



Advanced Quantitative Methods: Statistical estimators

Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

Johan A. Elkink

University College Dublin

2 February 2017



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

- 1 Introduction
- 2 Sampling distributions
- 3 Finite sample properties
- 4 Asymptotic properties



Outline

- 1 Introduction
- 2 Sampling distributions
- 3 Finite sample properties
- 4 Asymptotic properties

Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

“Problems of estimation are those in which it is required to estimate the value of one or more of the population parameters from a random sample of the population.”

(Fisher, 1922, 310)



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

- Ordinary Least Squares (OLS)
- Generalized Least Squares (GLS)
- Maximum Likelihood (ML)
- Simulated Maximum Likelihood (SML)
- General Method of Moments (GMM)
- Bayesian (usually Markov Chain Monte Carlo) (MCMC)
- etc.

Estimator criteria

Given that there are often many possible estimators for a particular population value, we want to be able to **evaluate** which one is most appropriate.

We can make a distinction between two types of criteria:

- **Finite sample** properties - how well does the estimator do given a limited sample size?
- **Asymptotic** properties - how well does the estimator do as the sample size gets infinitely large?



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

Estimator criteria



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

Given that there are often many possible estimators for a particular population value, we want to be able to **evaluate** which one is most appropriate.

We can make a distinction between two types of criteria:

- **Finite sample** properties - how well does the estimator do given a limited sample size?
- **Asymptotic** properties - how well does the estimator do as the sample size gets infinitely large?

(We will assume single parameter estimations (β), rather than multivariate ones (β) for the remainder of these slides.)



Outline

- 1 Introduction
- 2 Sampling distributions**
- 3 Finite sample properties
- 4 Asymptotic properties

Outline

Introduction

**Sampling
distributions**

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

Probabilities: definition



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

Frequentist definition:

$$P(x) = \lim_{n \rightarrow \infty} \frac{f(x)}{n},$$

where P is the probability, n the number of trials and f the frequency of event x occurring.

Probabilities: properties



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

For event E in **sample space** S :

$$0 \leq P(E) \leq 1$$

$$P(S) = 1$$

$$P(\neg E) = 1 - P(E)$$

$P(E_1 \cup E_2) = P(E_1) + P(E_2)$ if E_1 and E_2 are mutually exclusive

A sample space is the set of all possible outcomes, while an event is a specific subset of the sample space.



Outline

Introduction

**Sampling
distributions**

Finite sample
properties

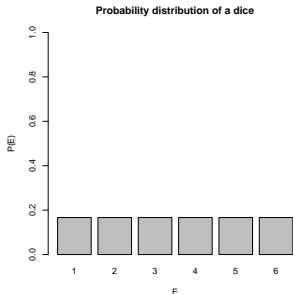
Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

A probability distribution of a discrete variable presents the probabilities for each possible outcome.

Because $P(S) = 1$, the bars add up to 1.





Outline

Introduction

Sampling
distributions

Finite sample
properties

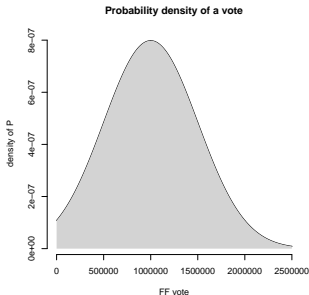
Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

A probability distribution of a continuous variable presents the probability density for each possible outcome. The surface under the plot represents the probability of the outcome being within a particular range.

Because $P(S) = 1$, the surface under the entire plot is 1.





Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness

Efficiency

Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness

Consistency

Efficiency

Imagine, that instead of having one sample, we take many samples.

If we do the same estimation in each of those randomly selected samples, we would get different results each time.

The distribution of these different estimates is the **sampling distribution** of the estimate.

Sampling distribution



Outline

Introduction

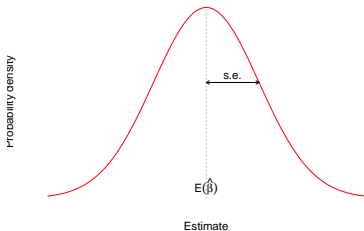
**Sampling
distributions**

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency



The sampling distribution is a probability density function, with as mean the expected value of $\hat{\beta}$.

