frac{1}{\eta}}$$

42 Appendix A: The Gamma and Beta Functions

42.1 Introduction

In statistical calculations for standard statistical distributions such as the normal (or Gaussian) distribution, the Student's t-distribution, the chi-squared distribution, and the F-distribution one often encounters the so called Gamma and Beta functions. More specifically in calculating the probability content for these distributions the incomplete Gamma and Beta functions occur. In the following we briefly define these functions and give numerical methods on how to calculate them. Also connections to the different statistical distributions are given. The main references for this has been [39–41] for the formalism and [9] for the numerical methods.

42.2 The Gamma Function

The Gamma function is normally defined as

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

where z is a complex variable with $Re(z) > 0$. This is the so called Euler's integral form for the Gamma function. There are, however, two other definitions worth mentioning. Firstly Euler's infinite limit form

$$\Gamma(z) = \lim_{n\to\infty} \frac{1 \cdot 2 \cdot 3 \cdots n}{z(z+1)(z+2)\cdots(z+n)}\, n^z \qquad z \neq 0, -1, -2, \ldots$$

and secondly the infinite product form sometimes attributed to Euler and sometimes to Weierstrass

$$\frac{1}{\Gamma(z)} = z^{\gamma z} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-\frac{z}{n}} \qquad |z| < \infty$$

where $\gamma \approx 0.5772156649$ is Euler's constant.

In figure 33 we show the Gamma function for real arguments from -5 to 5. Note the singularities at $x = 0, -1, -2, \ldots$.

For z a positive real integer n we have the well known relation to the factorial function

$$n! = \Gamma(n+1)$$

and, as the factorial function, the Gamma function satisfies the recurrence relation

$$\Gamma(z+1) = z\Gamma(z)$$

In the complex plane $\Gamma(z)$ has a pole at $z = 0$ and at all negative integer values of z. The reflection formula

$$\Gamma(1-z) = \frac{\pi}{\Gamma(z)\sin(\pi z)} = \frac{\pi z}{\Gamma(z+1)\sin(\pi z)}$$

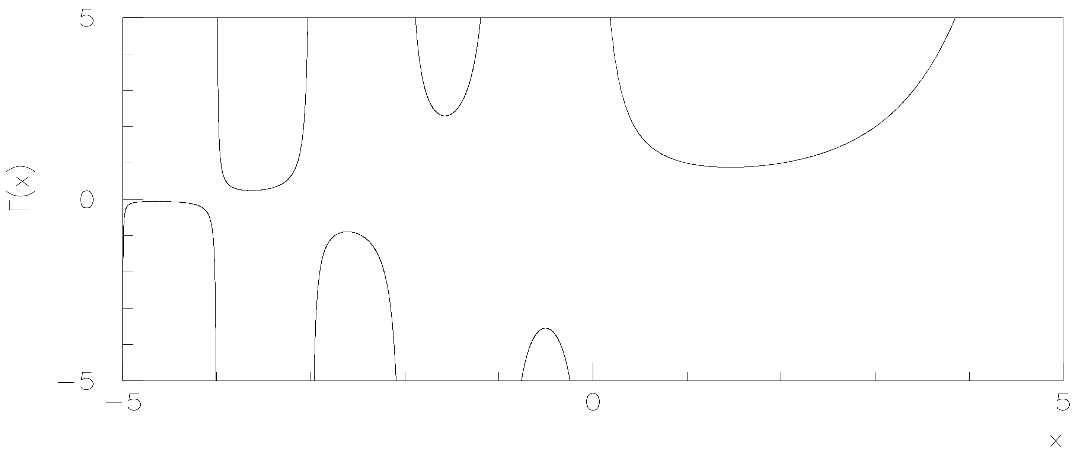


Figure 33: The Gamma function

may be used in order to get function values for $Re(z) < 1$ from values for $Re(z) > 1$.

A well known approximation to the Gamma function is Stirling's formula

$$\Gamma(z) = z^z e^{-z} \sqrt{\frac{2\pi}{z}} \left(1 + \frac{1}{12z} + \frac{1}{288z^2} - \frac{139}{51840z^3} - \frac{571}{2488320z^4} + \ldots\right)$$

for $|\arg z| < \pi$ and $|z| \to \infty$ and where often only the first term (1) in the series expansion is kept in approximate calculations. For the faculty of a positive integer n one often uses the approximation

$$n! \approx \sqrt{2\pi n}\; n^n\; e^{-n}$$

which has the same origin and also is called Stirling's formula.

42.2.1 Numerical Calculation

There are several methods to calculate the Gamma function in terms of series expansions etc. For numerical calculations, however, the formula by Lanczos is very useful [9]

$$\Gamma(z+1) = \left(z + \gamma + \tfrac{1}{2}\right)^{z+\frac{1}{2}} e^{-\left(z+\gamma+\frac{1}{2}\right)} \sqrt{2\pi} \left[c_0 + \frac{c_1}{z+1} + \frac{c_2}{z+2} + \cdots + \frac{c_n}{z+n} + \epsilon\right]$$

for $z > 0$ and an optimal choice of the parameters γ, n, and c_0 to c_n. For $\gamma = 5$, $n = 6$ and a certain set of c's the error is smaller than $|\epsilon| < 2 \cdot 10^{-10}$. This bound is true for all complex z in the half complex plane $Re(z) > 0$. The coefficients normally used are $c_0 = 1$, $c_1 = 76.18009173$, $c_2 = -86.50532033$, $c_3 = 24.01409822$, $c_4 = -1.231739516$, $c_5 = 0.00120858003$, and $c_6 = -0.00000536382$. Use the reflection formula given above to obtain results for $Re(z) < 1$ *e.g.* for negative real arguments. Beware, however, to avoid the singularities. While implementing routines for the Gamma function it is recommendable to evaluate the natural logarithm in order to avoid numerical overflow.

An alternative way of evaluating $\ln\Gamma(z)$ is given in references [42,43], giving formulæ which also are used in order to evaluate the Digamma function below. The expressions used for $\ln\Gamma(z)$ are

$$\ln\Gamma(z) = \begin{cases} \left(z-\frac{1}{2}\right)\ln z - z + \frac{1}{2}\ln 2\pi + z\sum\limits_{k=1}^{K}\frac{B_{2k}}{2k(2k-1)}z^{-2k} + R_K(z) & \text{for } 0 < x_0 \le x \\ \ln\Gamma(z+n) - \ln\prod\limits_{k=0}^{n-1}(z+k) & \text{for } 0 \le x < x_0 \\ \ln\pi + \ln\Gamma(1-z) - \ln\sin\pi z & \text{for } x < 0 \end{cases}$$

Here $n = [x_0] - [x]$ (the difference of integer parts, where x is the real part of $z = x + \imath y$) and *e.g.* $K = 10$ and $x_0 = 7.0$ gives excellent accuracy *i.e.* small R_K. Note that Kölbig [43] gives the wrong sign on the (third) constant term in the first case above.

42.2.2 Formulæ

Below we list some useful relations concerning the Gamma function, faculties and semi-faculties (denoted by two exclamation marks here). For a more complete list consult *e.g.* [40].

$$\begin{aligned}
\Gamma(z) &= \int\limits_0^1 \left(\ln\frac{1}{t}\right)^{z-1} dt \\
\Gamma(z+1) &= z\Gamma(z) = z! \\
\Gamma(z) &= \alpha^z \int\limits_0^\infty t^{z-1}e^{-\alpha t}dt \quad \text{for} \quad Re(z) > 0, \quad Re(\alpha) > 0 \\
\Gamma(k) &= (k-1)! \quad \text{for} \quad k \ge 1 \;\; (\text{integer}, \; 0! = 1) \\
z! &= \Gamma(z+1) = \int\limits_0^\infty e^{-t}t^z dt \quad \text{for} \quad Re(z) > -1 \\
\Gamma\left(\tfrac{1}{2}\right) &= \sqrt{\pi} \\
\Gamma\left(n+\tfrac{1}{2}\right) &= \frac{(2n-1)!!}{2^n}\sqrt{\pi} \\
\Gamma(z)\Gamma(1-z) &= \frac{\pi}{\sin\pi z} \\
z!(-z)! &= \frac{\pi z}{\sin\pi z} \\
(2m)!! &= 2\cdot 4\cdot 6\cdots 2m = 2^m m! \\
(2m-1)!! &= 1\cdot 3\cdot 5\cdots(2m-1) \\
(2m)! &= (2m)!!(2m-1)!! = 2^m m!(2m-1)!!
\end{aligned}$$

42.3 Digamma Function

It is often convenient to work with the logarithm of the Gamma function in order to avoid numerical overflow in the calculations. The first derivatives of this function

$$\psi(z) = \frac{d}{dz}\ln\Gamma(z) = \frac{1}{\Gamma(z)}\frac{d\Gamma(z)}{dz}$$

is known as the Digamma, or Psi, function. A series expansion of this function is given by

$$\psi(z+1) = -\gamma - \sum_{n=1}^{\infty}\left(\frac{1}{z+n} - \frac{1}{n}\right) \quad \text{for} \quad z \neq 0, -1, -2, -3, \ldots$$

where $\gamma \approx 0.5772156649$ is Euler's constant which is seen to be equal to $-\psi(1)$. If the derivative of the Gamma function itself is required we may thus simply use $d\Gamma(z)/dz = \Gamma(z)\cdot\psi(z)$. Note that some authors write $\psi(z) = \frac{d}{dz}\ln\Gamma(z+1) = \frac{d}{dz}z!$ for the Digamma function, and similarly for the polygamma functions below, thus shifting the argument by one unit.

In figure 34 we show the Digamma function for real arguments from -5 to 5. Note the singularities at $x = 0, -1, -2, \ldots$.

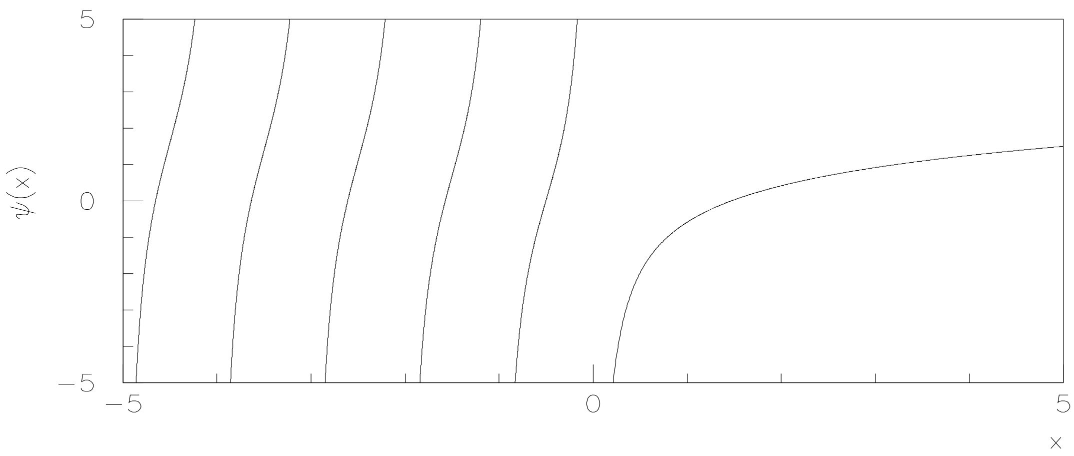


Figure 34: The Digamma, or Psi, function

For integer values of z we may write

$$\psi(n) = -\gamma + \sum_{m=1}^{n-1}\frac{1}{m}$$

which is efficient enough for numerical calculations for not too large values of n. Similarly for half-integer values we have

$$\psi\left(n+\tfrac{1}{2}\right) = -\gamma - 2\ln 2 + 2\sum_{m=1}^{n}\frac{1}{2m-1}$$

However, for arbitrary arguments the series expansion above is unusable. Following the recipe given in an article by K. S. Kölbig [43] we use

$$\psi(z) = \begin{cases} \ln z - \frac{1}{2z} - \sum\limits_{k=1}^{K} \frac{B_{2k}}{2k} z^{-2k} + R_K(z) & \text{for } 0 < x_0 \le x \\ \psi(z+n) - \sum\limits_{k=0}^{n-1} \frac{1}{z+k} & \text{for } 0 \le x < x_0 \\ \psi(-z) + \frac{1}{z} + \pi \cot \pi z & \text{for } x < 0 \end{cases}$$

Here $n = [x_0] - [x]$ (the difference of integer parts, where x is the real part of $z = x + \imath y$) and we have chosen $K = 10$ and $x_0 = 7.0$ which gives a very good accuracy (*i.e.* small R_K, typically less than 10^{-15}) for double precision calculations. The main interest in statistical calculations is normally function values for $\psi(x)$ for real positive arguments but the formulæ above are valid for any complex argument except for the singularities along the real axis at $z = 0, -1, -2, -3, \ldots$. The B_{2k} are Bernoulli numbers given by $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}$, $B_8 = -\frac{1}{30}$, $B_{10} = \frac{5}{66}$, $B_{12} = -\frac{691}{2730}$, $B_{14} = \frac{7}{6}$, $B_{16} = -\frac{3617}{510}$, $B_{18} = \frac{43867}{798}$, $B_{20} = -\frac{174611}{330}$ $\ldots$

