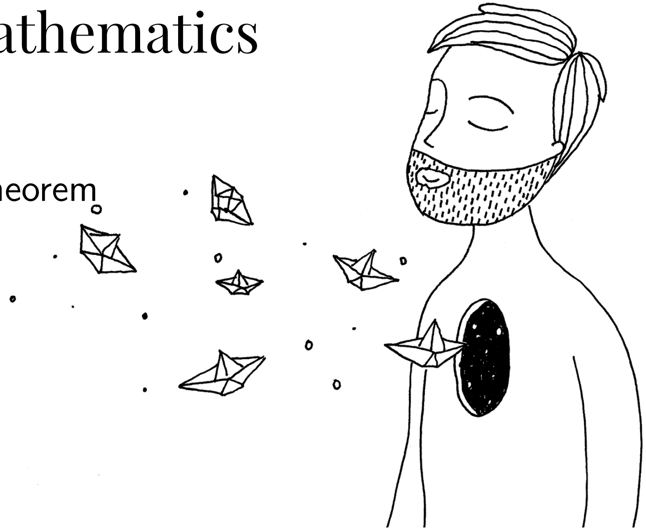
4509 - Bridging Mathematics

Lecture 11: Central Limit Theorem



PAULO FAGANDINI

NOVA SCHOOL OF BUSINESS & ECONOMICS

Consider a sequence $X_1, X_2, \ldots$ of independent and identical random variables with $\mu_X = \mu$ and $\sigma_X^2 = \sigma^2$.

Further, from now on, let $S_n = \sum_{i=1}^{n} X_i$, and $M_n = \frac{S_n}{n}$ (the **sample mean**).

Proposition

The variance of S_n is $n\sigma^2$.

Proposition

The variance of the sample mean is $\frac{\sigma^2}{n}$.

As a consequence we have that, the greater our sample... the lower the variance of the sample mean! See, it is trivial to show that $E[M_n] = \mu$, so if the variance goes to zero, it means that big samples should give very accurate estimates of the mean of the distribution!.

Important Inequalities

- **Markov:** If X is a nonnegative r.v., then $P(X \geq a) \leq \frac{E[X]}{a}$ for any $a > 0$.
- **Chebyshev:** If X is a r.v. with mean μ and variance σ^2, then $P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$ for any $c > 0$.

Definition

Let $Y_1, Y_2, \ldots$ be a sequence of r.v. (not necessarily indep), and let $a \in \mathbb{R}$. Y_n is said to **converge to a in probability** if for every $\epsilon > 0$ holds that:

$$\lim_{n \to \infty} P(|Y_n - a| \geq \epsilon) = 0$$

We can adapt conveniently the previous definition as...
For every $\epsilon > 0$, and for every $\delta > 0$, there is n_0 such that

$$P(|Y_n - a| < \epsilon) \leq \delta, \quad \forall n \geq n_0$$

ϵ is the accuracy and δ the confidence level. So with a confidence level δ we can say that Y_n is around a with an accuracy of ϵ.

One can think that if a sequence Y_n converges in probability to some constant c, then $E[Y_n]$ must also converge to c, however, this need not be the case!

Consider the discrete sequence of random variables Y_n with the following distribution:

$$P(Y_n = y) = \begin{cases} 1 - \frac{1}{n} & \text{, for } y = 0 \\ \frac{1}{n} & \text{, for } y = n^2 \\ 0 & \text{, elsewhere} \end{cases}$$

For every $\epsilon > 0$ we have:

$$\lim_{n \to \infty} P(|Y_n| \geq \epsilon) = \lim_{n \to \infty} \frac{1}{n} = 0$$

so Y_n converges to 0 in probability.
However,

$$E[Y_n] = \frac{n^2}{n} = n$$

which goes to ∞ as $n \to \infty$.

Theorem (Weak Law of Large Numbers)

Let $X_1, X_2, ...$ be independent and identically distributed (i.i.d) r.v. with mean μ. For every $\epsilon > 0$ we have

$$P(|M_n - \mu| \geq \epsilon) \to 0, \quad n \to \infty$$

From the WLLN we get that $Var[M_n] \to 0$ as $n \to \infty$. However, that is not the case for S_n. Note that S_n, which was the sum of the X's does not necessarily converge.Note:

$$E[S_n] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E\left[X_i\right] = \sum_{i=1}^{n} M_n = nM_n$$

Which clearly goes to infinity as $n \to \infty$. Furthermore, we cannot say anything about the distribution of S_n as $n \to \infty$

Let's define another variable.