The spread of this distribution is indicated by the **standard error** (s.e.).

$$se_{\hat{\beta}} = \sqrt{\text{var}(\hat{\beta})}$$

Outline



Outline

Introduction

Sampling
distributions

**Finite sample
properties**

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

- 1 Introduction
- 2 Sampling distributions
- 3 Finite sample properties**
- 4 Asymptotic properties

Unbiasedness



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness

Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

The **bias** of an estimator is the difference between the expected value of the sample distribution and the true value of the parameter to be estimated:

$$\text{bias}_{\hat{\beta}} = E(\hat{\beta}) - \beta$$

So an **unbiased estimator** is an estimator where $E(\hat{\beta}) = \beta$.

Bias: sampling distribution



Outline

Introduction

Sampling
distributions

Finite sample
properties

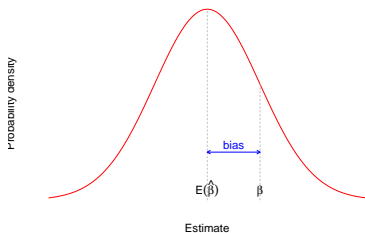
Unbiasedness

Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

For a biased estimator,
 $E(\hat{\beta}) \neq \beta$, and the bias is
 $E(\hat{\beta}) - \beta$.



Example: variance estimation



It can be shown (see notes) that, if s^2 is the **sample variance** and σ^2 the **population variance**, that:

$$E(s^2) = \frac{n-1}{n}\sigma^2,$$

in other words, s^2 is a **biased estimator** of σ^2 .

$$E\left(\frac{n}{n-1}s^2\right) = \frac{n}{n-1}E(s^2) = \frac{n}{n-1} \cdot \frac{n-1}{n}\sigma^2 = \sigma^2,$$

so when we estimate the population variance, we calculate $\frac{1}{n-1} \sum_i^n (x_i - \bar{x})^2$ instead of $\frac{1}{n} \sum_i^n (x_i - \bar{x})^2$.

Outline

Introduction

Sampling
distributionsFinite sample
properties**Unbiasedness**Efficiency
Mean Square
ErrorAsymptotic
propertiesLimits
Asymptotic
unbiasedness
Consistency
Efficiency

Unbiasedness: vector of coefficients



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness

Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

If β is a vector of coefficients, rather than a single parameter β , it simply still holds that an unbiased estimator is one where $\hat{\beta} = \beta$ and a biased estimator where it is not.



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness

Efficiency

Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness

Consistency

Efficiency

The estimator whose sampling distribution has the lowest variance is the more efficient estimator.

The most efficient unbiased estimator is called the **best unbiased** estimator.

Efficiency: sampling distribution



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness

Efficiency

Mean Square
Error

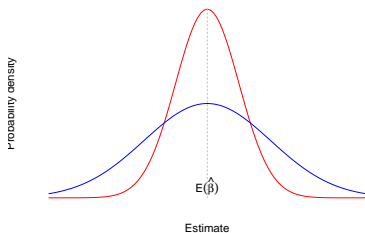
Asymptotic
properties

Limits

Asymptotic
unbiasedness

Consistency

Efficiency



$E(\hat{\beta})$ is the same for both estimators, but the estimator with the red sampling distribution is more efficient than the one with the blue sampling distribution.

$$se_{\hat{\beta}_{blue}} > se_{\hat{\beta}_{red}}$$



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness

Efficiency

Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness

Consistency

Efficiency

In the context of linear models, we often talk of the **best linear unbiased estimator** (BLUE), which is the estimator which is linear, unbiased, and has the lowest sampling variance of all possible unbiased linear estimators.

When the assumptions underlying OLS hold, OLS is BLUE.