42.4 Polygamma Function

Higher order derivatives of $\ln\Gamma(z)$ are called Polygamma functions[11]

$$\psi^{(n)}(z) = \frac{d^n}{dz^n}\psi(z) = \frac{d^{n+1}}{dz^{n+1}} \ln\Gamma(z) \quad \text{for} \quad n = 1, 2, 3, \ldots$$

Here a series expansion is given by

$$\psi^{(n)}(z) = (-1)^{n+1} n! \sum_{k=0}^{\infty} \frac{1}{(z+k)^{n+1}} \quad \text{for} \quad z \neq 0, -1, -2, \ldots$$

For numerical calculations we have adopted a technique similar to what was used to evaluate $\ln\Gamma(z)$ and $\psi(z)$.

$$\psi^{(n)}(z) = \begin{cases} (-1)^{n-1} \left[t_1 + \frac{n!}{2z^{n+1}} + \sum\limits_{k=1}^{K} B_{2k} \frac{(2k+n-1)!}{(2k)! z^{2k+n}} + R_K(z) \right] & \text{for } 0 < x_0 \le x \\ \psi^{(n)}(z+m) - (-1)^n n! \sum\limits_{k=0}^{m-1} \frac{1}{(z+k)^{n+1}} & \text{for } 0 \le x < x_0 \end{cases}$$

where $t_1 = -\ln z$ for $n = 0$ and $t_1 = (n-1)!/z^n$ for $n > 0$. Here $m = [x_0] - [x]$ *i.e.* the difference of integer parts, where x is the real part of $z = x + \imath y$. We treat primarily the case for real positive arguments x and if complex arguments are required one ought to add a third reflection formula as was done in the previous cases. Without any special optimization we have chosen $K = 14$ and $x_0 = 7.0$ which gives a very good accuracy, *i.e.*

[11]Sometimes the more specific notation tri-, tetra-, penta- and hexagamma functions are used for ψ', ψ'', $\psi^{(3)}$ and $\psi^{(4)}$, respectively.

small R_K, typically less than 10^{-15}, even for double precision calculations except for higher orders and low values of x where the function value itself gets large.[12]

For more relations on the Polygamma (and the Digamma) functions see *e.g.* [40]. Two useful relations used in this document in finding cumulants for some distributions are

$$\begin{aligned}\psi^{(n)}(1) &= (-1)^{n+1}n!\zeta(n+1)\\ \psi^{(n)}(\tfrac{1}{2}) &= (-1)^{n+1}n!(2^{n+1}-1)\zeta(n+1) = (2^{n+1}-1)\psi^{(n)}(1)\end{aligned}$$

where ζ is Riemann's zeta function (see page 60 and [27]).

42.5 The Incomplete Gamma Function

For the incomplete Gamma function there seem to be several definitions in the literature. Defining the two integrals

$$\gamma(a,x) = \int_0^x t^{a-1}e^{-t}dt \quad \text{and} \quad \Gamma(a,x) = \int_x^\infty t^{a-1}e^{-t}dt$$

with $Re(a) > 0$ the incomplete Gamma function is normally defined as

$$P(a,x) = \frac{\gamma(a,x)}{\Gamma(a)}$$

but sometimes also $\gamma(a,x)$ and $\Gamma(a,x)$ is referred to under the same name as well as the complement to $P(a,x)$

$$Q(a,x) = 1 - P(a,x) = \frac{\Gamma(a,x)}{\Gamma(a)}$$

Note that, by definition, $\gamma(a,x) + \Gamma(a,x) = \Gamma(a)$.

In figure 35 the incomplete Gamma function $P(a,x)$ is shown for a few a-values (0.5, 1, 5 and 10).

42.5.1 Numerical Calculation

For numerical evaluations of P two formulæ are useful [9]. For values $x < a+1$ the series

$$\gamma(a,x) = e^{-x}x^a \sum_{n=0}^{\infty} \frac{\Gamma(a)}{\Gamma(a+n+1)} x^n$$

converges rapidly while for $x \geq a+1$ the continued fraction

$$\Gamma(a,x) = e^{-x}x^a \left(\frac{1}{x+} \frac{1-a}{1+} \frac{1}{x+} \frac{2-a}{1+} \frac{2}{x+} \cdots \right)$$

is a better choice.

[12] For this calculation we need a few more Bernoulli numbers not given on page 160 above namely $B_{22} = \frac{854513}{138}$, $B_{24} = -\frac{236364091}{2730}$, $B_{26} = \frac{8553103}{6}$, and $B_{28} = -\frac{23749461029}{870}$

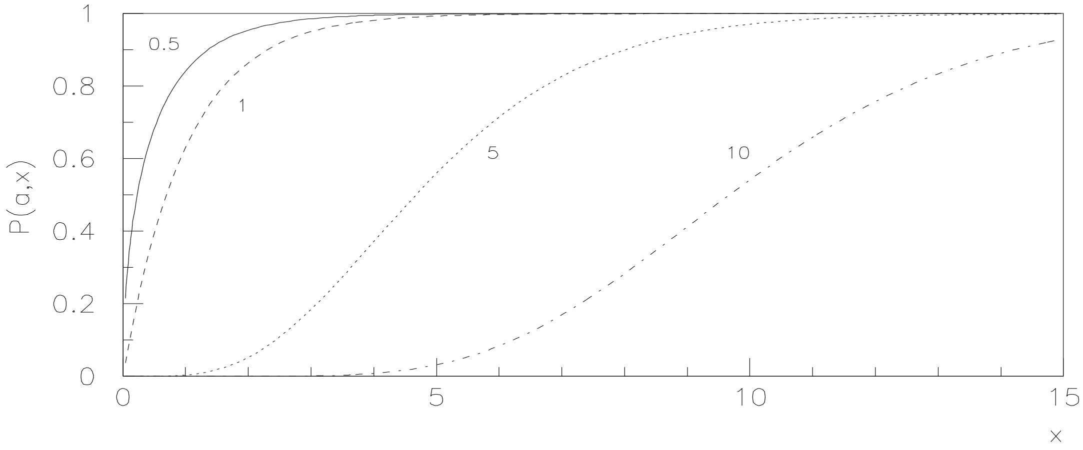


Figure 35: The incomplete Gamma function

42.5.2 Formulæ

Below we list some relations concerning the incomplete Gamma function. For a more complete list consult *e.g.* [40].

$$
\begin{aligned}
\Gamma(a) &= \gamma(a,x) + \Gamma(a,x) \\
\gamma(a,x) &= \int_0^x e^{-t} t^{a-1} dt \quad \text{for} \quad Re(a) > 0 \\
\gamma(a+1,x) &= a\gamma(a,x) - x^a e^{-x} \\
\gamma(n,x) &= (n-1)! \left[1 - e^{-x} \sum_{r=0}^{n-1} \frac{x^r}{r!} \right] \\
\Gamma(a,x) &= \int_x^\infty e^{-t} t^{a-1} dt \\
\Gamma(a+1,x) &= a\Gamma(a,x) - x^a e^{-x} \\
\Gamma(n,x) &= (n-1)! e^{-x} \sum_{r=0}^{n-1} \frac{x^r}{r!} \quad n = 1, 2, \ldots
\end{aligned}
$$

42.5.3 Special Cases

The usage of the incomplete Gamma function $P(a,x)$ in calculations made in this document often involves integer or half-integer values for a. These cases may be solved by the following formulæ

$$
P(n,x) = 1 - e^{-x} \sum_{k=0}^{n-1} \frac{x_k}{k!}
$$

$$
\begin{aligned}
P\left(\tfrac{1}{2},x\right) &= \mathrm{erf}\sqrt{x} \\
P(a+1,x) &= P(a,x)-\frac{x^a e^{-x}}{\Gamma(a+1)} = P(a,x)-\frac{x^a e^{-x}}{a\Gamma(a)} \\
P\left(\tfrac{n}{2},x\right) &= \mathrm{erf}\sqrt{x}-\sum_{k=1}^{\frac{n-1}{2}}\frac{x^{\frac{2k-1}{2}}e^{-x}}{\Gamma\left(\frac{2k+1}{2}\right)} = \mathrm{erf}\sqrt{x}-2e^{-x}\sqrt{\frac{x}{\pi}}\sum_{k=1}^{\frac{n-1}{2}}\frac{(2x)^{k-1}}{(2k-1)!!}
\end{aligned}
$$

the last formula for odd values of n.

42.6 The Beta Function

The Beta function is defined through the integral formula

$$
B(a,b)=B(b,a)=\int_0^1 t^{a-1}(1-t)^{b-1}dt
$$

and is related to the Gamma function by

$$
B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}
$$

The most straightforward way to calculate the Beta function is by using this last expression and a well optimized routine for the Gamma function. In table 9 on page 181 expressions for the Beta function for low integer and half-integer arguments are given.

Another integral, obtained by the substitution $x=t/(1-t)$, yielding the Beta function is

$$
B(a,b)=\int_0^\infty \frac{x^{a-1}}{(1+x)^{a+b}}dx
$$

42.7 The Incomplete Beta Function

The incomplete Beta function is defined as

$$
I_x(a,b)=\frac{B_x(a,b)}{B(a,b)}=\frac{1}{B(a,b)}\int_0^x t^{a-1}(1-t)^{b-1}dt
$$

for $a,b>0$ and $0\le x\le 1$.

The function $B_x(a,b)$, often also called the incomplete Beta function, satisfies the following formula

$$
\begin{aligned}
B_x(a,b) &= \int_0^{\frac{x}{1-x}} \frac{u^{a-1}}{(1+u)^{a+b}}du = B_1(b,a)-B_{1-x}(b,a)= \\
&= x^a\left[\frac{1}{a}+\frac{1-b}{a+1}x+\frac{(1-b)(2-b)}{2!(a+2)}x^2+\right. \\
&\qquad \left.\cdots+\frac{(1-b)(2-b)\cdots(n-b)}{n!(a+n)}x^n+\cdots\right]
\end{aligned}
$$

In figure 36 the incomplete Beta function is shown for a few (a,b)-values. Note that by symmetry the $(1,5)$ and $(5,1)$ curves are reflected around the diagonal. For large values of a and b the curve rises sharply from near zero to near one around $x = a/(a+b)$.

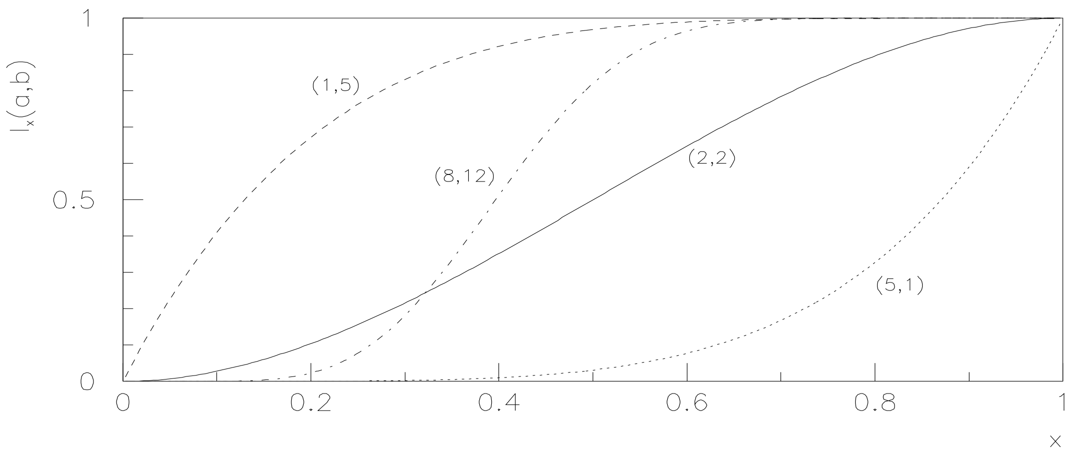


Figure 36: The incomplete Beta function

42.7.1 Numerical Calculation

In order to obtain $I_x(a,b)$ the series expansion

$$I_x(a,b) = \frac{x^a(1-x)^b}{aB(a,b)}\left[1 + \sum_{n=0}^{\infty} \frac{B(a+1,n+1)}{B(a+b,n+1)} x^{n+1}\right]$$

is not the most useful formula for computations. The continued fraction formula

$$I_x(a,b) = \frac{x^a(1-x)^b}{aB(a,b)}\left[\frac{1}{1+}\frac{d_1}{1+}\frac{d_2}{1+}\cdots\right]$$

turns out to be a better choice [9]. Here

$$d_{2m+1} = -\frac{(a+m)(a+b+m)x}{(a+2m)(a+2m+1)} \quad \text{and} \quad d_{2m} = \frac{m(b-m)x}{(a+2m-1)(a+2m)}$$

and the formula converges rapidly for $x < (a+1)/(a+b+1)$. For other x-values the same formula may be used after applying the symmetry relation

$$I_x(a,b) = 1 - I_{1-x}(b,a)$$

42.7.2 Approximation

For higher values of a and b, well already from $a+b>6$, the incomplete Beta function may be approximated by

- For $(a+b+1)(1-x) \leq 0.8$ using an approximation to the chi-square distribution in the variable $\chi^2 = (a+b-1)(1-x)(3-x) - (1-x)(b-1)$ with $n = 2b$ degrees of freedom.

