M1 INTERMEDIATE ECONOMETRICS

ASYMPTOTICS FOR SAMPLE MEANS

Koen Jochmans

September 15, 2025

1. Sample-mean theorem

Consider a random sample $X_1, X_2, \dots, X_n$ on a (say scalar) random variable X. Let

$$\theta = \mathbb{E}(X), \qquad \sigma^2 = \text{var}(X) = \mathbb{E}((X-\theta)^2).$$

We will assume that $\mathbb{E}(X^2) < +\infty$, so that the mean and variance are well defined.

The sample mean is the random variable

$$\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i.$$

The sample-mean theorem states that

(i) $\mathbb{E}(\bar{X}_n) = \theta$,

(ii) $\text{var}(\bar{X}_n) = n^{-1}\sigma^2$.

This is always true. For (i),

$$\mathbb{E}(\bar{X}_n) = \mathbb{E}\left(n^{-1}\sum_{i=1}^{n} X_i\right) = n^{-1}\sum_{i=1}^{n} \mathbb{E}\left(X_i\right) = n^{-1}\sum_{i=1}^{n} \mathbb{E}(X) = \theta.$$

For (ii),

$$\text{var}(\bar{X}_n) = \text{var}\left(n^{-1}\sum_{i=1}^{n} X_i\right) = n^{-2}\sum_{i=1}^{n} \text{var}\left(X_i\right) = n^{-2}\sum_{i=1}^{n} \text{var}\left(X\right) = n^{-1}\sigma^2.$$

In words, the sample-mean theorem states that the random variable $\bar{X}_n$ is unbiased for θ and has a variance around θ that shrinks to zero at the rate n^{-1}.

2. Law of large numbers

Beyond the mean and variance we cannot make general statements about the distribution of $\bar{X}_n$ for a given sample size n. We can say that this distribution collapses at θ as n grows large.

The probability that $\bar{X}_n$ lies further away than some $\epsilon > 0$ from θ equals

$$\mathbb{P}(\bar{X}_n > \theta + \varepsilon) + \mathbb{P}(\bar{X}_n < \theta - \varepsilon) = \mathbb{P}(|\bar{X}_n - \theta| > \varepsilon).$$

By Markov/Chebychev's inequality,

$$\mathbb{P}(|\bar{X}_n - \theta| > \varepsilon) = \mathbb{P}(|\bar{X}_n - \theta|^2 > \varepsilon^2) \leq \frac{\mathbb{E}((\bar{X}_n - \theta)^2)}{\varepsilon^2} = \frac{n^{-1}\sigma^2}{\varepsilon^2},$$

which goes to zero as $n \to \infty$. We write

$$\bar{X}_n \underset{p}{\to} \theta$$

to indicate that $\bar{X}_n$ converges in probability to θ (as $n \to \infty$).

This result does not require $\mathbb{E}(X^2) < +\infty$ (but it does require that $\mathbb{E}(|X|) < +\infty$, but showing this to be the case requires a more complicated proof.

3. Continuous-mapping theorem

It is useful to note that the law of large numbers implies that sample averages of transformations of the $X_1, X_2, \ldots, X_n$ also satisfy the law of large numbers.

Suppose that we care about the variable $Y = \varphi(X)$ for some (nonrandom) transformation φ. Then

$$\bar{Y}_n = n^{-1} \sum_{i=1}^{n} Y_i = n^{-1} \sum_{i=1}^{n} \varphi(X_i) \underset{p}{\rightarrow} \mathbb{E}(\varphi(X)) = \mathbb{E}(Y)$$

provided of course that $\mathbb{E}(|\varphi(X)|) < +\infty$. Examples are raw moments of X, that is, $\mathbb{E}(X^q)$ for some integer q.

A different situation is one where we care about a transformation of θ, say $\varphi(\theta)$. The continuous-mapping theorem (given here without proof) states that

$$\varphi(\bar{X}_n) \underset{p}{\rightarrow} \varphi(\theta)$$

as long as the function φ is continuous at θ. A relevant example is the inverse transformation $\varphi : x \mapsto x^{-1}$. In this case an estimator of $\theta^{-1} = \mathbb{E}(X)^{-1}$ takes the form $\bar{X}_n^{-1}$. Note that this estimator is not unbiased for θ^{-1}, in general, as

$$\mathbb{E}(\bar{X}_n^{-1}) \neq \mathbb{E}(\bar{X}_n)^{-1} = \theta^{-1}.$$

We nonetheless have that

$$\bar{X}_n^{-1} \underset{p}{\rightarrow} \theta^{-1}$$

as long as $\theta \neq 0$ (so that the inverse is well defined at θ). We will encounter other examples.

4. Central limit theorem

By virtue of the sample-mean theorem the standardized sample mean

$$\bar{Z}_n = \frac{\bar{X}_n - \theta}{\sigma/\sqrt{n}}$$

has mean zero and variance one for any sample size. Beyond these two moments, its cumulative distribution function,

$$F_n(z) = \mathbb{P}(Z_n \leq z)$$

depends on n in a complicated way.

