

Confidence interval

From Wikipedia, the free encyclopedia

In statistics, a **confidence interval** (**CI**) is a particular kind of interval estimate of a population parameter and is used to indicate the reliability of an estimate. It is an interval that frequently includes the parameter of interest, if the experiment is repeated. How frequently the interval contains the parameter is determined by the **confidence level** or confidence coefficient. Confidence intervals are used in frequentist statistics; the equivalent in Bayesian statistics is the *credible interval*.

The confidence level of a confidence interval does *not* indicate the probability that the given confidence interval contains the true value of the parameter. The only such claim that can be made is that the method of generating confidence intervals results in an interval that usually contains the true value; this is due to the fact that the endpoints of the confidence interval (the **confidence limits**) are random variables, while the true value of the parameter is fixed. The meaning of the term "confidence level" is that, if confidence intervals are constructed across many separate data analyses of repeated (and possibly different) experiments, the proportion of such intervals that contain the true value of the parameter will approximately match the confidence level; this is guaranteed by the reasoning underlying the construction of confidence intervals.

Contents

- 1 Conceptual basis
 - 1.1 Introduction
 - 1.2 Relationship with other statistical topics
- 2 Statistical theory
 - 2.1 Definition
 - 2.1.1 Confidence intervals as random intervals
 - 2.1.2 Confidence intervals for inference
 - 2.1.3 Intervals for random outcomes
 - 2.1.4 Approximate confidence intervals
 - 2.1.5 Comparison to Bayesian interval estimates
 - 2.2 Desirable properties
 - 2.3 Methods of derivation
- 3 Examples
 - 3.1 Practical example
 - 3.2 Theoretical example
- 4 Relation to hypothesis testing
- 5 Meaning and interpretation
 - 5.1 Meaning of the term "confidence"
- 6 Alternatives and critiques
 - 6.1 Philosophical issues
- 7 Confidence intervals for proportions and related quantities
- 8 See also
- 9 Online calculators
- 10 Notes

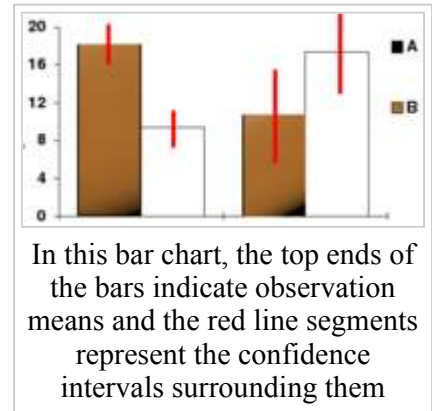
- 11 References
- 12 External links

Conceptual basis

Introduction

Interval estimates can be contrasted with point estimates. A point estimate is a single value given as the estimate of a population parameter that is of interest, for example the mean of some quantity. An interval estimate specifies instead a range within which the parameter is estimated to lie. Confidence intervals are commonly reported in tables or graphs along with point estimates of the same parameters, to show the reliability of the estimates.

For example, a confidence interval can be used to describe how reliable survey results are. In a poll of election voting-intentions, the result might be that 40% of respondents intend to vote for a certain party. A 90% confidence interval for the proportion in the whole population having the same intention on the survey date might be 38% to 42%. From the same data one may calculate a 95% confidence interval, which might in this case be 36% to 44%. A major factor determining the length of a confidence interval is the size of the sample used in the estimation procedure, for example the number of people taking part in a survey.



Relationship with other statistical topics

Confidence intervals are closely related to statistical significance testing. For example, if one wants to test the null hypothesis that some estimated parameter θ , $\theta = 0$ against the alternative that $\theta \neq 0$, then this test can be performed by finding if the confidence interval for θ contains 0.

More generally, given the availability of a hypothesis testing procedure that can test the null hypothesis $\theta = \theta_0$ against the alternative that $\theta \neq \theta_0$ for any value of θ_0 , then a confidence interval with confidence level $\gamma = 1 - \alpha$ can be defined as containing any number θ_0 for which the corresponding null hypothesis is not rejected at significance level α .^[1]

In consequence, if the estimates of two parameters (for example, the mean values of a variable in two independent groups of objects) have confidence intervals at a given γ value that do not overlap, then the difference between the two values is significant at the corresponding value of α . However, this test is too conservative. If two confidence intervals overlap, the difference between the two means still may be significantly different.^{[2][3]}

Confidence regions generalize the confidence interval concept to deal with multiple quantities. Such regions can indicate not only the extent of likely sampling errors but can also reveal whether (for example) it is the case that if the estimate for one quantity is unreliable then the other is also likely to be unreliable. See also confidence bands.