Efficiency: vector of coefficients



Outline

Introduction

Sampling
distributionsFinite sample
properties

Unbiasedness

EfficiencyMean Square
ErrorAsymptotic
properties

Limits

Asymptotic
unbiasedness

Consistency

Efficiency

Often, β is a vector instead of just one parameter β . Instead of just the variance of β , we have a **variance-covariance matrix**:

$$\text{var}(\beta) = \begin{bmatrix} \text{var}(\beta_1) & \text{cov}(\beta_1, \beta_2) & \dots & \text{cov}(\beta_1, \beta_k) \\ \text{cov}(\beta_2, \beta_1) & \text{var}(\beta_2) & \dots & \text{cov}(\beta_2, \beta_k) \\ \dots & \dots & \dots & \dots \\ \text{cov}(\beta_k, \beta_1) & \text{cov}(\beta_k, \beta_2) & \dots & \text{var}(\beta_k) \end{bmatrix}$$

Efficiency: vector of coefficients

In this case we can consider various criteria for efficiency:

- smallest trace
- smallest determinant
- smallest variance of any linear combination of its elements
- some weighted sum of the variances and covariances

$$E[(\hat{\beta} - \beta)' \mathbf{W} (\hat{\beta} - \beta)]$$

If \mathbf{W} is selected such that this equation cannot lead to a negative result, minimizing this expectation leads to the most efficient estimator on all the above grounds.



Mean Square Error



Sometimes an estimator can be slightly biased, but be so much more efficient, that it becomes preferable.

So an estimator which is BLUE is not necessarily the most efficient estimator - it is simply the most efficient *unbiased* estimator.

The **Mean Square Error** (MSE) is the sum of the variance and the square of the bias of an estimator:

$$\begin{aligned}MSE_{\hat{\beta}} &= E[(\hat{\beta} - \beta)^2] \\ &= \text{var}(\hat{\beta}) + E(\hat{\beta} - \beta)^2\end{aligned}$$

Outline

Introduction

Sampling
distributionsFinite sample
propertiesUnbiasedness
Efficiency
Mean Square
ErrorAsymptotic
properties

Limits

Asymptotic
unbiasedness
Consistency
Efficiency

MSE: vector of coefficients



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness

Efficiency

**Mean Square
Error**

Asymptotic
properties

Limits

Asymptotic
unbiasedness

Consistency

Efficiency

When β is a vector of coefficients instead of a single parameter β , we could look at the **MSE matrix**:

$$\mathbf{MSE} = E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'$$

Or we could just look at the trace of this matrix.



Outline

- 1 Introduction
- 2 Sampling distributions
- 3 Finite sample properties
- 4 Asymptotic properties

Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

**Asymptotic
properties**

Limits
Asymptotic
unbiasedness
Consistency
Efficiency



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

**Asymptotic
properties**

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

The **asymptotic properties** of an estimator concern the estimator's sampling distribution in extremely (or infinitely) large samples.

We need these properties when we cannot provide the same proofs of unbiasedness, efficiency, etc. for small samples. To what extent the properties hold for small samples is then left for further exploration (e.g. through simulation).



Outline

Introduction

Sampling
distributionsFinite sample
propertiesUnbiasedness
Efficiency
Mean Square
ErrorAsymptotic
properties**Limits**Asymptotic
unbiasedness
Consistency
Efficiency

“The real-valued sequence $\{x_n\}$ has the real number x^* for its **limit**, or converges to x^* , if for any positive ε , no matter how small, it is possible to find a positive integer N such that for all integers n greater than N , $|x_n - x^*| < \varepsilon$ ”.

We can write:

$$\lim_{n \rightarrow \infty} x_n = x^*.$$

(Davidson and MacKinnon, 1993, 102)



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness
Consistency
Efficiency

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0$$



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness
Consistency
Efficiency

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0$$

$$\lim_{n \rightarrow \infty} \frac{2n + 1}{n - 1} = 2$$



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness
Consistency
Efficiency

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0$$

$$\lim_{n \rightarrow \infty} \frac{2n + 1}{n - 1} = 2$$

$$\lim_{n \rightarrow \infty} \frac{n - 1}{n} \sigma^2 = \sigma^2$$



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness
Consistency
Efficiency

$$P(\|\mathbf{x}_n - \mathbf{x}^*\| > \varepsilon) < \delta,$$

whereby ε and δ are arbitrarily small, positive real numbers.

