Lecture 12: Convergence

$$F_n(y) = \mathbb{P}(Y_n \le y)$$

$$= \vdots$$

$$= P(X_i \le y)^n$$

$$= F_{X_i}(y)^n$$

$$= \begin{cases} 0^n, & y \le 0 \\ y^n, & 0 \le y \le 1 \\ 1^n, & y \ge 1 \end{cases}$$

$$= \begin{cases} 0 & y \le 0 \\ y^n & 0 \le y \le 1 \\ 1 & y \ge 1 \end{cases}$$

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

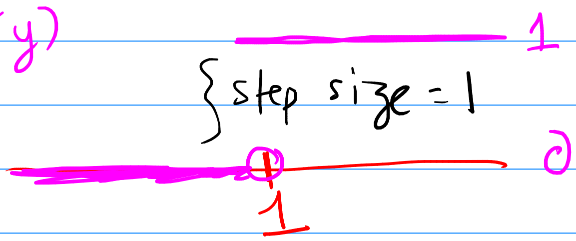
$Y_n = \max\limits_{i=1,\dots,n} X_n$

$$F_{X_i}(y) = \begin{cases} 0 & y \le 0 \\ y & 0 \le y \le 1 \\ 1 & y \ge 1 \end{cases}$$

as $n \to \infty$

$$\to \begin{cases} 0 & y \le 0 \\ 0 & 0 \le y \le 1 \\ 1 & y \ge 1 \end{cases}$$

$F(y)$



$\uparrow$ CDF of limit 1

So $Y_n \xrightarrow{d} 1$.

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

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

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

Let's show $Z_n \xrightarrow{d} Z$.

$$F_n(z) = \mathbb{P}(Z_n \le z)$$
$$= \mathbb{P}(n(1-Y_n) \le z)$$
$$= \mathbb{P}(Y_n \geq 1 - z/n)$$
$$= 1 - \mathbb{P}(Y_n \le 1 - z/n)$$
$$= 1 - \mathbb{P}(X_1 \le 1 - z/n, \dots, X_n \le 1 - z/n)$$
$$= 1 - \mathbb{P}(X_i \le 1 - z/n)^n$$

$$= \begin{cases} \cancel{0, \quad 1 - z/n \le 0} \\ 1 - z/n, \quad 0 \le 1 - z/n \le 1 \\ \cancel{1, \quad 1 - z/n \geq 1} \end{cases}$$

$$F_n(x) = 1 - (1 - \overset{c}{z}/n)^n$$

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

$\longrightarrow$??? as $n \to \infty$?

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

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

For seqs of numbers have algebraic rules

e.g. $x_n \to x$ and $y_n \to y$

1. $x_n + y_n \to x + y$
2. $x_n y_n \to xy$
3. $x_n / y_n \to x/y$
4. $a x_n + b y_n \to ax + by$

Theorem: Algebraic Props

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

and the convergence $\to$ is i.p. or a.s.
(not in dist)

then
1. $aX_n + bY_n \to aX + bY$
2. $X_n Y_n \to XY$ $\quad (X_n / Y_n \to X/Y)$

If consts are degenerate RVs then if
$C_n \to C$ (as numbers)

then $C_n \xrightarrow{a.s.} C$ (as RVs).

So if $C_n \to C$ (as numbers) and

$X_n \to X$ (i.p. or a.s.) then

1) $aX_n + C_n \to aX + C$

2) $C_n X_n \to CX$.

What about convergence in dist?

Theorem: Slutsky's Theorem

RVs

a const in $\mathbb{R}$

If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} C$ then

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

2) $X_n Y_n \xrightarrow{d} cX$

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

Theorem: Partial Converse

If $X_n \xrightarrow{d} C$ ← a const.

then $X_n \xrightarrow{p} C$.

Theorem: Continuous Mapping Theorem

any type of convergence

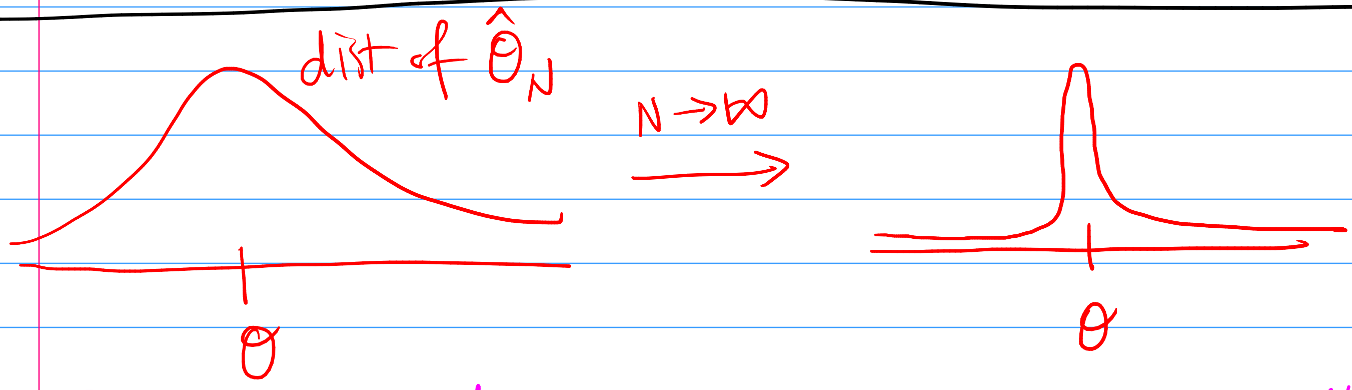
If $g: \mathbb{R} \to \mathbb{R}$ is continuous and $X_n \to X$