- For $(a+b+1)(1-x) \geq 0.8$ using an approximation to the standard normal distribution in the variable
$$z = \frac{3\left[w_1\left(1-\frac{1}{9b}\right) - w_2\left(1-\frac{1}{9a}\right)\right]}{\sqrt{\frac{w_1^2}{b}+\frac{w_2^2}{a}}}$$
where $w_1 = \sqrt[3]{bx}$ and $w_2 = \sqrt[3]{a(1-x)}$

In both cases the maximum difference to the true cumulative distribution is below 0.005 all way down to the limit where $a+b=6$ [25].

42.8 Relations to Probability Density Functions

The incomplete Gamma and Beta functions, $P(a,x)$ and $I_x(a,b)$ are related to many standard probability density functions or rather to their cumulative (distribution) functions. We give very brief examples here. For more details on these distributions consult any book in statistics.

42.8.1 The Beta Distribution

The cumulative distribution for the Beta distribution with parameters p and q is given by

$$F(x) = \frac{1}{B(p,q)} \int_0^x t^{p-1}(1-t)^{q-1} dt = \frac{B_x(p,q)}{B(p,q)} = I_x(p,q)$$

i.e. simply the incomplete Beta function.

42.8.2 The Binomial Distribution

For the binomial distribution with parameters n and p

$$\sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j} = I_p(k, n-k+1)$$

i.e. the cumulative distribution may be obtained by

$$P(k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} = I_{1-p}(n-k, k+1)$$

However, be careful to evaluate $P(n)$, which obviously is unity, using the incomplete Beta function since this is not defined for arguments which are less or equal to zero.

42.8.3 The Chi-squared Distribution

The cumulative chi-squared distribution for n degrees of freedom is given by

$$
\begin{aligned}
F(x) &= \frac{1}{2\Gamma\left(\frac{n}{2}\right)} \int_0^x \left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{2\Gamma\left(\frac{n}{2}\right)} \int_0^{\frac{x}{2}} y^{\frac{n}{2}-1} e^{-y} 2dy = \\
&= \frac{\gamma\left(\frac{n}{2}, \frac{x}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = P\left(\frac{n}{2}, \frac{x}{2}\right)
\end{aligned}
$$

where x is the chi-squared value sometimes denoted χ^2. In this calculation we made the simple substitution $y = x/2$ in simplifying the integral.

42.8.4 The F-distribution

The cumulative F-distribution with m and n degrees of freedom is given by

$$
\begin{aligned}
F(x) &= \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \int_0^x \frac{m^{\frac{m}{2}} n^{\frac{n}{2}} F^{\frac{m}{2}-1}}{(mF+n)^{\frac{m+n}{2}}} dF = \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \int_0^x \left(\frac{mF}{mF+n}\right)^{\frac{m}{2}} \left(\frac{n}{mF+n}\right)^{\frac{n}{2}} \frac{dF}{F} = \\
&= \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \int_0^{\frac{mx}{mx+n}} y^{\frac{m}{2}} (1-y)^{\frac{n}{2}} \frac{dy}{y(1-y)} = \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \int_0^{\frac{mx}{mx+n}} y^{\frac{m}{2}-1} (1-y)^{\frac{n}{2}-1} dy = \\
&= \frac{B_z\left(\frac{m}{2}, \frac{n}{2}\right)}{B\left(\frac{m}{2}, \frac{n}{2}\right)} = I_z\left(\frac{m}{2}, \frac{n}{2}\right)
\end{aligned}
$$

with $z = mx/(n + mx)$. Here we have made the substitution $y = mF/(mF + n)$, leading to $dF/F = dy/y(1 - y)$, in simplifying the integral.

42.8.5 The Gamma Distribution

Not surprisingly the cumulative distribution for the Gamma distribution with parameters a and b is given by an incomplete Gamma function.

$$
\begin{aligned}
F(x) &= \int_0^x f(x)dx = \frac{a^b}{\Gamma(b)} \int_0^x u^{b-1} e^{-au} du = \frac{a^b}{\Gamma(b)} \int_0^{ax} \left(\frac{v}{a}\right)^{b-1} e^{-v} \frac{dv}{a} = \\
&= \frac{1}{\Gamma(b)} \int_0^{ax} v^{b-1} e^{-v} dv = \frac{\gamma(b, ax)}{\Gamma(b)} = P(b, ax)
\end{aligned}
$$

42.8.6 The Negative Binomial Distribution

The negative binomial distribution with parameters n and p is related to the incomplete Beta function via the relation

$$
\sum_{s=a}^{n} \binom{n+s-1}{s} p^n (1-p)^s = I_{1-p}(a, n)
$$

Also the geometric distribution, a special case of the negative binomial distribution, is connected to the incomplete Beta function, see summary below.

42.8.7 The Normal Distribution

The cumulative normal, or Gaussian, distribution is given by[13]

$$F(x) = \begin{cases} \frac{1}{2} + \frac{1}{2} P\left(\frac{1}{2}, \frac{x^2}{2}\right) & \text{if } x \geq 0 \\ \frac{1}{2} - \frac{1}{2} P\left(\frac{1}{2}, \frac{x^2}{2}\right) & \text{if } x < 0 \end{cases}$$

where $P\left(\frac{1}{2}, \frac{x^2}{2}\right)$ is the incomplete Gamma function occurring as twice the integral of the standard normal curve from 0 to x since

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt &= \frac{1}{2\sqrt{\pi}} \int_0^{\frac{x^2}{2}} \frac{e^{-u}}{\sqrt{u}} du = \frac{1}{2\Gamma\left(\frac{1}{2}\right)} \int_0^{\frac{x^2}{2}} u^{-\frac{1}{2}} e^{-u} du = \\ &= \frac{\gamma\left(\frac{1}{2}, \frac{x^2}{2}\right)}{2\Gamma\left(\frac{1}{2}\right)} = \frac{1}{2} P\left(\frac{1}{2}, \frac{x^2}{2}\right) \end{aligned}$$

The so called error function may be expressed in terms of the incomplete Gamma function

$$\text{erf } x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = P\left(\tfrac{1}{2}, x^2\right)$$

as is the case for the complementary error function

$$\text{erfc } x = 1 - \text{erf } x = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt = 1 - P\left(\tfrac{1}{2}, x^2\right)$$

defined for $x \geq 0$, for $x < 0$ use $\text{erf}(-x) = -\text{erf}(x)$ and $\text{erfc}(-x) = 1 + \text{erf}(x)$. See also section 13.

There are also other series expansions for erf x like

$$\begin{aligned} \text{erf } x &= \frac{2}{\sqrt{\pi}} \left[x - \frac{x^3}{3 \cdot 1!} + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7 \cdot 3!} + \ldots \right] = \\ &= 1 - \frac{e^{-x^2}}{\sqrt{\pi} x} \left[1 - \frac{1}{2x^2} + \frac{1 \cdot 3}{(2x^2)^2} - \frac{1 \cdot 3 \cdot 5}{(2x^2)^3} + \ldots \right] \end{aligned}$$

42.8.8 The Poisson Distribution

Although the Poisson distribution is a probability density function in a discrete variable the cumulative distribution may be expressed in terms of the incomplete Gamma function. The probability for outcomes from zero to $k-1$ inclusive for a Poisson distribution with parameter (mean) μ is

$$P_\mu(< k) = \sum_{n=0}^{k-1} \frac{\mu^n e^{-\mu}}{n!} = 1 - P(k, \mu) \quad \text{for} \quad k = 1, 2, \ldots$$

[13]Without loss of generality it is enough to regard the standard normal density.

42.8.9 Student's t-distribution

The symmetric integral of the t-distribution with n degrees of freedom, often denoted $A(t|n)$, is given by

$$
\begin{aligned}
A(t|n) &= \frac{1}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)} \int\limits_{-t}^{t} \left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} dx = \frac{2}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)} \int\limits_{0}^{t} \left(\frac{n}{n+x^2}\right)^{\frac{n+1}{2}} dx = \\
&= \frac{2}{\sqrt{n}B\left(\frac{1}{2},\frac{n}{2}\right)} \int\limits_{0}^{\frac{t^2}{n+t^2}} (1-y)^{\frac{n+1}{2}} \frac{n}{2}\left(\frac{1}{1-y}\right)^2 \sqrt{\frac{1-y}{ny}} dy = \\
&= \frac{1}{B\left(\frac{1}{2},\frac{n}{2}\right)} \int\limits_{0}^{\frac{t^2}{n+t^2}} y^{-\frac{1}{2}}(1-y)^{\frac{n}{2}-1} dy = \frac{B_z\left(\frac{1}{2},\frac{n}{2}\right)}{B\left(\frac{1}{2},\frac{n}{2}\right)} = I_z\left(\frac{1}{2},\frac{n}{2}\right)
\end{aligned}
$$

with $z = t^2/(n+t^2)$.

42.8.10 Summary

The following table summarizes the relations between the cumulative, distribution, functions of some standard probability density functions and the incomplete Gamma and Beta functions.

Distribution	Parameters	Cumulative distribution	Range
Beta	p, q	$F(x) = I_x(p,q)$	$0 \leq x \leq 1$
Binomial	n, p	$P(k) = I_{1-p}(n-k, k+1)$	$k = 0, 1, \ldots, n$
Chi-squared	n	$F(x) = P\left(\frac{n}{2}, \frac{x}{2}\right)$	$x \geq 0$
F	m, n	$F(x) = I_{\frac{mx}{n+mx}}\left(\frac{m}{2}, \frac{n}{2}\right)$	$x \geq 0$
Gamma	a, b	$F(x) = P(b, ax)$	$x \geq 0$
Geometric	p	$P(k) = I_p(1, k)$	$k = 1, 2, \ldots$
Negative binomial	n, p	$P(k) = I_p(n, k+1)$	$k = 0, 1, \ldots$
Standard normal		$F(x) = \frac{1}{2} - \frac{1}{2}P\left(\frac{1}{2}, \frac{x^2}{2}\right)$	$-\infty < x < 0$
		$F(x) = \frac{1}{2} + \frac{1}{2}P\left(\frac{1}{2}, \frac{x^2}{2}\right)$	$0 \leq x < \infty$
Poisson	μ	$P(k) = 1 - P(k+1, \mu)$	$k = 0, 1, \ldots$
Student	n	$F(x) = \frac{1}{2} - \frac{1}{2}I_{\frac{x^2}{n+x^2}}\left(\frac{1}{2}, \frac{n}{2}\right)$	$-\infty < x < 0$
		$F(x) = \frac{1}{2} + \frac{1}{2}I_{\frac{x^2}{n+x^2}}\left(\frac{1}{2}, \frac{n}{2}\right)$	$0 \leq x < \infty$

43 Appendix B: Hypergeometric Functions

43.1 Introduction

The hypergeometric and the confluent hypergeometric functions has a central role inasmuch as many standard functions may be expressed in terms of them. This appendix is based on information from [39,44,45] in which much more detailed information on the hypergeometric and confluent hypergeometric function may be found.

43.2 Hypergeometric Function

The hypergeometric function, sometimes called Gauss's differential equation, is given by [39,44]

$$x(1-x)\frac{\partial^2 f(x)}{\partial x^2} + [c-(a+b+1)x]\frac{\partial f(x)}{\partial x} - abf(x) = 0$$

One solution is

$$f(x) = {}_2F_1(a,b,c;x) = 1 + \frac{ab}{c}\frac{x}{1!} + \frac{a(a+1)b(b+1)}{c(c+1)}\frac{x^2}{2!} + \cdots \qquad c \neq 0,-1,-2,-3,\ldots$$

The range of convergence is $|x| < 1$ and $x = 1$, for $c > a+b$, and $x = -1$, for $c > a+b-1$. Using the so called Pochhammer symbol

$$(a)_n = a(a+1)(a+2)\cdots(a+n-1) = \frac{(a+n-1)!}{(a-1)!} = \frac{\Gamma(a+n)}{\Gamma(a)}$$

with $(a)_0 = 1$ this solution may be written[14] as

$${}_2F_1(a,b,c;x) = \sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{x^n}{n!} = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{n=0}^{\infty}\frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)}\frac{x^n}{n!}$$

By symmetry ${}_2F_1(a,b,c;x) = {}_2F_1(b,a,c;x)$ and sometimes the indices are dropped and when the risk for confusion is negligible one simply writes $F(a,b,c;x)$.

