Topic 6: Convergence and Limit Theorems

- Sum of random variables
- Laws of large numbers
- Central limit theorem
- Convergence of sequences of RVs

Sum of random variables

Let $X_1, X_2, ..., X_n$ be a sequence of random variables. Define S_n as

$$S_n = X_1 + X_2 + \cdots + X_n$$

- The mean and variance of S become

$$\begin{aligned} E[S_n] &= E[X_1] + E[X_2] + \cdots + E[X_n] \\ \text{var}(S_n) &= \sum_{k=1}^{n} \text{var}(X_k) + \sum_{\substack{j=1 \\ j \neq k}}^{n} \sum_{k=1}^{n} \text{cov}(X_j, X_k) \end{aligned}$$

- If $X_1, X_2, ..., X_n$ are *independent* random variables, then

$$\text{var}(S_n) = \sum_{k=1}^{n} \text{var}(X_k)$$

The characteristic function can be used to calculate the joint pdf as

$$\begin{aligned} \Phi_{S_n}(\omega) &= E\left[e^{j\omega S_n}\right] = \Phi_{X_1}(\omega) \cdots \Phi_{X_n}(\omega) \\ f_{S_n}(x) &= \mathscr{F}^{-1}\left\{\Phi_{X_1}(\omega) \cdots \Phi_{X_n}(\omega)\right\} \end{aligned}$$

Sum of a random number of independent RVs

- Consider the sum of N i.i.d. RVs X_i with finite mean and variance

$$S_N = \sum_{k=1}^{N} X_k$$

where N is a random variable independent of the X_k.

- Using conditional expectation, the mean and variance of S_N are

$$\begin{aligned} E[S_N] &= E\left[E[S_N|N]\right] = E[NE[X]] = E[N]E[X] \\ \text{var}(S_N) &= \text{var}(N)E[X]^2 + E[N]\text{var}(X) \end{aligned}$$

- The characteristic function of S_N is

$$\begin{aligned} \Phi_{S_N}(\omega) &= E\left[E[e^{j\omega S_N}|N]\right] &= E\left[\Phi_X(\omega)^N\right] \\ &= E\left[z^N\right]\Big|_{z=\Phi_X(\omega)} &= G_N\left(\Phi_X(\omega)\right) \end{aligned}$$

which is the generating function of N evaluated at $z = \Phi(\omega)$.

- Example:
  - Number of jobs N submitted to the CPU is a geometric RV with parameter p.
  - The excution time of each job is an exponential RV with mean λ.

  Find the pdf of the total execution time.

Laws of large numbers

Let $X_1, X_2, ..., X_n$ be *independent, identically distributed* (iid) random variables with mean $E[X_j] = \mu$, $(\mu < \infty)$.

- The sample mean of the sequence is defined as

$$M_n = \frac{1}{n}\sum_{j=1}^{n} X_j$$

- For large n, M_n can be used to estimate μ since

$$\begin{aligned} E[M_n] &= \frac{1}{n}\sum_{j=1}^{n} E[X_j] = \mu \\ \text{var}(M_n) &= \frac{1}{n^2}\text{var}(S_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

  - From Chebyshev inequality,

$$\begin{aligned} & P[|M_n - \mu| \geq \varepsilon] &\leq& \frac{\sigma^2}{n\varepsilon^2} \\ \text{or} \quad & P[|M_n - \mu| < \varepsilon] &\geq& 1 - \sigma^2/n\varepsilon^2 \end{aligned}$$

  As $n \to \infty$, we have $\text{var}(M_n) \to 0$ and $\sigma^2/n\varepsilon^2 \to 0$.

- The Weak Law of Large Numbers (WLLN)

$$\lim_{n\to\infty} P\left[|M_n - \mu| < \varepsilon\right] = 1 \qquad \text{for any } \epsilon > 0$$

  The WLLN implies that for a large (fixed) value of n, the sample mean will be within ϵ of the true mean with high probability.

- The Strong Law of Large Numbers (SLLN)

$$P\left[\lim_{n\to\infty} M_n = \mu\right] = 1$$

  The SLLN implies that, with probability 1, every sequence of sample means will approach and stay close to the true mean.

Example:

- Given an event A, we can estimate $p = P[A]$ by
  - performing a sequence of N Bernoulli trials
  - observing the relative frequency of A occurring $f_A(N)$

  How large should N be to have

$$P\left[|f_A(N) - p| \leq 0.01\right] \geq 0.95 \ ?$$

  i.e., a 0.95 chance that the relative frequency is within 0.01 of $P[A]$?

The Central Limit Theorem

- Let $X_1, X_2, ..., X_n$ be i.i.d. RVs with finite mean and variance

$$\begin{aligned} E[X_i] &= \mu < \infty \\ \mathrm{var}(X_i) &= \sigma^2 < \infty \end{aligned}$$

- Let $S_n = \sum_{i=1}^{n} X_i$, and define Z_n as

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}},$$

  Z_n has zero-mean and unit-variance.

