36710 - 36752

ADVANCED PROBABILITY OVERVIEW

FALL 2020

LECTURE 14 : MON , OCT 19 , 2020

LAST TIME :

Def (CONVERGENCE IN PROB.) LET $\{X_n\}_{n=1}^{\infty}$ BE A SEQUENCE OF R.V.'S DEFINED ON THE SAME PROB. SPACE $(\Omega, \mathcal{F}, P)$. LET X BE ANOTHER RV ON THE SAME SPACE. THEN WE SAY THAT $\{X_n\}_n$ CONVERGES IN PROBABILITY TO X, $X_n \xrightarrow{P} X$, WHEN

$$\forall \varepsilon > 0, \quad P\left(\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}\right) \to 0 \quad \text{AS } n \to \infty$$

Remark IF X_n'S AND X TAKES VALUES ON SOME METRIC SPACE $(\mathcal{X}, d)$, THEN THE ABOVE DEFINITION CAN BE EXTENDED TO

METRIC

$$\forall \varepsilon > 0, \quad \lim_{n \to \infty} P\left(\{\omega : d(X_n(\omega), X(\omega)) > \varepsilon\}\right) = 0$$

- CONVERGENCE IN PROBABILITY TELLS US SOMETHING ABOUT THE JOINT DISTR. OF X_n AND X FOR EACH n. SEE EXAMPLE FROM LAST TIME

Remark : EQUALITY IN DISTRIBUTION. TWO R.V.'S X AND Y ARE EQUAL IN DISTRIBUTION, $X \stackrel{d}{=} Y$, WHEN THEY HAVE THE

SAME DISTRIBUTION. EXAMPLE $\Omega = [0,1]$, $\mathcal{F}$ BOREL σ-FIELD ON $[0,1]$

$P = \lambda$ ON $[0,1]$

$$X(\omega) = 1_{[0,1/2)}(\omega)$$

$$Y(\omega) = 1_{[0,1/4)\cup[3/4,1)}(\omega)$$

$$X \overset{d}{=} Y \sim \text{Bernoulli}(1/2)$$

BUT, IN GENERAL, $X(\omega) \neq Y(\omega)$

- IF $\{X_n\}_n$ IS A SEQUENCE OF RANDOM VECTORS IN $\mathbb{R}^d$ AND X IS A RANDOM VECTOR IN $\mathbb{R}^d$ THEN $X_n \xrightarrow{P} X$ AS $n \to \infty$ WHEN

$$\lim_{n\to\infty} \Pr\left(\|X_n - X\| > \varepsilon\right) = 0 \qquad \forall \varepsilon > 0$$

ANY NORM ON $\mathbb{R}^d$

<u>Claim</u> $X_n \xrightarrow{P} X$ IIF $X_{n(i)} \xrightarrow{P} X_{(i)}$ ALL $i = 1, \dots, d$

i-th COORDINATE OF X_n

<u>op (LITTLE OH-P) NOTATION</u>: RECALL THAT, FOR SEQUENCES $\{x_n\}$ AND $\{y_n\}$ OF NUMBERS, $x_n = o(y_n)$ MEANS THAT $\forall \varepsilon > 0$

LITTLE-OH

$\exists\, n_0(\varepsilon)$ S.T. $\forall n > n_0$

$$\left|\frac{x_n}{y_n}\right| < \varepsilon.$$

IN PARTICULAR $x_n = o(1) \iff x_n \to 0$ AS $n \to \infty$

LET $\{X_n\}$ BE A SEQUENCE OF RANDOM VARIABLES AND $\{r_n\}$ A SEQUENCE OF POSITIVE NUMBERS. THEN $X_n = o_p(r_n)$ MEANS THAT $\forall \varepsilon > 0$ $\exists\, n_0(\varepsilon)$ S.T. FOR $n > n_0$

$$\Pr\left(\frac{|X_n|}{r_n} \geq \varepsilon\right) < \varepsilon \iff \frac{X_n}{r_n} \xrightarrow{P} 0$$

WEAK FORM OF

Thm (WLLN) LET $\{X_n\}$ BE A SEQUENCE OF S.T. $\mathbb{E}[X_n] = \mu$

$V[X_n] = \sigma_n^2$

ALL n

AND $\mathrm{Cov}(X_n, X_{n'}) = 0$ ALL $n \neq n'$.

ASSUME THAT $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} = 0$. THEN

$$\frac{1}{n} \sum_{i=1}^{n} X_i := \frac{S_n}{n} \xrightarrow{P} \mu .$$

Pf/

$$P\left(\left| \frac{S_n}{n} - \mu \right| > \varepsilon \right) \leq \frac{1}{\varepsilon^2} \mathrm{Var}\left[\frac{S_n}{n} \right] = \frac{1}{\varepsilon^2} \sum_{i=1}^{n} \frac{\sigma_i^2}{n^2} \to 0$$

USING o_p - NOTATION

$$\frac{S_n}{n} = \mu + \underbrace{o_p(1)}$$

RANDOM FLUCTUATION THAT VANISHES IN PROB.