Another independent solution to the hypergeometric equation is

$$f(x) = x^{1-c}\ {}_2F_1(a+1-c, b+1-c, 2-c; x) \qquad c \neq 2,3,4,\ldots$$

The n:th derivative of the hypergeometric function is given by

$$\frac{d^n}{dx^n}\ {}_2F_1(a,b,c;x) = \frac{(a)_n(b)_n}{(c)_n}\ {}_2F_1(a+n,b+n,c+n;x)$$

and

$${}_2F_1(a,b,c;x) = (1-x)^{c-a-b}\ {}_2F_1(c-a,c-b,c;x)$$

Several common mathematical function may be expressed in terms of the hypergeometric function such as, the incomplete Beta function $B_x(a,b)$, the complete elliptical integrals

[14] The notation ${}_2F_1$ indicates the presence of two Pochhammer symbols in the numerator and one in the denominator.

K and E, the Gegenbauer functions $T_n^\beta(x)$, the Legendre functions $P_n(x)$, $P_n^m(x)$ and $Q_\nu(x)$ (second kind), and the Chebyshev functions $T_n(x)$, $U_n(x)$ and $V_n(x)$

$$\begin{aligned}
(1-z)^{-a} &= {}_2F_1(a,b,b;z) \\
\ln(1+z) &= x\cdot {}_2F_1(1,1,2;-z) \\
\arctan z &= z\cdot {}_2F_1\left(\tfrac{1}{2},1,\tfrac{3}{2};-z^2\right) \\
\arcsin z &= z\cdot {}_2F_1\left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{3}{2};z^2\right) = z\sqrt{1-z^2}\ {}_2F_1\left(1,1,\tfrac{3}{2};z^2\right) \\
B_x(a,b) &= \frac{x^a}{a}\ {}_2F_1(a,1-b,a+1;x) \\
K &= \int_0^{\frac{\pi}{2}}(1-k^2\sin^2\theta)^{-\frac{1}{2}}d\theta = \frac{\pi}{2}\ {}_2F_1\left(\tfrac{1}{2},\tfrac{1}{2},1;k^2\right) \\
E &= \int_0^{\frac{\pi}{2}}(1-k^2\sin^2\theta)^{\frac{1}{2}}d\theta = \frac{\pi}{2}\ {}_2F_1\left(\tfrac{1}{2},-\tfrac{1}{2},1;k^2\right) \\
T_n^\beta(x) &= \frac{(n+2\beta)!}{2^\beta n!\beta!}\ {}_2F_1\left(-n,n+2\beta+1,1+\beta;\tfrac{1-x}{2}\right) \\
P_n(x) &= {}_2F_1\left(-n,n+1,1;\tfrac{1-x}{2}\right) \\
P_n^m(x) &= \frac{(n+m)!}{(n-m)!}\frac{(1-x^2)^{\frac{m}{2}}}{2^m m!}\ {}_2F_1\left(m-n,m+n+1,m+1;\tfrac{1-x}{2}\right) \\
P_{2n}(x) &= (-1)^n\frac{(2n-1)!!}{(2n)!!}\ {}_2F_1\left(-n,n+\tfrac{1}{2},\tfrac{1}{2};x^2\right) \\
P_{2n+1}(x) &= (-1)^n\frac{(2n+1)!!}{(2n)!!}\ x\cdot {}_2F_1\left(-n,n+\tfrac{3}{2},\tfrac{3}{2};x^2\right) \\
Q_\nu(x) &= \frac{\sqrt{\pi}\nu!}{\left(\nu+\frac{1}{2}\right)!(2x)^{\nu+1}}\ {}_2F_1\left(\tfrac{\nu+1}{2},\tfrac{\nu}{2}+1,\tfrac{\nu+3}{2};\tfrac{1}{x^2}\right) \\
T_n(x) &= {}_2F_1\left(-n,n,\tfrac{1}{2};\tfrac{1-x}{2}\right) \\
U_n(x) &= (n+1)\cdot {}_2F_1\left(-n,n+2,\tfrac{3}{2};\tfrac{1-x}{2}\right) \\
V_n(x) &= n\sqrt{1-x^2}\ {}_2F_1\left(-n+1,n+1,\tfrac{3}{2};\tfrac{1-x}{2}\right)
\end{aligned}$$

for $Q_\nu(x)$ the conditions are $|x|>1, |\arg x|<\pi$, and $\nu\neq -1,-2,-3,\ldots$. See [44] for many more similar and additional formulæ.

43.3 Confluent Hypergeometric Function

The confluent hypergeometric equation, or Kummer's equation as it is often called, is given by [39,45]

$$x\frac{\partial^2 f(x)}{\partial x^2} + (c-x)\frac{\partial f(x)}{\partial x} - af(x) = 0$$

One solution to this equation is

$$f(x) = {}_1F_1(a,c;x) = M(a,c;x) = \sum_{n=0}^{\infty} \frac{(a)_n}{(c)_n} \frac{x^n}{n!} \qquad c \neq 0,-1,-2,\ldots$$

This solution is convergent for all finite real x (or complex z). Another solution is given by

$$f(x) = x^{1-c} M(a+1-c, 2-c; x) \qquad c \neq 2,3,4,\ldots$$

Often a linear combination of the first and second solution is used

$$U(a,c;x) = \frac{\pi}{\sin \pi c}\left[\frac{M(a,c;x)}{(a-c)!(c-1)!} - \frac{x^{1-c}M(a+1-c,2-c;x)}{(a-1)!(1-c)!}\right]$$

The confluent hypergeometric functions M and U may be expressed in integral form as

$$\begin{aligned}
M(a,c;x) &= \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 e^{xt} t^{a-1} (1-t)^{c-a-1} dt \qquad Re\ c > 0,\ Re\ a > 0 \\
U(a,c;x) &= \frac{1}{\Gamma(a)} \int_0^\infty e^{-xt} t^{a-1} (1+t)^{c-a-1} dt \qquad Re\ x > 0,\ Re\ a > 0
\end{aligned}$$

Useful formulæ are the Kummer transformations

$$\begin{aligned}
M(a,c;x) &= e^x M(c-a, c; -x) \\
U(a,c;x) &= x^{1-c} U(a-c+1, 2-c; x)
\end{aligned}$$

The n:th derivatives of the confluent hypergeometric functions are given by

$$\begin{aligned}
\frac{d^n}{dz^n} M(a,b;z) &= \frac{(a)_n}{(b)_n} M(a+n, b+n; z) \\
\frac{d^n}{dz^n} U(a,b;z) &= (-1)^n (a)_n U(a+n, b+n; z)
\end{aligned}$$

Several common mathematical function may be expressed in terms of the hypergeometric function such as the error function, the incomplete Gamma function $\gamma(a,x)$, Bessel functions $J_\nu(x)$, modified Bessel functions of the first kind $I_\nu(x)$, Hermite functions $H_n(x)$, Laguerre functions $L_n(x)$, associated Laguerre functions $L_n^m(x)$, Whittaker functions $M_{k\mu}(x)$ and $W_{k\mu}(x)$, Fresnel integrals $C(x)$ and $S(x)$, modified Bessel function of the second kind $K_\nu(x)$

$$\begin{aligned}
e^z &= M(a,a;z) \\
\mathrm{erf}(x) &= \frac{2}{\sqrt{\pi}} x M\left(\tfrac{1}{2}, \tfrac{3}{2}; -x^2\right) = \frac{2}{\sqrt{\pi}} x e^{-x^2} M\left(1, \tfrac{3}{2}; x^2\right) \\
\gamma(a,x) &= \frac{x^a}{a} M(a, a+1; -x) \qquad Re\ a > 0 \\
J_\nu(x) &= \frac{e^{-\imath x}}{\nu!} \left(\frac{x}{2}\right)^\nu M\left(\nu+\tfrac{1}{2}, 2\nu+1; 2\imath x\right)
\end{aligned}$$

$$
\begin{aligned}
I_\nu(x) &= \frac{e^{-x}}{\nu!}\left(\frac{x}{2}\right)^\nu M\left(\nu+\tfrac{1}{2}, 2\nu+1; 2x\right) \\
H_{2n}(x) &= (-1)^n \frac{(2n)!}{n!} M\left(-n, \tfrac{1}{2}; x^2\right) \\
H_{2n+1}(x) &= (-1)^n \frac{2(2n+1)!}{n!} x M\left(-n, \tfrac{3}{2}; x^2\right) \\
L_n(x) &= M(-n, 1; x) \\
L_n^m(x) &= (-1)^m \frac{\partial^m}{\partial x^m} L_{n+m}(x) = \frac{(n+m)!}{n!m!} M(-n, m+1; x) \\
M_{k\mu}(x) &= e^{-\frac{x}{2}} x^{\mu+\frac{1}{2}} M\left(\mu-k+\tfrac{1}{2}, 2\mu+1; x\right) \\
W_{k\mu}(x) &= e^{-\frac{x}{2}} x^{\mu+\frac{1}{2}} U\left(\mu-k+\tfrac{1}{2}, 2\mu+1; x\right) \\
C(x) + \imath S(x) &= x M\left(\frac{1}{x}, \frac{3}{2}; \frac{\imath \pi x^2}{2}\right) \\
K_\nu(x) &= \sqrt{\pi} e^{-x} (2x)^\nu U\left(\nu+\tfrac{1}{2}, 2\nu+1; 2x\right)
\end{aligned}
$$

See [45] for many more similar and additional formulæ.

Table 1: Percentage points of the chi-square distribution

	$1-\alpha$							
n	0.5000	0.8000	0.9000	0.9500	0.9750	0.9900	0.9950	0.9990
1	0.4549	1.6424	2.7055	3.8415	5.0239	6.6349	7.8794	10.828
2	1.3863	3.2189	4.6052	5.9915	7.3778	9.2103	10.597	13.816
3	2.3660	4.6416	6.2514	7.8147	9.3484	11.345	12.838	16.266
4	3.3567	5.9886	7.7794	9.4877	11.143	13.277	14.860	18.467
5	4.3515	7.2893	9.2364	11.070	12.833	15.086	16.750	20.515
6	5.3481	8.5581	10.645	12.592	14.449	16.812	18.548	22.458
7	6.3458	9.8032	12.017	14.067	16.013	18.475	20.278	24.322
8	7.3441	11.030	13.362	15.507	17.535	20.090	21.955	26.124
9	8.3428	12.242	14.684	16.919	19.023	21.666	23.589	27.877
10	9.3418	13.442	15.987	18.307	20.483	23.209	25.188	29.588
11	10.341	14.631	17.275	19.675	21.920	24.725	26.757	31.264
12	11.340	15.812	18.549	21.026	23.337	26.217	28.300	32.909
13	12.340	16.985	19.812	22.362	24.736	27.688	29.819	34.528
14	13.339	18.151	21.064	23.685	26.119	29.141	31.319	36.123
15	14.339	19.311	22.307	24.996	27.488	30.578	32.801	37.697
16	15.338	20.465	23.542	26.296	28.845	32.000	34.267	39.252
17	16.338	21.615	24.769	27.587	30.191	33.409	35.718	40.790
18	17.338	22.760	25.989	28.869	31.526	34.805	37.156	42.312
19	18.338	23.900	27.204	30.144	32.852	36.191	38.582	43.820
20	19.337	25.038	28.412	31.410	34.170	37.566	39.997	45.315
21	20.337	26.171	29.615	32.671	35.479	38.932	41.401	46.797
22	21.337	27.301	30.813	33.924	36.781	40.289	42.796	48.268
23	22.337	28.429	32.007	35.172	38.076	41.638	44.181	49.728
24	23.337	29.553	33.196	36.415	39.364	42.980	45.559	51.179
25	24.337	30.675	34.382	37.652	40.646	44.314	46.928	52.620
26	25.336	31.795	35.563	38.885	41.923	45.642	48.290	54.052
27	26.336	32.912	36.741	40.113	43.195	46.963	49.645	55.476
28	27.336	34.027	37.916	41.337	44.461	48.278	50.993	56.892
29	28.336	35.139	39.087	42.557	45.722	49.588	52.336	58.301
30	29.336	36.250	40.256	43.773	46.979	50.892	53.672	59.703
40	39.335	47.269	51.805	55.758	59.342	63.691	66.766	73.402
50	49.335	58.164	63.167	67.505	71.420	76.154	79.490	86.661
60	59.335	68.972	74.397	79.082	83.298	88.379	91.952	99.607
70	69.334	79.715	85.527	90.531	95.023	100.43	104.21	112.32
80	79.334	90.405	96.578	101.88	106.63	112.33	116.32	124.84
90	89.334	101.05	107.57	113.15	118.14	124.12	128.30	137.21
100	99.334	111.67	118.50	124.34	129.56	135.81	140.17	149.45

