Econometrics. 2022.4.29

Statistical Inference.

properties

probability theory

sample properties

data generation process (population) $\xrightarrow{\text{sampling}}$ observed data (sample)

parameter

estimation
hypothesis testing

sample statistic

statistics (statistical inference).

- Frequentist approach : parameters are fixed.
- Bayesian approach : parameters are random.
-

- Parametric approach : population distribution is assumed. (maximum likelihood).
- Semi-parametric approach : " is not assumed.
- Non-parametric approach : inferring distribution rather than parameters. (histogram).

- **Statistic**

A statistic is any function computed from data in a sample.

E.g. the sample mean $\bar{x} = \frac{1}{n} \sum x_i$

the sample variance $s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

- **Estimator**

An estimator is a rule (function) or strategy for using the data to estimate the parameter.

Any statistic can be an estimator.

Estimation.

	To find	To evaluate
Estimator	. method of moments.	★ mean squared error.
.	★ maximum likelihood.	. unbiasedness.
.		. minimum variance.
.		
.		
.		

Method of moments.

Population moments : $\mu_k = E[X^k]$

Sample moments : $\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$.

⇒ Express parameters using moments, and estimate with sample moments.

parameters.

E.g. $X \sim N(\mu, \sigma^2)$, $\underline{\mu} = \underset{\sim}{\mu_1}$, $\underline{\sigma}^2 = \underset{\sim}{\mu_2} - \underset{\sim}{\mu_1}^2$

$\hat{\mu} = \hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$, $\hat{\sigma}^2 = \hat{\mu}_2 - \hat{\mu}_1^2$

$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Maximum likelihood estimator (MLE).

Parameter θ, sample $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$.

The likelihood function of θ is

$$L(\theta | x) = f(x | \theta) \Leftarrow \text{joint density of } x.$$

$\Rightarrow$ The MLE of θ is

$$\hat{\theta}(X) = \arg\max_{\theta} L(\theta | X) = \arg\max_{\theta} \log L(\theta | X).$$

E.g. $X_1, \cdots, X_n \sim$ i.i.d. $N(\mu, \sigma^2)$.

$$L(\mu, \sigma | x) = f(x | \mu, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x_i - \mu)^2}{\sigma^2}\right)$$

Mean squared error (MSE)

The MSE of W as an estimator of parameter θ is

$$E[(W-\theta)^2] = E[(W - E[W] + E[W] - \theta)^2]$$

$$= \cdots$$

$$= \underset{\text{variance}}{Var[W]} + (\underbrace{E[W]-\theta}_{\text{bias}})^2.$$

bias^2.

$\Rightarrow$ minimum variance $\underbrace{\text{unbiased}}_{E[W]-\theta=0}$ estimator

$\Rightarrow$ The MLE is minimum variance unbiased for large sample.