CHAP 7

* Note on Convergence Concept

We introduce five convergence Modes involving sequences of random variables.

- Def: A random sequence is a sequence of random variables $X_1(s), X_2(s), \cdots$

  outcome

For a given outcome s_0, $X_1(s_0), X_2(s_0), \cdots$ is a deterministic sequence which may or may not converge.

- Convergence Everywhere (e)

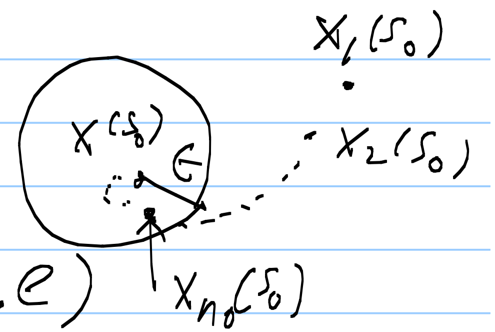
A random sequence is said to converge everywhere to $X(s)$ if, given $\epsilon > 0$, we can find a number n_0, such that

$$|X_n(s) - X(s)| < \epsilon, \ \forall s \quad \left(\begin{array}{l} \lim_{n\to\infty} X_n(s) = X(s) \\ \text{if } s \text{ is given} \end{array} \right)$$

for every $n > n_0$. We use

$$\lim_{n\to\infty} X_n(s) = X(s)$$

to denote this mode.



- Convergence Almost Everywhere (a.e)

If the event $\{s \mid \lim_{n\to\infty} X_n(s) = X(s)\}$ exist and has probability 1, then we say that the sequence converges almost everywhere to $X(s)$ (or with probability one). We denote the mode by $P\left[\lim_{n\to\infty} X_n(s) = X(s)\right] = 1$.

- Convergence in the Mean Square (MS) Sense

The sequence $X_1(s), X_2(s), \cdots$ converges to $X(s)$ in the MS sense if

$$\lim_{n\to\infty} E\left[\left|X_n(s) - X(s)\right|^2\right] = 0.$$

We denote the mode by

$$\underset{n\to\infty}{\mathrm{l.i.m.}}\, X_n(s) = X(s)$$

where l.i.m. means "limit in the MS sense."

- Convergence in Probability (p)

If $\lim_{n\to\infty} P\left[|X_n(s) - X(s)| > \epsilon \right] = 0$

for any $\epsilon > 0$, then the random sequence $X_1(s), X_2(s), \cdots$ converges to $X(s)$ in probability.

- Convergence in Distribution (d)

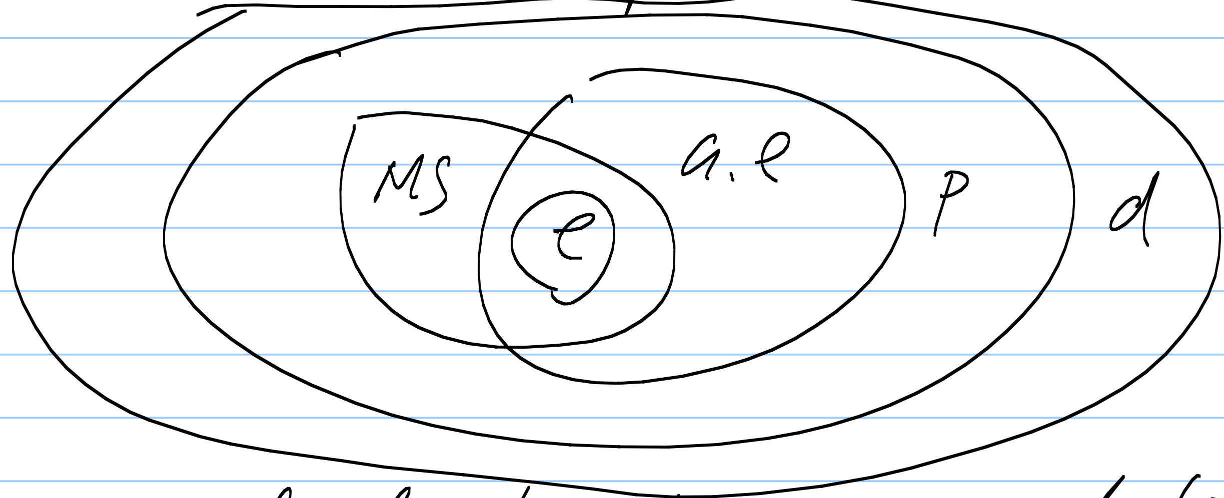
If $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$

for every point x of continuity of $F_X(x)$, then the random sequence $X_1(s), X_2(s), \cdots$

Converges to $X(s)$ in distribution.