Table 2: Extreme confidence levels for the chi-square distribution

	Chi-square Confidence Levels (as χ^2 values)											
d.f.	0.1	0.01	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}	10^{-11}	10^{-12}
1	2.71	6.63	10.8	15.1	19.5	23.9	28.4	32.8	37.3	41.8	46.3	50.8
2	4.61	9.21	13.8	18.4	23.0	27.6	32.2	36.8	41.4	46.1	50.7	55.3
3	6.25	11.3	16.3	21.1	25.9	30.7	35.4	40.1	44.8	49.5	54.2	58.9
4	7.78	13.3	18.5	23.5	28.5	33.4	38.2	43.1	47.9	52.7	57.4	62.2
5	9.24	15.1	20.5	25.7	30.9	35.9	40.9	45.8	50.7	55.6	60.4	65.2
6	10.6	16.8	22.5	27.9	33.1	38.3	43.3	48.4	53.3	58.3	63.2	68.1
7	12.0	18.5	24.3	29.9	35.3	40.5	45.7	50.8	55.9	60.9	65.9	70.8
8	13.4	20.1	26.1	31.8	37.3	42.7	48.0	53.2	58.3	63.4	68.4	73.5
9	14.7	21.7	27.9	33.7	39.3	44.8	50.2	55.4	60.7	65.8	70.9	76.0
10	16.0	23.2	29.6	35.6	41.3	46.9	52.3	57.7	62.9	68.2	73.3	78.5
11	17.3	24.7	31.3	37.4	43.2	48.9	54.4	59.8	65.2	70.5	75.7	80.9
12	18.5	26.2	32.9	39.1	45.1	50.8	56.4	61.9	67.3	72.7	78.0	83.2
13	19.8	27.7	34.5	40.9	46.9	52.7	58.4	64.0	69.5	74.9	80.2	85.5
14	21.1	29.1	36.1	42.6	48.7	54.6	60.4	66.0	71.6	77.0	82.4	87.8
15	22.3	30.6	37.7	44.3	50.5	56.5	62.3	68.0	73.6	79.1	84.6	90.0
16	23.5	32.0	39.3	45.9	52.2	58.3	64.2	70.0	75.7	81.2	86.7	92.2
17	24.8	33.4	40.8	47.6	54.0	60.1	66.1	71.9	77.6	83.3	88.8	94.3
18	26.0	34.8	42.3	49.2	55.7	61.9	68.0	73.8	79.6	85.3	90.9	96.4
19	27.2	36.2	43.8	50.8	57.4	63.7	69.8	75.7	81.6	87.3	92.9	98.5
20	28.4	37.6	45.3	52.4	59.0	65.4	71.6	77.6	83.5	89.3	94.9	101
25	34.4	44.3	52.6	60.1	67.2	73.9	80.4	86.6	92.8	98.8	105	111
30	40.3	50.9	59.7	67.6	75.0	82.0	88.8	95.3	102	108	114	120
35	46.1	57.3	66.6	74.9	82.6	89.9	97.0	104	110	117	123	129
40	51.8	63.7	73.4	82.1	90.1	97.7	105	112	119	125	132	138
45	57.5	70.0	80.1	89.1	97.4	105	113	120	127	134	140	147
50	63.2	76.2	86.7	96.0	105	113	120	128	135	142	149	155
60	74.4	88.4	99.6	110	119	127	135	143	150	158	165	172
70	85.5	100	112	123	132	141	150	158	166	173	181	188
80	96.6	112	125	136	146	155	164	172	180	188	196	204
90	108	124	137	149	159	169	178	187	195	203	211	219
100	118	136	149	161	172	182	192	201	209	218	226	234
120	140	159	174	186	198	209	219	228	237	246	255	263
150	173	193	209	223	236	247	258	268	278	288	297	306
200	226	249	268	283	297	310	322	333	344	355	365	374
300	332	360	381	400	416	431	445	458	471	483	495	506
400	437	469	493	514	532	549	565	580	594	607	620	632
500	541	576	603	626	646	665	682	698	714	728	742	756
600	645	684	713	737	759	779	798	815	832	847	862	877
800	852	896	929	957	982	1005	1026	1045	1064	1081	1098	1114
1000	1058	1107	1144	1175	1202	1227	1250	1272	1292	1311	1330	1348

Table 3: Extreme confidence levels for the chi-square distribution (as χ^2/d.f. values)

	Chi-square Confidence Levels (as χ^2/d.f. values)											
d.f.	0.1	0.01	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}	10^{-11}	10^{-12}
1	2.71	6.63	10.83	15.14	19.51	23.93	28.37	32.84	37.32	41.82	46.33	50.84
2	2.30	4.61	6.91	9.21	11.51	13.82	16.12	18.42	20.72	23.03	25.33	27.63
3	2.08	3.78	5.42	7.04	8.63	10.22	11.80	13.38	14.95	16.51	18.08	19.64
4	1.94	3.32	4.62	5.88	7.12	8.34	9.56	10.77	11.97	13.17	14.36	15.55
5	1.85	3.02	4.10	5.15	6.17	7.18	8.17	9.16	10.14	11.11	12.08	13.05
6	1.77	2.80	3.74	4.64	5.52	6.38	7.22	8.06	8.89	9.72	10.54	11.35
7	1.72	2.64	3.47	4.27	5.04	5.79	6.53	7.26	7.98	8.70	9.41	10.12
8	1.67	2.51	3.27	3.98	4.67	5.34	6.00	6.65	7.29	7.92	8.56	9.18
9	1.63	2.41	3.10	3.75	4.37	4.98	5.57	6.16	6.74	7.31	7.88	8.45
10	1.60	2.32	2.96	3.56	4.13	4.69	5.23	5.77	6.29	6.82	7.33	7.85
11	1.57	2.25	2.84	3.40	3.93	4.44	4.94	5.44	5.92	6.41	6.88	7.35
12	1.55	2.18	2.74	3.26	3.76	4.24	4.70	5.16	5.61	6.06	6.50	6.93
13	1.52	2.13	2.66	3.14	3.61	4.06	4.49	4.92	5.34	5.76	6.17	6.58
14	1.50	2.08	2.58	3.04	3.48	3.90	4.31	4.72	5.11	5.50	5.89	6.27
15	1.49	2.04	2.51	2.95	3.37	3.77	4.16	4.54	4.91	5.28	5.64	6.00
16	1.47	2.00	2.45	2.87	3.27	3.65	4.01	4.37	4.73	5.08	5.42	5.76
17	1.46	1.97	2.40	2.80	3.17	3.54	3.89	4.23	4.57	4.90	5.22	5.55
18	1.44	1.93	2.35	2.73	3.09	3.44	3.78	4.10	4.42	4.74	5.05	5.36
19	1.43	1.90	2.31	2.67	3.02	3.35	3.67	3.99	4.29	4.59	4.89	5.18
20	1.42	1.88	2.27	2.62	2.95	3.27	3.58	3.88	4.17	4.46	4.75	5.03
25	1.38	1.77	2.10	2.41	2.69	2.96	3.21	3.47	3.71	3.95	4.19	4.42
30	1.34	1.70	1.99	2.25	2.50	2.73	2.96	3.18	3.39	3.60	3.80	4.00
35	1.32	1.64	1.90	2.14	2.36	2.57	2.77	2.96	3.15	3.34	3.52	3.69
40	1.30	1.59	1.84	2.05	2.25	2.44	2.62	2.80	2.97	3.13	3.29	3.45
45	1.28	1.55	1.78	1.98	2.16	2.34	2.50	2.66	2.82	2.97	3.12	3.26
50	1.26	1.52	1.73	1.92	2.09	2.25	2.41	2.55	2.70	2.84	2.97	3.11
60	1.24	1.47	1.66	1.83	1.98	2.12	2.25	2.38	2.51	2.63	2.75	2.86
70	1.22	1.43	1.60	1.75	1.89	2.02	2.14	2.25	2.37	2.48	2.58	2.68
80	1.21	1.40	1.56	1.70	1.82	1.94	2.05	2.15	2.26	2.35	2.45	2.54
90	1.20	1.38	1.52	1.65	1.77	1.87	1.98	2.07	2.17	2.26	2.35	2.43
100	1.18	1.36	1.49	1.61	1.72	1.82	1.92	2.01	2.09	2.18	2.26	2.34
120	1.17	1.32	1.45	1.55	1.65	1.74	1.82	1.90	1.98	2.05	2.12	2.19
150	1.15	1.29	1.40	1.49	1.57	1.65	1.72	1.79	1.85	1.92	1.98	2.04
200	1.13	1.25	1.34	1.42	1.48	1.55	1.61	1.67	1.72	1.77	1.82	1.87
300	1.11	1.20	1.27	1.33	1.39	1.44	1.48	1.53	1.57	1.61	1.65	1.69
400	1.09	1.17	1.23	1.28	1.33	1.37	1.41	1.45	1.48	1.52	1.55	1.58
500	1.08	1.15	1.21	1.25	1.29	1.33	1.36	1.40	1.43	1.46	1.48	1.51
600	1.07	1.14	1.19	1.23	1.27	1.30	1.33	1.36	1.39	1.41	1.44	1.46
800	1.06	1.12	1.16	1.20	1.23	1.26	1.28	1.31	1.33	1.35	1.37	1.39
1000	1.06	1.11	1.14	1.17	1.20	1.23	1.25	1.27	1.29	1.31	1.33	1.35

Table 4: Exact and approximate values for the Bernoulli numbers

Bernoulli numbers

n	N/D	$=$	$B_n/10^k$	k
0	1/1	=	1.00000 00000	0
1	−1/2	= −	5.00000 00000	−1
2	1/6	=	1.66666 66667	−1
4	−1/30	= −	3.33333 33333	−2
6	1/42	=	2.38095 23810	−2
8	−1/30	= −	3.33333 33333	−2
10	5/66	=	7.57575 75758	−2
12	−691/2730	= −	2.53113 55311	−1
14	7/6	=	1.16666 66667	0
16	−3 617/510	= −	7.09215 68627	0
18	43 867/798	=	5.49711 77945	1
20	−174 611/330	= −	5.29124 24242	2
22	854 513/138	=	6.19212 31884	3
24	−236 364 091/2 730	= −	8.65802 53114	4
26	8 553 103/6	=	1.42551 71667	6
28	−23 749 461 029/870	= −	2.72982 31068	7
30	8 615 841 276 005/14 322	=	6.01580 87390	8
32	−7 709 321 041 217/510	= −	1.51163 15767	10
34	2 577 687 858 367/6	=	4.29614 64306	11
36	−26 315 271 553 053 477 373/1 919 190	= −	1.37116 55205	13
38	2 929 993 913 841 559/6	=	4.88332 31897	14
40	−261 082 718 496 449 122 051/13 530	= −	1.92965 79342	16
42	1 520 097 643 918 070 802 691/1 806	=	8.41693 04757	17
44	−27 833 269 579 301 024 235 023/690	= −	4.03380 71854	19
46	596 451 111 593 912 163 277 961/282	=	2.11507 48638	21
48	−5 609 403 368 997 817 686 249 127 547/46 410	= −	1.20866 26522	23
50	495 057 205 241 079 648 212 477 525/66	=	7.50086 67461	24
52	−801 165 718 135 489 957 347 924 991 853/1 590	= −	5.03877 81015	26
54	29 149 963 634 884 862 421 418 123 812 691/798	=	3.65287 76485	28
56	−2 479 392 929 313 226 753 685 415 739 663 229/870	= −	2.84987 69302	30
58	84 483 613 348 880 041 862 046 775 994 036 021/354	=	2.38654 27500	32
60	−1 215 233 140 483 755 572 040 304 994 079 820 246 041 491/56 786 730	= −	2.13999 49257	34
62	12 300 585 434 086 858 541 953 039 857 403 386 151/6	=	2.05009 75723	36
64	−106 783 830 147 866 529 886 385 444 979 142 647 942 017/510	= −	2.09380 05911	38
66	1 472 600 022 126 335 654 051 619 428 551 932 342 241 899 101/64 722	=	2.27526 96488	40
68	−78 773 130 858 718 728 141 909 149 208 474 606 244 347 001/30	= −	2.62577 10286	42
70	1 505 381 347 333 367 003 803 076 567 377 857 208 511 438 160 235/4 686	=	3.21250 82103	44

