

MA 6.101

Probability and Statistics

**Tejas Bodas**

Assistant Professor, IIIT Hyderabad

# Convergence of Random Variables

# Pointwise Convergence

- ▶ When do we say that  $\{x_n\}$  converges to  $x \in \mathbb{R}$  ?

We say that  $\{x_n\}$  converges to  $x \in \mathbb{R}$  (denoted by  $x_n \rightarrow x$ ) if for every  $\epsilon > 0$ , we can find an  $N(\epsilon) \in \mathbb{N}$  such that for  $|x_n - x| < \epsilon$  for  $n > N(\epsilon)$ .

- ▶ What about convergence of functions?
- ▶ When do we say that a sequence of functions  $F_n(\cdot)$  converge to  $F(\cdot)$  on the domain  $\mathbb{R}$ ?

We say that the sequence of function  $F_n(\cdot)$  converge to  $F(\cdot)$  pointwise if the sequence  $\{F_n(x)\}$  converges to  $F(x)$  ( $F_n(x) \rightarrow F(x)$ ) for all  $x \in \mathbb{R}$ .

# Uniform Convergence

We say that the sequence of function  $F_n(\cdot)$  converge to  $F(\cdot)$  pointwise if the sequence  $\{F_n(x)\}$  converges to  $F(x)$  ( $F_n(x) \rightarrow F(x)$ ) for all  $x \in \mathbb{R}$ .

- ▶ For every  $x$ , the sequence  $\{F_n(x)\}$  converges to  $F(x)$ .
- ▶ For every  $\epsilon$ , there exists  $N(\epsilon, x)$  which can depend on  $x$ .
- ▶ Only those  $F_n(x)$  are  $\epsilon$  close to  $F(x)$  for which  $n > N(\epsilon, x)$ .

If  $N(\epsilon, x) = N(\epsilon)$  (i.e., independent of  $x$ ) for every  $x \in \mathbb{R}$ , then such convergence of  $F_n(\cdot)$  to  $F(\cdot)$  is called as uniform convergence.

# Convergence of Sequence of random variables

- ▶ We will now be interested in the convergence properties of an infinite sequence of random variables  $\{X_n\}$  to some limiting random variable  $X$ .
- ▶ What does the convergence  $X_n \rightarrow X$  even mean ?
- ▶ When you perform the random experiment once, you get a sequence of realizations  $\{x_n\}$  and  $x$ .
- ▶ If you are 'lucky', maybe  $x_n \rightarrow x$ .
- ▶ But if you were to perform the experiment again, you may not be so 'lucky' and get a different sequence  $\{x'_n\}$  which may not converge to  $x'$ .
- ▶ We will come up with notions of convergence that depend on how often you see the sequence of realizations converging.

# Convergence of Sequence of random variables

- ▶ Convergence of  $X_n \rightarrow X$
- ▶ Here  $X$  could even be a deterministic number.
- ▶  $X'_n$ 's could be dependent on each other.
- ▶ Each random variable  $X_n$  could have a different law (pmf/pdf).

# Modes of Convergence ( $X_n \rightarrow X$ )

Pointwise or Sure convergence

$\{X_n, n \geq 0\}$  converges to  $X$  pointwise or surely if for all  $\omega \in \Omega$  we have  $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$

► Consider  $\Omega = \{H, T\}$ .

► Further,  $X_n = \begin{cases} \frac{1}{n} & \text{if } \omega = H \\ 1 + \frac{1}{n} & \text{if } \omega = T. \end{cases}$  and  $X = \begin{cases} 0 & \text{if } \omega = H \\ 1 & \text{if } \omega = T. \end{cases}$

# Almost sure convergence

$X_n$  converges to  $X$  almost surely if

$$P(\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)) = 1.$$

- ▶ The set of outcomes where the convergence does not happen has measure 0.  $P\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) \neq X(\omega)\} = 0.$
- ▶ Consider  $\Omega = [0, 1]$  where you pick a number uniformly in  $[0, 1]$ . Let  $X_n(\omega) = \omega^n$  for all  $\omega \in \Omega$  and  $X(\omega) = 0$  for all  $\omega$ .
- ▶  $X_n(\omega) \rightarrow X(\omega)$  for  $\omega \in [0, 1)$ .
- ▶  $X_n(\omega) \nrightarrow X(\omega)$  for  $\omega = 1$  and  $\mathbb{P}\{\omega = 1\}.$
- ▶ This is almost sure convergence as  $\mathbb{P}\{[0, 1)\} = 1.$



# Almost sure (a.s.) convergence

$X_n$  converges to  $X$  almost surely if

$$P(\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)) = 1.$$

- ▶ Example 2: Strong law of large numbers (SLLN).

Let  $\{X_n, n \geq 0\}$  denote a sequence of i.i.d random variables with mean  $\mu$  and denote  $S_n = \sum_{i=1}^n X_i$ . Then  $\frac{S_n}{n} \rightarrow \mu$  a.s.

- ▶ Toss a biased coin (probability of head is  $\mu$ ) repeatedly. What is  $\omega$  and  $\Omega$ ?
- ▶ Let  $X_i$  denote the outcome of the  $i^{\text{th}}$  toss and  $S_n$  denotes the number of heads in  $n$  tosses.
- ▶ The empirical mean is given by  $\frac{S_n}{n}$ .

## Detour: Incremental formula for sample mean

- ▶ Now that we know  $\frac{S_n}{n} \rightarrow \mu$  we can use  $\hat{\mu}_n := \frac{S_n}{n}$  as an 'estimator' for the mean especially in cases when the underlying distribution is not known.
- ▶ Note that the estimator  $\hat{\mu}_n$  is a random variable. What is its cdf? what is its mean & Variance?
- ▶  $\hat{\mu}_n = \frac{S_n}{n}$  is an 'unbiased estimator' since  $E[\hat{\mu}_n] = \mu$ .
- ▶  $Var(\hat{\mu}_n) = \frac{\sigma^2}{n}$
- ▶ We will soon see CLT that will tell the CDF of  $\hat{\mu}_n$  without any information on the law of  $X_i$ .

## Detour: Incremental formula for sample mean

- ▶ Now given  $\hat{\mu}_n$ , suppose you see an additional sample  $X_{n+1}$ .
- ▶ How will you compute  $\hat{\mu}_{n+1}$ ?
- ▶ Naive way :  $\hat{\mu}_{n+1} = \frac{\sum_{i=1}^{n+1} X_i}{n+1}$ .
- ▶ There is an incremental formula that uses  $\hat{\mu}_n$ .

$$\hat{\mu}_{n+1} = \hat{\mu}_n + \frac{1}{n+1} [X_{n+1} - \hat{\mu}_n]$$

- ▶ Such averaging formulas are used extensively in Reinforcement learning.