In applied practice, confidence intervals are typically stated at the 95% confidence level.^[4] However, when presented graphically, confidence intervals can be shown at several confidence levels, for example 50%, 95% and 99%.

Statistical theory

Definition

Confidence intervals as random intervals

A *confidence interval* with confidence level γ for a parameter θ of a probability distribution is an interval with random endpoints $(u(X), v(X))$ (or its observed value $(u(x), v(x))$), based on a random sample X (or its observed value x) from that distribution, with the property:

$$\gamma = \Pr_{\theta}(u(X) < \theta < v(X)).$$

The confidence interval is specified by the pair of statistics (i.e., observable random variables) $u(X)$ and $v(X)$.

Confidence intervals for inference

For the above to provide a viable means to statistical inference, something further is required: a tie between the quantity being estimated and the probability distribution of the outcome X . Suppose that this probability distribution is characterised by the unobservable parameter θ , which is a quantity to be estimated, and by other unobservable parameters φ which are not of immediate interest. These other quantities φ in which there is no immediate interest are called nuisance parameters, as statistical theory still needs to find some way to deal with them.

The definition of a confidence interval for θ for any number α between 0 and 1 is an interval

$$(u(X), v(X))$$

for which

$$\Pr_{\theta, \varphi}(u(X) < \theta < v(X)) = 1 - \alpha \text{ for all } (\theta, \varphi)$$

and $u(X)$ and $v(X)$ are *observable* random variables, i.e. one need not know the value of the unobservable quantities θ, φ in order to know the values of $u(X)$ and $v(X)$.

The number $1 - \alpha$ (sometimes reported as a percentage $100\% \cdot (1 - \alpha)$) is the **confidence level**, sometimes called the **confidence coefficient**. Most standard books adopt this convention, where α will be a small number. Here $\Pr_{\theta, \varphi}$ is used to indicate the probability when the random variable X has the distribution characterised by (θ, φ) . An important part of this specification is that the random interval (U, V) covers the unknown value θ with a high probability no matter what the true value of θ actually is.

Note that here $\Pr_{\theta, \varphi}$ need not refer to an explicitly given parameterised family of distributions, although it often does. Just as the random variable X notionally corresponds to other possible realisations of x from the same population or from the same version of reality, the parameters (θ, φ) indicate that we need to consider other versions of reality in which the distribution of X might have different characteristics.

Intervals for random outcomes

Confidence intervals can be defined for random quantities as well as for fixed quantities as in the above. See prediction interval. For this, consider an additional single-valued random variable Y which may or may not be statistically dependent on X . Then the rule for constructing the interval $(u(x), v(x))$ provides a confidence interval for the as-yet-to-be observed value y of Y if

$$\Pr_{\theta, \varphi}(u(X) < Y < v(X)) = 1 - \alpha \text{ for all } (\theta, \varphi).$$

Here $\Pr_{\theta, \varphi}$ is used to indicate the probability over the joint distribution of the random variables (X, Y) when this is characterised by parameters (θ, φ) .

Approximate confidence intervals

For non-standard applications it is sometimes not possible to find rules for constructing confidence intervals that have exactly the required properties. But practically useful intervals can still be found. The coverage probability $c(\theta, \varphi)$ for a random interval is defined by

$$\Pr_{\theta, \varphi}(u(X) < \theta < v(X)) = c(\theta, \varphi)$$

and the rule for constructing the interval may be accepted as providing a confidence interval if

$$c(\theta, \varphi) \cong 1 - \alpha \text{ for all } (\theta, \varphi)$$

to an acceptable level of approximation.

Comparison to Bayesian interval estimates

A Bayesian interval estimate is called a credible interval. Using much of the same notation as above, the definition of a credible interval for the unknown true value of θ is, for a given α ^[5],

$$\Pr(u(x) < \Theta < v(x) | X = x) = 1 - \alpha.$$

Here Θ is used to emphasize that the unknown value of θ is being treated as a random variable. The definitions of the two types of intervals may be compared as follows.