Table 5: Percentage points of the F-distribution

$\alpha = 0.10$

m \ n	1	2	3	4	5	10	20	50	100	∞
1	39.86	49.50	53.59	55.83	57.24	60.19	61.74	62.69	63.01	63.33
2	8.526	9.000	9.162	9.243	9.293	9.392	9.441	9.471	9.481	9.491
3	5.538	5.462	5.391	5.343	5.309	5.230	5.184	5.155	5.144	5.134
4	4.545	4.325	4.191	4.107	4.051	3.920	3.844	3.795	3.778	3.761
5	4.060	3.780	3.619	3.520	3.453	3.297	3.207	3.147	3.126	3.105
10	3.285	2.924	2.728	2.605	2.522	2.323	2.201	2.117	2.087	2.055
20	2.975	2.589	2.380	2.249	2.158	1.937	1.794	1.690	1.650	1.607
50	2.809	2.412	2.197	2.061	1.966	1.729	1.568	1.441	1.388	1.327
100	2.756	2.356	2.139	2.002	1.906	1.663	1.494	1.355	1.293	1.214
∞	2.706	2.303	2.084	1.945	1.847	1.599	1.421	1.263	1.185	1.000

$\alpha = 0.05$

m \ n	1	2	3	4	5	10	20	50	100	∞
1	161.4	199.5	215.7	224.6	230.2	241.9	248.0	251.8	253.0	254.3
2	18.51	19.00	19.16	19.25	19.30	19.40	19.45	19.48	19.49	19.50
3	10.13	9.552	9.277	9.117	9.013	8.786	8.660	8.581	8.554	8.526
4	7.709	6.944	6.591	6.388	6.256	5.964	5.803	5.699	5.664	5.628
5	6.608	5.786	5.409	5.192	5.050	4.735	4.558	4.444	4.405	4.365
10	4.965	4.103	3.708	3.478	3.326	2.978	2.774	2.637	2.588	2.538
20	4.351	3.493	3.098	2.866	2.711	2.348	2.124	1.966	1.907	1.843
50	4.034	3.183	2.790	2.557	2.400	2.026	1.784	1.599	1.525	1.438
100	3.936	3.087	2.696	2.463	2.305	1.927	1.676	1.477	1.392	1.283
∞	3.841	2.996	2.605	2.372	2.214	1.831	1.571	1.350	1.243	1.000

$\alpha = 0.01$

m \ n	1	2	3	4	5	10	20	50	100	∞
1	4052	5000	5403	5625	5764	6056	6209	6303	6334	6366
2	98.50	99.00	99.17	99.25	99.30	99.40	99.45	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.23	26.69	26.35	26.24	26.13
4	21.20	18.00	16.69	15.98	15.52	14.55	14.02	13.69	13.58	13.46
5	16.26	13.27	12.06	11.39	10.97	10.05	9.553	9.238	9.130	9.020
10	10.04	7.559	6.552	5.994	5.636	4.849	4.405	4.115	4.014	3.909
20	8.096	5.849	4.938	4.431	4.103	3.368	2.938	2.643	2.535	2.421
50	7.171	5.057	4.199	3.720	3.408	2.698	2.265	1.949	1.825	1.683
100	6.895	4.824	3.984	3.513	3.206	2.503	2.067	1.735	1.598	1.427
∞	6.635	4.605	3.782	3.319	3.017	2.321	1.878	1.523	1.358	1.000

Table 6: Probability content from $-z$ to z of Gauss distribution in %

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.00	0.80	1.60	2.39	3.19	3.99	4.78	5.58	6.38	7.17
0.1	7.97	8.76	9.55	10.34	11.13	11.92	12.71	13.50	14.28	15.07
0.2	15.85	16.63	17.41	18.19	18.97	19.74	20.51	21.28	22.05	22.82
0.3	23.58	24.34	25.10	25.86	26.61	27.37	28.12	28.86	29.61	30.35
0.4	31.08	31.82	32.55	33.28	34.01	34.73	35.45	36.16	36.88	37.59
0.5	38.29	38.99	39.69	40.39	41.08	41.77	42.45	43.13	43.81	44.48
0.6	45.15	45.81	46.47	47.13	47.78	48.43	49.07	49.71	50.35	50.98
0.7	51.61	52.23	52.85	53.46	54.07	54.67	55.27	55.87	56.46	57.05
0.8	57.63	58.21	58.78	59.35	59.91	60.47	61.02	61.57	62.11	62.65
0.9	63.19	63.72	64.24	64.76	65.28	65.79	66.29	66.80	67.29	67.78
1.0	68.27	68.75	69.23	69.70	70.17	70.63	71.09	71.54	71.99	72.43
1.1	72.87	73.30	73.73	74.15	74.57	74.99	75.40	75.80	76.20	76.60
1.2	76.99	77.37	77.75	78.13	78.50	78.87	79.23	79.59	79.95	80.29
1.3	80.64	80.98	81.32	81.65	81.98	82.30	82.62	82.93	83.24	83.55
1.4	83.85	84.15	84.44	84.73	85.01	85.29	85.57	85.84	86.11	86.38
1.5	86.64	86.90	87.15	87.40	87.64	87.89	88.12	88.36	88.59	88.82
1.6	89.04	89.26	89.48	89.69	89.90	90.11	90.31	90.51	90.70	90.90
1.7	91.09	91.27	91.46	91.64	91.81	91.99	92.16	92.33	92.49	92.65
1.8	92.81	92.97	93.12	93.27	93.42	93.57	93.71	93.85	93.99	94.12
1.9	94.26	94.39	94.51	94.64	94.76	94.88	95.00	95.12	95.23	95.34
2.0	95.45	95.56	95.66	95.76	95.86	95.96	96.06	96.15	96.25	96.34
2.1	96.43	96.51	96.60	96.68	96.76	96.84	96.92	97.00	97.07	97.15
2.2	97.22	97.29	97.36	97.43	97.49	97.56	97.62	97.68	97.74	97.80
2.3	97.86	97.91	97.97	98.02	98.07	98.12	98.17	98.22	98.27	98.32
2.4	98.36	98.40	98.45	98.49	98.53	98.57	98.61	98.65	98.69	98.72
2.5	98.76	98.79	98.83	98.86	98.89	98.92	98.95	98.98	99.01	99.04
2.6	99.07	99.09	99.12	99.15	99.17	99.20	99.22	99.24	99.26	99.29
2.7	99.31	99.33	99.35	99.37	99.39	99.40	99.42	99.44	99.46	99.47
2.8	99.49	99.50	99.52	99.53	99.55	99.56	99.58	99.59	99.60	99.61
2.9	99.63	99.64	99.65	99.66	99.67	99.68	99.69	99.70	99.71	99.72
3.0	99.73	99.74	99.75	99.76	99.76	99.77	99.78	99.79	99.79	99.80
3.1	99.81	99.81	99.82	99.83	99.83	99.84	99.84	99.85	99.85	99.86
3.2	99.86	99.87	99.87	99.88	99.88	99.88	99.89	99.89	99.90	99.90
3.3	99.90	99.91	99.91	99.91	99.92	99.92	99.92	99.92	99.93	99.93
3.4	99.93	99.94	99.94	99.94	99.94	99.94	99.95	99.95	99.95	99.95
3.5	99.95	99.96	99.96	99.96	99.96	99.96	99.96	99.96	99.97	99.97
3.6	99.97	99.97	99.97	99.97	99.97	99.97	99.97	99.98	99.98	99.98
3.7	99.98	99.98	99.98	99.98	99.98	99.98	99.98	99.98	99.98	99.98
3.8	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99
3.9	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99	99.99

Table 7: Standard normal distribution z-values for a specific probability content from $-z$ to z. Read column-wise and add marginal column and row z. Read column-wise and add marginal column and row figures to find probabilities.

Prob.	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.000	0.000	0.125	0.253	0.385	0.524	0.674	0.841	1.036	1.282	1.645
0.002	0.002	0.128	0.256	0.388	0.527	0.677	0.845	1.041	1.287	1.655
0.004	0.005	0.130	0.258	0.390	0.530	0.681	0.849	1.045	1.293	1.665
0.006	0.007	0.133	0.261	0.393	0.533	0.684	0.852	1.049	1.299	1.675
0.008	0.010	0.135	0.263	0.396	0.536	0.687	0.856	1.054	1.305	1.685
0.010	0.012	0.138	0.266	0.398	0.538	0.690	0.859	1.058	1.311	1.696
0.012	0.015	0.141	0.268	0.401	0.541	0.693	0.863	1.063	1.317	1.706
0.014	0.017	0.143	0.271	0.404	0.544	0.696	0.867	1.067	1.323	1.717
0.016	0.020	0.146	0.274	0.407	0.547	0.700	0.870	1.071	1.329	1.728
0.018	0.022	0.148	0.276	0.409	0.550	0.703	0.874	1.076	1.335	1.740
0.020	0.025	0.151	0.279	0.412	0.553	0.706	0.878	1.080	1.341	1.751
0.022	0.027	0.153	0.281	0.415	0.556	0.709	0.881	1.085	1.347	1.763
0.024	0.030	0.156	0.284	0.417	0.559	0.712	0.885	1.089	1.353	1.775
0.026	0.033	0.158	0.287	0.420	0.562	0.716	0.889	1.094	1.360	1.787
0.028	0.035	0.161	0.289	0.423	0.565	0.719	0.893	1.099	1.366	1.800
0.030	0.038	0.163	0.292	0.426	0.568	0.722	0.896	1.103	1.372	1.812
0.032	0.040	0.166	0.295	0.428	0.571	0.725	0.900	1.108	1.379	1.825
0.034	0.043	0.168	0.297	0.431	0.574	0.729	0.904	1.112	1.385	1.839
0.036	0.045	0.171	0.300	0.434	0.577	0.732	0.908	1.117	1.392	1.853
0.038	0.048	0.173	0.302	0.437	0.580	0.735	0.911	1.122	1.399	1.867
0.040	0.050	0.176	0.305	0.439	0.582	0.739	0.915	1.126	1.405	1.881
0.042	0.053	0.179	0.308	0.442	0.585	0.742	0.919	1.131	1.412	1.896
0.044	0.055	0.181	0.310	0.445	0.588	0.745	0.923	1.136	1.419	1.911
0.046	0.058	0.184	0.313	0.448	0.591	0.749	0.927	1.141	1.426	1.927
0.048	0.060	0.186	0.316	0.451	0.594	0.752	0.931	1.146	1.433	1.944
0.050	0.063	0.189	0.318	0.453	0.597	0.755	0.935	1.150	1.440	1.960
0.052	0.065	0.191	0.321	0.456	0.600	0.759	0.938	1.155	1.447	1.978
0.054	0.068	0.194	0.323	0.459	0.603	0.762	0.942	1.160	1.454	1.996
0.056	0.070	0.196	0.326	0.462	0.606	0.765	0.946	1.165	1.461	2.015
0.058	0.073	0.199	0.329	0.464	0.609	0.769	0.950	1.170	1.469	2.034
0.060	0.075	0.202	0.331	0.467	0.612	0.772	0.954	1.175	1.476	2.054
0.062	0.078	0.204	0.334	0.470	0.615	0.775	0.958	1.180	1.484	2.075
0.064	0.080	0.207	0.337	0.473	0.619	0.779	0.962	1.185	1.491	2.097
0.066	0.083	0.209	0.339	0.476	0.622	0.782	0.966	1.190	1.499	2.121
0.068	0.085	0.212	0.342	0.478	0.625	0.786	0.970	1.195	1.507	2.145
0.070	0.088	0.214	0.345	0.481	0.628	0.789	0.974	1.200	1.514	2.171
0.072	0.090	0.217	0.347	0.484	0.631	0.792	0.978	1.206	1.522	2.198
0.074	0.093	0.219	0.350	0.487	0.634	0.796	0.982	1.211	1.530	2.227
0.076	0.095	0.222	0.353	0.490	0.637	0.799	0.986	1.216	1.539	2.258
0.078	0.098	0.225	0.355	0.493	0.640	0.803	0.990	1.221	1.547	2.291
0.080	0.100	0.227	0.358	0.495	0.643	0.806	0.994	1.227	1.555	2.327
0.082	0.103	0.230	0.361	0.498	0.646	0.810	0.999	1.232	1.564	2.366
0.084	0.105	0.232	0.363	0.501	0.649	0.813	1.003	1.237	1.572	2.409
0.086	0.108	0.235	0.366	0.504	0.652	0.817	1.007	1.243	1.581	2.458
0.088	0.110	0.237	0.369	0.507	0.655	0.820	1.011	1.248	1.590	2.513
0.090	0.113	0.240	0.371	0.510	0.659	0.824	1.015	1.254	1.599	2.576
0.092	0.115	0.243	0.374	0.513	0.662	0.827	1.019	1.259	1.608	2.652
0.094	0.118	0.245	0.377	0.515	0.665	0.831	1.024	1.265	1.617	2.748
0.096	0.120	0.248	0.379	0.518	0.668	0.834	1.028	1.270	1.626	2.879
0.098	0.123	0.250	0.382	0.521	0.671	0.838	1.032	1.276	1.636	3.091

