CE 211
Mathematics for Engineers

Lecture 9

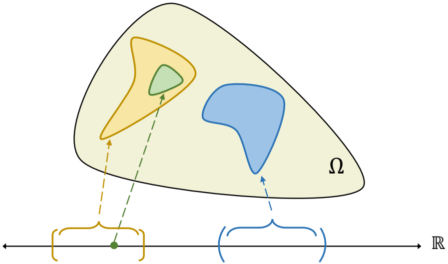
Limit Theorems

Previously on Mathematics for Engineers

A random variable is an alternate way of constructing events. Defining random variables allows us to translate events of interest into probabilities more easily.

Definition (Random Variable)

A real-valued random variable is a function or mapping $X : \Omega \to \mathbb{R}$ such that for all $S \subset \mathbb{R}$, $X^{-1}(S) \in \mathcal{F}$.



*Technically, there are some restrictions on S just like valid events, but we'll ignore those details.

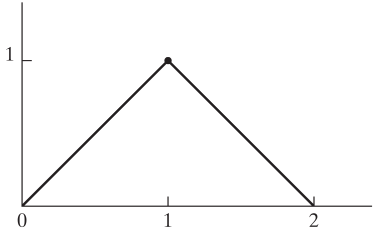
Previously on Mathematics for Engineers

The CDFs can be differentiated to get the PDFs of the sums.

$$f_Z(z) = \frac{d}{dz}\int_{-\infty}^{\infty} F_X(z-y)f_Y(y)dy$$
$$= \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy$$

Use this expression to derive the PDF of the sum of two uniforms between 0 and 1. If $z \in [0,1]$, $f_Z(z) = \int_0^z dy$ and if $z \in [1,2]$, $f_Z(z) = \int_{z-1}^{1} dy$. Thus,

$$f_Z(z) = \begin{cases} z & \text{if } 0 \leq z \leq 1 \\ 2-z & \text{if } 1 < z \leq 2 \\ 0 & \text{otherwise} \end{cases}$$



Previously on Mathematics for Engineers

Suppose X and Y are two standard normal random variables. What is the distribution of $Z = X + Y$? Using the earlier formula,

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-y)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy \end{aligned}$$

which when simplified leads to

$$\frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}(z)^2}$$

which is the PDF of a normal random variable with mean 0 and variance 2.

Previously on Mathematics for Engineers

Claim

Suppose $X_1, \ldots, X_n$ *denotes a vector of random variables such that* $|\mathbb{E}(X_i)| < \infty$. *Then,*

$$\mathbb{E}(X_1 + \ldots + X_n) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n)$$

Note that for this result we **do not need** the random variable to

- Be independent
- Have the same distribution

The variance of a sum of random variables can be written as

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j)$$

Previously on Mathematics for Engineers

Claim (Markov's Inequality)

Let X be non-negative and assume $\mathbb{E}(X)$ is finite. For any $t > 0$,

$$\mathbb{P}(X > t) \leq \frac{\mathbb{E}(X)}{t}$$

Claim (Chebyshev's Inequality)

Let X be a random variable with finite μ and σ^2. For any $t > 0$,

$$\mathbb{P}\Big(|X - \mu| \geq t\Big) \leq \frac{\sigma^2}{t^2}$$

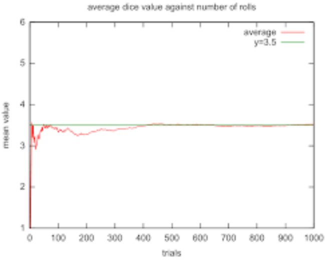
Lecture Outline

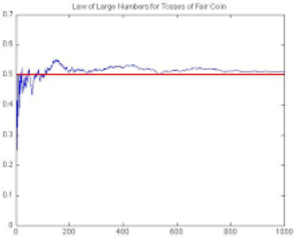
1. Measuring Convergence
2. Limit Theorems

Lecture Outline

Measuring Convergence

We have the interpreted expected value of a random variable as the average of the realizations when we perform a large number of experiments. For instance, in the case of rolling a die and tossing a coin





Note that the above pictures indicate the convergence of two different quantities. The expected value in one case and the probability of H in the other.