- The definition of a confidence interval involves probabilities calculated from the distribution of X for given (θ, φ) (or conditional on these values) and the condition needs to hold for all values of (θ, φ) .
- The definition of a credible interval involves probabilities calculated from the distribution of Θ conditional on the observed values of $X = x$ and marginalised (or averaged) over the values of Φ , where this last quantity is the random variable corresponding to the uncertainty about the nuisance parameters in φ .

Note that the treatment of the nuisance parameters above is often omitted from discussions comparing confidence and credible intervals but it is markedly different between the two cases.

In some simple standard cases, the intervals produced as confidence and credible intervals from the same data set can be identical. They are very different if informative prior information is included in the Bayesian analysis.

Desirable properties

When applying standard statistical procedures, there will often be standard ways of constructing confidence intervals. These will have been devised so as to meet certain desirable properties, which will hold given that the assumptions on which the procedure rely are true. These desirable properties may be described as: validity, optimality and invariance. Of these "validity" is most important, followed closely by "optimality". "Invariance" may be considered as a property of the method of derivation of a confidence interval rather than of the rule for constructing the interval. In non-standard applications, the same desirable properties would be sought.

- *Validity*. This means that the nominal coverage probability (confidence level) of the confidence interval should hold, either exactly or to a good approximation.
- *Optimality*. This means that the rule for constructing the confidence interval should make as much use of the information in the data-set as possible. Recall that one could throw away half of a dataset and still be able to derive a valid confidence interval. One way of assessing optimality is by the length of the interval, so that a rule for constructing a confidence interval is judged better than another if it leads to intervals whose lengths are typically shorter.
- *Invariance*. In many applications the quantity being estimated might not be tightly defined as such. For example, a survey might result in an estimate of the median income in a population, but it might equally be considered as providing an estimate of the logarithm of the median income, given that this is a common scale for presenting graphical results. It would be desirable that the method used for constructing a confidence interval for the median income would give equivalent results when applied to constructing a confidence interval for the logarithm of the median income: specifically the values at the ends of the latter interval would be the logarithms of the values at the ends of former interval.

Methods of derivation

For non-standard applications, there are several routes that might be taken to derive a rule for the construction of confidence intervals. Established rules for standard procedures might be justified or explained via several of these routes. Typically a rule for constructing confidence intervals is closely tied to a particular way of finding a point estimate of the quantity being considered.

Statistics

This is closely related to the method of moments for estimation. A simple example arises where the quantity to be estimated is the mean, in which case a natural estimate is the sample mean. The usual arguments indicate that the sample variance can be used to estimate the variance of the sample mean. A naive confidence interval for the true mean can be constructed centered on the sample mean with a width which is a multiple of the square root of the sample variance.

Likelihood theory

Where estimates are constructed using the maximum likelihood principle, the theory for this provides two ways of constructing confidence intervals or confidence regions for the estimates.

Estimating equations

The estimation approach here can be considered as both a generalization of the method of moments and a generalization of the maximum likelihood approach. There are corresponding generalizations of the results of maximum likelihood theory that allow confidence intervals to be constructed based on estimates derived from estimating equations.

Via significance testing

If significance tests are available for general values of a parameter, then confidence intervals/regions can be constructed by including in the 100p% confidence region all those points for which the significance test of the null hypothesis that the true value is the given value is not rejected at a significance level of (1-p).^[1]

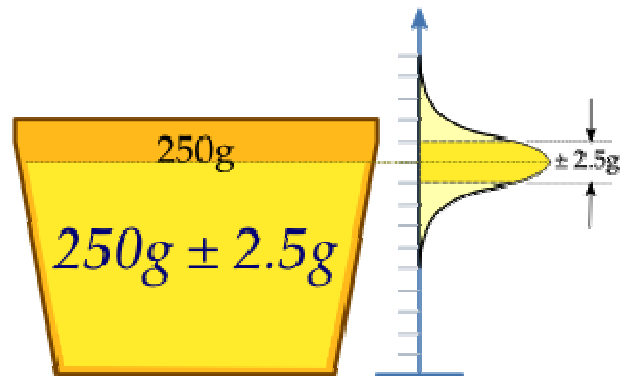
Bootstrapping

In situations where the distributional assumptions for that above methods are uncertain or violated, resampling methods allow construction of confidence intervals or prediction intervals. The observed data distribution and the internal correlations are used as the surrogate for the correlations in the wider population.