Table 8: Percentage points of the t-distribution

	$1-\alpha$									
n	0.60	0.70	0.80	0.90	0.95	0.975	0.990	0.995	0.999	0.9995
1	0.325	0.727	1.376	3.078	6.314	12.71	31.82	63.66	318.3	636.6
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925	22.33	31.60
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841	10.21	12.92
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	0.255	0.528	0.849	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660	3.232	3.460
70	0.254	0.527	0.847	1.294	1.667	1.994	2.381	2.648	3.211	3.435
80	0.254	0.526	0.846	1.292	1.664	1.990	2.374	2.639	3.195	3.416
90	0.254	0.526	0.846	1.291	1.662	1.987	2.368	2.632	3.183	3.402
100	0.254	0.526	0.845	1.290	1.660	1.984	2.364	2.626	3.174	3.390
110	0.254	0.526	0.845	1.289	1.659	1.982	2.361	2.621	3.166	3.381
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Table 9: Expressions for the Beta function $B(m,n)$ for integer and half-integer arguments

$n \rightarrow$ / $m \downarrow$	$\frac{1}{2}$	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3	$\frac{7}{2}$	4	$\frac{9}{2}$	5
$\frac{1}{2}$	π									
1	2	1								
$\frac{3}{2}$	$\frac{1}{2}\pi$	$\frac{2}{3}$	$\frac{1}{8}\pi$							
2	$\frac{4}{3}$	$\frac{1}{2}$	$\frac{4}{15}$	$\frac{1}{6}$						
$\frac{5}{2}$	$\frac{3}{8}\pi$	$\frac{2}{5}$	$\frac{1}{16}\pi$	$\frac{4}{35}$	$\frac{3}{128}\pi$					
3	$\frac{16}{15}$	$\frac{1}{3}$	$\frac{16}{105}$	$\frac{1}{12}$	$\frac{16}{315}$	$\frac{1}{30}$				
$\frac{7}{2}$	$\frac{5}{16}\pi$	$\frac{2}{7}$	$\frac{5}{128}\pi$	$\frac{4}{63}$	$\frac{3}{256}\pi$	$\frac{16}{693}$	$\frac{5}{1024}\pi$			
4	$\frac{32}{35}$	$\frac{1}{4}$	$\frac{32}{315}$	$\frac{1}{20}$	$\frac{32}{1155}$	$\frac{1}{60}$	$\frac{32}{3003}$	$\frac{1}{140}$		
$\frac{9}{2}$	$\frac{35}{128}\pi$	$\frac{2}{9}$	$\frac{7}{256}\pi$	$\frac{4}{99}$	$\frac{7}{1024}\pi$	$\frac{16}{1287}$	$\frac{5}{2048}\pi$	$\frac{32}{6435}$	$\frac{35}{32768}\pi$	
5	$\frac{256}{315}$	$\frac{1}{5}$	$\frac{256}{3465}$	$\frac{1}{30}$	$\frac{256}{15015}$	$\frac{1}{105}$	$\frac{256}{45045}$	$\frac{1}{280}$	$\frac{256}{109395}$	$\frac{1}{630}$
$\frac{11}{2}$	$\frac{63}{256}\pi$	$\frac{2}{11}$	$\frac{21}{1024}\pi$	$\frac{4}{143}$	$\frac{9}{2048}\pi$	$\frac{16}{2145}$	$\frac{45}{32768}\pi$	$\frac{32}{12155}$	$\frac{35}{65536}\pi$	$\frac{256}{230945}$
6	$\frac{512}{693}$	$\frac{1}{6}$	$\frac{512}{9009}$	$\frac{1}{42}$	$\frac{512}{45045}$	$\frac{1}{168}$	$\frac{512}{153153}$	$\frac{1}{504}$	$\frac{512}{415701}$	$\frac{1}{1260}$
$\frac{13}{2}$	$\frac{231}{1024}\pi$	$\frac{2}{13}$	$\frac{33}{2048}\pi$	$\frac{4}{195}$	$\frac{99}{32768}\pi$	$\frac{16}{3315}$	$\frac{55}{65536}\pi$	$\frac{32}{20995}$	$\frac{77}{262144}\pi$	$\frac{256}{440895}$
7	$\frac{2048}{3003}$	$\frac{1}{7}$	$\frac{2048}{45045}$	$\frac{1}{56}$	$\frac{2048}{255255}$	$\frac{1}{252}$	$\frac{2048}{969969}$	$\frac{1}{840}$	$\frac{2048}{2909907}$	$\frac{1}{2310}$
$\frac{15}{2}$	$\frac{429}{2048}\pi$	$\frac{2}{15}$	$\frac{429}{32768}\pi$	$\frac{4}{255}$	$\frac{143}{65536}$	$\frac{16}{4845}$	$\frac{143}{262144}\pi$	$\frac{32}{33915}$	$\frac{91}{524288}\pi$	$\frac{256}{780045}$

$n \rightarrow$ / $m \downarrow$	$\frac{11}{2}$	6	$\frac{13}{2}$	7	$\frac{15}{2}$
$\frac{11}{2}$	$\frac{63}{262144}\pi$				
6	$\frac{512}{969969}$	$\frac{1}{2772}$			
$\frac{13}{2}$	$\frac{63}{524288}\pi$	$\frac{512}{2028117}$	$\frac{231}{4194304}\pi$		
7	$\frac{2048}{7436429}$	$\frac{1}{5544}$	$\frac{2048}{16900975}$	$\frac{1}{12012}$	
$\frac{15}{2}$	$\frac{273}{4194304}\pi$	$\frac{512}{3900225}$	$\frac{231}{8388608}\pi$	$\frac{2048}{35102025}$	$\frac{429}{33554432}\pi$

Mathematical Constants

Introduction

It is handy to have available the values of different mathematical constants appearing in many expressions in statistical calculations. In this section we list, with high precision, many of those which may be needed. In some cases we give, after the tables, basic expressions which may be nice to recall. Note, however, that this is not full explanations and consult the main text or other sources for details.

Some Basic Constants

exact	approx.
π	3.14159 26535 89793 23846
e	2.71828 18284 59045 23536
γ	0.57721 56649 01532 86061
$\sqrt{\pi}$	1.77245 38509 05516 02730
$1/\sqrt{2\pi}$	0.39894 22804 01432 67794

$$
\begin{aligned}
e &= \sum_{n=0}^{\infty} \frac{1}{n!} = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \\
\gamma &= \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \ln n\right)
\end{aligned}
$$

Gamma Function

	exact	approx.
$\Gamma(\frac{1}{2})$	$\sqrt{\pi}$	1.77245 38509 05516 02730
$\Gamma(\frac{3}{2})$	$\frac{1}{2}\sqrt{\pi}$	0.88622 69254 52758 01365
$\Gamma(\frac{5}{2})$	$\frac{3}{4}\sqrt{\pi}$	1.32934 03881 79137 02047
$\Gamma(\frac{7}{2})$	$\frac{15}{8}\sqrt{\pi}$	3.32335 09704 47842 55118
$\Gamma(\frac{9}{2})$	$\frac{105}{16}\sqrt{\pi}$	11.63172 83965 67448 92914

$$
\begin{aligned}
\Gamma(z) &= \int_0^{\infty} t^{z-1}e^{-t}dt \\
n! &= \Gamma(n+1) = n\Gamma(n) \\
\Gamma\left(n+\tfrac{1}{2}\right) &= \frac{(2n-1)!!}{2^n}\Gamma\left(\tfrac{1}{2}\right) = \\
&= \frac{(2n)!}{2^{2n}n!}\sqrt{\pi}
\end{aligned}
$$

See further section 42.2 on page 156 and reference [40] for more details.

Beta Function

For exact expressions for the Beta function for half-integer and integer values see table 9 on page 181.

$$
\begin{aligned}
B(a,b) &= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \\
&= \int_0^1 x^{a-1}(1-x)^{b-1}dx = \\
&= \int_0^{\infty} \frac{x^{a-1}}{(1+x)^{a+b}}dx
\end{aligned}
$$

See further section 42.6 on page 163.

Digamma Function

	exact	approx.
$\psi(\frac{1}{2})$	$-\gamma - 2\ln 2$	−1.96351 00260 21423 47944
$\psi(\frac{3}{2})$	$\psi(\frac{1}{2}) + 2$	0.03648 99739 78576 52056
$\psi(\frac{5}{2})$	$\psi(\frac{1}{2}) + \frac{8}{3}$	0.70315 66406 45243 18723
$\psi(\frac{7}{2})$	$\psi(\frac{1}{2}) + \frac{46}{15}$	1.10315 66406 45243 18723
$\psi(\frac{9}{2})$	$\psi(\frac{1}{2}) + \frac{352}{105}$	1.38887 09263 59528 90151
$\psi(\frac{11}{2})$	$\psi(\frac{9}{2}) + \frac{2}{9}$	1.61109 31485 81751 12373
$\psi(\frac{13}{2})$	$\psi(\frac{11}{2}) + \frac{2}{11}$	1.79291 13303 99932 94192
$\psi(\frac{15}{2})$	$\psi(\frac{13}{2}) + \frac{2}{13}$	1.94675 74842 46086 78807
$\psi(\frac{17}{2})$	$\psi(\frac{15}{2}) + \frac{2}{15}$	2.08009 08175 94201 21402
$\psi(\frac{19}{2})$	$\psi(\frac{17}{2}) + \frac{2}{17}$	2.19773 78764 02949 53317
$\psi(1)$	$-\gamma$	−0.57721 56649 01532 86061
$\psi(2)$	$1-\gamma$	0.42278 43350 98467 13939
$\psi(3)$	$\frac{3}{2}-\gamma$	0.92278 43350 98467 13939
$\psi(4)$	$\frac{11}{6}-\gamma$	1.25611 76684 31800 47273
$\psi(5)$	$\frac{25}{12}-\gamma$	1.50611 76684 31800 47273
$\psi(6)$	$\frac{137}{60}-\gamma$	1.70611 76684 31800 47273
$\psi(7)$	$\frac{49}{20}-\gamma$	1.87278 43350 98467 13939
$\psi(8)$	$\frac{363}{140}-\gamma$	2.01564 14779 55609 99654
$\psi(9)$	$\frac{761}{280}-\gamma$	2.14064 14779 55609 99654
$\psi(10)$	$\frac{7129}{2520}-\gamma$	2.25175 25890 66721 10765

$$
\begin{aligned}
\psi(z) &= \frac{d}{dz}\ln\Gamma(z) = \frac{1}{\Gamma(z)}\frac{d\Gamma(z)}{dz} \\
\psi(z+1) &= \psi(z) + \tfrac{1}{z} \\
\psi(n) &= -\gamma + \sum_{m=1}^{n-1}\frac{1}{m} \\
\psi(n+\tfrac{1}{2}) &= -\gamma - 2\ln 2 + 2\sum_{m=1}^{n}\frac{1}{2m-1}
\end{aligned}
$$

See further section 42.3 on page 159.

Polygamma Function

	exact	approx.
$\psi^{(1)}(\frac{1}{2})$	$\pi^2/2$	4.93480 22005 44679 30942
$\psi^{(2)}(\frac{1}{2})$	$-14\zeta_3$	–16.82879 66442 34319 99560
$\psi^{(3)}(\frac{1}{2})$	π^4	97.40909 10340 02437 23644
$\psi^{(4)}(\frac{1}{2})$	$-744\zeta_5$	–771.47424 98266 67225 19054
$\psi^{(5)}(\frac{1}{2})$	$8\pi^6$	7691.11354 86024 35496 24176
$\psi^{(1)}(1)$	ζ_2	1.64493 40668 48226 43647
$\psi^{(2)}(1)$	$-2\zeta_3$	–2.40411 38063 19188 57080
$\psi^{(3)}(1)$	$6\zeta_4$	6.49393 94022 66829 14910
$\psi^{(4)}(1)$	$-24\zeta_5$	–24.88626 61234 40878 23195
$\psi^{(5)}(1)$	$120\zeta_6$	122.08116 74381 33896 76574

$$\begin{aligned}
\psi^{(n)}(z) &= \frac{d^n}{dz^n}\psi(z) = \frac{d^{n+1}}{dz^{n+1}}\ln\Gamma(z) \\
\psi^{(n)}(z) &= (-1)^{n+1}n!\sum_{k=0}^{\infty}\frac{1}{(z+k)^{n+1}} \\
\psi^{(n)}(1) &= (-1)^{n+1}n!\zeta_{n+1} \\
\psi^{(n)}(\tfrac{1}{2}) &= (2^{n+1}-1)\psi^{(n)}(1) \\
\psi^{(m)}(n+1) &= (-1)^m m!\Big[-\zeta_{m+1}+1+ \\
&\quad +\frac{1}{2^{m+1}}+\ldots+\frac{1}{n^{m+1}}\Big] = \\
&= \psi^{(m)}(n)+(-1)^m m!\frac{1}{n^{m+1}}
\end{aligned}$$

See further section 42.4 on page 160.

Bernoulli Numbers

See table 4 on page 176.

