9.0 Random Sequences and Convergence Concepts

9.1 Random Sequences

Often we are presented with a sequence of random variables. For example, in statistics $X_1, X_2, \cdots, X_n$ might be an i.i.d. sample from some distribution and characteristics of the underlying distribution (such as its mean or variance) is estimated using the samples.

Definition: A random sequence is a collection of random variables $X_n(u) = X(u,n)$ defined on some probability space $(\mathcal{U}, \mathcal{F}, \mathcal{P})$ with $n \in \mathbb{Z}^+ = 1, 2, \cdots$ or $n \in \mathbb{Z} = \cdots, -2, -1, 0, 1, 2, \cdots$

Any finite number of random variables in the random sequence is described statistically by the joint distribution functions of these random variables.

Definition: The mean function of $\mathbf{X}_n(u)$ is

$$\mu_{X_n}(u) = E[\mathbf{X}_n(u)]$$

and the correlation function is

$$\mathbf{R}_X(n_1, n_2) = E[\mathbf{X}_{n_1}(u)\mathbf{X}^*_{n_2}(u)]$$

and the covariance function is

$$\begin{aligned}\mathbf{K}_X(n_1, n_2) &= E[(\mathbf{X}_{n_1}(u) - \mu_X(n_1))(\mathbf{X}^*_{n_2}(u) - \mu^*_X(n_2))] \\ &= \mathbf{R}_X(n_1, n_2) - \mu_X(n_1)\mu^*_X(n_2)\end{aligned}$$

Definition: A function $f(n,m)$ is said to be Hermitian symmetric if

$$f(n,m) = f^*(m,n)$$

Definition: A Hermitian symmetric function $f(n,m)$ is said to be non-negative definite if for any set of k elements $n_1, \cdots, n_k$ in the index set and any k complex numbers $a_1, \cdots, a_k$, we have

$$\sum_{l=1}^{k}\sum_{m=1}^{k} a_l a^*_m f(n_l, n_m) \geq 0$$

Theorem: $\mathbf{R}_X(n,m)$ is Hermitian symmetric and non-negative definite.
Proof: $\mathbf{R}_X^*(n,m) = [E[\mathbf{X}(u,n)\mathbf{X}^*(u,m)]]^*$

$$= E[\mathbf{X}(u,m)\mathbf{X}^*(u,n)] = \mathbf{R}_X(m,n)$$

$$\Longrightarrow \qquad \mathbf{R}_X(n,m) \textit{ is Hermitian symmetric.}$$

$$\sum_l \sum_m a_l a_m^* \mathbf{R}_X(n_l,n_m) = E\left[\sum_l \sum_m a_l \mathbf{X}(u,n_l) a_m^* \mathbf{X}^*(u,n_m)\right]$$

$$= E\left[\sum_l a_l \mathbf{X}(u,n_l) \sum_m a_m^* \mathbf{X}^*(u,n_m)\right]$$

$$= E\left[\left|\sum_l a_l \mathbf{X}(u,n_l)\right|^2\right] \geq 0$$

$$\Longrightarrow \qquad \mathbf{R}_X(n,m) \textit{ is non-negative definite.}$$

9.2 Convergence Concepts

Definition: Let $X_n(u), n = 1,2,\cdots$ be a sequence of random variables with respective distribution functions $F_n(x)$. Let $X(u)$ be a random variable with distribution function $F(x)$. If $F_n \Longrightarrow F$, then X_n is said to converge in distribution or in law to X, written $X_n \stackrel{\mathcal{L}}{\Longrightarrow} X$.

Thus, $X_n \stackrel{\mathcal{L}}{\Longrightarrow} X$ *iff* $\lim_{n\to\infty} P[X_n \leq x] = P[X \leq x]$ for every x such that $P[X = x] = 0$ (*CLT deals with this.*)

Note: we also have this concept when the random variable X_n may be defined on different probability spaces. In this case it is more appropriate to write

$$X_n \stackrel{\mathcal{L}}{\Longrightarrow} X \quad \textit{iff} \quad \lim_{n\to\infty} P_n[X_n \leq x] = P[X \leq x] \textit{ for every } x$$

s.t. $P[X = x] = 0$.