Definition (Convergence in distribution)

A sequence of random variables X_1, X_2, ..., converges in distribution to a random variable X if:

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

at all points where $F_X(x)$ is continuous. We denote it as $X_n \xrightarrow{d} X$.

Theorem (Continuity Theorem)

Let X_n be a sequence of random variables with cumulative distribution function $F_n(x)$ and corresponding to moment generating functions $M_n(s)$ (transform). Let X be a random variable with cumulative distribution function $F(x)$ and moment generating function $M(s)$.

If $M_n(s) \to M(s)$ for any s in an open interval containing 0, then $F_n(x) \to F(x)$ at all continuity points of F, i.e. $X_n \xrightarrow{d} X$.

Definition (Transform)

The transform of the distribution of a random variable X (also referred to as the moment generating function of X) is a function $M_X(s)$ of a free parameter s, defined by:

$$M_X(s) = E\left[e^{sX}\right]$$

Moreover, we have that:

$$M(s) = \sum_x e^{sx} p_X(x)$$

for a discrete random variable, and

$$M(s) = \int_{-\infty}^{\infty} e^{sx} f_X(x) dx$$

for a continuous r.v.

A special case

The transform of a Normal Random Variable:
Let X be a normal random variable with mean μ and variance σ^2. To compute the corresponding transform, consider the special case of the standard normal r.v. $Y \sim N(0, 1)$. The *pdf* of the standard normal is

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}}$$

and its transform is:

A special case

$$
\begin{aligned}
M_Y(s) &= \int_{\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} e^{sy} dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} e^{\frac{-y^2}{2} + sy} dy \\
&= e^{\frac{s^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} e^{\frac{-y^2}{2} + sy - \frac{s^2}{2}} dy \\
&= e^{\frac{s^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} e^{\frac{-(y-s)^2}{2}} dy \\
&= e^{\frac{s^2}{2}}
\end{aligned}
$$

From transform to moments

Recall that

$$M(s) = \int_{-\infty}^{\infty} e^{sx} f_X(x) dx$$

Taking derivatives with respect to s in both sides we get:

$$\begin{aligned}
\frac{d}{ds} M(s) &= \frac{d}{ds} \int_{-\infty}^{\infty} e^{sx} f_X(x) dx \\
&= \int_{-\infty}^{\infty} \frac{d}{ds} e^{sx} f_X(x) dx \\
&= \int_{-\infty}^{\infty} x e^{sx} f_X(x) dx
\end{aligned}$$

If $s = 0$ we have:

$$\left. \frac{d}{ds} M(s) \right|_{s=0} = \int_{-\infty}^{\infty} x f_X(x) dx = E[X]$$

From transform to moments

Generalizing, taking the *n-th* derivative with respect to s in both sides we get:

$$\frac{d^n}{ds^n}M(s) = \int_{-\infty}^{\infty} x^n e^{sx} f_X(x)dx$$

If $s = 0$ we have:

$$\frac{d^n}{ds^n}M(s)\Big|_{s=0} = \int_{-\infty}^{\infty} x^n f_X(x)dx = E[X^n]$$

Let $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$

Proposition

$E[Z_n] = 0$ *and* $var(Z_n) = 1$.

Proof.

$E[Z_n] = \frac{1}{\sigma\sqrt{n}}(E[S_n] - n\mu) = \frac{1}{\sigma\sqrt{n}}(n\mu - n\mu) = 0$
$var(Z_n) = \frac{1}{n\sigma^2}var(S_n - n\mu) = \frac{1}{n\sigma^2}var(S_n) = \frac{1}{n\sigma^2}n\sigma^2 = 1$ □

The Central Limit Theorem

Theorem

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v. with common mean μ and variance σ^2. Define:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

then the cumulative distribution function of Z_n converges to the standard normal cumulative distribution function

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} dx$$

in the sense that

$$\lim_{n \to \infty} P(Z_n \leq z) = \Phi(z), \quad \forall z$$

Proof CLM

Let X_1, X_2... be a sequence of i.i.d. r.v. with mean 0 and variance σ^2, and associated to the transform $M_X(s)$. Assume that $M_X(s)$ is finite when $-d < s < d$, for d some positive number.
Now, let

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{S_n}{\sigma\sqrt{n}}$$

and let's rewrite $S_n = \sum_{i=1}^{n} X_i$.