Riemann's Zeta-function

	exact	approx.
ζ_0	-1	
ζ_1	∞	
ζ_2	$\pi^2/6$	1.64493 40668 48226 43647
ζ_3		1.20205 69031 59594 28540
ζ_4	$\pi^4/90$	1.08232 32337 11138 19152
ζ_5		1.03692 77551 43369 92633
ζ_6	$\pi^6/945$	1.01734 30619 84449 13971
ζ_7		1.00834 92773 81922 82684
ζ_8	$\pi^8/9450$	1.00407 73561 97944 33938
ζ_9		1.00200 83928 26082 21442
ζ_{10}	$\pi^{10}/93555$	1.00099 45751 27818 08534

$$\begin{aligned}
\zeta_n &= \sum_{k=1}^{\infty}\frac{1}{k^n} \\
\zeta_{2n} &= \frac{2^{2n-1}\pi^{2n}|B_{2n}|}{(2n)!}
\end{aligned}$$

See also page 60 and for details reference [27].

ERRATA et ADDENDA

Errors in this report are corrected as they are found but for those who have printed an early version of the hand-book we list here errata. These are thus already obsolete in this copy. Minor errors in language etc are not listed. Note, however, that a few additions (subsections and tables) have been made to the original report (see below).

- Contents part now having roman page numbers thus shifting arabic page numbers for the main text.
- A new section 6.2 on conditional probability density for binormal distribution has been added after the first edition
- Section 42.6, formula, line 2, ν changed into λ giving

$$f(x;\mu,\lambda) = \frac{\lambda}{2} e^{-\lambda|x-\mu|}$$

- Section 10.3, formula 2, line 4 has been corrected

$$\phi_x(t) = E(e^{\imath tx}) = e^{\imath t\mu} E(e^{\imath t(x-\mu)}) = e^{\imath t\mu} \phi_{x-\mu}(t) = e^{\imath t\mu} \frac{\lambda^2}{\lambda^2 + t^2}$$

- Section 14.4, formula, line 2 changed to

$$\phi(t) = E(e^{\imath tx}) = \frac{1}{\alpha} \int\limits_0^\infty e^{(\imath t - \frac{1}{\alpha})x} dx = \frac{1}{1 - \imath t\alpha}$$

- Section 18.1, figure 14 was erroneous in early editions and should look as is now shown in figure 74.
- Section 27.2, line 12: change νr_i to νp_i.
- Section 27.6 on significance levels for the multinomial distribution has been added after the first edition.
- Section 27.7 on the case with equal group probabilities for a multinomial distribution has been added after the first edition.
- A small paragraph added to section 28.1 introducing the multinormal distribution.
- A new section 28.2 on conditional probability density for the multinormal distribution has been added after the first edition.
- Section 36.4, first formula, line 5, should read:

$$P(r) = \sum_{k=0}^{r} \frac{\mu^k e^{-\mu}}{k!} = 1 - P(r+1, \mu)$$

- Section 36.4, second formula, line 9, should read:

$$P(r) = \sum_{k=0}^{r} \frac{\mu^k e^{-\mu}}{k!} = 1 - \int_0^{2\mu} f(x; \nu = 2r + 2) dx$$

- and in the next line it should read $f(x; \nu = 2r + 2)$.
- Section 42.5.2, formula 3, line 6, should read:

$$\Gamma(z) = \alpha^z \int_0^{\infty} t^{z-1} e^{-\alpha t} dt \quad \text{for} \quad Re(z) > 0, \quad Re(\alpha) > 0$$

- Section 42.6, line 6: a reference to table 9 has been added (cf below).
- Table 9 on page 181, on the Beta function $B(m, n)$ for integer and half-integer arguments, has been added after the first version of the paper.

These were, mostly minor, changes up to the 18th of March 1998 in order of apperance. In October 1998 the first somewhat larger revision was made:

- Some text concerning the coefficient of kurtosis added in section 2.2.
- Figure 6 for the chi-square distribution corrected for a normalization error for the $n = 10$ curve.
- Added figure 8 for the chi distribution on page 45.
- Added section 11 for the doubly non-central F-distribution and section 12 for the doubly non-central t-distribution.
- Added figure figure 12 for the F-distribution on page 62.
- Added section 30 on the non-central Beta-distribution on page 109.
- For the non-central chi-square distribution we have added figure 22 and subsections 31.4 and 31.6 for the cumulative distribution and random number generation, respectively.
- For the non-central F-distribution figure 23 has been added on page 114. Errors in the formulæ for $f(F'; m, n, \lambda)$ in the introduction and z_1 in the section on approximations have been corrected. Subsections 32.2 on moments, 32.3 for the cumulative distribution, and 32.5 for random number generation have been added.
- For the non-central t-distribution figure 24 has been added on page 117, some text altered in the first subsection, and an error corrected in the denominator of the approximation formula in subsection 33.5. Subsections 33.2 on the derivation of the distribution, 33.3 on its moments, 33.4 on the cumulative distribution, and 33.6 on random number generation have been added.

- A new subsection 34.8.9 has been added on yet another method, using a ratio between two uniform deviates, to achieve standard normal random numbers. With this change three new references [36–38] were introduced.

- A comparison of the efficiency for different algorithms to obtain standard normal random numbers have been introduced as subsection 34.8.10.

- Added a comment on factorial moments and cumulants for a Poisson distribution in section 36.2.

- This list of "Errata et Addenda" for past versions of the hand-book has been added on page 184 and onwards.

- Table 2 on page 174 and table 3 on page 175 for extreme significance levels of the chi-square distribution have been added thus shifting the numbers of several other tables. This also slightly affected the text in section 8.10.

- The Bernoulli numbers used in section 15.4 now follow the same convention used *e.g.* in section 42.3. This change also affected the formula for κ_{2n} in section 23.4. Table 4 on 176 on Bernoulli numbers was introduced at the same time shifting the numbers of several other tables.

- A list of some mathematical constants which are useful in statistical calculations have been introduced on page 182.

References

[1] *The Advanced Theory of Statistics* by M. G. Kendall and A. Stuart, Vol. 1, Charles Griffin & Company Limited, London 1958.

[2] *The Advanced Theory of Statistics* by M. G. Kendall and A. Stuart, Vol. 2, Charles Griffin & Company Limited, London 1961.

[3] *An Introduction to Mathematical Statistics and Its Applications* by Richard J. Larsen and Morris L. Marx, Prentice-Hall International, Inc. (1986).

[4] *Statistical Methods in Experimental Physics* by W. T. Eadie, D. Drijard, F. E. James, M. Roos and B. Sadoulet, North-Holland Publishing Company, Amsterdam-London (1971).

[5] *Probability and Statistics in Particle Physics* by A. G. Frodesen, O. Skjeggestad and H. Tøfte, Universitetsforlaget, Bergen-Oslo-Tromsø (1979).

[6] *Statistics for Nuclear and Particle Physics* by Louis Lyons, Cambridge University Press (1986).

[7] *Statistics – A Guide to the Use of Statistical Methods in the Physical Sciences* by Roger J. Barlow, John Wiley & Sons Ltd., 1989.

[8] *Numerical Recipes (The Art of Scientific Computing)* by William H. Press, Brian P. Flannery, Saul A. Teukolsky and William T. Vetterling, Cambridge University Press, 1986.

[9] *Numerical Recipes in Fortran (The Art of Scientific Computing), second edition* by William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery, Cambridge University Press, 1992.

[10] *Handbook of Mathematical Functions – with Formulas, Graphs, and Mathematical Tables*, edited by Milton Abramowitz and Irene A. Stegun, Dover Publications, Inc., New York, 1965.

[11] *Statistical Distributions* by N. A. J. Hastings and J. B. Peacock, Butterworth & Co (Publishers) Ltd, 1975.

[12] *A Monte Carlo Sampler* by C. J. Everett and E. D. Cashwell, LA-5061-MS Informal Report, October 1972, Los Alamos Scientific Laboratory of the University of California, New Mexico.

[13] *Random Number Generation* by Christian Walck, USIP Report 87-15, Stockholm University, December 1987.

[14] *A Review of Pseudorandom Number Generators* by F. James, Computer Physics Communications **60** (1990) 329–344.

[15] *Toward a Universal Random Number Generator* by George Marsaglia, Arif Zaman and Wai Wan Tsang, Statistics & Probability Letters **9** (1990) 35–39.

[16] *Implementation of a New Uniform Random Number Generator (including benchmark tests)* by Christian Walck, Internal Note SUF–PFY/89–01, Particle Physics Group, Fysikum, Stockholm University, 21 December 1989.

[17] *A Random Number Generator for PC's* by George Marsaglia, B. Narasimhan and Arif Zaman, Computer Physics Communications **60** (1990) 345–349.

[18] *A Portable High-Quality Random Number Generator for Lattice Field Theory Simulations* by Martin Lüscher, Computer Physics Communications **79** (1994) 100–110.

[19] *Implementation of Yet Another Uniform Random Number Generator* by Christian Walck, Internal Note SUF–PFY/94–01, Particle Physics Group, Fysikum, Stockholm University, 17 February 1994.

[20] *Random number generation* by Birger Jansson, Victor Pettersons Bokindustri AB, Stockholm, 1966.

[21] J. W. Butler, Symp. on Monte Carlo Methods, Whiley, New York (1956) 249–264.

[22] J. C. Butcher, Comp. J. **3** (1961) 251–253.

[23] *Ars Conjectandi* by Jacques Bernoulli, published posthumously in 1713.

[24] G. E. P. Box and M. E. Muller in Annals of Math. Stat. **29** (1958) 610–611.

[25] *Probability Functions* by M. Zelen and N. C. Severo in *Handbook of Mathematical Functions*, ed. M. Abramowitz and I. A. Stegun, Dover Publications, Inc., New York, 1965, 925.

[26] *Error Function and Fresnel Integrals* by Walter Gautschi in *Handbook of Mathematical Functions*, ed. M. Abramowitz and I. A. Stegun, Dover Publications, Inc., New York, 1965, 295.

[27] *Bernoulli and Euler Polynomials — Riemann Zeta Function* by Emilie V. Haynsworth and Karl Goldberg in *Handbook of Mathematical Functions*, ed. M. Abramowitz and I. A. Stegun, Dover Publications, Inc., New York, 1965, 803.

[28] *Computer Methods for Sampling From the Exponential and Normal Distributions* by J. H. Ahrens and U. Dieter, Communications of the ACM **15** (1972) 873.

[29] J. von Neumann, Nat. Bureau Standards, AMS **12** (1951) 36.

[30] G. Marsaglia, Ann. Math. Stat. **32** (1961) 899.

[31] *Multiplicity Distributions in Strong Interactions: A Generalized Negative Binomial Model* by S. Hegyi, Phys. Lett. **B387** (1996) 642.

[32] *Probability, Random Variables and Stochastic Processes* by Athanasios Papoulis, McGraw-Hill book company (1965).

[33] *Theory of Ionization Fluctuation* by J. E. Moyal, Phil. Mag. **46** (1955) 263.

[34] *A New Empirical Regularity for Multiplicity Distributions in Place of KNO Scaling* by the UA5 Collaboration: G. J. Alner et al., Phys. Lett. **B160** (1985) 199.

[35] G. Marsaglia, M. D. MacLaren and T. A. Bray. Comm. ACM **7** (1964) 4–10.

[36] *Computer Generation of Random Variables Using the Ratio of Uniform Deviates* by A. J. Kinderman and John F. Monahan, ACM Transactions on Mathematical Software **3** (1977) 257–260.

[37] *A Fast Normal Random Number Generator* by Joseph L. Leva, ACM Transactions on Mathematical Software **18** (1992) 449–453.

[38] *Algorithm 712: A Normal Random Number Generator* by Joseph L. Leva, ACM Transactions on Mathematical Software **18** (1992) 454–455.

[39] *Mathematical Methods for Physicists* by George Arfken, Academic Press, 1970.

[40] *Gamma Function and Related Functions* by Philip J. Davis in *Handbook of Mathematical Functions*, ed. M. Abramowitz and I. A. Stegun, Dover Publications, Inc., New York, 1965, 253.

[41] *Tables of Integrals, Series, and Products* by I. S. Gradshteyn and I. M. Ryzhik, Fourth Edition, Academic Press, New York and London, 1965.

[42] *The Special Functions and their Approximations* by Yudell L. Luke, Volume 1, Academic Press, New York and London, 1969.

[43] *Programs for Calculating the Logarithm of the Gamma Function, and the Digamma Function, for Complex Argument* by K. S. Kölbig, Computer Physics Communications **4** (1972) 221–226.

[44] *Hypergeometric Functions* by Fritz Oberhettinger in *Handbook of Mathematical Functions*, ed. M. Abramowitz and I. A. Stegun, Dover Publications, Inc., New York, 1965, 555.

[45] *Confluent Hypergeometric Functions* by Lucy Joan Slater in *Handbook of Mathematical Functions*, ed. M. Abramowitz and I. A. Stegun, Dover Publications, Inc., New York, 1965, 503.

Index