Examples

Practical example

A machine fills cups with margarine, and is supposed to be adjusted so that the content of the cups is 250 g of margarine. As the machine cannot fill every cup with exactly 250 g, the content added to individual cups shows some variation, and is considered a random variable X . This variation is assumed to be normally distributed around the desired average of 250 g, with a standard deviation of 2.5 g. To determine if the machine is adequately calibrated, a sample of $n = 25$ cups of margarine is chosen at random and the cups are weighed. The resulting measured masses of margarine are X_1, \dots, X_{25} , a random sample from X .



To get an impression of the expectation μ , it is sufficient to give an estimate. The appropriate estimator is the sample mean:

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

The sample shows actual weights x_1, \dots, x_{25} , with mean:

$$\bar{x} = \frac{1}{25} \sum_{i=1}^{25} x_i = 250.2 \text{ grams.}$$

If we take another sample of 25 cups, we could easily expect to find mass values like 250.4 or 251.1 grams. A sample mean value of 280 grams however would be extremely rare if the mean content of the cups is in fact close to 250 grams. There is a whole interval around the observed value 250.2 grams of the sample mean within which, if the whole population mean actually takes a value in this range, the observed data would not be considered particularly unusual. Such an interval is called a confidence interval for the parameter μ . How do we calculate such an interval? The endpoints of the interval have to be calculated from the sample, so they are statistics, functions of the sample X_1, \dots, X_{25} and hence random variables themselves.

In our case we may determine the endpoints by considering that the sample mean \bar{X} from a normally distributed sample is also normally distributed, with the same expectation μ , but with standard error $\sigma/\sqrt{n} = 0.5$ (grams). By standardizing, we get a random variable

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{0.5}$$

dependent on the parameter μ to be estimated, but with a standard normal distribution independent of the parameter μ . Hence it is possible to find numbers $-z$ and z , independent of μ , where Z lies in between with probability $1 - \alpha$, a measure of how confident we want to be. We take $1 - \alpha = 0.95$. So we have:

$$P(-z \leq Z \leq z) = 1 - \alpha = 0.95.$$

The number z follows from the cumulative distribution function:

$$\Phi(z) = P(Z \leq z) = 1 - \frac{\alpha}{2} = 0.975,$$

$$z = \Phi^{-1}(\Phi(z)) = \Phi^{-1}(0.975) = 1.96,$$

and we get:

$$\begin{aligned} 0.95 = 1 - \alpha &= P(-z \leq Z \leq z) = P\left(-1.96 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 1.96\right) \\ &= P\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\right) \\ &= P(\bar{X} - 1.96 \times 0.5 \leq \mu \leq \bar{X} + 1.96 \times 0.5) \\ &= P(\bar{X} - 0.98 \leq \mu \leq \bar{X} + 0.98). \end{aligned}$$

This might be interpreted as: with probability 0.95 we will find a confidence interval in which we will meet the parameter μ between the stochastic endpoints

$$\bar{X} - 0.98$$

and

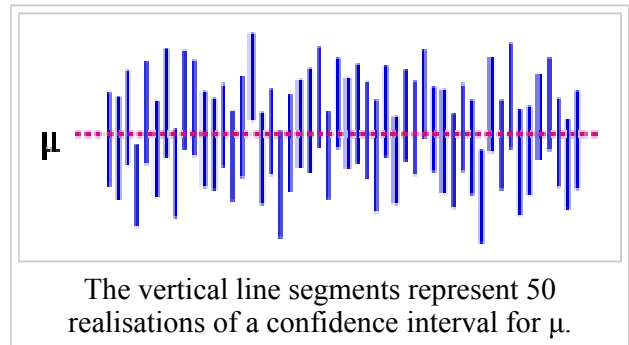
$$\bar{X} + 0.98.$$

This does not mean that there is 0.95 probability of meeting the parameter μ in the calculated interval. Every time the measurements are repeated, there will be another value for the mean \bar{X} of the sample. In 95% of the cases μ will be between the endpoints calculated from this mean, but in 5% of the cases it will not be. The actual confidence interval is calculated by entering the measured masses in the formula. Our 0.95 confidence interval becomes:

$$(\bar{x} - 0.98; \bar{x} + 0.98) = (250.2 - 0.98; 250.2 + 0.98) = (249.22; 251.18).$$

As the desired value 250 of μ is within the resulted confidence interval, there is no reason to believe the machine is wrongly calibrated.

