WIKIPEDIA

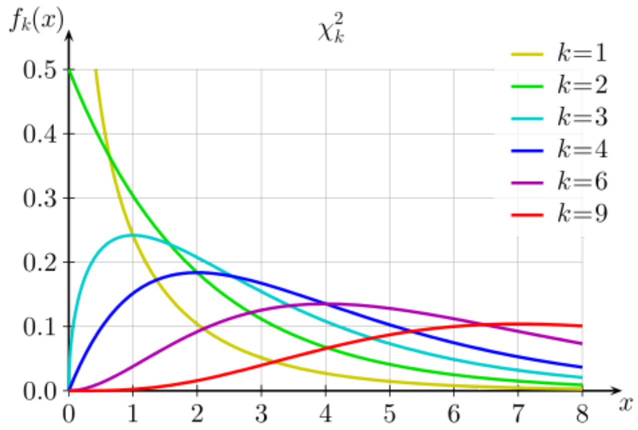
Chi-square distribution

In probability theory and statistics, the **chi-square distribution** (also **chi-squared** or **χ^2-distribution**) with k degrees of freedom is the distribution of a sum of the squares of k independent standard normal random variables. The chi-square distribution is a special case of the gamma distribution and is one of the most widely used probability distributions in inferential statistics, notably in hypothesis testing and in construction of confidence intervals.[2][3][4][5] This distribution is sometimes called the **central chi-square distribution**, a special case of the more general noncentral chi-square distribution.

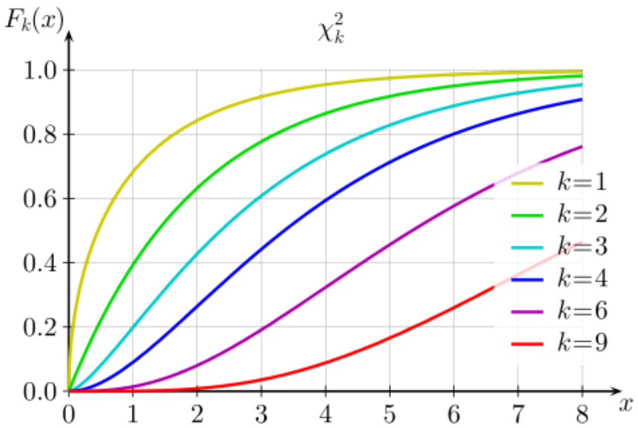
The chi-square distribution is used in the common chi-square tests for goodness of fit of an observed distribution to a theoretical one, the independence of two criteria of classification of qualitative data, and in confidence interval estimation for a population standard deviation of a normal distribution from a sample standard deviation. Many other statistical tests also use this distribution, such as Friedman's analysis of variance by ranks.

chi-square

Probability density function



Cumulative distribution function



Notation	$\chi^2(k)$ or χ_k^2
Parameters	$k \in \mathbb{N}^*$ (known as "degrees of freedom")
Support	$x \in (0, +\infty)$ if $k = 1$, otherwise $x \in [0, +\infty)$
PDF	$\frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$
CDF	$\frac{1}{\Gamma(k/2)} \gamma\left(\frac{k}{2}, \frac{x}{2}\right)$
Mean	k
Median	$\approx k\left(1 - \frac{2}{9k}\right)^3$
Mode	$\max(k-2, 0)$
Variance	$2k$
Skewness	$\sqrt{8/k}$
Ex.	

Contents

kurtosis	$\frac{12}{k}$
Entropy	$\frac{k}{2}+\ln(2\Gamma(\frac{k}{2})) + (1-\frac{k}{2})\psi(\frac{k}{2})$
MGF	$(1-2t)^{-k/2}$ for $t<\frac{1}{2}$
CF	$(1-2it)^{-k/2}$ [1]
PGF	$(1-2\ln t)^{-k/2}$ for $0<t<\sqrt{e}$

Definitions

If $Z_1, ..., Z_k$ are independent, standard normal random variables, then the sum of their squares,

$$Q\ = \sum_{i=1}^k Z_i^2,$$

is distributed according to the chi-square distribution with k degrees of freedom. This is usually denoted as

$$Q\ \sim\ \chi^2(k)\ \text{ or }\ Q\ \sim\ \chi_k^2.$$

The chi-square distribution has one parameter: a positive integer k that specifies the number of degrees of freedom (the number of Z_i s).

Introduction

The chi-square distribution is used primarily in hypothesis testing, and to a lesser extent for confidence intervals for population variance when the underlying distribution is normal. Unlike more widely known distributions such as the normal distribution and the exponential distribution, the chi-square distribution is not as often applied in the direct modeling of natural phenomena. It arises in the following hypothesis tests, among others:

- Chi-square test of independence in contingency tables
- Chi-square test of goodness of fit of observed data to hypothetical distributions
- Likelihood-ratio test for nested models
- Log-rank test in survival analysis
- Cochran–Mantel–Haenszel test for stratified contingency tables

It is also a component of the definition of the t-distribution and the F-distribution used in t-tests, analysis of variance, and regression analysis.