- Note: The relationship between various convergence modes is depicted by

$$e \Rightarrow a.e \Rightarrow p \Rightarrow d$$
$$\Rightarrow MS \Rightarrow p \Rightarrow d$$



Notes: (i) Central Limit Theorem is defined in the distribution sense.

(2) Laws of Large Number are defined in the probability sense (weak) or in probability-one sense (strong).

Parameter : mean, moments

Consider an event $\underline{A}$. We define the indicator random variable

$$X_A = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise.} \end{cases}$$

X_A is a Bernoulli rv with success probability $P[A]$ and $E[X_A] = P[A]$.

Consider the iid random sequence $X_1, X_2, \dots$
$X_{A1}, X_{A2}, \dots$

The sample mean

$$\underbrace{M_n(X)} = \frac{\sum_{i=1}^{n} X_{A_i}^{\xi}}{n} = \hat{P}_n(A)$$

The sequence $M_1(X), M_2(X), \ldots$ may converge in certain sense.

Define $\hat{R}_n$ as a function of $X_1, X_2, \ldots, X_n$

$\hat{R}_n$ is used to serve as a point estimate for certain parameter r of the probability model for X.

Consider $r = \mathrm{Var}[X]$.

If $E[X] = 0$, $\mathrm{Var}[X] = E[X^2]$. We can define $Y_n = X_n^2$. Then,

$M_n(X)$

$$M_n(Y) \triangleq \frac{1}{n} \sum_{i=1}^{n} Y_n = \frac{1}{n} \sum_{i=1}^{n} X_n^2$$

converges in probability to $E[Y] = \mathrm{Var}[X]$ (according to the weak law of large number).

If $E[X] = \mu_X$, then we can form $\underline{Y_n = (X_n - \mu_X)^2}$ and have that $M_n(Y)$ converges to $\mathrm{Var}[X]$ in probability.

$$M_n(Y) = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n} (X_i - E[X])^2$$

Consider $X_1, X_2, \cdots, X_n$ which are iid with known probability density form, however with an unknown parameter γ.
We form $Y_n \triangleq g(X_1, X_2, \cdots, X_n)$ to provide an estimate of the unknown parameter γ.
There are four measures used to characterize the estimation performance.

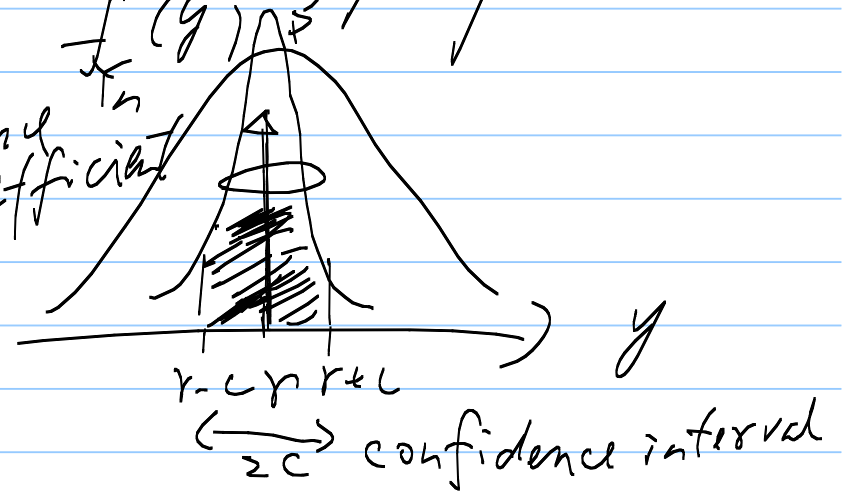
① $E[Y_n]$

② $E[(Y_n - \gamma)^2]$

③ $P[\gamma - c < Y_n < \gamma + c] \triangleq 1-\alpha$

$2c$ point estimate

confidence coefficient



(4) $P[\underbrace{Y_n}-c < \gamma < Y_n + c] \geq 1-\alpha$

Confidence Interval $= (Y_n - c, Y_n + c)$ is random

$P[\gamma \in (Y_n - c, Y_n + c)] \geq 1-\alpha$

The length of confidence interval is $2c$, which is deterministic!