The calculated interval has fixed endpoints, where μ might be in between (or not). Thus this event has probability either 0 or 1. One **cannot** say: "with probability $(1 - \alpha)$ the parameter μ lies in the confidence interval." One only knows that by repetition in $100(1 - \alpha)$ % of the cases, μ will be in the calculated interval. In 100α % of the cases however it does not. And unfortunately one does not know in which of the cases this happens. That is why one can say: "with **confidence level** $100(1 - \alpha)$ %, μ lies in the confidence interval."



The figure on the right shows 50 realisations of a confidence interval for a given population mean μ . If we randomly choose one realisation, the probability is 95% we end up having chosen an interval that contains the parameter; however we may be unlucky and have picked the wrong one. We will never know; we are stuck with our interval.

Theoretical example

Suppose X_1, \dots, X_n are an independent sample from a normally distributed population with (parameters) mean μ and variance σ^2 . Let

$$\bar{X} = (X_1 + \dots + X_n) / n,$$
$$S^2 = \frac{1}{n - 1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

be the well known statistics, sample mean and sample variance.

Then

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has a Student's t-distribution with $n - 1$ degrees of freedom. Note that the distribution of T does not depend on the values of the unobservable parameters μ and σ^2 ; i.e., it is a pivotal quantity. Suppose we wanted to calculate a 90% confidence interval for μ . Then, denoting c as the 95th percentile of this distribution,

$$\Pr(-c < T < c) = 0.9.$$

(Note: "95th" and "0.9" are correct in the preceding expressions. There is a 5% chance that T will be less than $-c$ and a 5% chance that it will be larger than $+c$. Thus, the probability that T will be between $-c$ and $+c$ is 90%.)

Consequently

$$\Pr\left(\bar{X} - \frac{cS}{\sqrt{n}} < \mu < \bar{X} + \frac{cS}{\sqrt{n}}\right) = 0.9$$

and we have a theoretical (stochastic) 90% confidence interval for μ .

After observing the sample we find values \bar{x} for \bar{X} and s for S , from which we compute the confidence interval

$$\left[\bar{x} - \frac{cs}{\sqrt{n}}, \bar{x} + \frac{cs}{\sqrt{n}}\right],$$

an interval with fixed numbers as endpoints, of which we can no more say there is a certain probability it contains the parameter μ . Either μ is in this interval or isn't.

Relation to hypothesis testing

While the formulations of the notions of confidence intervals and of statistical hypothesis testing are distinct they are in some senses related and to some extent complementary. While not all confidence intervals are constructed in this way, one general purpose approach to constructing confidence intervals is to define a $100(1 - \alpha)\%$ confidence interval to consist of all those values θ_0 for which a test of the hypothesis $\theta = \theta_0$ is not rejected at a significance level of $100\alpha\%$. Such an approach may not always be available since it presupposes the practical availability of an appropriate significance test. Naturally, any assumptions required for the significance test would carry over to the confidence intervals.

It may be convenient to make the general correspondence that parameter values within a confidence interval are equivalent to those values that would not be rejected by an hypothesis test, but this would be dangerous. In many instances the confidence intervals that are quoted are only approximately valid, perhaps derived from "plus or minus twice the standard error", and the implications of this for the supposedly corresponding hypothesis tests are usually unknown.

Meaning and interpretation

For users of frequentist methods, various interpretations of a confidence interval can be given.

- The confidence interval can be expressed in terms of samples (or repeated samples): "*Were this procedure to be repeated on multiple samples, the calculated confidence interval (which would differ for each sample) would encompass the true population parameter 90% of the time.*"^[6] Note that this need not be repeated sampling from the same population, just repeated sampling^[7].
- The explanation of a confidence interval can amount to something like: "*The confidence interval represents values for the population parameter for which the difference between the parameter and the observed estimate is not statistically significant at the 10% level*"^[8]. In fact, this relates to one particular way in which a confidence interval may be constructed.
- The probability associated with a confidence interval may also be considered from a pre-experiment point of view, in the same context in which arguments for the random allocation of treatments to study items are made. Here the experimenter sets out the way in which they intend to calculate a confidence interval and know, before they do the actual experiment, that the interval they will end up calculating has a certain chance of covering the true but unknown value.^[9] This is very similar to the "repeated sample" interpretation above, except that it avoids relying on considering hypothetical repeats of a sampling procedure that may not be repeatable in any meaningful sense. See Neyman construction.