The primary reason that the chi-square distribution is used extensively in hypothesis testing is its relationship to the normal distribution. Many hypothesis tests use a test statistic, such as the t-statistic in a t-test. For these hypothesis tests, as the sample size, n, increases, the sampling distribution of the test statistic approaches the normal distribution (central limit theorem). Because the test statistic (such as t) is asymptotically normally distributed, provided the sample size is sufficiently large, the distribution used for hypothesis testing may be approximated by a normal distribution. Testing hypotheses using a normal distribution is well understood and relatively easy. The simplest chi-square distribution is the square of a standard normal distribution. So wherever a normal distribution could be used for a hypothesis test, a chi-square distribution could be used.

Suppose that Z is a random variable sampled from the standard normal distribution, where the mean equals to 0 and the variance equals to 1: $Z \sim N(0,1)$. Now, consider the random variable $Q = Z^2$. The distribution of the random variable Q is an example of a chi-square distribution: $Q\ \sim\ \chi_1^2$. The subscript 1 indicates that this particular chi-square distribution is constructed from only 1 standard normal distribution. A chi-square distribution constructed by squaring a single standard normal distribution is said to have 1 degree of freedom. Thus, as the sample size for a hypothesis test increases, the distribution of the test statistic approaches a normal distribution. Just as extreme values of the normal distribution have low probability (and give small p-values), extreme values of the chi-square distribution have low probability.

An additional reason that the chi-square distribution is widely used is that it turns up as the large sample distribution of generalized likelihood ratio tests (LRT).[6] LRT's have several desirable properties; in particular, simple LRT's commonly provide the highest power to reject the null hypothesis (Neyman–Pearson lemma) and this leads also to optimality properties of generalised LRTs. However, the normal and chi-square approximations are only valid asymptotically. For this reason, it is preferable to use the t distribution rather than the normal approximation or the chi-square approximation for a small sample size. Similarly, in analyses of contingency tables, the chi-square approximation will be poor for a small sample size, and it is preferable to use Fisher's exact test. Ramsey shows that the exact binomial test is always more powerful than the normal approximation.[7]

Lancaster shows the connections among the binomial, normal, and chi-square distributions, as follows.[8] De Moivre and Laplace established that a binomial distribution could be approximated by a normal distribution. Specifically they showed the asymptotic normality of the random variable

$$\chi = \frac{m - Np}{\sqrt{Npq}}$$

where m is the observed number of successes in N trials, where the probability of success is p, and $q = 1 - p$.

Squaring both sides of the equation gives

$$\chi^2 = \frac{(m - Np)^2}{Npq}$$

Using $N = Np + N(1-p)$, $N = m + (N - m)$, and $q = 1 - p$, this equation simplifies to

$$\chi^2 = \frac{(m - Np)^2}{Np} + \frac{(N - m - Nq)^2}{Nq}$$

The expression on the right is of the form that Karl Pearson would generalize to the form:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

where

χ^2 = Pearson's cumulative test statistic, which asymptotically approaches a χ^2 distribution.
O_i = the number of observations of type i.
$E_i = Np_i$ = the expected (theoretical) frequency of type i, asserted by the null hypothesis that the fraction of type i in the population is p_i
n = the number of cells in the table.

In the case of a binomial outcome (flipping a coin), the binomial distribution may be approximated by a normal distribution (for sufficiently large n). Because the square of a standard normal distribution is the chi-square distribution with one degree of freedom, the probability of a result such as 1 heads in 10 trials can be approximated either by using the normal distribution directly, or the chi-square distribution for the normalised, squared difference between observed and expected value. However, many problems involve more than the two possible outcomes of a binomial, and instead require 3 or more categories, which leads to the multinomial distribution. Just as de Moivre and Laplace sought for and found the normal approximation to the binomial, Pearson sought for and found a degenerate multivariate normal approximation to the multinomial distribution (the numbers in each category add up to the total sample size, which is considered fixed). Pearson showed that the chi-square distribution arose from such a multivariate normal approximation to the multinomial distribution, taking careful account of the statistical dependence (negative correlations) between numbers of observations in different categories. [8]

Probability density function

The probability density function (pdf) of the chi-square distribution is

$$f(x;\,k) = \begin{cases} \dfrac{x^{\frac{k}{2}-1}e^{-\frac{x}{2}}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}, & x > 0; \\ 0, & \text{otherwise}. \end{cases}$$

where $\Gamma(k/2)$ denotes the gamma function, which has closed-form values for integer k.