These type of results are also called the Law of Large Numbers (LLN) and it will be the first limit theorem that we will see.

Measuring Convergence

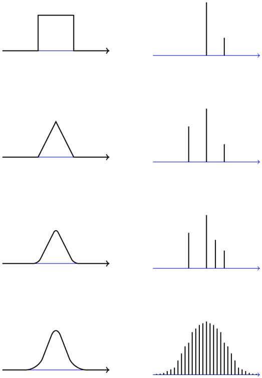
Introduction

Consider the problem of measuring the length of an object. Measurement errors are common and hence every time we use a certain technique we can get different results.

The difference between the true and the measured value indicates the errors and the sum of all errors tends to have a normal distribution.

Laplace noticed that the histograms of the sums such as binomial distributions had a certain resemblance to the normal distribution.

The figure below shows the PDF of sums of uniform and binomial random variables $Z_n = X_1 + X_2 + \ldots + X_n$.



This resemblance to the normal distribution is captured in the second limit theorem that we will see called the Central Limit Theorem (CLT).

Measuring Convergence

Sequence of Reals

A sequence of real numbers is represented as $(x_n)_{n\in\mathbb{N}}$ or $\{x_n\}_{n\in\mathbb{N}}$ or simply as (x_n) or $\{x_n\}$.

Finding the limits of a sequence for large n is a fundamental question in real analysis. For example, the sequence $x_n = 1/n$ converges to 0 as n goes to ∞.

Definition

A sequence $\{x_n\}$ is said to converge to a real number x, and is denoted as $x_n \to x$, if for every $\epsilon > 0 \, \exists \, N > 0$ such that $|x_n - x| < \epsilon$ for all $n \geq N$

We also write this as $\lim_{n\to\infty} x_n = x$.

Measuring Convergence

Sequence of Reals

We can try to extend these ideas to sequences of random variables but then probabilities, random variables, PMFs, and PDFs are all functions!

We would like to see if a sequence of random variables $X_1, X_2, \ldots$ gets closer to a random variable X as $n \to \infty$. To measure 'closeness' new notions of convergence are needed.

Measuring Convergence

Example 1

Consider the tossing of a single unbiased coin. Suppose, we define random variables $X_1, X_2, \ldots$ on this sample space $\Omega = \{H, T\}$ as follows

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

- Are the random variables $X_1, X_2, \ldots$ independent and identically distributed (iid)?
- Sketch the PMF and CDF of these random variables.
- Is the support of the sequence of the random variables same or different?

What do these random variables appear to converge to, i.e., if $X_n \to X$. What is X?

Measuring Convergence

Example 2

Instead, consider tossing a coin multiple times and let $X_1, X_2, \ldots$ be defined as follows:

$$X_n(\omega) = \begin{cases} 1 & \text{if } \omega = H \\ 0 & \text{otherwise} \end{cases}$$

- Are the random variables $X_1, X_2, \ldots$ independent and identically distributed (iid)?
- Sketch the PMF and CDF of these random variables.
- Is the support of the sequence of the random variables same or different?

What do these random variables appear to converge to, i.e., if $X_n \to X$. What is X?

Consider the tossing of multiple coins. In the nth toss, we toss a biased coin with H probability $1 - \frac{1}{n}$ and T probability of $\frac{1}{n}$. If each $X_1, X_2, \ldots$ are defined as

$$X_n(\omega) = \begin{cases} 1 & \text{if } \omega = H \\ 0 & \text{otherwise} \end{cases}$$

- Are the random variables $X_1, X_2, \ldots$ independent and identically distributed (iid)?
- Sketch the PMF and CDF of these random variables.
- Is the support of the sequence of the random variables same or different?

