Consistent estimator

by Marco Taboga, PhD

An estimator of a given parameter is said to be consistent if it converges in probability to the true value of the parameter as the sample size tends to infinity.

The main elements of an estimation problem

Before providing a definition of consistent estimator, let us briefly recall the main elements of a parameter estimation problem:

- a **sample** of data drawn from an unknown probability distribution; we denote the sample by ξ_n, where the subscript n is the sample size, that is, the number of observations in the sample;
- a **parameter** of the unknown data-generating distribution, denoted by θ_0 (e.g., the mean of a univariate distribution or the correlation coefficient of a bivariate distribution);
- an **estimator**, which is a function that associates an estimate $\widehat{\theta}_n$ to each sample ξ_n that could possibly be observed.

Sampling variability

Before being observed, the sample ξ_n is regarded as random.

Therefore, $\widehat{\theta}_n$, which depends on ξ_n, is a random variable.

When needed, we write

$$\widehat{\theta}_n = \widehat{\theta}_n(\xi_n)$$

to highlight the fact that the estimator $\widehat{\theta}_n$ is a function of the sample ξ_n.

Definition

Now, imagine that we are able to collect new data and increase our sample size n indefinitely, so as to obtain a sequence of samples $\{\xi_n\}$ and a sequence of estimators $\{\widehat{\theta}_n\}$.

If this "imaginary" sequence of estimators converges in probability to the true parameter value, then it is said to be consistent.

Definition A sequence of estimators $\{\widehat{\theta}_n\}$ is said to be consistent if and only if

$$\operatorname*{plim}_{n\to\infty} \widehat{\theta}_n = \theta_0$$

(sample size increases)

where plim denotes convergence in probability.

Terminology

Note that we have defined "consistent sequences of estimators".

But what do we mean by "consistent estimator"? The latter locution is informally used to mean that:

1. the same predefined rule is used to generate all the estimators in the sequence;
2. the terms of the sequence converge in probability to the true parameter value.

Thus, the concept of consistency extends from the sequence of estimators to the rule used to generate it.

For instance, suppose that the rule is to "compute the sample mean", so that $\{\widehat{\theta}_n\}$ is a sequence of sample means over samples of increasing size.

If $\{\widehat{\theta}_n\}$ converges in probability to the mean of the distribution that generated the samples, then we say that $\{\widehat{\theta}_n\}$ is consistent.

By a slight abuse of language, we also say that the sample mean is a consistent estimator.

Examples

The following table contains examples of consistent estimators (with links to lectures where consistency is proved).

Estimator	Estimated parameter	Lecture where proof can be found
Sample mean	Expected value	Estimation of the mean
Sample variance	Variance	Estimation of the variance
OLS estimator	Coefficients of a linear regression	Properties of the OLS estimator
Maximum likelihood estimator	Any parameter of a distribution	Maximum likelihood

Inconsistent estimator

An estimator which is not consistent is said to be inconsistent.

Consistent and asymptotically normal

You will often read that a given estimator is not only consistent but also asymptotically normal, that is, its distribution converges to a normal distribution as the sample size increases.

You might think that convergence to a normal distribution is at odds with the fact that consistency implies convergence in probability to a constant (the true parameter value).

In other words, you might ask yourself: "Is convergence to a constant or to a distribution?"

To answer this question, we should give a more precise definition of asymptotic normality.

Consider the ratio

$$\frac{\widehat{\theta}_n - \theta_0}{\operatorname{Std}\left[\widehat{\theta}_n\right]}$$

When $\{\widehat{\theta}_n\}$ is consistent, both the difference $\widehat{\theta}_n - \theta_0$ and the standard deviation $\operatorname{Std}\left[\widehat{\theta}_n\right]$ converge to zero as n tends to infinity. However, their ratio can converge to a distribution. When it converges to a standard normal distribution, then the sequence $\{\widehat{\theta}_n\}$ is said to be asymptotically normal.

The practical consequence of asymptotic normality is that, when n is large, we can approximate the above ratio with a standard normal distribution.

It follows that $\widehat{\theta}_n$ can be approximated by a normal distribution with mean θ_0 and standard deviation $\operatorname{Std}\left[\widehat{\theta}_n\right]$. But the latter converges to zero, so that the distribution becomes more and more concentrated around the mean, ultimately converging to a constant.

More details

Consistency is discussed in more detail in the lecture on Point estimation.

Keep reading the glossary

Previous entry: Conditional probability mass function

Next entry: Convergence criterion

How to cite

Please cite as:

Taboga, Marco (2021). "Consistent estimator", Lectures on probability theory and mathematical statistics. Kindle Direct Publishing. Online appendix. https://www.statlect.com/glossary/consistent-estimator.