We can write:

$$\text{plim}_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}^*.$$

(Davidson and MacKinnon, 1993, 103)

Convergence in distribution



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness
Consistency
Efficiency

$$\lim_{n \rightarrow \infty} P(\mathbf{x}_n \leq \mathbf{b}) = P(\mathbf{x}^* \leq \mathbf{b})$$

for any arbitrary \mathbf{b} . In other words, the probability distribution of $\{\mathbf{x}_n\}$ approaches that of \mathbf{x}^* as n increases. This can be written as:

$$\mathbf{x}_n \xrightarrow{D} \mathbf{x}^* .$$

(Davidson and MacKinnon, 1993, 107)

Asymptotic unbiasedness



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits

**Asymptotic
unbiasedness**

Consistency
Efficiency

Just as we can look at the bias in finite samples, we can also talk of the **asymptotic bias** in infinitely large samples:

$$asy.bias = \lim_{n \rightarrow \infty} (E(\hat{\beta}) - \beta)$$

(but see Greene, 2003, 917)

Asymptotic unbiasedness: example



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits

**Asymptotic
unbiasedness**

Consistency

Efficiency

$$E(s^2) = \frac{n-1}{n}\sigma^2,$$

so s^2 is a **biased** estimator of σ^2 .

$$\lim_{n \rightarrow \infty} E(s^2) = \lim_{n \rightarrow \infty} \frac{n-1}{n}\sigma^2 = \sigma^2,$$

so s^2 is an **asymptotically unbiased** estimator of σ^2 .

(see for the asymptotic variance of this estimator Greene, 2003, 917–918)



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

“A statistic satisfies the criterion of consistency, if, when it is calculated from the whole population, it is equal to the required parameter.”

(Fisher, 1922, 309)



Outline

Introduction

Sampling
distributionsFinite sample
propertiesUnbiasedness
Efficiency
Mean Square
ErrorAsymptotic
properties

Limits

Asymptotic
unbiasedness**Consistency**

Efficiency

An estimator is **consistent** iff:

$$\text{plim}_{n \rightarrow \infty} \hat{\beta} = \beta$$

In other words, the **asymptotic variance** of $\hat{\beta}$ becomes very small and $\hat{\beta}$ is **asymptotically unbiased**, so that the probability distribution of $\hat{\beta}$ “collapses” to a very tight distribution around β .

Consistency: sample mean



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

The **Central Limit Theorem** states:

$$E(\bar{x}) = \mu_x \quad \text{var}(\bar{x}) = \frac{\sigma^2}{n}$$

therefore:

$$\text{plim}_{n \rightarrow \infty} E(\bar{x}) = \mu_x \quad \text{plim}_{n \rightarrow \infty} \text{var}(\bar{x}) = 0$$

The mean of a random sample (\bar{x}) is a **consistent estimator** of the mean of the population (μ_x).



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

“The criterion of efficiency is satisfied by those statistics which, when derived from large samples, tend to a normal distribution with the least possible standard deviation.”

(Fisher, 1922, 310)



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness

Efficiency

Mean Square
Error

Asymptotic
properties

Limits

Asymptotic
unbiasedness

Consistency

Efficiency

The **asymptotic variance** of an estimator is:

$$\lim_{n \rightarrow \infty} \text{var}(\hat{\beta})$$



Outline

Introduction

Sampling
distributions

Finite sample
properties

Unbiasedness
Efficiency
Mean Square
Error

Asymptotic
properties

Limits
Asymptotic
unbiasedness
Consistency
Efficiency

“An estimator is **asymptotically efficient** if it is consistent, asymptotically normally distributed, and has an asymptotic covariance matrix that is not larger than the asymptotic covariance matrix of any other consistent, asymptotically normally distributed estimator.”

(Greene, 2003, 71)



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

Appendix



Other criteria

Least squares
Likelihood
Sufficiency
Computation

References

5 Other criteria

Least squares



We can look at the difference between the predicted and the observed values of the dependent variable:

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$$

and we want this prediction error to be as small as possible.