Proof CLM

$$
\begin{aligned}
M_{Z_n}(s) &= E\left[e^{sZ_n}\right] \\
&= E\left[e^{\frac{s\sum_{i=1}^{n} X_i}{\sigma\sqrt{n}}}\right] \\
&= \prod_{i=1}^{n} E\left[e^{\frac{sX_i}{\sigma\sqrt{n}}}\right] \\
&= \prod_{i=1}^{n} M_X\left(\frac{s}{\sigma\sqrt{n}}\right) \\
&= \left(M_X\left(\frac{s}{\sigma\sqrt{n}}\right)\right)^n
\end{aligned}
$$

Proof CLM

Use a Taylor expansion around $s = 0$ to write

$$M_X(s) = M_X(0) + \frac{d}{ds} M_X(0)(s-0) + \frac{1}{2}\frac{d^2}{ds^2} M_X(0)(s-0)^2 + o(s^2)$$

where $\lim_{s \to 0} \frac{o(s^2)}{s^2} = 0$
And remember that $M_X(0) = E[e^0] = 1$, $\frac{d}{ds} M_X(s)\Big|_{s=0} = E[X] = 0$, and
$\frac{1}{2}\frac{d^2}{ds^2} M_X(s)\Bigg|_{s=0} = E[X^2] = \frac{\sigma^2}{2}$.

$$M_{Z_n}(s) = \left(M_X \left(\frac{s}{\sigma\sqrt{n}} \right) \right)^n = \left(1 + \frac{s^2}{2n} + o\left(\frac{s^2}{\sigma^2 n} \right) \right)^n$$

Noting that the error goes to zero, we have the following limit:

$$\lim_{n \to \infty} M_{Z_n}(s) = \lim_{n \to \infty} \left(1 + \frac{s^2/2}{n}\right)^n = e^{s^2/2}$$

Which is exactly the same as the transform of the Normal Distribution.

Using the Continuity Theorem, we get that our Z_n's distribution converges to the Normal Distribution.

Definition (Almost surely convergence)

Let $X_1, X_2, \ldots$ be a sequence of r.v. (not necessarily independent) associated with the same probability model. Let $c \in \mathbb{R}$. We say that $X_n \to c$ **almost surely** if

$$P\left(\lim_{n\to\infty} X_n = c\right) = 1$$

Theorem (Strong Law of Large Numbers)

*Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v. with mean μ. Then the sequence of sample means M_n converges to μ, **with probability 1**, in the sense that:*

$$P\left(\lim_{n\to\infty} M_n = \mu\right) = 1$$

It is not easy to see the difference between *SLLN* and *WLLN*. So first thing, they **are** very close. Indeed, almost sure convergence implies convergence in probability, however the opposite is not true.

Lemma (Second Borel-Cantelli Lemma)

If $\sum_{i=1}^{\infty} Pr(E_n) = \infty$*, with* $\{E_n\}_{n=1}^{\infty}$ *being independent events, then,* $Pr(\limsup_{n\to\infty} E_n) = 1$*.*

In words, if the sum of the probabilities of an event goes to infinity, then that event happens infinitely many times, and therefore with a non zero probability.

Consider the r.v. X_n as $P(X_n = 1) = \frac{1}{n}$ and $P(X_n = 0) = 1 - P(X_n = 1) = 1 - \frac{1}{n}$.

For any $\epsilon \in (0, 1)$, $X_n < \epsilon$ only if $X_n = 0$, which happens with probability $1 - \frac{1}{n}$, and as $n \to \infty$, then $P(X_n < \epsilon) \to 1$, so $X_n \to 0$ in probability.

Because $\sum_{i=1}^{n} P(X_n = 1) = \infty$, so there are a lot of elements in Ω that make $X_n = 1$, and therefore the sequence does not converge *almost surely*.

References

- B. E. Hansen, *Econometrics*, Ch. 6 and 7 (asymptotic theory).
- H. White, *Asymptotic Theory for Econometricians*, Ch. 2, 3 and 4.
- F. Hayashi, *Econometrics*, Ch. 2 (Sections 2.1 to 2.12).