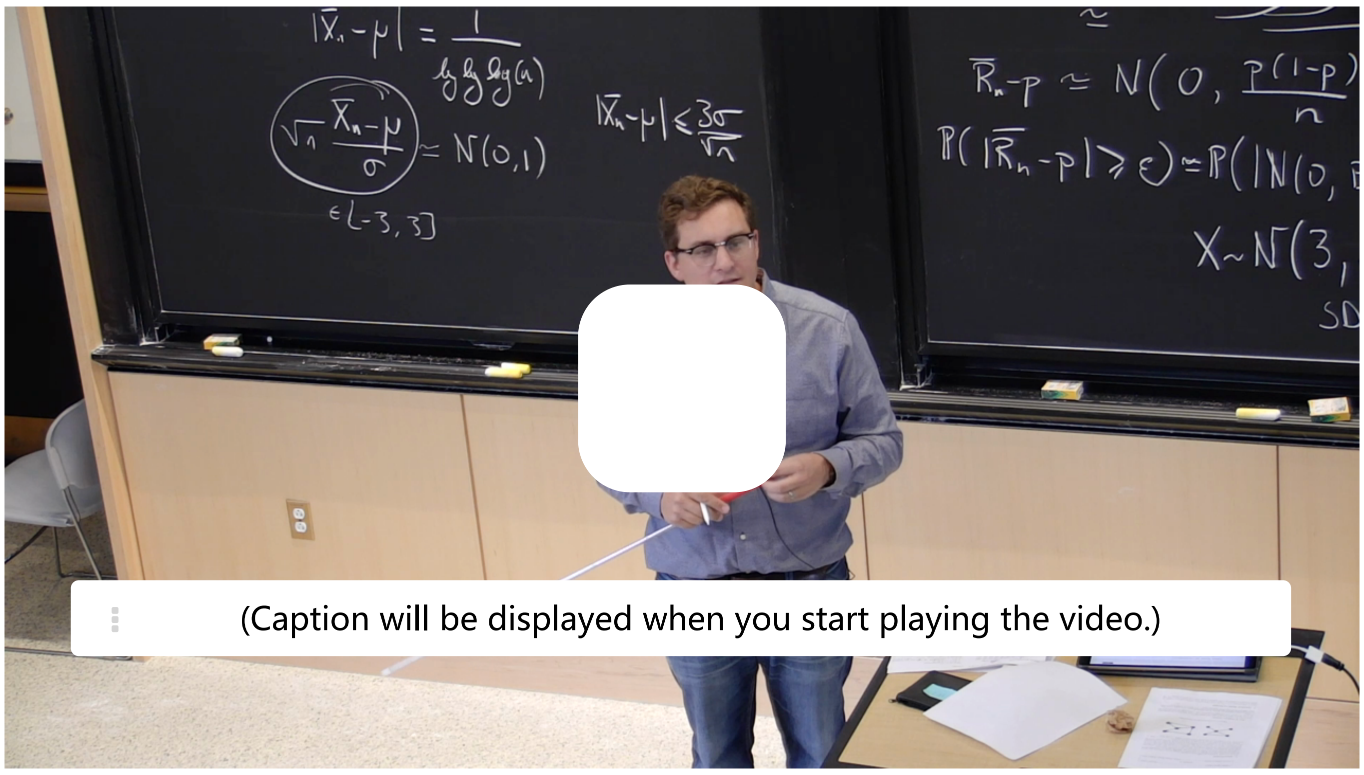
7. Modes of Convergence

Convergence almost surely, in probability, and in distribution



Start of transcript. Skip to the end.

So I mentioned three types of convergence.

This section is not particularly important for this class.

All you have to remember is that what

you want to do kind of works.

But I want to make sure that we're not making any mistake.

And since I'm trying to give you a little bit

of mathematical formalism, I want

to make sure that we understand what we're

视频

下载视频文件

字幕

下载 SubRip (.srt) file

下载 Text (.txt) file

Note: We did not study modes of convergence in great detail in *6.431x Probability–the Science of Uncertainty and Data*. This is one of the most theoretical topic in this course. Basic understanding of their differences will be expected, but in the rest of this course, convergence will mostly be discussed in the context of the laws of large numbers and central limit theorem. **You will not be tested directly on modes of convergence in any exam**.

Equivalent definition of convergence in distribution for real r.v.s

Convergence in distribution is also known as **convergence in law** and **weak convergence** .