For derivations of the pdf in the cases of one, two and k degrees of freedom, see Proofs related to chi-square distribution.

Cumulative distribution function

Its cumulative distribution function is:

$$F(x;\,k) = \frac{\gamma(\frac{k}{2},\,\frac{x}{2})}{\Gamma(\frac{k}{2})} = P\left(\frac{k}{2},\,\frac{x}{2}\right),$$

where $\gamma(s,t)$ is the lower incomplete gamma function and $P(s,t)$ is the regularized gamma function.

In a special case of k = 2 this function has a simple form:

$$F(x;\,2) = 1 - e^{-x/2}$$

and the integer recurrence of the gamma function makes it easy to compute for other small even k.

Tables of the chi-square cumulative distribution function are widely available and the function is included in many spreadsheets and all statistical packages.

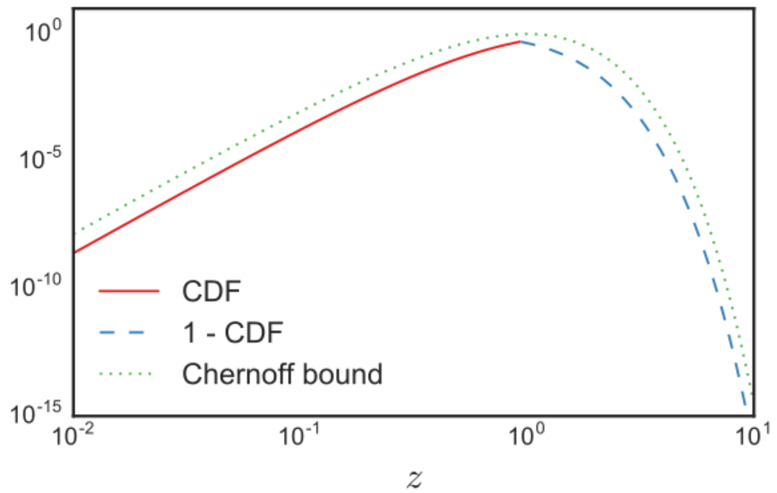
Letting $z \equiv x/k$, Chernoff bounds on the lower and upper tails of the CDF may be obtained.[9] For the cases when $0 < z < 1$ (which include all of the cases when this CDF is less than half):

$$F(zk;\, k) \leq (ze^{1-z})^{k/2}.$$

The tail bound for the cases when $z > 1$, similarly, is

$$1 - F(zk;\, k) \leq (ze^{1-z})^{k/2}.$$

For another approximation for the CDF modeled after the cube of a Gaussian, see under Noncentral chi-square distribution.



Chernoff bound for the CDF and tail (1-CDF) of a chi-square random variable with ten degrees of freedom (k = 10)

Properties

Sum of squares of i.i.d normals minus their mean

If $Z_1, ..., Z_k$ are independent, standard normal random variables, then

$$\sum_{i=1}^{k}(Z_i - \overline{Z})^2 \sim \chi^2_{k-1}$$

where

$$\overline{Z} = \frac{1}{k}\sum_{i=1}^{k} Z_i.$$

Additivity

It follows from the definition of the chi-square distribution that the sum of independent chi-square variables is also chi-square distributed. Specifically, if $X_i, i = \overline{1,n}$ are independent chi-square variables with k_i, $i = \overline{1,n}$ degrees of freedom, respectively, then $Y = X_1 + \ldots + X_n$ is chi-square distributed with $k_1 + \ldots + k_n$ degrees of freedom.

Sample mean

The sample mean of n i.i.d. chi-square variables of degree k is distributed according to a gamma distribution with shape α and scale θ parameters:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim \operatorname{Gamma}(\alpha = n\,k/2, \theta = 2/n) \qquad \text{where } X_i \sim \chi^2(k)$$

Asymptotically, given that for a scale parameter α going to infinity, a Gamma distribution converges towards a normal distribution with expectation $\mu = \alpha \cdot \theta$ and variance $\sigma^2 = \alpha\,\theta^2$, the sample mean converges towards:

$$\overline{X} \xrightarrow{n\to\infty} N(\mu = k, \sigma^2 = 2\,k/n)$$

Note that we would have obtained the same result invoking instead the central limit theorem, noting that for each chi-square variable of degree k the expectation is k , and its variance $2\,k$ (and hence the variance of the sample mean $\overline{X}$ being $\sigma^2 = \frac{2k}{n}$).

Entropy

The differential entropy is given by

$$h = \int_0^\infty f(x;\,k)\ln f(x;\,k)\,dx = \frac{k}{2} + \ln\left[2\,\Gamma\left(\frac{k}{2}\right)\right] + \left(1 - \frac{k}{2}\right)\,\psi\left[\frac{k}{2}\right],$$

where $\psi(x)$ is the Digamma function.

