Exercises due Aug 3, 2021 19:59 EDT

Concept Checks: Empirical CDF

3/3 points (graded)

Let $X_1, \ldots, X_n \overset{iid}{\sim} X$, with (true) cdf $F(t)$, and let $F_n(t)$ be the empirical cdf of $X_1, \ldots, X_n$.

What is the domain of F_n? That is, what are all the values of t for which F_n is defined.

- (●) $0 \le t \le 1$
- ( ) $-\infty \le t \le \infty$

For any t (in the domain of F_n), the empirical cdf $F_n(t)$ is

- ( ) random
- (●) deterministic

For any t (in the domain of F_n), the true cdf $F(t)$ is

- (●) random
- ( ) deterministic

Submit You have used 1 of 1 attempt

Pointwise and Uniform Convergence of Functions

A sequence of functions $g_n(x)$ **converges pointwise** to a function $g(x)$ if for each x,

$$\lim_{n\to\infty} g_n(x) = g(x).$$

Example: In the region $x > 1$, $g_n(x) = \frac{1}{x^n}$ converges **pointwise** to $g(x) = 0$. For any fixed $x > 1$, $\dfrac{1}{x^n} \xrightarrow[n\to\infty]{} 0$.

A sequence of functions $g_n(x)$ **converges uniformly** to a function $g(x)$ if $\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |g_n(x) - g(x)| = 0.$ That is, for every $M > 0$, there exists an n_M such that $\sup_x |g_n(x) - g(x)| < M$ for all $n \geq n_M$.

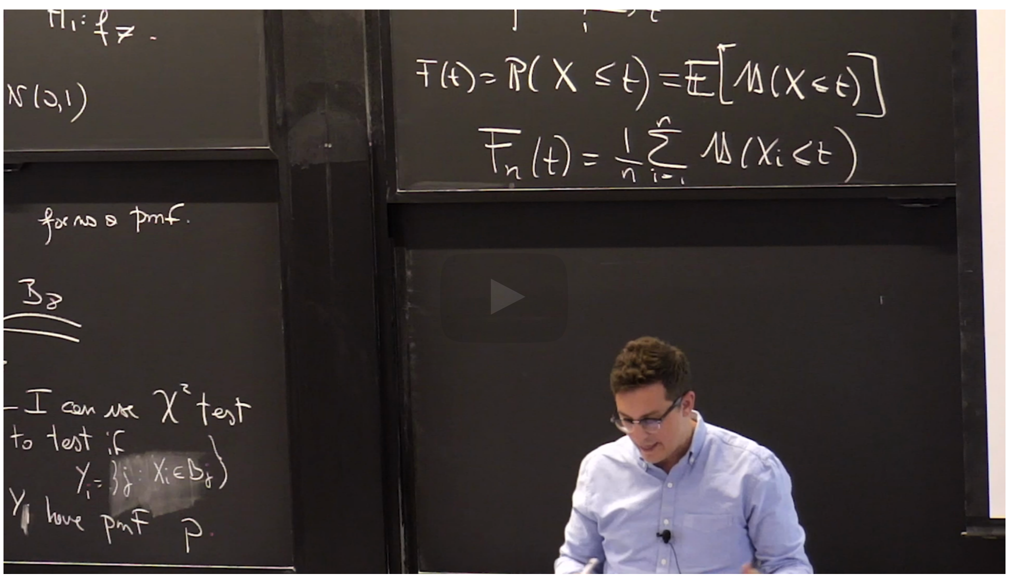
Example: In the region $x > \color{blue}{2}$, $g_n(x) = \frac{1}{x^n}$ converges **uniformly** to $g(x) = 0,$ since

$$\sup_{x>2} g_n(x) = \sup_{x>2} \frac{1}{x^n} = \frac{1}{2^n} \xrightarrow[n\to\infty]{} 0.$$

Example of pointwise but not uniform convergence:

The sequence of functions $g_n(x) = \frac{1}{x^n}$ does **not** converge uniformly to $g(x) = 0$ in the region $x > 1,$ since $\sup_{x>1} g_n(x) = \sup_{x>1} \dfrac{1}{x^n} = 1,$ which does not converge to 0 as $n \to \infty$.

Consistency of Empirical CDF, Uniform versus Pointwise Convergence, Fundamental Theorem of Statistics



0:00 / 0:00 1.50x

Start of transcript. Skip to the end.

OK.

So now I have this thing.

Hopefully it's going to be close,

and I want to know how close they are.

OK, I would like to have a way of measuring

how the step function and the true function I'm actually

trying to test are close, because I

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Consistency of the Empirical cdf

2 points possible (graded)

Let $X_1, \ldots, X_n \overset{iid}{\sim} X$ be i.i.d. random variables with cdf $F(t)$.

Recall the empirical cdf is the random function

$$\begin{aligned} F_n : \mathbb{R} &\to [0,1] \\ t &\mapsto \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \leq t). \end{aligned}$$

Then following convergence holds almost surely:

$$F_n(0) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \leq 0) \xrightarrow[n\to\infty]{a.s.} L$$

for some value L. What is L?
(Choose all that apply.)

- [ ] 0
- [ ] 1
- [ ] $F(0)$
- [ ] $F(1)$
- [ ] $\mathbb{E}[\mathbf{1}(X \leq 0)]$

What result is invoked to obtain the value of L?

central limit theorem

(strong) law of large numbers

Slutsky's theroem

Submit You have used 0 of 3 attempts

Let $X_1, \ldots, X_n \overset{iid}{\sim} X$ be i.i.d. random variables with cdf $F(t)$ and empirical cdf $F_n(t)$.

The **Glivenko-Cantelli theorem** , also known as the **Fundamental Theorem of Statistics** , states that

$$\sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \xrightarrow[n \to \infty]{a.s.} 0.$$

This is a stronger result than the one in the problem above in that the convergence happens **uniformly** over t. This means for all large enough n and for any $\delta > 0$, the difference $|F_n(t) - F(t)|$ is bounded above by δ for all t. Almost sure convergence means that for all $\delta > 0$ and $\epsilon > 0$, there exists $N = N(\delta, \epsilon)$ such that the event $\sup_t |F_n(t) - F(t)| < \delta$ occurs with probability at least $1 - \epsilon$ for all $n > N$. In other words, with probability approaching 1, the function F_n is a close L_∞ (the sup-norm) approximation of F.

Discussion

Hide Discussion

Topic: Unit 4 Hypothesis testing:Lecture 16: Goodness of Fit Tests Continued: Kolmogorov-Smirnov test, Kolmogorov-Lilliefors test, Quantile-Quantile Plots / 4. Consistency of the Empirical CDF

Add a Post

Show all posts by recent activity

There are no posts in this topic yet.

✖