For a sequence $(T_n)_{n\geq 1}$ of random variables that take values in $\mathbb{R}$, the definition of convergence in distribution given in lecture is equivalent to the definition we have learned in the course *6.431x: Probability–the Science of Uncertainty and Data*. That is, the following two notions are equivalent:

1. For all continuous and bounded function f,

$$T_n \xrightarrow[n\to\infty]{(d)} T \quad \text{iff} \quad \mathbb{E}\left[f\left(T_n\right)\right] \xrightarrow[n\to\infty]{} \mathbb{E}\left[f\left(T\right)\right]$$

2. For all $x \in \mathbb{R}$ at which the cdf of T is continuous,

$$T_n \xrightarrow[n\to\infty]{(d)} T \quad \text{iff} \quad \mathbf{P}\left[T_n \leq x\right] \xrightarrow[n\to\infty]{} \mathbf{P}\left[T \leq x\right]$$

The first definition generalizes to random variables in higher dimensional spaces. This will be used in the multivariate central limit theorem later.

Hide

Convergence in probability and in distribution 1

0.5/2.0 points (graded)
Let $(T_n)_{n\geq 1} = T_1, T_2, \ldots$ be a sequence of r.v.s such that

$$T_n \sim \text{Unif}\left(5 - \frac{1}{2n}, 5 + \frac{1}{2n}\right).$$

Given an arbitrary fixed number $0 < \delta < 1,$ find the smallest number N (in terms of δ) such that $\mathbf{P}\left(\{|T_n - 5| > \delta\}\right) = 0$ whenever $n > N$.

$N =$ 2 ✖ **Answer:** 1/(2*delta)

2

Does $(T_n)_{n\geq 1}$ converge in probability to a constant? If so, what is the limiting value? Enter **DNE** if $(\{X_n\})$ does not converge in probability.

$(T_n)_{n\geq 1} \xrightarrow{\mathbf{P}}$ DNE ✖ **Answer:** 5

DNE

Does $(T_n)_{n\geq 1}$ converge in distribution?

- (selected) Yes ✔
- No

Let $F_n(t)$ be the cdf of T_n and $F(t)$ be the cdf of the constant limit. For which values of t does $\lim_{n\to\infty} F_n(t) = F(t)$? (Choose all that apply.)

- [ ] $t < 5$ ✔
- [x] $t = 5$
- [ ] $t > 5$ ✔

✖

STANDARD NOTATION

Solution:

- Given a fixed $0 < \delta < 1$, since $T_n \sim \text{Unif}\left(5 - \frac{1}{2n}, 5 + \frac{1}{2n}\right)$, we know that $\mathbf{P}\left(|T_n - 5| > \delta\right) = 0$ whenever $\frac{1}{2n} < \delta$, or equivalently, for all $n > \frac{1}{2\delta}$.
- By the definition of convergence in probability, $T_n \xrightarrow[n\to\infty]{\mathbf{P}} 5.$
- Since convergence in probability implies convergence in distribution, $T_n \xrightarrow[n\to\infty]{\text{d.}} 5.$
- The cdf F_n of T_n is a piecewise linear function with value 0 for all $t \leq 5 - \frac{1}{2n}$, 1 for all $t \geq 5 + \frac{1}{2n}$, and a line connecting the points $(t, F_n(t)) = \left(5 - \frac{1}{2n}, 0\right)$ and $(t, F_n(t)) = \left(5 + \frac{1}{2n}, 1\right)$ in the interval $5 - \frac{1}{2n} \leq t \leq 5 + \frac{1}{2n}$. In particular, $F_n(5) = \frac{1}{2}$ for all n.

On the other hand, the cdf F of the contant 5 is $F(t) = 0$ when $t < 5$, and $F(t) = 1$ when $t \geq 5$. Therefore, $F_n(t) \xrightarrow[n\to\infty]{} F(t)$ for all $t \neq 5$.

Remark: We have just verified that T_n indeed converges in distribution to the deterministic limit 5, that is, $F_n(t) \xrightarrow[n\to\infty]{} F(t)$ for all t where $F(t)$ is continuous.

提交 你已经尝试了3次（总共可以尝试3次）

Answers are displayed within the problem

Convergence in probability and in distribution 2

3/4 points (graded)

Let $(Y_n)_{n\geq 1}$ be a sequence of i.i.d. random variables with $Y_n \sim \text{Unif}(0,1)$.