The chi-square distribution is the maximum entropy probability distribution for a random variate X for which $\operatorname{E}(X) = k$ and $\operatorname{E}(\ln(X)) = \psi(k/2) + \ln(2)$ are fixed. Since the chi-square is in the family of gamma distributions, this can be derived by substituting appropriate values in the Expectation of the log moment of gamma. For derivation from more basic principles, see the derivation in moment-generating function of the sufficient statistic.

Noncentral moments

The moments about zero of a chi-square distribution with k degrees of freedom are given by[10][11]

$$\operatorname{E}(X^m) = k(k+2)(k+4)\cdots(k+2m-2) = 2^m \frac{\Gamma\left(m + \frac{k}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}.$$

Cumulants

The cumulants are readily obtained by a (formal) power series expansion of the logarithm of the characteristic function:

$$\kappa_n = 2^{n-1}(n-1)!\,k$$

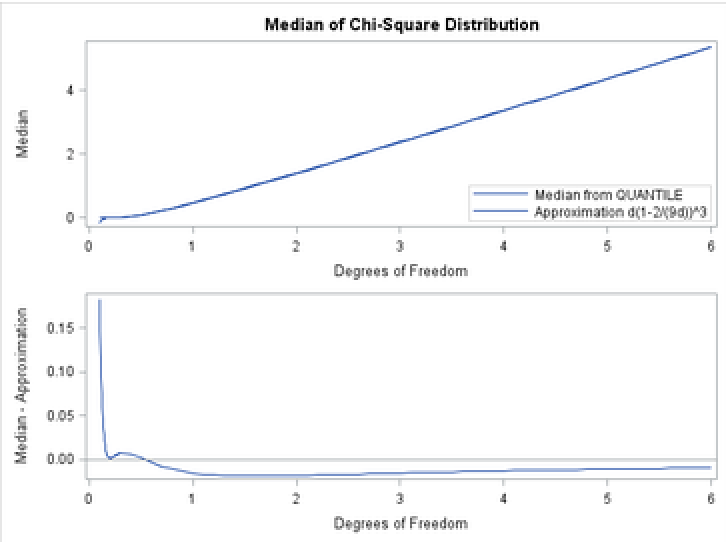
Asymptotic properties

By the central limit theorem, because the chi-square distribution is the sum of k independent random variables with finite mean and variance, it converges to a normal distribution for large k. For many practical purposes, for $k > 50$ the distribution is sufficiently close to a normal distribution for the difference to be ignored.[12] Specifically, if $X \sim \chi^2(k)$, then as k tends to infinity, the distribution of $(X-k)/\sqrt{2k}$ tends to a standard normal distribution. However, convergence is slow as the skewness is $\sqrt{8/k}$ and the excess kurtosis is $12/k$.

The sampling distribution of $\ln(\chi^2)$ converges to normality much faster than the sampling distribution of χ^2,[13] as the logarithm removes much of the asymmetry.[14] Other functions of the chi-square distribution converge more rapidly to a normal distribution. Some examples are:

- If $X \sim \chi^2(k)$ then $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2k-1}$ and unit variance (1922, by R. A. Fisher, see (18.23), p. 426 of Johnson.[4]
- If $X \sim \chi^2(k)$ then $\sqrt[3]{X/k}$ is approximately normally distributed with mean $1-\frac{2}{9k}$ and variance $\frac{2}{9k}$.[15] This is known as the Wilson–Hilferty transformation, see (18.24), p. 426 of Johnson.[4]
  - This normalizing transformation leads directly to the commonly used median approximation $k\left(1-\frac{2}{9k}\right)^3$ by back-transforming from the mean, which is also the median, of the normal distribution.

Related distributions



Approximate formula for median compared with numerical quantile (top). Difference between numerical quantile and approximate formula (bottom).

- As $k \to \infty$, $(\chi^2_k - k)/\sqrt{2k} \xrightarrow{d} N(0,1)$ (normal distribution)
- $\chi^2_k \sim {\chi'}^2_k(0)$ (noncentral chi-square distribution with non-centrality parameter $\lambda = 0$)
- If $Y \sim \mathrm{F}(\nu_1, \nu_2)$ then $X = \lim_{\nu_2 \to \infty} \nu_1 Y$ has the chi-square distribution $\chi^2_{\nu_1}$
  - As a special case, if $Y \sim \mathrm{F}(1, \nu_2)$ then $X = \lim_{\nu_2 \to \infty} Y$ has the chi-square distribution χ^2_1
- $\|\boldsymbol{N}_{i=1,\ldots,k}(0,1)\|^2 \sim \chi^2_k$ (The squared norm of *k* standard normally distributed variables is a chi-square distribution with *k* degrees of freedom)
- If $X \sim \chi^2(\nu)$ and $c > 0$, then $cX \sim \Gamma(k = \nu/2, \theta = 2c)$. (gamma distribution)
- If $X \sim \chi^2_k$ then $\sqrt{X} \sim \chi_k$ (chi distribution)
- If $X \sim \chi^2(2)$, then $X \sim \mathrm{Exp}(1/2)$ is an exponential distribution. (See gamma distribution for more.)
- If $X \sim \chi^2(2k)$, then $X \sim \mathrm{Erlang}(k, 1/2)$ is an Erlang distribution.
- If $X \sim \mathrm{Erlang}(k, \lambda)$, then $2\lambda X \sim \chi^2_{2k}$