Remark ALGEBRA OF o_p TERMS:

LET $X_n = o_p(1)$. THEN $c \cdot X_n = o_p(1)$ ALL $c \in \mathbb{R}$

$Y_n = o_p(1)$. THEN $X_n + Y_n = o_p(1)$

WE CAN ALSO WRITE $X_n = o_p(Y_n)$ IF $\frac{X_n}{Y_n} = o_p(1)$

TAYLOR EXPANSION OF A FUNCTION AT A RANDOM POINT:

LET f BE A REAL VALUED FUNCTION WITH K DERIVATIVES AT θ_0

AND SUPPOSE $X_n \xrightarrow{P} \theta_0$; THEN $\exists\, Y_n$ S.T. $Y_n = o_p(1)$ AND

$$f(X_n) = f(\theta_0) + (X_n - \theta_0) f'(\theta_0) + \cdots + \underbrace{\frac{(X_n - \theta_0)^k}{k!} \left(f^{(k)}(\theta_0) - Y_n \right)}_{\frac{(X_n - \theta_0)^k}{k!} f^{(k)}(\theta_0) + o_p(|X_n - \theta_0|^k)}$$

$$o_p(|X_n - \theta_0|^k) = \underbrace{o_p}_{o_p(1)}$$

ASIDE: IN HIGH-DIM STATISTICS LITERATURE, CONSISTENCY IS ESTABLISHED OFTEN VIA FINITE SAMPLE BOUNDS. THIS DOES NOT REQUIRE THAT OUR RV'S ARE DEFINED ON A COMMON PROB. SPACE.

RECALL THAT AN ESTIMATOR IS CONSISTENT IF

$$\hat{\theta}_n \xrightarrow{P} \theta_0$$

$\frac{S_n}{n}$ IS A CONSISTENT ESTIMATOR OF μ IN WLLN

EXAMPLE $\quad y_i = \langle \theta_0, x_i \rangle + \varepsilon_i \qquad i=1 \ldots n$

$x_i \in \mathbb{R}^d$

WANT TO ESTIMATE θ_0 USING OLS $\hat{\theta}_n$

$\varepsilon_i \sim (0, \sigma^2)$ INDEPENDENT.

ONE CAN SHOW THAT, WITH PROB $\geq 1 - \frac{1}{n}$

$$\|\hat{\theta}_n - \theta_0\| \leq C \sigma \sqrt{\frac{d + \log n}{n}} \qquad \text{FOR EACH } n.$$

WEAK LAW OF LARGE NUMBERS.

Thm LET $X, X_1, X_2, \ldots$ BE A SEQUENCE OF INDEPENDENT AND IDENTICALLY DISTRIBUTED RV'S WITH $\mathbb{E}[X] = \mu$. THEN

$S_n = \sum_{i=1}^{n} X_i \quad \leftarrow \quad \frac{S_n}{n} \xrightarrow{P} \mu$

Pf/ IF $\mathbb{E}[X^2]$ EXISTS, THEN THE RESULT FOLLOWS FROM CHEBYSHEV'S INEQ.

OTHERWISE, WE USE TRUNCATION. LET $t > 0$ AND DEFINE

$$X_{tk} = X_k \mathbb{1}_{|X_k| \leq t}$$

$$Y_{tk} = X_k \mathbb{1}_{|X_k| > t}$$

$\Rightarrow X_k = X_{tk} + Y_{tk}$ ALL k.

SO

$$\frac{S_n}{n} = \frac{1}{n} \sum_{k=1}^{n} X_{tk} + \frac{1}{n} \sum_{k=1}^{n} Y_{tk} := U_{tn} + V_{tn}$$

NEXT $$\mathbb{E}\left[|V_{tn}|\right] \le \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}|Y_{tk}| = \mathbb{E}\left[|X| \mathbb{1}_{|X|\ge t}\right]$$

BY DCT $\mathbb{E}\left[|X| \mathbb{1}_{|X|\ge t}\right] \to 0$ AS $t \to \infty$

LET $\varepsilon \in (0,1)$. FOR ANY $\delta \in (0,1)$ LET $t = t(\varepsilon,\delta)$ BE

LARGE ENOUGH SO THAT

$$\mathbb{E}\left[|X| \mathbb{1}_{|X|\ge t}\right] = \mathbb{E}\left[|Y_{t1}|\right] \le \varepsilon\delta/6 \to \text{NUMBER SIX}$$

LET $\mu_t = \mathbb{E}[X_{t,1}]$. THEN

$$\triangle \quad |\mu_t - \mu| \le \mathbb{E}\left[|Y_{t1}|\right] \le \varepsilon\delta/6 < \varepsilon/3 \qquad (\delta<1)$$

AT THIS POINT WRITE:

$$\left|\frac{S_n}{n} - \mu\right| \le |U_{tn} - \mu_t| + |V_{tn}| + |\mu_t - \mu|$$

(WE ADDED AND SUBTRACTED μ_t AND USED TRIANGLE INEQ.).

