Lecture 13 - More Convergence

Almost Sure:

$$P(\{s \mid X_n(s) \to X(s)\}) = 1$$

In Probability

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0$$

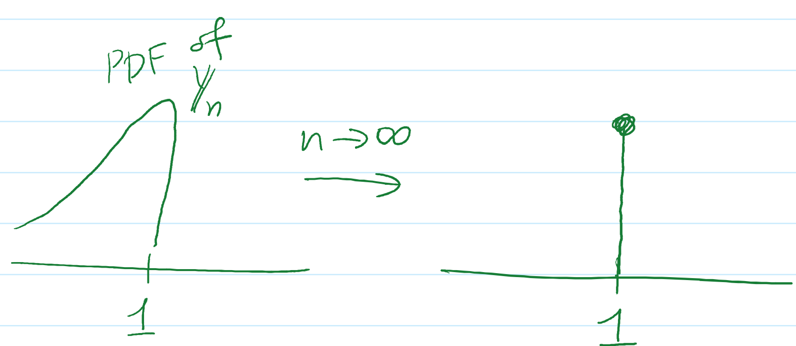
In Distribution:

$$F_n \longrightarrow F$$

CDF of X_n CDF of X

Ex. $X_i \overset{iid}{\sim} \mathcal{U}(0,1)$

$$Y_n = \max_{i=1,\dots,n} X_i$$



$$Z_n = n(1 - Y_n)$$

distributional limit?

$$F_n(z) = \mathbb{P}(Z_n \leq z)$$

$$= \mathbb{P}(n(1 - Y_n) \leq z)$$

maximum

$$= \mathbb{P}(Y_n \geq 1 - z/n)$$

$$= P(Y_n \geq 1 - z/n)$$

$$= 1 - P(Y_n \leq 1 - z/n)$$

maxi...

$$= 1 - P(X_1 \leq 1 - z/n, X_2 \leq 1 - z/n, \dots, X_n \leq 1 - z/n)$$

$$= 1 - \prod_{i=1}^{n} P(X_i \leq 1 - z/n)$$

$X_i \sim U(0,1)$

CDF

$$= 1 - (1 - z/n)^n = F_n(z)$$

<u>Recall</u>: $\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = e$

$$\lim_{n \to \infty} \left(1 + \frac{c}{n}\right)^n = e^c$$

$$\lim_{n \to \infty} F_n(z) = \lim_{n \to \infty} 1 - (1 - z/n)^n$$

$$= 1 - \lim_{n \to \infty} (1 - z/n)^n$$

$$F(z) = 1 - e^{-z}$$ CDF of Exp(1)

So $Z_n \xrightarrow{d} \text{Exp}(1)$

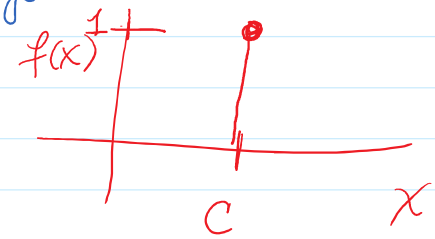
<u>Theorem</u>: a.s. $\Rightarrow$ p $\Rightarrow$ d

<u>Partial Converse</u>:

... is a constant

Partial Converse:

If $X_n \xrightarrow{d} c$ — c is a constant i.e. a degenerate RV



then $X_n \xrightarrow{P} c$.

For a seq of numbers:

$x_n, y_n \in \mathbb{R}$, and $x_n \to x$, $y_n \to y$

then

$\to x_n + y_n \to x + y$

$\to x_n y_n \to xy$

$\to x_n / y_n \to x/y$

$\to a x_n + b y_n \to ax + by$

Theorem:

Let $X_n \to X$, $Y_n \to Y$ and $a, b \in \mathbb{R}$

and the convergence is either a.s. or p

(NOT d).

Then

① $aX_n + bY_n \to aX + bY$

② $X_n Y_n \to XY$ (or $X_n / Y_n \to X/Y$)

Note that:

We can treat a seq $c_n \in \mathbb{R}$ as a seq of degenerate RVs

So if $c_n \to c$ (as numbers)

then $c_n \xrightarrow{a.s.} c$ (as RVs)

Punchline: $c_n \to c$ and $X_n \to X$ (a.s, or i.p)

then

(1) $aX_n + bc_n \to aX + bc$

(2) $c_n X_n \to cX$.