In each of the above, the following applies: If the true value of the parameter lies outside the 90% confidence interval once it has been calculated, then an event has occurred which had a probability of 10% (or less) of happening by chance.

Meaning of the term "confidence"

There is a difference in meaning between the common usage of the word "confidence" and its statistical usage, which is often confusing to the layman, and this is one of the critiques of confidence intervals, namely that in application by non-statisticians, the term "confidence" is misleading.

In common usage, a claim to 95% confidence in something is normally taken as indicating virtual certainty. In statistics, a claim to 95% confidence simply means that the researcher has seen something occur that only happens one time in 20 or less. If one were to roll two dice and get double six (which happens 1/36th of the time, or about 3%), few would claim this as proof that the dice were fixed, although statistically speaking one could have 97% confidence that they were. Similarly, the finding of a statistical link at 95% confidence is not proof, nor even very good evidence, that there is any real connection between the things linked.

When a study involves multiple statistical tests, some laymen assume that the confidence associated with individual tests is the confidence one should have in the results of the study itself. In fact, the results of all the statistical tests conducted during a study must be judged as a whole in determining what confidence one may place in the positive links it produces. For example, say a study is conducted which involves 40 statistical tests at 95% confidence, and which produces 3 positive results. Each test has a 5% chance of producing a false positive, so such a study will produce 3 false positives about two times in three. Thus the confidence one can have that any of the study's positive conclusions are correct is only about 32%, well below the 95% the researchers have set as their standard of acceptance.

Alternatives and critiques

Main articles: Interval estimation, Credible interval, and Prediction interval

Confidence intervals are one method of interval estimation, and the most widely used in frequentist statistics. An analogous concept in Bayesian statistics is credible intervals, while an alternative frequentist method is that of prediction intervals which, rather than estimating *parameters*, estimate the outcome of *future* samples. For other approaches to expressing uncertainty using intervals, see interval estimation.

There is disagreement about which of these methods produces the most useful results: the mathematics of the computations are rarely in question – confidence intervals being based on sampling distributions, credible intervals being based on Bayes' theorem – but the application of these methods, the utility and interpretation of the produced statistics, is debated.

Users of Bayesian methods, if they produced an interval estimate, would in contrast to confidence intervals, want to say "*My degree of belief that the parameter is in fact in this interval is 90%*,"^[10] while users of prediction intervals would instead say "*I predict that the next sample will fall in this interval 90% of the time.*"^[citation needed]

Confidence intervals are an expression of probability and are subject to the normal laws of probability. If several statistics are presented with confidence intervals, each calculated separately on the assumption of independence, that assumption must be honoured or the calculations will be rendered invalid. For example, if a researcher generates a set of statistics with intervals and selects some of them as significant, the act of selecting invalidates the calculations used to generate the intervals.

Philosophical issues

The principle behind confidence intervals was formulated to provide an answer to the question raised in statistical inference of how to deal with the uncertainty inherent in results derived from data that are themselves only a randomly selected subset of an entire statistical population of possible datasets. There are other answers, notably that provided by Bayesian inference in the form of credible intervals. The idea of confidence intervals is that they correspond to a chosen rule for determining the confidence bounds, where this rule is essentially determined before any data are obtained, or before an experiment is done. The criterion for choosing this rule is that, over all possible datasets that might be obtained, there is a high probability that the interval determined by the rule will include the true value of the quantity under consideration. That is a fairly straightforward and reasonable way of specifying a rule for determining uncertainty intervals. The Bayesian approach appears to offer intervals that can, subject to acceptance of an interpretation of "probability" as Bayesian probability, be interpreted as meaning that the specific interval calculated from a given dataset has a certain probability of including the true value, conditional on the data and other information available. The confidence interval approach does not allow this, since in this formulation and at this same stage, both the bounds of interval and the true values are fixed values and there is no randomness involved.