Let

$$M_n = \max(Y_1, Y_2, \ldots, Y_n).$$

For any fixed number $0 < \delta < 1$, find $\mathbf{P}(|M_n - 1| > \delta\})$. (Type **delta** for δ.)

$\mathbf{P}(|M_n - 1| > \delta\}) =$ delta

δ

✖ **Answer:** (1-delta)^n

Does the sequence $(M_n)_{n\geq 1}$ converge in probability to a constant? If yes, enter the value of the constant limit; if no, enter **DNE**.

$(M_n)_{n\geq 1} \xrightarrow{\mathbf{P}}$ 1

1

✔ **Answer:** 1

Find the CDF $F_{M_n}(x)$ for $0 \leq x \leq 1$.

$F_{M_n}(x) = P(M_n \leq x) =$ x^n

x^n

✔ **Answer:** x^n

Does $(M_n)_{n\geq 1}$ converge in distribution?

- (●) Yes ✔
- ( ) No

STANDARD NOTATION

Solution:

Note that M_n is always at most one, so $|M_n - 1| \geq \delta$ can be replaced with $1 - M_n \geq \delta$.

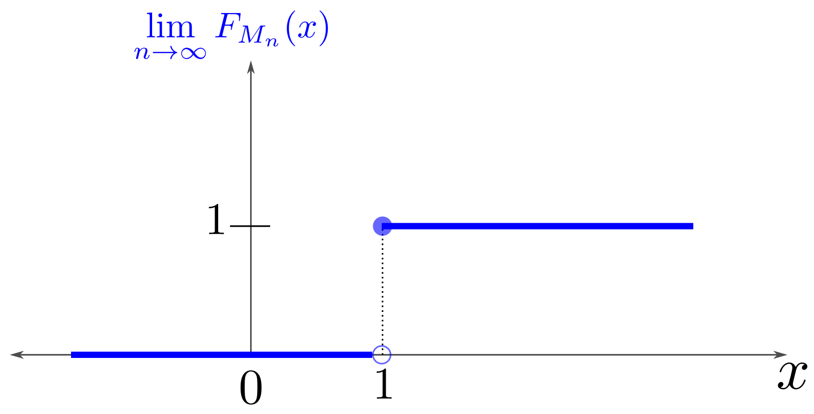
$$\begin{aligned}
\mathbf{P}(|M_n - 1| \geq \delta) &= \mathbf{P}(1 - M_n \geq \delta) \\
&= \mathbf{P}(M_n \leq 1 - \delta) \\
&= \mathbf{P}(Y_1 \leq 1-\delta)\,\mathbf{P}(Y_2 \leq 1-\delta)\cdots\mathbf{P}(Y_n \leq 1-\delta) \qquad \text{since } Y_i \text{ independent} \\
&= (1-\delta)^n \xrightarrow[n\to\infty]{} 0 \qquad \text{since } 0 < (1-\delta) < 1.
\end{aligned}$$

Hence, the sequence $(M_n)_{n\geq 1}$ converges in probability to the deterministic limit $M = 1$. This implies that is also converges in distribution to the same limit.

The CDF is computed in the same way:

$$F_{M_n}(x) = \mathbf{P}(M_n \le x) = \mathbf{P}(Y_1 \le x) \cdots \mathbf{P}(Y_n \le x) = x^n \qquad \text{for } 0 \le x \le 1.$$

We already know that $(M_n)_{n\ge1}$ converges in distribution to M; here we check directly through definition. As $n \to \infty$, $F_{M_n}(x)$ approaches the step function shown below



which coincides with the CDF of the limit M we found above; hence, indeed $M_n \xrightarrow[n\to\infty]{(d.)} M$.

Remark: In general, for a sequence $(T_n)_{n\ge1}$, if $\mathbb{E}[T_n] \xrightarrow[n\to\infty]{} \mu$ and Var$(T_n) \xrightarrow[n\to\infty]{} 0$, then $T_n \xrightarrow[n\to\infty]{\mathbf{P}} \mu$. Both this problem and the previous one satisfy these conditions.