WE KNOW THAT $|\mu_t - \mu| < \varepsilon/3$. WE NEED TO

SHOW THAT WITH PROB AT LEAST $1-\delta$,

$$(*) \quad |U_{tn} - \mu_t| < \varepsilon/3 \quad \text{AND} \quad |V_{tn}| < \varepsilon/3$$

THIS WILL IMPLY THAT, WITH PROB $\ge 1-\delta$,

$$\left|\frac{S_n}{n} - \mu\right| < \varepsilon,$$

AND THE RESULT WILL FOLLOW BECAUSE ε, δ ARE ARBITRARILY SMALL.

TO SHOW $(*)$, LET $B_n = \{|U_{tn} - \mu_t| \ge \varepsilon/3\}$

$$C_n = \{|V_{tn}| \ge \varepsilon/3\}$$

WE CAN USE THE WEAK FORM OF WLLN TO SHOW THAT

$$\Pr(B_n) = \Pr\left(|U_{tn} - \mu_t| \ge \varepsilon/3\right) < \delta/2$$

FOR ALL n LARGER THAN SOME $n_0 = n_0(\varepsilon, t, \delta)$

NOW, USING MARKOV'S INEQUALITY,

$$\Pr(C_n) = \Pr\left(|V_{t_n}| \geq \varepsilon/3\right) \leq \frac{\mathbb{E}\left[|V_{t_n}|\right]}{\varepsilon/3}$$

$$= \frac{3\,\mathbb{E}\left[|Y_{t_2}|\right]}{\varepsilon} < \delta/2$$

BY $\triangle$.

SO, ON THE EVENT $(B_n \cup C_n)^c$,

$$\left|\frac{S_n}{n} - \mu\right| < \varepsilon$$

BUT
$$\Pr\left((B_n \cup C_n)^c\right) = 1 - \Pr(B_n \cup C_n)$$

$$\leq 1 - \left(\frac{\delta}{2} + \frac{\delta}{2}\right) = 1 - \delta$$

UNION BOUND $\left[\Pr(B_n \cup C_n) \leq \Pr(B_n) + \Pr(C_n)\right]$

OTHER MODES OF STOCHASTIC CONVERGENCE

TO BE CLEAR: WE ARE DEALING WITH RANDOM VARIABLES DEFINED ON $(\Omega, \mathcal{F}, P)$

Def (CONVERGENCE WITH PROB ONE OR ALMOST SURE CONVERGENCE OR ALMOST EVERYWHERE CONVERGENCE)

$\{X_n\}$ AND X DEFINED ON $(\Omega, \mathcal{F}, P)$. $X_n \xrightarrow{a.s.} X$

WHEN
$$P\left(\left\{\omega : \lim_{n \to \infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

↳ LIMIT IS NOT UNIFORM! IN Ω

EQUIVALENTLY, $\forall \varepsilon > 0$ $\Pr\left(|X_n - X| > \varepsilon \text{ i.o.}\right) = 0$

NOTICE THIS IS THE SAME AS $P\left(\limsup_n A_n\right) = 0$

$$A_n = \{ |X_n - X| > \varepsilon \}$$

<u>Def</u> (L^p CONVERGENCE)

$$\| X_n - X \|_p \to 0 \quad \text{AS } n \to \infty$$

$$\| X \|_p = \left(\mathbb{E}[|X|^p] \right)^{1/p} \qquad p \geq 1$$

TYPICALLY, IN STATISTICS WE RELY ON L^2 CONVERGENCE.

<u>EXAMPLE</u> LET $\hat{\theta}_n$ BE AN ESTIMATOR OF θ_0

$$\hat{\theta}_n \xrightarrow{L^2} \theta_0 \quad \text{IFF} \quad \underbrace{\mathbb{E}[\hat{\theta}_n] \to \theta_0}_{\text{bias}(\hat{\theta}_n) \to 0} \text{ AND } V[\hat{\theta}_n] \to 0$$

<u>Def (CONVERGENCE IN MEASURE)</u> $(\Omega, \mathcal{F}, \mu)$ MEASURE SPACE.

$\{f_n\}$ AND f MEASURABLE FUNCTIONS ON Ω. $f_n \xrightarrow{\mu} f$ WHEN

$$\mu\left(\{ \omega : | f_n(\omega) - f(\omega) | > \varepsilon \} \right) \to 0 \quad \text{AS } n \to \infty.$$

GENERAL MEASURE (POSSIBLY INFINITY)