What do these random variables appear to converge to, i.e., if $X_n \to X$. What is X? Note that the probability space is changing in this example.

Definition (Convergence in Distribution)

A sequence of random variables $\{X_n\}$ converges to X in distribution and is denoted as $X_n \xrightarrow{d} X$ or $X_n \xrightarrow{\mathcal{D}} X$ if

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

for all x where $F_X(x)$ is continuous.

The above limit is easy to interpret as for a given x, $F_{X_1}(x), F_{X_2}(x), \ldots$ is a sequence of real numbers.

Hence, we can use the ϵ definition to say that for every x and a given $\epsilon > 0 \, \exists \, N > 0$ such that $|F_{X_n}(x) - F_X(x)| < \epsilon \, \forall \, n \geq N$.

Convergence in distribution is also called **convergence in law** or we say X_n **converges weakly** to X.

Measuring Convergence

Convergence in Distribution

Consider the Bernoulli random variable from Example 2.

$$X_n(\omega) = \begin{cases} 1 & \text{if } \omega = H \\ 0 & \text{otherwise} \end{cases}$$

It is clear that $X_n \xrightarrow{d} X$, where X is another Bernoulli random variable defined like X_n. Instead, imagine another random variable Y

$$Y(\omega) = \begin{cases} 1 & \text{if } \omega = T \\ 0 & \text{otherwise} \end{cases}$$

Does $X_n \xrightarrow{d} Y$? Convergence in distribution only requires the limiting distributions to match but the limiting random variable can look very counter intuitive. For this reason, it is a weak form of convergence.

Definition (Convergence in Probability)

A sequence of random variables $\{X_n\}$ converges to X in probability and is denoted as $X_n \xrightarrow{p} X$ if

$$\lim_{n \to \infty} \mathbb{P}(|X_n - X| \geq \epsilon) = 0$$

for all $\epsilon > 0$.

Think of the event A_n that represents the set of all outcomes for which $|X_n(\omega) - X(\omega)| \geq \epsilon$. The probability of such events must keep shrinking as $n \to \infty$.

Convergence in probability is stronger than convergence in distribution.

Let us revisit the earlier example and check for convergence using this new measure

$$X_n(\omega) = \begin{cases} 1 & \text{if } \omega = H \\ 0 & \text{otherwise} \end{cases} \qquad Y(\omega) = \begin{cases} 1 & \text{if } \omega = T \\ 0 & \text{otherwise} \end{cases}$$

This experiment has two outcomes H and T and by the definition of the above random variables

$$|X_n(H) - Y(H)| = |X_n(T) - Y(T)| = 1$$

Thus, for every $\epsilon \in (0, 1)$, the event corresponding to $|X_n - Y| \geq \epsilon$ is $\{H, T\}$. Hence, the associated probability is not 0 and therefore $X_n \nrightarrow Y$ in probability.

It turns out that convergence in probability is also not that strong. For instance, consider the problem of throwing a dart on (0,1). Suppose, we define a sequence of indicator random variables

$$\mathbf{1}_{(0,1)}, \mathbf{1}_{(0,1/2)}, \mathbf{1}_{(1/2,1)}, \mathbf{1}_{(0,1/4)}, \mathbf{1}_{(1/4,1/2)}, \mathbf{1}_{(1/2,3/4)}, \mathbf{1}_{(3/4,1)}, \ldots$$

Do you think this sequence of random variables should converge? If so, to what? Does this converge in probability?

Definition (Almost Sure Convergence)

A sequence of random variables $\{X_n\}$ converges almost surely to X and is denoted as $X_n \xrightarrow{a.s.} X$ if

$$\mathbb{P}\Big(\Big\{ \omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = X(\omega) \Big\} \Big) = 1$$

Again notice that for a given ω, the sequence $\{X_n(\omega)\}$ is just a sequence of reals and hence the limit is easy to interpret.