提交 你已经尝试了3次（总共可以尝试3次）

Answers are displayed within the problem

Expectations and convergence in probability

3/3 points (graded)

Let $(T_n)_{n\ge1}$ be a sequence of r.v.s such that for each n, T_n takes only two possible values 0 and 2^n with the following probabilities:

$$\mathbf{P}(T_n = 0) = 1 - \frac{1}{n}$$
$$\mathbf{P}(T_n = 2^n) = \frac{1}{n}.$$

Does the sequence $(T_n)_{n\ge1}$ converge in probability to a constant? If so, enter the limiting value; if not, enter **DNE**.

$T_n \xrightarrow{\mathbf{P}}$ 0 ✔ Answer: 0

0

Compute $\mathbb{E}[T_n]$ in terms of n.

$\mathbb{E}[T_n] =$ 2^n/n ✔ Answer: (2^n)/n

$\frac{2^n}{n}$

Does the sequence of expectations $\mathbb{E}[T_n]$ converge? If so, enter the limiting value; if not, enter **DNE**.

$\lim_{n\to\infty} \mathbb{E}[T_n] =$ DNE ✔ Answer: DNE

DNE

STANDARD NOTATION

Solution:

For any $\epsilon > 0$,

$$\mathbf{P}(|X_n - 0| > \epsilon) = \frac{1}{n} \longrightarrow 0.$$

THerefore, $(X_i)_{n\geq 1}$ converges in probability to the deterministic limit 0.
However,

$$\mathbb{E}[X_n] \;=\; \frac{2^n}{n} \longrightarrow \infty.$$

Hence, the sequence $(\mathbb{E}[X_n])_{n\geq 1}$ does not converge.
Remark: Convergence in probability does not imply convergence of expectation values.

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

Convergence almost surely (a.s) is also known as **convergence with probability 1 (w.p.1)** and **strong convergence** . We will not discuss this type of convergence much beyond this lecture.

Probability review: the (Strong) Law of Large Numbers

1/1 point (graded)
A digital signal receiver decodes bits of incoming signal as 0s or 1 and makes an error in decoding a bit with probability 10^{-4}.

Assuming decoding success is independent for different bits, as the receiver receives more and more signals, what is the fraction of erroneously decoded bits?

Fraction of errors: 10^-4 ✔ **Answer:** 10^(-4)

Solution:

The transmission of each bit of data can be modelled as independent Bernoulli random variable X_i with expectation of error $\mathbb{E}[X_i] = 10^{-4}$.
The (strong/weak) law of large number states that

$$\frac{\sum_{i=1}^n X_i}{n} \xrightarrow[n\to\infty]{(\text{a.s.}/\mathbf{P})} \mathbb{E}[X].$$

In particular, the strong law says that **with probability 1**, the fraction of errors approach 10^{-4}.

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

(Optional theoretical material) Distinguishing different types of convergences

In the following examples, the explicit definition of a random variable as a function on a probability space to $\mathbb{R}$, rather than just its distribution, will be needed to establish the type of convergence.

Convergence in distribution but NOT in probability

Let $X_1, X_2, \ldots, X_i, \ldots$ be a sequence of random variables For i odd, $X_i \sim f_{\text{odd}}(x)$ where $f_{\text{odd}}(x) = x$ in $[0,1]$.
Let $X_2 \sim f_2(x)$ where $f_{\text{even}} = 1 - x$ in $[0,1]$.

Then for all i, $X_i \sim$ Unif $(0,1)$. Since the distribution of all X_i is the same, the sequence converges in distribution to Unif $(0,1)$.

However, $\{X_i\}$ does not converge in probability. There is no random variable X (i.e. no function from $[0,1]$ to $\mathbb{R}$) such that X_i. $\mathbf{P}(|X_i - X| > \epsilon)$ the probability (or the "size" of the set in the probability space) that

Convergence in probability but NOT almost surely

As discussed in the lecture, a sequence $X_n \sim \text{Ber}(1/n)$ converges in probability to 0. However, depending on how the random variables are defined as functions on the underlying probability space, (and different random variables can have the same distribution), the sequence can converge almost surely or not.