For example, in the poll example outlined in the introduction, one common-sense interpretation of a "95% interval" might be that readers of this information can be 95% confident that the actual number of voters intending to vote for the party in question is between 36% to 44%. However, this is technically incorrect. The actual meaning of confidence levels and confidence intervals is rather more subtle. In the above case, a correct interpretation would be as follows: If the polling a large were repeated a number of times, each time generating a 95% confidence interval from the poll sample, then approximately 95% of the generated intervals would contain the true percentage of voters who intend to vote for the given party. Each time the polling is repeated, a *different* confidence interval is produced; hence, it is not possible to make precise statements about probabilities for any one given interval. For more information, see the section on meaning and interpretation.

The questions of how an interval expressing uncertainty in an estimate might be formulated, and of how such intervals might be interpreted, are not strictly a mathematical problems and are philosophically problematic.^[11] Mathematics can take over once the basic principles of an approach to inference have been established, but it has only a limited role in saying why one approach should be preferred to another.

Confidence intervals for proportions and related quantities

See also: Margin of error and Binomial proportion confidence interval

An approximate confidence interval for a population mean can be constructed for random variables that are not normally distributed in the population, relying on the central limit theorem, if the sample sizes and counts are big enough. The formulae are identical to the case above (where the sample mean is actually normally distributed about the population mean). The approximation will be quite good with only a few dozen observations in the sample if the probability distribution of the random variable is not too different from the normal distribution (e.g. its cumulative distribution function does not have any discontinuities and its skewness is moderate).

One type of sample mean is the mean of an indicator variable, which takes on the value 1 for true and the value 0 for false. The mean of such a variable is equal to the proportion that have the variable equal to one (both in the population and in any sample). This is a useful property of indicator variables, especially for hypothesis testing. To apply the central limit theorem, one must use a large enough sample. A rough rule of thumb is that one should see at least 5 cases in which the indicator is 1 and at least 5 in which it is 0. Confidence intervals constructed using the above formulae may include negative numbers or numbers greater than 1, but proportions obviously cannot be negative or exceed 1. Additionally, sample proportions can only take on a finite number of values, so the central limit theorem and the normal distribution are not the best tools for building a confidence interval. See "Binomial proportion confidence interval" for better methods which are specific to this case.

See also

- p-value
- Error bar
- Tolerance interval
- Robust confidence intervals

Online calculators

- TAMU's Confidence Interval Calculators (<http://www.stat.tamu.edu/~jhardin/applets/index.html>)
- GraphPad QuickCalcs (<http://www.graphpad.com/quickcalcs/index.cfm>)

Notes

1. ^{a b} Cox DR, Hinkley DV. (1974) Theoretical Statistics, Chapman & Hall, Section 7.2(iii)
2. ^a Goldstein, H., & Healey, M.J.R. (1995). "The graphical presentation of a collection of means." (<http://www.jstor.org/stable/view/2983411>) *Journal of the Royal Statistical Society*, **158**, 175–77.
3. ^a Wolfe R, Hanley J (Jan 2002). "If we're so different, why do we keep overlapping? When 1 plus 1 doesn't make 2" (<http://www.cmaj.ca/cgi/pmidlookup?view=long&pmid=11800251>) . *CMAJ* **166** (1): 65–6.

- PMID 11800251 (<http://www.ncbi.nlm.nih.gov/pubmed/11800251>) . PMC 99228 (<http://www.pubmedcentral.gov/articlerender.fcgi?tool=pmcentrez&artid=99228>) . <http://www.cmaj.ca/cgi/pmidlookup?view=long&pmid=11800251>.
4. ^ Zar, J.H. (1984) *Biostatistical Analysis*. Prentice Hall International, New Jersey. pp 43–45
 5. ^ Bernardo JE, Smith, Adrian (2000). *Bayesian theory*. New York: Wiley. pp. 259. ISBN 0-471-49464-X.
 6. ^ Cox DR, Hinkley DV. (1974) *Theoretical Statistics*, Chapman & Hall, p49, 209
 7. ^ Kendall, M.G. and Stuart, D.G. (1973) *The Advanced Theory of Statistics*. Vol 2: Inference and Relationship, Griffin, London. Section 20.4
 8. ^ Cox DR, Hinkley DV. (1974) *Theoretical Statistics*, Chapman & Hall, p214, 225, 233
 9. ^ Neyman, J. (1937) "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability" (<http://links.jstor.org/sici?sici=0080-4614%2819370830%29236%3A767%3C333%3A00ATOS%3E2.0.CO%3B2-6>) , *Philosophical Transactions of the Royal Society of London A*, 236, 333–380.
 10. ^ Cox DR, Hinkley DV. (1974) *Theoretical Statistics*, Chapman & Hall, p390
 11. ^ T. Seidenfeld, *Philosophical Problems of Statistical Inference: Learning from R.A. Fisher*, Springer-Verlag, 1979