Let $Z \sim N(0,1)$ be a standard-normal random variable. So

$$\mathbb{P}(Z \leq z) = \Phi(z),$$

the standard-normal distribution function at z. The central limit theorem states that, as $n \to \infty$,

$$F_n(z) \to \Phi(z)$$

for any $z \in \mathbb{R}$. The standardized sample mean converges in distribution to a standard-normal variable. We usually write this as $\bar{Z}_n \underset{d}{\to} Z$ or even more simply as

$$\bar{Z}_n \underset{d}{\to} N(0,1).$$

Equivalently,

$$\sqrt{n}(\bar{X}_n - \theta) \underset{d}{\to} N(0, \sigma^2),$$

in light of the definition of $\bar{Z}_n$. In practice, this means that, when n is not too small, we can treat $\bar{X}_n$ as (approximately) normal with mean θ and variance $n^{-1}\sigma^2$.

5. Delta method

If ultimate interest does not lie in the mean θ but some transformation $\varphi(\theta)$ we may wonder how the distribution of $\varphi(\bar{X}_n)$ behaves. To this end we presume that φ is continuously differentiable and do a mean-value expansion

$$\varphi(\bar{X}_n) = \varphi(\theta) + \left.\frac{\varphi(t)}{\partial t}\right|_{t=t_n^*} (\bar{X}_n - \theta),$$

where t_n^* is a random variable that lies between $\bar{X}_n$ and θ (there may be many such values t_n^*; this is not important). What matters is that, because t_n^* lies no further away from θ than does $\bar{X}_n$ and the latter converges in probability to θ,

$$\mathbb{P}(|t_n^* - \theta| > \varepsilon) \leq \mathbb{P}(|\bar{X}_n - \theta| > \varepsilon) \to 0,$$

and so

$$\left.\frac{\varphi(t)}{\partial t}\right|_{t=t^*} \underset{p}{\to} \left.\frac{\varphi(t)}{\partial t}\right|_{t=\theta}$$

by the assumed continuity of the derivative. Therefore, by Slutsky's theorem,

$$\sqrt{n}(\varphi(\bar{X}_n) - \varphi(\theta)) \underset{p}{\to} \left.\frac{\varphi(t)}{\partial t}\right|_{t=\theta} \sqrt{n}(\bar{X}_n - \theta) \underset{d}{\to} N(0, \sigma^2({}^{\varphi(t)}/_{\partial t}|_{t=\theta})^2).$$

So, we can treat $\varphi(\bar{X}_n)$ as approximately normal with mean $\varphi(\theta)$ and variance $n^{-1}\sigma^2({}^{\varphi(t)}/_{\partial t}|_{t=\theta})^2$.

6. Slutsky theorems

Suppose that we have convergence results of the form $A_n \underset{p}{\to} c$ and $B_n \underset{d}{\to} B$, where A_n and B_n are sequences of random variables, c is some constant, and B a random variable that does not depend on n. However, we care about the behavior of the product or sum of the variables. Slutsky theorems are useful here. They state that

$A_n + B_n \underset{d}{\to} c + B,$

$A_n\, B_n \underset{d}{\to} c\, B.$

One application of the latter result has a random sample $X_1, X_2, \ldots, X_n$ from before, and looks at the "t-statistic"

$$\frac{\bar{X}_n - \theta}{s/\sqrt{n}}, \qquad s^2 = n^{-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.$$

Here,

$$A_n = \frac{\sigma}{s}, \qquad B_n = \frac{\bar{X}_n - \theta}{\sigma/\sqrt{n}} = \bar{Z}_n.$$

The central limit theorem immediately gives $B_n \underset{d}{\to} N(0,1)$ but does not allow to make the same claim for the t-statistic. However, for the sample variance s^2 we have

$$\begin{aligned}
s^2 &= n^{-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \\
&= n^{-1} \sum_{i=1}^{n} ((X_i - \theta) - (\bar{X}_n - \theta))^2 \\
&= n^{-1} \sum_{i=1}^{n} (X_i - \theta)^2 + n^{-1} \sum_{i=1}^{n} (\bar{X}_n - \theta)^2 - 2n^{-1} \sum_{i=1}^{n} (X_i - \theta)(\bar{X}_n - \theta) \\
&= n^{-1} \sum_{i=1}^{n} (X_i - \theta)^2 - (\bar{X}_n - \theta)^2 \\
&\underset{p}{\to} \sigma^2
\end{aligned}$$

because $n^{-1} \sum_{i=1}^{n} (X_i - \theta)^2 \underset{p}{\to} \sigma^2$ by a direct application of the law of large numbers and $(\bar{X}_n - \theta)^2 \underset{p}{\to} 0$ by the continuous-mapping theorem. Therefore, $s \underset{p}{\to} \sigma$ or, equivalently, $A_n \underset{p}{\to} 1$. So,

$$\frac{\bar{X}_n - \theta}{s/\sqrt{n}} = \frac{\sigma}{s} \frac{\bar{X}_n - \theta}{\sigma/\sqrt{n}} = A_n B_n \underset{d}{\to} Z \sim N(0,1).$$

So, replacing the unknown variance σ^2 by an estimator s^2 does not affect the limit distribution.