- If $X \sim \text{Rayleigh}(1)$ (Rayleigh distribution) then $X^2 \sim \chi^2(2)$
- If $X \sim \text{Maxwell}(1)$ (Maxwell distribution) then $X^2 \sim \chi^2(3)$
- If $X \sim \chi^2(\nu)$ then $\tfrac{1}{X} \sim \text{Inv-}\chi^2(\nu)$ (Inverse-chi-square distribution)
- The chi-square distribution is a special case of type III Pearson distribution
- If $X \sim \chi^2(\nu_1)$ and $Y \sim \chi^2(\nu_2)$ are independent then $\tfrac{X}{X+Y} \sim \text{Beta}(\tfrac{\nu_1}{2}, \tfrac{\nu_2}{2})$ (beta distribution)
- If $X \sim \text{U}(0,1)$ (uniform distribution) then $-2\log(X) \sim \chi^2(2)$
- $\chi^2(6)$ is a transformation of Laplace distribution
- If $X_i \sim \text{Laplace}(\mu, \beta)$ then $\displaystyle\sum_{i=1}^n \frac{2|X_i - \mu|}{\beta} \sim \chi^2(2n)$
- If X_i follows the generalized normal distribution (version 1) with parameters μ, α, β then $\displaystyle\sum_{i=1}^n \frac{2|X_i - \mu|^\beta}{\alpha} \sim \chi^2\left(\frac{2n}{\beta}\right)$ [16]
- chi-square distribution is a transformation of Pareto distribution
- Student's t-distribution is a transformation of chi-square distribution
- Student's t-distribution can be obtained from chi-square distribution and normal distribution
- Noncentral beta distribution can be obtained as a transformation of chi-square distribution and Noncentral chi-square distribution
- Noncentral t-distribution can be obtained from normal distribution and chi-square distribution

A chi-square variable with k degrees of freedom is defined as the sum of the squares of k independent standard normal random variables.

If Y is a k-dimensional Gaussian random vector with mean vector μ and rank k covariance matrix C, then $X = (Y - \mu)^T C^{-1} (Y - \mu)$ is chi-square distributed with k degrees of freedom.

The sum of squares of statistically independent unit-variance Gaussian variables which do *not* have mean zero yields a generalization of the chi-square distribution called the noncentral chi-square distribution.

If Y is a vector of k i.i.d. standard normal random variables and A is a $k \times k$ symmetric, idempotent matrix with rank $k - n$, then the quadratic form $Y^T AY$ is chi-square distributed with $k - n$ degrees of freedom.

If Σ is a $p \times p$ positive-semidefinite covariance matrix with strictly positive diagonal entries, then for $X \sim N(0, \Sigma)$ and w a random p-vector independent of X such that $w_1 + \cdots + w_p = 1$ and $w_i \geq 0, i = 1, \cdots, p$, it holds that

$$\frac{1}{\left(\frac{w_1}{X_1}, \cdots, \frac{w_p}{X_p}\right) \Sigma \left(\frac{w_1}{X_1}, \cdots, \frac{w_p}{X_p}\right)^\top} \sim \chi_1^2.$$ [14]

The chi-square distribution is also naturally related to other distributions arising from the Gaussian. In particular,

- Y is F-distributed, $Y \sim F(k_1, k_2)$ if $Y = \dfrac{X_1/k_1}{X_2/k_2}$, where $X_1 \sim \chi^2(k_1)$ and $X_2 \sim \chi^2(k_2)$ are statistically independent.

- If $X_1 \sim \chi^2(k_1)$ and $X_2 \sim \chi^2(k_2)$ are statistically independent, then $X_1 + X_2 \sim \chi^2(k_1 + k_2)$. If X_1 and X_2 are not independent, then $X_1 + X_2$ is not chi-square distributed.

Generalizations

The chi-square distribution is obtained as the sum of the squares of *k* independent, zero-mean, unit-variance Gaussian random variables. Generalizations of this distribution can be obtained by summing the squares of other types of Gaussian random variables. Several such distributions are described below.

