

Consistent estimator

In [statistics](#), a **consistent estimator** or **asymptotically consistent estimator** is an [estimator](#)—a rule for computing estimates of a parameter θ_0 —having the property that as the number of data points used increases indefinitely, the resulting sequence of estimates converges in probability to θ_0 . This means that the distributions of the estimates become more and more concentrated near the true value of the parameter being estimated, so that the probability of the estimator being arbitrarily close to θ_0 converges to one.

In practice one constructs an estimator as a function of an available sample of size n , and then imagines being able to keep collecting data and expanding the sample *ad infinitum*. In this way one would obtain a sequence of estimates indexed by n , and consistency is a property of what occurs as the sample size “grows to infinity”. If the sequence of estimates can be mathematically shown to converge in probability to the true value θ_0 , it is called a consistent estimator; otherwise the estimator is said to be **inconsistent**.

Consistency as defined here is sometimes referred to as *weak consistency*. When we replace convergence in probability with [almost sure convergence](#), then the estimator is said to be *strongly consistent*. Consistency is related to [bias](#); see [bias versus consistency](#).

Contents

Definition

Examples

Sample mean of a normal random variable

Establishing consistency

Bias versus consistency

Unbiased but not consistent

Biased but consistent

See also

Notes

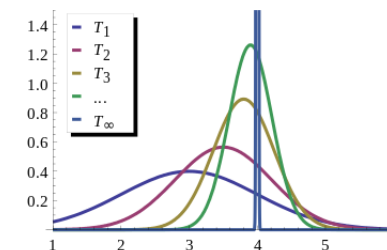
References

External links

Definition

Formally speaking, an [estimator](#) T_n of parameter θ is said to be **consistent**, if it [converges in probability](#) to the true value of the parameter:^[1]

$$\operatorname{plim}_{n \rightarrow \infty} T_n = \theta.$$



$\{T_1, T_2, T_3, \dots\}$ is a sequence of estimators for parameter θ_0 , the true value of which is 4. This sequence is consistent: the estimators are getting more and more concentrated near the true value θ_0 ; at the same time, these estimators are biased. The limiting distribution of the sequence is a degenerate random variable which equals θ_0 with probability 1.

A more rigorous definition takes into account the fact that θ is actually unknown, and thus the convergence in probability must take place for every possible value of this parameter. Suppose $\{p_\theta: \theta \in \Theta\}$ is a family of distributions (the parametric model), and $X^\theta = \{X_1, X_2, \dots : X_i \sim p_\theta\}$ is an infinite sample from the distribution p_θ . Let $\{T_n(X^\theta)\}$ be a sequence of estimators for some parameter $g(\theta)$. Usually T_n will be based on the first n observations of a sample. Then this sequence $\{T_n\}$ is said to be (weakly) **consistent** if^[2]

$$\text{plim}_{n \rightarrow \infty} T_n(X^\theta) = g(\theta), \text{ for all } \theta \in \Theta.$$

This definition uses $g(\theta)$ instead of simply θ , because often one is interested in estimating a certain function or a sub-vector of the underlying parameter. In the next example we estimate the location parameter of the model, but not the scale:

Examples

Sample mean of a normal random variable

Suppose one has a sequence of observations $\{X_1, X_2, \dots\}$ from a normal $N(\mu, \sigma^2)$ distribution. To estimate μ based on the first n observations, one can use the sample mean: $T_n = (X_1 + \dots + X_n)/n$. This defines a sequence of estimators, indexed by the sample size n .

From the properties of the normal distribution, we know the sampling distribution of this statistic: T_n is itself normally distributed, with mean μ and variance σ^2/n . Equivalently, $(T_n - \mu)/(\sigma/\sqrt{n})$ has a standard normal distribution:

$$\Pr[|T_n - \mu| \geq \varepsilon] = \Pr\left[\frac{\sqrt{n}|T_n - \mu|}{\sigma} \geq \sqrt{n}\varepsilon/\sigma\right] = 2\left(1 - \Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right)\right) \rightarrow 0$$

as n tends to infinity, for any fixed $\varepsilon > 0$. Therefore, the sequence T_n of sample means is consistent for the population mean μ (recalling that Φ is the cumulative distribution of the normal distribution).

Establishing consistency

The notion of asymptotic consistency is very close, almost synonymous to the notion of convergence in probability. As such, any theorem, lemma, or property which establishes convergence in probability may be used to prove the consistency. Many such tools exist:

- In order to demonstrate consistency directly from the definition one can use the inequality^[3]

$$\Pr[h(T_n - \theta) \geq \varepsilon] \leq \frac{\mathbf{E}[h(T_n - \theta)]}{h(\varepsilon)},$$

the most common choice for function h being either the absolute value (in which case it is known as Markov inequality), or the quadratic function (respectively Chebyshev's inequality).