This concept is useful in the sense that $F(x)$ may be complicated but each

$F_n(x)$ may be relatively simple.

Definition: Let $X_n(u), n = 1, 2, \cdots$ be a sequence of random variables. If for any $\epsilon > 0$

$$\lim_{n\to\infty} P[|X_n(u) - X(u)| \geq \epsilon] = 0$$

then we say X_n converges to X in probability. Write $\lim_{n\to\infty} X_n(u) = X(u)$ in probability, or $X_n \overset{\mathcal{P}}{\to} X$.

$|X_n(u) - X(u)| \geq \epsilon_{i.o.}$ is the limit superior of the events

$$[\underbrace{|X_n - X| \geq \epsilon}_{A_n}] = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k, \quad A_k = \{u : |X_k(u) - X(u)| \geq \epsilon\}$$

If $P[|X_n(u) - X(u)| \geq \epsilon_{i.o.}] = 0$ *(i.o.=infinitely often)*,
we say X_n converges to X with probability 1 and this occurs *if and only if*

$$P[sup_{m\geq n}|X_m(u) - X_n(u)| \geq \epsilon] \to 0 \qquad as \quad n \to \infty.$$

Notes:

convergence with probability 1$\Longrightarrow$ convergence in probability
convergence in probability $\Longrightarrow$ convergence in distribution

The converse of these implications does not hold in general.

Example Let X and Y be independent and

$$P(X = 0) = P(Y = 0) = \frac{1}{2}$$

$$P(X = 1) = P(Y = 1) = \frac{1}{2}$$

Let $X_n = Y \quad \forall n$. Note $P(|X - Y| = 1) = \frac{1}{2}$. So, $X_n \overset{\mathcal{L}}{\Longrightarrow} X$ but not in prob., since $P(|X_n - X| = 1) = \frac{1}{2}$ (take $0 < \epsilon < 1$).

Definition: For a random sequence, let

$$S := \{u \in U : \{X(u, n)\} \text{ is convergent}\}$$

Then, $X(u, n)$ is said to be surely convergent if $S = U$, it is said to be almost surely convergent if $P(S) = 1$ (this is the same as convergent with probability 1).

sure convergence $\Longrightarrow$ almost sure convergence $\Longrightarrow$ convergence in probability.

Let $X_n(u)$ be a sequence of i.i.d. random variables with finite mean μ and finite variance σ^2. The sample mean is defined as

$$Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Theorem

$$\lim_{n\to\infty} Y_n = \mu \quad \textit{in probability}$$

This is called the Weak Law of Large Numbers (WLLN).

Actually, $P[\lim_{n\to\infty} Y_n = \mu] = 1$. This is the Strong Law of Large Numbers.

Proof: WLLN.

$$E[Y_n] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \mu$$

$$E[|Y_n - \mu|^2] = \sigma^2_{Y_n} = \frac{\sigma^2}{n}$$

$$\left[Var[Y_n] = \frac{1}{n^2} \sum_{i=1}^{n} Var[X_i] = \frac{\sigma^2}{n} \right]$$

$$P[|Y_n - \mu| \geq \epsilon] \leq \frac{\sigma^2_{Y_n}}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \to 0 \textit{ as } n \to \infty$$

$$\Longrightarrow Y_n \xrightarrow{\mathcal{P}} \mu$$

Definition: If for the random sequence X_n

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

then we say X_n converges to X in the mean square sense. Write

$$\lim_{n\to\infty} X_n(u) = X(u) \ \textit{(m.s.)}$$

Mean Square Convergence $\Longrightarrow$ Convergence in Probability.