Many options:

- minimizing absolute errors, $\min(|\mathbf{e}|)$
- minimizing squared errors, $\min(\mathbf{e}'\mathbf{e})$
- using weights, $\min(\mathbf{e}'\mathbf{W}\mathbf{e})$
- etc.

OLS minimizes $\mathbf{e}'\mathbf{e}$.

Other criteria

Least squares

Likelihood

Sufficiency

Computation

References



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

R^2 represents the explained variance in \mathbf{y} , as a proportion of the total variance in \mathbf{y} .



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

Only appropriate when:

- using OLS for estimation
- explained variation is linear only
- there is an intercept in the model



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

R^2 represents the explained variance in \mathbf{y} , as a proportion of the total variance in \mathbf{y} .

- OLS maximizes R^2 implicitly



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

R^2 represents the explained variance in \mathbf{y} , as a proportion of the total variance in \mathbf{y} .

- OLS maximizes R^2 implicitly
- High R^2 can be caused by:
 - Using a lot of regressors
 - Including lagged dependent variable



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

R^2 represents the explained variance in \mathbf{y} , as a proportion of the total variance in \mathbf{y} .

- OLS maximizes R^2 implicitly
- High R^2 can be caused by:
 - Using a lot of regressors
 - Including lagged dependent variable
- No good equivalent for non-linear of models



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

R^2 represents the explained variance in \mathbf{y} , as a proportion of the total variance in \mathbf{y} .

- OLS maximizes R^2 implicitly
- High R^2 can be caused by:
 - Using a lot of regressors
 - Including lagged dependent variable
- No good equivalent for non-linear of models
- When interested in causal effect, high R^2 is not all too interesting



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

“These measures of goodness of fit have a fatal attraction. Although it is generally conceded among insiders that they do not mean a thing, high values are still a source of pride and satisfaction to their authors, however hard they may try to conceal these feelings.”

(Cramer 1987, 253, 45-51, as cited in Kennedy 2008, 27)



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

Too much emphasis on these criteria can lead to **overfitting**, where you get excellent results for the sample at hand, but not if you would use the same estimates on any other data.

One way to reduce the problem of overfitting is to **split the sample** in two, use one half for estimation, and then use the estimated values to predict the dependent variable in the other half of the sample, and check errors and R^2 .



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

“The likelihood that any parameter (or set of parameters) should have any assigned value (or set of values) is proportional to the probability that if this were so, the totality of observations should be that observed.”

(Fisher, 1922, 310)

Maximum Likelihood



The likelihood is proportional to the probability of observing the data you have (give or take some arbitrarily small deviation), given some parameter estimate:

$$L(\beta|\mathbf{y}, \mathbf{X}) = \alpha P(\mathbf{y}, \mathbf{X}|\beta)$$

The **likelihood function** is thus proportional to the **probability density function** of the given sample, as a function of the parameter values.

The estimator that maximizes this function also maximizes this probability density function and is the **Maximum Likelihood Estimator** (MLE or ML).



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

“A statistic satisfies the criterion of sufficiency when no other statistic which can be calculated from the same sample provides any additional information as to the value of the parameter to be estimated.”

(Fisher, 1922, 310)



Other criteria

Least squares

Likelihood

Sufficiency

Computation

References

Used to be a major problem ...

... not so much anymore.

Still worth considering for very large datasets.

E.g. $|(\mathbf{I} - \hat{\rho}\mathbf{W})^{-1}|$ in a spatial regression model.



Other criteria

Least squares
Likelihood
Sufficiency
Computation

References

- Cramer, Jan Solomon. 1987. "Mean and variance of R^2 in small and moderate samples." *Journal of econometrics* 35(2):253–266.
- Davidson, Russell and James G. MacKinnon. 1993. *Estimation and inference in econometrics*. Oxford: Oxford University Press.
- Fisher, Ronald A. 1922. "On the Mathematical Foundations of Theoretical Statistics." *Philosophical Transactions of the Royal Society of London, Series A* 222:309–368.
- Greene, William H. 2003. *Econometric analysis*. 5th ed. Upper Saddle River: Prentice Hall.
- Kennedy, Peter. 2008. *A guide to econometrics*. 6th ed. Malden, MA: Blackwell.