- Another useful result is the continuous mapping theorem: if T_n is consistent for θ and $g(\cdot)$ is a real-valued function continuous at point θ , then $g(T_n)$ will be consistent for $g(\theta)$:^[4]

$$T_n \xrightarrow{p} \theta \quad \Rightarrow \quad g(T_n) \xrightarrow{p} g(\theta)$$

- Slutsky's theorem can be used to combine several different estimators, or an estimator with a non-random convergent sequence. If $T_n \xrightarrow{d} \alpha$, and $S_n \xrightarrow{p} \beta$, then ^[5]

$$\begin{aligned} T_n + S_n &\xrightarrow{d} \alpha + \beta, \\ T_n S_n &\xrightarrow{d} \alpha\beta, \\ T_n/S_n &\xrightarrow{d} \alpha/\beta, \text{ provided that } \beta \neq 0 \end{aligned}$$

- If estimator T_n is given by an explicit formula, then most likely the formula will employ sums of random variables, and then the law of large numbers can be used: for a sequence $\{X_n\}$ of random variables and under suitable conditions,

$$\frac{1}{n} \sum_{i=1}^n g(X_i) \xrightarrow{p} \mathbf{E}[g(X)]$$

- If estimator T_n is defined implicitly, for example as a value that maximizes certain objective function (see extremum estimator), then a more complicated argument involving stochastic equicontinuity has to be used.^[6]

Bias versus consistency

Unbiased but not consistent

An estimator can be unbiased but not consistent. For example, for an iid sample $\{x_1, \dots, x_n\}$ one can use $T_n(X) = x_n$ as the estimator of the mean $\mathbf{E}[x]$. Note that here the sampling distribution of T_n is the same as the underlying distribution (for any n , as it ignores all points but the last), so $\mathbf{E}[T_n(X)] = \mathbf{E}[x]$ and it is unbiased, but it does not converge to any value.

However, if a sequence of estimators is unbiased *and* converges to a value, then it is consistent, as it must converge to the correct value.

Biased but consistent

Alternatively, an estimator can be biased but consistent. For example, if the mean is estimated by $\frac{1}{n} \sum x_i + \frac{1}{n}$ it is biased, but as $n \rightarrow \infty$, it approaches the correct value, and so it is consistent.

Important examples include the sample variance and sample standard deviation. Without Bessel's correction (that is, when using the sample size n instead of the degrees of freedom $n - 1$), these are both negatively biased but consistent estimators. With the correction, the corrected sample variance is unbiased, while the corrected sample standard deviation is still biased, but less so, and both are still consistent: the correction factor converges to 1 as sample size grows.

Here is another example. Let T_n be a sequence of estimators for θ .

$$\Pr(T_n) = \begin{cases} 1 - 1/n, & \text{if } T_n = \theta \\ 1/n, & \text{if } T_n = n\delta + \theta \end{cases}$$

We can see that $T_n \xrightarrow{p} \theta$, $\mathbf{E}[T_n] = \theta + \delta$, and the bias does not converge to zero.

See also

- [Efficient estimator](#)
- [Fisher consistency](#) — alternative, although rarely used concept of consistency for the estimators
- [Regression dilution](#)
- [Statistical hypothesis testing](#)

Notes

- [Amemiya 1985](#), Definition 3.4.2.
- [Lehman & Casella 1998](#), p. 332.
- [Amemiya 1985](#), equation (3.2.5).
- [Amemiya 1985](#), Theorem 3.2.6.
- [Amemiya 1985](#), Theorem 3.2.7.
- [Newey & McFadden 1994](#), Chapter 2.

References

- [Amemiya, Takeshi \(1985\). *Advanced Econometrics* \(https://archive.org/details/advancedeconomet00amem\). Harvard University Press. ISBN 0-674-00560-0.](#)
- [Lehmann, E. L.; Casella, G. \(1998\). *Theory of Point Estimation* \(2nd ed.\). Springer. ISBN 0-387-98502-6.](#)
- [Newey, W. K.; McFadden, D. \(1994\). "Chapter 36: Large sample estimation and hypothesis testing" \(https://pdfs.semanticscholar.org/2fd6/c467097b104f4437097141e5e6cb774f9a2d.pdf\) \(PDF\). In Robert F. Engle and Daniel L. McFadden \(eds.\). *Handbook of Econometrics*. 4. Elsevier Science. ISBN 0-444-88766-0.](#)
- [Nikulin, M. S. \(2001\) \[1994\], "Consistent estimator" \(https://www.encyclopediaofmath.org/index.php?title=C/c025240\), in Hazewinkel, Michiel \(ed.\), *Encyclopedia of Mathematics*, Springer Science+Business Media B.V. / Kluwer Academic Publishers, ISBN 978-1-55608-010-4](#)
- [Sober, E. \(1988\), "Likelihood and convergence", *Philosophy of Science*, **55**: 228–237, doi:10.1086/289429 \(https://doi.org/10.1086%2F289429\).](#)

External links

- [Econometrics lecture \(topic: unbiased vs. consistent\) \(https://www.youtube.com/watch?v=TfuqBxRgRTU&list=PLD15D38DC7AA3B737&index=3#t=32m00m\) on YouTube](https://www.youtube.com/watch?v=TfuqBxRgRTU&list=PLD15D38DC7AA3B737&index=3#t=32m00m) by [Mark Thoma](#)

Retrieved from "https://en.wikipedia.org/w/index.php?title=Consistent_estimator&oldid=914634782"

This page was last edited on 8 September 2019, at 14:19 (UTC).

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.