Statistical and Econometric Models

Lectures: Lennart Oelschläger · Tutorials: Sebastian Büscher ✉

Problem Set – Week 2 (for tutorial on April 29, 2024)
Solutions

Estimator Properties

Problem 2.1:

Let $x_1, x_2, \dots, x_N \in \mathbb{R}$ be realisations (a random sample, that is an independent identically distributed sequence) of a random Variable x with $\mathbb{E}x = \mu < \infty$ and $\text{Var}(x) = \sigma^2 < \infty$. In the lecture, the sample mean $\bar{\boldsymbol{x}} = \frac{1}{N}\sum_{i=1}^{N} x_i$ was introduced as an estimator for the expectation $\mathbb{E}x$. For this estimator,

(a) show that it is unbiased,

Solution: Let $x_i \in \mathbb{R}$ be independent, identically distributed (iid) random realisations of a random variable x with $\mathbb{E}x_i = \mathbb{E}x = \mu < \infty$ and $\text{Var}(x_i) = \text{Var}(x) = \sigma^2 < \infty$.

$$
\begin{aligned}
\mathbb{E}\left(\bar{\boldsymbol{x}}\right) &= \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) && \mid \text{linearity} \\
&= \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\left(x_i\right) && \mid \mathbb{E}x_i = \mu \\
&= \frac{1}{N}\sum_{i=1}^{N} \mu = \frac{1}{N} N\mu = \mu \\
&= \mathbb{E}x.
\end{aligned}
$$

(b) calculate the variance.

Solution:

$$
\begin{aligned}
\text{Var}\left(\bar{\boldsymbol{x}}\right) &= \text{Var}\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) && \mid \text{Calculation rules for Var} \\
&= \frac{1}{N}\left(\sum_{i=1}^{N} \text{Var}\left(x_i\right)\right)\frac{1}{N} && \mid \text{Var}(x_i) = \sigma^2 \\
&= \frac{1}{N^2}\sum_{i=1}^{N} \sigma^2 = \frac{1}{N^2} N\sigma^2 = \frac{1}{N}\sigma^2 \\
&= \frac{1}{N}\text{Var}\left(x\right)
\end{aligned}
$$

Problem 2.2:

We want to visualise the results of Problem 2.1 using `R`. This can be done by following the steps:

(a) Draw $N = 1000$ (pseudo-)random realisations of a normally distributed random variable with mean $\mu = 5$ and variance $\sigma^2 = 4$. This can be done in `R` using the function `rnorm`[1].

(b) Calculate the sample mean of the drawn realisations.

(c) Repeat step (a) and (b) $M = 1000$ times. This can be done using a `for`-loop. For each time calculate the mean of the draws and store it in a vector. This way you obtain $M = 1000$ estimates of the mean of the distribution from which the samples are drawn.

(d) Calculate the mean and variance of the estimates. Compare the results with the estimator properties from Problem 2.1.

(e) Plot a histogram of the results using the function `hist` and compare the shape with a histogram of (pseudo)-random draws from a normal distribution with appropriate mean and variance.

This can also be done for random variables that follow a different distribution than the one given in (a), showing that the properties of the sample mean as an estimator for the expected value does not depend on the distribution of x.

(f) Try out the same steps (a) to (e) with different distributions. You can first variate the mean and variance of the normal distribution. You should also try out different types of distributions. Use the help function `?Distributions` for some inspiration. You may also vary N and M.

(g) Can you find a distribution for which the histogram of the distribution of the estimates of the mean (plot from part (e)) does not represent approximately a normal distribution? Why is this the case? Does this contradict the previous results from Problem 2.1?

(h) Which theorem formalises the results that we can observe here?

[1] Use the help function `?rnorm` to first familiarise yourself with the function.