References

- Fisher, R.A. (1956) *Statistical Methods and Scientific Inference*. Oliver and Boyd, Edinburgh. (See p. 32.)
- Freund, J.E. (1962) *Mathematical Statistics* Prentice Hall, Englewood Cliffs, NJ. (See pp. 227–228.)
- Hacking, I. (1965) *Logic of Statistical Inference*. Cambridge University Press, Cambridge. ISBN 0521051657
- Keeping, E.S. (1962) *Introduction to Statistical Inference*. D. Van Nostrand, Princeton, NJ.
- Kiefer, J. (1977) "Conditional Confidence Statements and Confidence Estimators (with discussion)" (<http://links.jstor.org/sici?sici=0162-1459%28197712%2972%3A360%3C789%3ACCSACE%3E2.0.CO%3B2-9>) *Journal of the American Statistical Association*, **72**, 789–827.
- Mayo, D. G. (1981) "In defence of the Neyman-Pearson theory of confidence intervals" (<http://www.phil.vt.edu/dmayo/PhilStatistics/In%20Defense%20of%20the%20Neyman-Pearson%20Theory%20of%20Confidence%20Intervals.pdf>) , *Philosophy of Science*, 48 (2), 269–280. JSTOR 187185 (<http://www.jstor.org/stable/187185>)
- Neyman, J. (1937) "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability" (<http://links.jstor.org/sici?sici=0080-4614%2819370830%29236%3A767%3C333%3A00ATOS%3E2.0.CO%3B2-6>) *Philosophical Transactions of the Royal Society of London A*, **236**, 333–380. (Seminal work.)
- Robinson, G.K. (1975) "Some Counterexamples to the Theory of Confidence Intervals." (<http://links.jstor.org/sici?sici=0006-3444%28197504%2962%3A1%3C155%3ASCTTTO%3E2.0.CO%3B2-4>) *Biometrika*, **62**, 155–161.
- Smithson, M. (2003) *Confidence intervals*. Quantitative Applications in the Social Sciences Series, No. 140. Belmont, CA: SAGE Publications. ISBN 9780761924999.

External links

- The Exploratory Software for Confidence Intervals tutorial programs that run under Excel (<http://www.latrobe.edu.au/psy/esci/>)
- Confidence interval calculators for R-Squares (<http://www.danielsoper.com/statcalc/calc28.aspx>) , Regression Coefficients (<http://www.danielsoper.com/statcalc/calc26.aspx>) , and Regression Intercepts (<http://www.danielsoper.com/statcalc/calc27.aspx>)
- Weisstein, Eric W., "Confidence Interval" (<http://mathworld.wolfram.com/ConfidenceInterval.html>) " from MathWorld.
- CAUSEweb.org (<http://www.causeweb.org>) Many resources for teaching statistics including Confidence Intervals.
- An interactive introduction to Confidence Intervals (<http://www.usablestats.com/tutorials/CI>)

- *Confidence Intervals: Confidence Level, Sample Size, and Margin of Error*
(<http://demonstrations.wolfram.com/ConfidenceIntervalsConfidenceLevelSampleSizeAndMarginOfError/>)
by Eric Schulz, the Wolfram Demonstrations Project.

Retrieved from "http://en.wikipedia.org/wiki/Confidence_interval"

Categories: [Statistical inference](#) | [Statistical terminology](#) | [Econometrics](#) | [Market research](#) | [Psychology](#) | [Biostatistics](#) | [Measurement](#) | [Intervals](#)

- This page was last modified on 19 November 2010 at 21:09.
- Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. See Terms of Use for details.
Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.