Linear combination

If $X_1, \ldots, X_n$ are chi square random variables and $a_1, \ldots, a_n \in \mathbb{R}_{>0}$, then a closed expression for the distribution of $X = \sum_{i=1}^{n} a_i X_i$ is not known. It may be, however, approximated efficiently using the property of characteristic functions of chi-square random variables.[17]

Chi-square distributions

Noncentral chi-square distribution

The noncentral chi-square distribution is obtained from the sum of the squares of independent Gaussian random variables having unit variance and *nonzero* means.

Generalized chi-square distribution

The generalized chi-square distribution is obtained from the quadratic form *z′Az* where *z* is a zero-mean Gaussian vector having an arbitrary covariance matrix, and *A* is an arbitrary matrix.

Gamma, exponential, and related distributions

The chi-square distribution $X \sim \chi_k^2$ is a special case of the gamma distribution, in that $X \sim \Gamma\left(\frac{k}{2}, \frac{1}{2}\right)$ using the rate parameterization of the gamma distribution (or $X \sim \Gamma\left(\frac{k}{2}, 2\right)$ using the scale parameterization of the gamma distribution) where *k* is an integer.

Because the exponential distribution is also a special case of the gamma distribution, we also have that if $X \sim \chi_2^2$, then $X \sim \operatorname{Exp}\left(\frac{1}{2}\right)$ is an exponential distribution.

The Erlang distribution is also a special case of the gamma distribution and thus we also have that if $X \sim \chi_k^2$ with even k, then X is Erlang distributed with shape parameter $\mathrm{k}/2$ and scale parameter $1/2$.

Occurrence and applications

The chi-square distribution has numerous applications in inferential statistics, for instance in chi-square tests and in estimating variances. It enters the problem of estimating the mean of a normally distributed population and the problem of estimating the slope of a regression line via its role in Student's t-distribution. It enters all analysis of variance problems via its role in the F-distribution, which is the distribution of the ratio of two independent chi-squared random variables, each divided by their respective degrees of freedom.

Following are some of the most common situations in which the chi-square distribution arises from a Gaussian-distributed sample.

- if $X_1, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$ random variables, then $\sum_{i=1}^{n}(X_i - \overline{X})^2 \sim \sigma^2 \chi^2_{n-1}$ where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
- The box below shows some statistics based on $X_i \sim N(\mu_i, \sigma_i^2), i = \overline{1, k}$ independent random variables that have probability distributions related to the chi-square distribution:

Name	Statistic
chi-square distribution	$\sum_{i=1}^{k}\left(\frac{X_i - \mu_i}{\sigma_i}\right)^2$
noncentral chi-square distribution	$\sum_{i=1}^{k}\left(\frac{X_i}{\sigma_i}\right)^2$
chi distribution	$\sqrt{\sum_{i=1}^{k}\left(\frac{X_i - \mu_i}{\sigma_i}\right)^2}$
noncentral chi distribution	$\sqrt{\sum_{i=1}^{k}\left(\frac{X_i}{\sigma_i}\right)^2}$

The chi-square distribution is also often encountered in magnetic resonance imaging.[18]

Computational methods

Table of χ^2 values vs *p*-values

The *p*-value is the probability of observing a test statistic *at least* as extreme in a chi-square distribution. Accordingly, since the cumulative distribution function (CDF) for the appropriate degrees of freedom *(df)* gives the probability of having obtained a value *less extreme* than this point, subtracting the CDF value from 1 gives the *p*-value. A low *p*-value, below the chosen significance level, indicates statistical significance, i.e., sufficient evidence to reject the null hypothesis. A significance level of 0.05 is often used as the cutoff between significant and non-significant results.

The table below gives a number of *p*-values matching to χ^2 for the first 10 degrees of freedom.

Degrees of freedom (df)	χ^2 value[19]										
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.63	10.83
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.61	5.99	9.21	13.82
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.81	11.34	16.27
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28	18.47
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52
6	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	16.81	22.46
7	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	18.48	24.32
8	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	20.09	26.12
9	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	21.67	27.88
10	3.94	4.87	6.18	7.27	9.34	11.78	13.44	15.99	18.31	23.21	29.59
P value (Probability)	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001

These values can be calculated evaluating the quantile function (also known as “inverse CDF” or “ICDF”) of the chi-square distribution;[20] e. g., the χ^2 ICDF for $p = 0.05$ and $\text{df} = 7$ yields $14.06714 \approx 14.07$ as in the table above.

History

This distribution was first described by the German statistician Friedrich Robert Helmert in papers of 1875–6,[21][22] where he computed the sampling distribution of the sample variance of a normal population. Thus in German this was traditionally known as the *Helmert'sche* ("Helmertian") or "Helmert distribution".