What about convergence in dist?

Theorem: Slutsky's Theorem

If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$ constant

then

(1) $X_n + Y_n \xrightarrow{d} X + c$

(2) $X_n Y_n \xrightarrow{d} Xc$

$\left(X_n / Y_n \xrightarrow{d} X/c\right)$

Theorem: Continuous Mapping Theorem

If $g: \mathbb{R} \to \mathbb{R}$ is a continuous function

If $g: \mathbb{R} \to \mathbb{R}$ is a continuous function

and $X_n \to X$ (all types of convergence)

then $g(X_n) \to g(X)$

Pf.

$$\lim_{x \to a} g(x) = g\left(\lim_{x \to a} x\right)$$

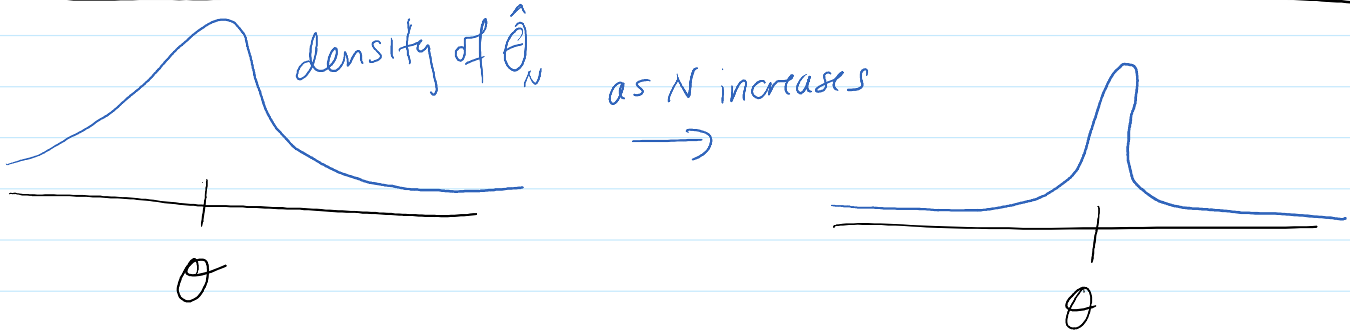
Defn : Consistent Estimator

depends on N samples

We say an estimator $\hat{\theta}_N$ is consistent for θ

if $\hat{\theta}_N \xrightarrow{P} \theta$.

a constant



Another way: consistency $\approx$ asymptotically unbiased (*)

(*) For nice enough dist's

Ex.

$$S^2 = \frac{1}{N-1} \sum_{n=1}^{N} (X_n - \overline{X}_N)^2, \quad \mathbb{E}S^2 = \sigma^2$$

$N-1 \quad n=1$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N}(X_n - \overline{X}_N)^2, \quad \mathbb{E}\,\hat{\sigma}^2 = \frac{N-1}{N}\sigma^2$$

<u>notice</u>: $\mathbb{E}\,\hat{\sigma}^2 \xrightarrow{n} \sigma^2$

<u>Theorem: MSE $\to 0$ then $\hat{\theta}$ is consistent</u>

If $\mathrm{MSE}(\hat{\theta}_N) \xrightarrow{N} 0$ then $\hat{\theta}_N \xrightarrow{P} \theta$.

<u>pf.</u>

Want to show:

$$\lim_{n\to\infty} \underbrace{P(|\hat{\theta}_N - \theta| \geq \varepsilon)} = 0$$

<u>Markov's Ineq</u>: $P(X \geq a) \leq \frac{\mathbb{E}X}{a} \qquad (X \geq 0)$

$$P(|\hat{\theta}_N - \theta| \geq \varepsilon) = \mathbb{P}(\overbrace{(\hat{\theta}_N - \theta)^2}^{X \geq 0} \geq \varepsilon^2) \quad \text{by Markov's}$$