Ex: Consider the sample mean $M_n(X)$ as an estimate of $E[X]$. From Tchebyshev's inequality,

$$P[\underbrace{|M_n(X) - E[X]|} \geq c] \leq \frac{Var[M_n(X)]}{c^2} = \frac{Var[X]}{nc^2}$$

In terms of "interval estimate" formulation,

$\rightsquigarrow_x$ $P\left[M_n(X) - c \le E[X] \le M_n(X) + c \right] \le \frac{Var[X]}{nc^2}$

where

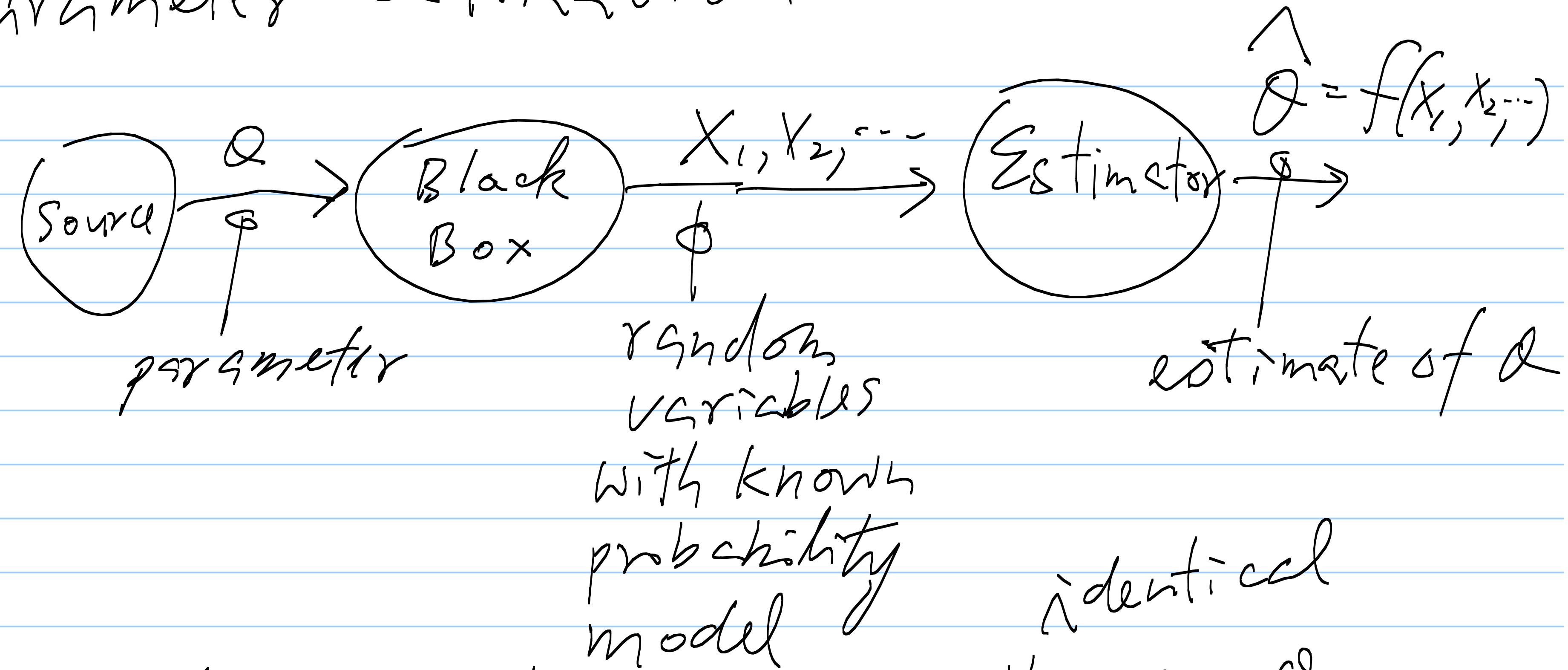
1) $[M_n(X) - c, M_n(X) + c]$ is the random confidence interval with fixed length $2c$

2) $\frac{Var[X]}{nc^2}$ is the confidence coefficient

since

$$P\left[E[X] \in [M_n(X) - c, M_n(X) + c] \right] \ge \underbrace{1 - \frac{Var[X]}{nc^2}}_{1-\alpha}$$

Parameter Estimation



Ex: X_i's are iid Gaussian with identical mean θ and identical variance 1. A good estimate is the sample mean, i.e., $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Theorem 7.1: $M_n(X)$ has

$$E[M_n(X)] = E[X], \quad Var[M_n(X)] = \frac{Var[X]}{n}$$