The distribution was independently rediscovered by the English mathematician Karl Pearson in the context of goodness of fit, for which he developed his Pearson's chi-square test, published in 1900, with computed table of values published in (Elderton 1902), collected in (Pearson 1914, pp. xxxi–xxxiii, 26–28, Table XII). The name "chi-square" ultimately derives from Pearson's shorthand for the exponent in a multivariate normal distribution with the Greek letter Chi, writing $-½\chi^2$ for what would appear in modern notation as $-½\mathbf{x}^T\Sigma^{-1}\mathbf{x}$ (Σ being the covariance matrix).[23] The idea of a family of "chi-square distributions", however, is not due to Pearson but arose as a further development due to Fisher in the 1920s.[21]

See also

- Chi distribution
- Cochran's theorem
- *F*-distribution
- Fisher's method for combining independent tests of significance
- Gamma distribution
- Generalized chi-square distribution
- Hotelling's *T*-square distribution
- Noncentral chi-square distribution
- Pearson's chi-square test
- Reduced chi-squared statistic
- Student's *t*-distribution
- Wilks's lambda distribution

- Wishart distribution

References

1. M.A. Sanders. "Characteristic function of the central chi-square distribution" (https://web.archive.org/web/20110715091705/http://www.planetmathematics.com/CentralChiDistr.pdf#) (PDF). Archived from the original (http://www.planetmathematics.com/CentralChiDistr.pdf) (PDF) on 2011-07-15. Retrieved 2009-03-06.
2. Abramowitz, Milton; Stegun, Irene Ann, eds. (1983) [June 1964]. "Chapter 26" (http://www.math.sfu.ca/~cbm/aands/page_940.htm). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Applied Mathematics Series. **55** (Ninth reprint with additional corrections of tenth original printing with corrections (December 1972); first ed.). Washington D.C.; New York: United States Department of Commerce, National Bureau of Standards; Dover Publications. p. 940. ISBN 978-0-486-61272-0. LCCN 64-60036 (https://lccn.loc.gov/64-60036). MR 0167642 (https://www.ams.org/mathscinet-getitem?mr=0167642). LCCN 65-12253 (https://lccn.loc.gov/65012253).
3. NIST (2006). Engineering Statistics Handbook – Chi-Squared Distribution (http://www.itl.nist.gov/div898/handbook/eda/section3/eda3666.htm)
4. Johnson, N. L.; Kotz, S.; Balakrishnan, N. (1994). "Chi-Square Distributions including Chi and Rayleigh". *Continuous Univariate Distributions*. **1** (Second ed.). John Wiley and Sons. pp. 415–493. ISBN 978-0-471-58495-7.
5. Mood, Alexander; Graybill, Franklin A.; Boes, Duane C. (1974). *Introduction to the Theory of Statistics* (Third ed.). McGraw-Hill. pp. 241–246. ISBN 978-0-07-042864-5.
6. Westfall, Peter H. (2013). *Understanding Advanced Statistical Methods*. Boca Raton, FL: CRC Press. ISBN 978-1-4665-1210-8.
7. Ramsey, PH (1988). "Evaluating the Normal Approximation to the Binomial Test". *Journal of Educational Statistics*. **13** (2): 173–82. doi:10.2307/1164752 (https://doi.org/10.2307%2F1164752). JSTOR 1164752 (https://www.jstor.org/stable/1164752).
8. Lancaster, H.O. (1969), *The Chi-squared Distribution*, Wiley
9. Dasgupta, Sanjoy D. A.; Gupta, Anupam K. (January 2003). "An Elementary Proof of a Theorem of Johnson and Lindenstrauss" (http://cseweb.ucsd.edu/~dasgupta/papers/jl.pdf) (PDF). *Random Structures and Algorithms*. **22** (1): 60–65. doi:10.1002/rsa.10073 (https://doi.org/10.1002%2Frsa.10073). Retrieved 2012-05-01.
10. Chi-squared distribution (http://mathworld.wolfram.com/Chi-SquaredDistribution.html), from MathWorld, retrieved Feb. 11, 2009
11. M. K. Simon, *Probability Distributions Involving Gaussian Random Variables*, New York: Springer, 2002, eq. (2.35), ISBN 978-0-387-34657-1
12. Box, Hunter and Hunter (1978). *Statistics for experimenters* (https://archive.org/details/statisticsforexp00geor/page/118). Wiley. p. 118 (https://archive.org/details/statisticsforexp00geor/page/118). ISBN 978-0471093152.
13. Bartlett, M. S.; Kendall, D. G. (1946). "The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation". *Supplement to the Journal of the Royal Statistical Society*. **8** (1): 128–138. doi:10.2307/2983618 (https://doi.org/10.2307%2F2983618). JSTOR 2983618 (https://www.jstor.org/stable/2983618).
14. Pillai, Natesh S. (2016). "An unexpected encounter with Cauchy and Lévy". *Annals of Statistics*. **44** (5): 2089–2097. arXiv:1505.01957 (https://arxiv.org/abs/1505.01957). doi:10.1214/15-aos1407 (https://doi.org/10.1214%2F15-aos1407).
15. Wilson, E. B.; Hilferty, M. M. (1931). "The distribution of chi-squared" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1076144). *Proc. Natl. Acad. Sci. USA*. **17** (12): 684–688. Bibcode:1931PNAS...17..684W (https://ui.adsabs.harvard.edu/abs/1931PNAS...17..684W). doi:10.1073/pnas.17.12.684 (https://doi.org/10.1073%2Fpnas.17.12.684). PMC 1076144 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1076144). PMID 16577411 (https://pubmed.ncbi.nlm.nih.gov/16577411).