$$\leq \frac{\mathbb{E}[(\hat{\theta}_N - \theta)^2]}{\varepsilon^2}$$

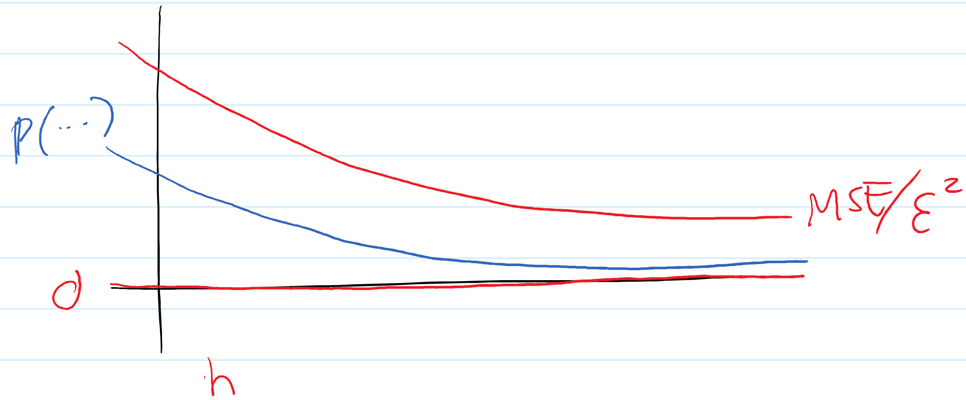
$$= \frac{\mathrm{MSE}(\hat{\theta}_N)}{\varepsilon^2}$$

So

$\mathrm{MSE}(\hat{\theta}$

$$0 \le \mathbb{P}(|\hat{\theta}_N - \theta| \ge \varepsilon) \le \frac{MSE(\hat{\theta}_N)}{\varepsilon^2}$$

If $MSE(\hat{\theta}_N) \to 0$ then so does my $\mathbb{P}(|\hat{\theta}_N - \theta| \ge \varepsilon) \to 0$



<u>Intuition:</u>

$\overline{X}_N$ this should be a "good" estimator of μ.

<u>Saw:</u> $\mathbb{E}\overline{X}_N = \mu$ and $Var\, \overline{X}_N = \sigma^2 / N$

<u>Theorem: Weak Law of Large Numbers</u> (WLLN)

If X_n s are uncorrelated and

1. $\mathbb{E}X_n = \mu$
2. $Var(X_n) = \sigma^2 < \infty$

(2) $Var(X_n) = 0 < \infty$

then $\bar{X}_N \xrightarrow{P} \mu$.

Weak: weak assumptions and converg. in prob.

Pf. $\boxed{\text{Chebyshev's} \quad \mu = \mathbb{E}Y \quad P(|Y-\mu| \geq \varepsilon) \leq \frac{Var(Y)}{\varepsilon^2}}$

Want to show:

$$P(|\bar{X}_N - \mu| \geq \varepsilon) \xrightarrow{N} 0$$

Let $Y = \bar{X}_N$ then Chebyshev's says

$$0 \leq \underbrace{P(|\bar{X}_N - \mu| \geq \varepsilon)}_{} \leq \frac{Var(\bar{X}_N)}{\varepsilon^2} = \frac{\sigma^2}{N\varepsilon^2}$$

$$\to 0$$

as $n \to \infty$.

Squeeze theorem

Ex. $X_n \overset{iid}{\sim} Pois(\lambda)$

$$\mathbb{E}X_n = \lambda \text{ and } Var(X_n) = \lambda < \infty$$

WLLN: $\bar{X}_N \xrightarrow{P} \lambda$.

WLLN: $\bar{X}_N \xrightarrow{P} \lambda$.

Can slightly generalize WLLN

Assume $Var(X_n) = \sigma_n^2$ but

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \sigma_n^2 < \infty$$

chebyshev's

then

$$\mathbb{P}(|\bar{X}_N - \mu| \geq \varepsilon) \leq \frac{\boxed{Var(\bar{X}_N)}}{\varepsilon^2}$$

$$= \frac{\boxed{\frac{1}{N^2} \sum_{n=1}^{N} \sigma_n^2}}{\varepsilon^2}$$

$< \infty$

$$= \frac{1}{N} \left(\frac{\frac{1}{N} \sum_n \sigma_n^2}{\varepsilon^2} \right) \longrightarrow 0$$

as $n \to \infty$

So $\bar{X}_N \xrightarrow{P} \mu$.