Parametric Statistics

Large Random Samples

Sofia Triantafillou

sof.triantafillou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

October 23, 2023

Lecture Summary

The Sample Mean

Definition

Let $X_1, \ldots X_n$ be random variables. Their average

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

is called their *sample mean*.

The Sample Mean

Definition

Let $X_1, \dots X_n$ be random variables. Their average

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

is called their *sample mean*.

Mean and variance of the sample mean

Let $X_1, \dots X_n$ be a random sample from a distribution with mean μ and variance σ^2. Then $E(\overline{X}_n) = \mu$, and $Var(\overline{X}_n) = \sigma^2/n$.

Example

- You toss a fair coin.
- How many heads do you expect if you toss it 10 times?

Example

- You toss a fair coin.
- How many heads do you expect if you toss it 10 times?

$$\Pr(X = 5) = \binom{10}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^5 = 0.2461.$$

Example

- You toss a fair coin.
- How many heads do you expect if you toss it 10 times?

$$\Pr(X = 5) = \binom{10}{5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^5 = 0.2461.$$

- How many heads do you expect if you toss it 100 times?

$$\Pr(Y = 50) = \binom{100}{50} \left(\frac{1}{2}\right)^{50} \left(\frac{1}{2}\right)^{50} = 0.0796.$$

Example

- You toss a fair coin.
- How many heads do you expect if you toss it 10 times?

$$\Pr(X = 5) = \binom{10}{5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^5 = 0.2461.$$

- How many heads do you expect if you toss it 100 times?

$$\Pr(Y = 50) = \binom{100}{50} \left(\frac{1}{2}\right)^{50} \left(\frac{1}{2}\right)^{50} = 0.0796.$$

But:

$$\Pr\left(0.4 \leq \frac{X}{10} \leq 0.6\right) = \Pr(4 \leq X \leq 6) = 0.6563.$$

$$\Pr\left(0.4 \leq \frac{Y}{100} \leq 0.6\right) = \Pr(40 \leq Y \leq 60) = 0.9648.$$

Weak Law of Large Numbers

Suppose that $X_1, \ldots, X_n$ form a random sample from a distribution (i.e., $X_1, \ldots, X_n$ are i.i.d.) for which the mean is μ and the variance is finite. Let $\overline{X_n}$ denote the sample mean. Then

$$\overline{X_n} \xrightarrow{p} \mu.$$

- The WLLN has to do with the large sample behavior of the (distribution of the) sample mean.

Sequence of Random Variables

A sequence of random variables is in fact a sequence of functions $X_n : S \to \mathbb{R}$.

Example

Consider the following random experiment: A fair coin is tossed once. Here, the sample space has only two elements $S = \{H, T\}$. We define a sequence of random variables $X_1, X_2, X_3, \cdots$ on this sample space as follows:

$$X_n(s) = \begin{cases} \frac{1}{n+1} & \text{if } s = H \\ 1 & \text{if } s = T \end{cases}$$

1. Are the X_i 's independent?
2. Find the PMF and CDF of $X_n, F_{X_n}(x)$ for $n = 1, 2, 3, \cdots$.
3. As n goes to infinity, what does $F_{X_n}(x)$ look like?

Convergence of Random Variables

Reminder: Arithmetic Convergence

A sequence $a_1, a_2, a_3, \cdots$ converges to a limit L if

$$\lim_{n\to\infty} a_n = L.$$

That is, for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - L| < \epsilon, \quad \text{for all } n > N.$$

Convergence of Random Variables

Convergence in Distribution

A sequence of random variables $X_1, X_2, X_3, \cdots$ converges in distribution to a random variable X, shown by $X_n \xrightarrow{d} X$, if

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

for all x at which $F_X(x)$ is continuous.

Example

Let $X_2, X_3, X_4, \cdots$ be a sequence of random variable such that

$$F_{X_n}(x) = \begin{cases} 1 - \left(1 - \frac{1}{n}\right)^{nx} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Show that X_n converges in distribution to Exponential(1).