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

Defn: Consistent Estimator

We say an estimator $\hat{\theta}_N$ is consistent for θ if $\hat{\theta}_N \xrightarrow{P} \theta$.

est. for sample size N



(*) essentially, consistency $\approx$ asymptotically unbiased

---

Ex. $S^2 = \frac{1}{N-1} \sum_{n=1}^{N} (X_n - \bar{X})^2$

Saw $\mathbb{E} S^2 = \sigma^2$ (unbiased)

also had

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{n=1}^{N} (X_n - \bar{X})^2$$

$$\mathbb{E}\hat{\sigma}^2 = \frac{N-1}{N} \sigma^2 < \sigma^2$$

as $N \to \infty$ then $\mathbb{E}\hat{\sigma}^2 \to \sigma^2$ (asymptotically unbiased)

Theorem: $\text{MSE}(\hat{\theta}_N) \to 0$ as $N \to \infty$ then $\hat{\theta}_N$ is consistent for θ.

Pf. $\forall \varepsilon > 0 \quad \mathbb{P}(|\hat{\theta}_N - \theta| \geq \varepsilon) \to 0$ as $N \to \infty$.

$$0 \leq \mathbb{P}(|\hat{\theta}_N - \theta| \geq \varepsilon) = \mathbb{P}\Big(\underbrace{(\hat{\theta}_N - \theta)^2}_{\geq 0} \geq \underbrace{\varepsilon^2}_{a}\Big)$$

Markov's: $X \geq 0$
$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}X}{a}$

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

$$= \text{MSE}(\hat{\theta}_N)/\varepsilon^2 \to 0$$

as $N \to 0$

Squeeze theorem

Ex. $S^2 = \frac{1}{N-1} \sum_n (X_n - \bar{X})^2 \qquad X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$$\text{MSE}(S^2) = \cancelto{0}{\text{Bias}(S^2)}^2 + \text{Var}(S^2)$$

prev. showed that $\text{Var}(S^2) = \frac{2\sigma^4}{N-1}$

So as $N \to \infty$

$$MSE(S^2) = Var(S^2) = \frac{2\sigma^4}{N-1} \to 0 \text{ as } N \to \infty$$

So S^2 is consistant for σ^2. $(S^2 \xrightarrow{P} \sigma^2)$

What about $\hat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N}(X_n - \bar{X})^2$?

Notice that $\hat{\sigma}^2 = \frac{N-1}{N} S^2 = c_n S^2$

where $c_n = \frac{N-1}{N}$

and $c_n \to 1$

So by my algebraic properties,

$$\hat{\sigma}^2 = c_n S^2 \xrightarrow{P} 1 \cdot \sigma^2 = \sigma^2$$

So $\hat{\sigma}^2$ is consistent for σ^2.

Intuition: that $\bar{X}_N$ should be a good est. of $\mu = \mathbb{E}X_n$.

Theorem: Weak Law of Large Numbers (WLLN)

If X_n are uncorrelated and

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

then $\overline{X}_N = \frac{1}{N}\sum_{n=1}^{N} X_n$ we have that

$$\overline{X}_N \xrightarrow{P} \mu .$$

weak

Pf. $MSE(\bar{x}) = Bias(\bar{x})^2 + Var(\bar{x})$

$E\overline{X} = \mu$, $Var(\bar{x}) = \sigma^2/N$

$= 0^2 + \sigma^2/N$

$= \sigma^2/N \to 0$ as $N \to \infty$.

So by prev. theorem $\overline{X} \xrightarrow{P} \mu$.

Ex. $X_n \overset{iid}{\sim} Pois(\lambda)$, $EX_n = \lambda = Var(X_n)$

then WLLN says that

$$\overline{X}_N \xrightarrow{P} \lambda .$$

Ex. $g(x) = \frac{1}{1+x^2}$ ← a continuous function,

so $g(\bar{x}_N) \xrightarrow{P} g(\lambda) = \frac{1}{1+\lambda^2}$