For example, imagine a single coin toss from which the adjacent sequence of random variables are defined

$$X_n(\omega) = \begin{cases} 1 & \text{if } \omega = H \\ (-1)^n & \text{otherwise} \end{cases}$$

What is $\mathbb{P}\Big(\{\omega \in \Omega : \lim_{n \to \infty} X_n(\omega) = 1\} \Big)$?

This type of convergence is the strongest form of convergence and is also called convergence with probability 1 or X_n converges strongly to X.

Does the sequence of indicator random variables defined earlier converge almost surely to 0? The limit of the random variable matches for most part but is different from $X(\omega)$ infinitely often (i.o.).

Measuring Convergence

Almost Sure Convergence

Almost sure convergence $\Rightarrow$ Convergence in probability $\Rightarrow$ Convergence in distribution

The other direction need not necessarily hold!

Limit Theorems

The first result is fundamental to all statistical models. It states that the random variable derived from the average of iid random variables is close to their mean.

This result comes in two forms, one involving convergence in probability and another with almost sure convergence.

The proof for the latter is more involved but the first can be proved for special cases using Chebyshev's inequality.

Claim (WLLN)

Let $\{X_n\}$ be a sequence of iid random variables with finite mean μ.

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{p} \mu$$

Proof.

For the finite variance case ($\sigma^2 < \infty$),

$$\mathbb{E}(\overline{X}_n) = \mu$$

$$V(\overline{X}_n) = \sigma^2/n$$

Using Chebyshev's inequality,

$$\mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2} \Rightarrow \lim_{n\to\infty} \mathbb{P}(|\overline{X}_n - \mu| \geq \epsilon) = 0$$

■

Claim (SLLN)

Let $\{X_n\}$ be a sequence of iid random variables with finite mean μ.

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{a.s.} \mu$$

There are some rare cases where the weak laws hold and the strong law does not.

We saw in one of the earlier classes that the sum of independent normal random variables is normal. The CLT goes a step further to state the sum of any iid random variables tends to normal distribution but the convergence is in distribution.

Claim (CLT)

Let $\{X_n\}$ be a sequence of iid random variables with expected value $\mu < \infty$ and variance $\sigma^2 < \infty$ and also suppose. $Z_n = X_1 + X_2 + \ldots + X_n$ Then,

$$\frac{Z_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)$$

Suppose 100 packages are loaded on a plane and the weight of each package is uniformly distributed between 5 and 50 kgs. What is the probability that the total weight will exceed 3000 kgs?

Note that the assumptions of CLT require finite mean and variance. Hence, you cannot apply it to a sequence of random variables that violate this assumption (the St. Petersburg paradox for instance).

Define a new sequence $\{Y_n\}$ such that $Y_n = \frac{X_n - \mu}{\sigma}$. Clearly, $\mathbb{E}(Y_n) = 0$ and $V(Y_n) = 1$.

By definition,

$$\frac{Z_n - n\mu}{\sigma\sqrt{n}} = \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sigma\sqrt{n}} = \frac{Y_1 + Y_2 + \ldots + Y_n}{\sqrt{n}}$$

Show that

$$M_{\frac{Z_n - n\mu}{\sigma\sqrt{n}}}(t) = \mathbb{E}\left(e^{t/\sqrt{n}Y_1}\right)^n = \left(1 + \frac{t}{\sqrt{n}}\mathbb{E}(Y_1) + \frac{1}{2!}\left(\frac{t}{\sqrt{n}}\right)^2 \mathbb{E}(Y_1^2) + \ldots\right)^n$$
$$\approx \left(1 + \frac{t^2/2}{n}\right)^n$$

Therefore,

$$\lim_{n \to \infty} M_{\frac{Z_n - n\mu}{\sigma\sqrt{n}}}(t) = e^{t^2/2}$$

Your Moment of Zen