Convergence of Random Variables

Convergence in Probability

A sequence of random variables $X_1, X_2, X_3, \cdots$ converges in probability to a random variable X, shown by $X_n \xrightarrow{p} X$, if

$$\lim_{n\to\infty} P\left(|X_n - X| \geq \epsilon\right) = 0, \qquad \text{for all } \epsilon > 0.$$

Example

Let $X_n \sim \text{Exponential}(n)$, show that $X_n \xrightarrow{p} 0$. That is, the sequence $X_1, X_2, X_3, \cdots$ converges in probability to the zero random variable X.

Convergence of Random Variables

Convergence in Mean

Let $r \geq 1$ be a fixed number. A sequence of random variables $X_1, X_2, X_3, \ldots$ converges in the r th mean or in the L^r norm to a random variable X, shown by $X_n \xrightarrow{L^r} X$, if

$$\lim_{n \to \infty} E\left(|X_n - X|^r\right) = 0$$

If $r = 2$, it is called the mean-square convergence, and it is shown by $X_n \xrightarrow{\text{m.s.}} X$.

Example

Let $X_n \sim \text{Uniform}\left(0, \frac{1}{n}\right)$. Show that $X_n \xrightarrow{m.s.} 0$

Convergence of Random Variables

Almost Sure Convergence

A sequence of random variables $X_1, X_2, X_3, \cdots$ converges almost surely to a random variable X, shown by $X_n \xrightarrow{\text{a.s.}} X$, if

$$P\left(\lim_{n\to\infty} |X_n - X| < \epsilon\right) = 1.$$

Theorem

Consider the sequence $X_1, X_2, X_3, \cdots$. If for all $\epsilon > 0$, we have

$$\sum_{n=1}^{\infty} P\left(|X_n - X| > \epsilon\right) < \infty,$$

then $X_n \xrightarrow{\text{a.s.}} X$.

Convergence of Random Variables

Convergence in probability implies convergence in distribution
If $X_n \xrightarrow{p} X$, then $X_n \xrightarrow{d} X$.

The converse is not always true, but special case:
If $X_n \xrightarrow{d} c$, where c is a constant, then $X_n \xrightarrow{p} c$.

Convergence in mean implies convergence in probability:
If $X_n \xrightarrow{L^r} X$ for some $r \geq 1$, then $X_n \xrightarrow{p} X$.

Almost sure convergence implies convergence in probability:
If $X_n \xrightarrow{a.s} X$, then $X_n \xrightarrow{p} X$.

Convergence of Random Variables

Continuous Mapping Theorem

Let $X_1, X_2, X_3, \cdots$ be a sequence of random variables. Let also $h : \mathbb{R} \mapsto \mathbb{R}$ be a continuous function. Then, the following statements are true:

1. If $X_n \xrightarrow{d} X$, then $h(X_n) \xrightarrow{d} h(X)$.
2. If $X_n \xrightarrow{p} X$, then $h(X_n) \xrightarrow{p} h(X)$.
3. If $X_n \xrightarrow{\text{a.s.}} X$, then $h(X_n) \xrightarrow{\text{a.s.}} h(X)$.

Inequalities

Markov Inequality

Let X be a random variable such that $P(X \geq 0) = 1$. Then for every real number t,

$$P(X \geq t) \leq \frac{E(X)}{t}.$$

Chebysev Inequality

Let X be a random variable for which $Var(X)$ exists. Then for every number $t > 0$,

$$P(|X - E(X)| \geq t) \leq \frac{Var(X)}{t^2}.$$

Weak Law of Large Numbers

Theorem (Weak Law of Large Numbers)

Suppose that $X_1, \ldots, X_n$ form a random sample from a distribution (i.e., $X_1, \ldots, X_n$ are i.i.d.) for which the mean is μ and the variance is finite. Let $\overline{X_n}$ denote the sample mean. Then

$$\overline{X_n} \xrightarrow{p} \mu.$$