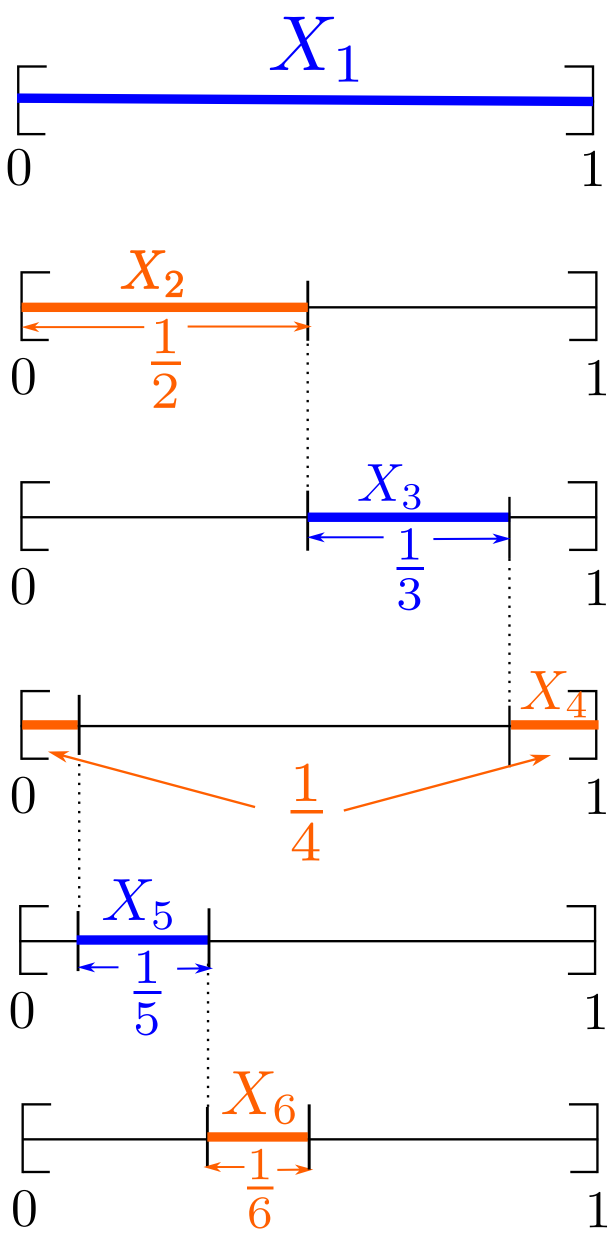
Examples:

1. $\{X_n\}$ converges almost surely: Define $X_n : [0,1] \longrightarrow \mathbb{R}$ by

$$X_n(\omega) = \begin{cases} 1 & \text{if } \omega \in [0, 1/n] \\ 0 & \text{otherwise.} \end{cases}$$

   Then for each $w \in (0,1]$, $X_n(\omega) \longrightarrow 0$; hence $\mathbf{P}(\{\omega : X_n(\omega) \longrightarrow 0\}) = 1$, i.e. $X_n \xrightarrow{a.s.} 0$.

2. $\{X_n\}$ does NOT converge almost surely: As above, each random variable X_n is a function $X_n : [0,1] \longrightarrow \mathbb{R}$. In this example, for each X_n, we specify a subinterval of $[0,1]$ of length $1/n$ where X_n takes value 1 and outside which X_n takes value 0 by the figure below:



   In the figure above,

$$\begin{aligned} X_1(\omega) &= 1 \quad \text{for all } \omega \in [0,1]; \\ X_2(\omega) &= 1 \quad \text{for all } \omega \in [0,1/2]; \\ X_3(\omega) &= 1 \quad \text{for all } \omega \in [1/2, 1/2+1/3] \\ X_4(\omega) &= 1 \quad \text{for all } \omega \in [1/2+1/3, 1] \cup [0, 1/4-(1-(1/2+1/3))] \end{aligned}$$

   and so on. The subinterval(s) $\{\omega : X_n(\omega) = 1\}$ is of total length $\frac{1}{n}$, lies immediately to the right of the subinterval $\{\omega : X_{n-1}(\omega) = 1\}$, is truncated at $\omega = 1$ with the “rest" of the length $\frac{1}{n}$ interval "cycled" back to the right of $\omega = 0$.

Because $\sum_{n=1}^{\infty} \frac{1}{n} \to \infty$ but the interval $[0,1]$ has finite length, this "cycling process" will continue, and each number in $[0,1]$ will lie in a subinterval $\{\omega : X_n(\omega) = 1\}$ for infinitely many n's. Hence, $\{\omega : X_n(\omega) \longrightarrow 0\} = \emptyset$, and consequently $\mathbf{P}(\{\omega : X_n(\omega) \longrightarrow 0\}) = 0$.

Hide

讨论

显示讨论

主题： Unit 1 Introduction to statistics:Lecture 2: Probability Redux / 7. Modes of Convergence

认证证书是什么?