16. Bäckström, T.; Fischer, J. (January 2018). "Fast Randomization for Distributed Low-Bitrate Coding of Speech and Audio" (https://aaltodoc.aalto.fi/handle/123456789/33466). *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. **26** (1): 19–30. doi:10.1109/TASLP.2017.2757601 (https://doi.org/10.1109%2FTASLP.2017.2757601).
17. Bausch, J. (2013). "On the Efficient Calculation of a Linear Combination of Chi-Square Random Variables with an Application in Counting String Vacua". *J. Phys. A: Math. Theor*. **46** (50): 505202. arXiv:1208.2691 (https://arxiv.org/abs/1208.2691). Bibcode:2013JPhA...46X5202B (https://ui.adsabs.harvard.edu/abs/2013JPhA...46X5202B). doi:10.1088/1751-8113/46/50/505202 (https://doi.org/10.1088%2F1751-8113%2F46%2F50%2F505202).
18. den Dekker A. J., Sijbers J., (2014) "Data distributions in magnetic resonance images: a review", *Physica Medica*, [1] (https://dx.doi.org/10.1016/j.ejmp.2014.05.002)
19. Chi-Squared Test (http://www2.lv.psu.edu/jxm57/irp/chisquar.html) Table B.2. Dr. Jacqueline S. McLaughlin at The Pennsylvania State University. In turn citing: R. A. Fisher and F. Yates, Statistical Tables for Biological Agricultural and Medical Research, 6th ed., Table IV. Two values have been corrected, 7.82 with 7.81 and 4.60 with 4.61
20. R Tutorial: Chi-squared Distribution (http://www.r-tutor.com/elementary-statistics/probability-distributions/chi-squared-distribution)
21. Hald 1998, pp. 633–692, 27. Sampling Distributions under Normality.
22. F. R. Helmert, "Ueber die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einige damit im Zusammenhange stehende Fragen (http://gdz.sub.uni-goettingen.de/dms/load/img/?PPN=PPN599415665_0021&DMDID=DMDLOG_0018)", *Zeitschrift für Mathematik und Physik* 21 (http://gdz.sub.uni-goettingen.de/dms/load/toc/?PPN=PPN599415665_0021), 1876, pp. 102–219
23. R. L. Plackett, *Karl Pearson and the Chi-Squared Test*, International Statistical Review, 1983, 61f. (https://www.jstor.org/stable/1402731?seq=3) See also Jeff Miller, Earliest Known Uses of Some of the Words of Mathematics (http://jeff560.tripod.com/c.html).

Further reading

- Hald, Anders (1998). *A history of mathematical statistics from 1750 to 1930*. New York: Wiley. ISBN 978-0-471-17912-2.
- Elderton, William Palin (1902). "Tables for Testing the Goodness of Fit of Theory to Observation" (https://zenodo.org/record/1431595). *Biometrika*. **1** (2): 155–163. doi:10.1093/biomet/1.2.155 (https://doi.org/10.1093%2Fbiomet%2F1.2.155).
- "Chi-squared distribution" (https://www.encyclopediaofmath.org/index.php?title=Chi-squared_distribution), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]

External links

- Earliest Uses of Some of the Words of Mathematics: entry on Chi squared has a brief history (http://jeff560.tripod.com/c.html)
- Course notes on Chi-Squared Goodness of Fit Testing (http://www.stat.yale.edu/Courses/1997-98/101/chigf.htm) from Yale University Stats 101 class.
- *Mathematica* demonstration showing the chi-squared sampling distribution of various statistics, e.g. Σx^2, for a normal population (http://demonstrations.wolfram.com/StatisticsAssociatedWithNormalSamples/)
- Simple algorithm for approximating cdf and inverse cdf for the chi-squared distribution with a pocket calculator (https://www.jstor.org/stable/2348373)
- Values of the Chi-squared distribution (https://www.medcalc.org/manual/chi-square-table.php)