- As $n \to \infty$ then $Z_n \to \mathcal{N}(0,1)$. That is

$$\lim_{n\to\infty} P[Z_n \leq z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx.$$

  – Convergence applies to *any* distribution of X with *finite mean* and *finite variance.*
  – This is the Central Limit Theorem (CLT) and is widely used in EE.

- Examples:
  1. Suppose that cell-phone call durations are iid RVs with $\mu = 8$ and $\sigma = 2$ (minutes).
     – Estimate the probability of 100 calls taking over 840 minutes.
     – After how many calls can we be 90% sure that the total time used is more than 1000 minutes?
  2. Does the CLT apply to Cauchy random variables?

Gaussian approximation for binomial probabilities

- A Binomial random variable is a sum of iid Bernoulli RVs.

$$X = \sum_{i=1}^{n} Z_i \ , \quad Z_i \sim \text{Bern}(p) \ \text{ are i.i.d.}$$

then $X \sim \text{binomial}(np)$.

- By CLT, the Binomial cdf $F_X(x)$ approaches a Gaussian cdf

$$p[X = k] \approx \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{ -\frac{(k-np)^2}{2np(1-p)} \right\}$$

The approximation is best for k near np.

- Example:
  - A digital communication link has bit-error probability p.
  - Estimate the probability that a n-bit received message has at least k bits in error.

Convergence of sequences of RVs

- Given a sequence of RVs $\{X_n(\omega)\}$:
  - $\{X_n(\omega)\}$ can be viewed as a sequence of *functions* of ω.
  - For each $\omega \in \Omega$, $\{X_n(\omega)\}$ is a sequence of numbers $\{x_1, x_2, x_3, \ldots\}$.
  - A sequence $\{x_n\}$ is said to converge to x if for any $\epsilon > 0$, there exists N such that

$$|x_n - x| < \epsilon \quad \text{for all } n > N.$$

  We write $x_n \to x$.

- In what sense does $\{X_n(\omega)\}$ converge to a random variable $X(\omega)$ as $n \to \infty$?

Types of convergence for a sequence of RVs:

- *Sure convergence*: $\{X_n(\omega)\}$ converges surely to $X(\omega)$ if

$$X_n(\omega) \to X(\omega) \quad \text{as } n \to \infty \quad \text{for all } \omega \in S$$

For every $\omega \in S$, the sequence $\{X_n(\omega)\}$ converges to $X(\omega)$ as $n \to \infty$.

- *Almost-sure convergence*: $\{X_n(\omega)\}$ converges almost surely $X(\omega)$ if

$$P[\omega : X_n(\omega) \to X(\omega) \quad \text{as } n \to \infty] = 1$$

$X_n(\omega)$ converges to $X(\omega)$ as $n \to \infty$ for all ω in S, except possibly on a set of *zero probability.*

  – The strong LLN is an example of almost-sure convergence.

- *Mean-square convergence*: $\{X_n(\omega)\}$ converges in the mean square sense to $X(\omega)$ if

$$E\left[(X_n(\omega) - X(\omega))^2\right] \to 0 \quad \text{as } n \to \infty$$

Here the convergence is in a sequence of *a function* of $X_n(\omega)$.

  – Cauchy criterion:
  $\{X_n(\omega)\}$ converges in the mean square sense if and only if

$$E\left[(X_n(\omega) - X_m(\omega))^2\right] \to 0 \quad \text{as } n \to \infty \text{ and } m \to \infty$$

- *Convergence in probability*: $\{X_n(\omega)\}$ converges in probability to $X(\omega)$ if, for any $\varepsilon > 0$,

$$P\left[|X_n(\omega) - X(\omega)| > \varepsilon\right] \to 0 \quad \text{as } n \to \infty$$

For each $\omega \in S$, the sequence $X_n(\omega)$ is not required to stay within $\pm\epsilon$ of $X(\omega)$ as $n \to \infty$, but only be within with high probability.

  – The WLLN is an example of convergence in probability.

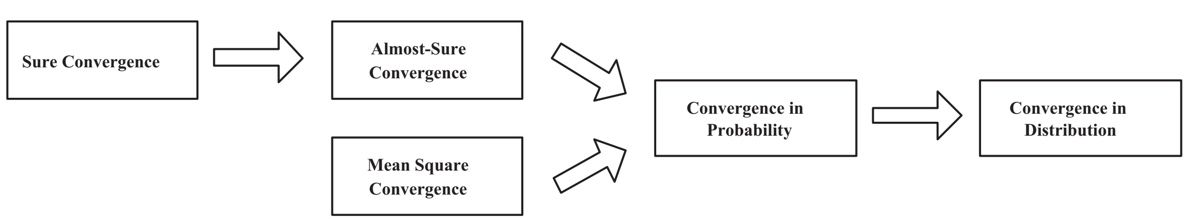
- *Convergence in distribution*: $\{X_n(\omega)\}$ with cdf $\{F_n(x)\}$ converges in distribution to X with cdf $F(x)$ if

$$F_n(x) \to F(x) \quad \text{as } n \to \infty$$

for all x at which $F(x)$ is continuous.

  – The CLT is an example of convergence in distribution.

- Relationship among different convergences



MS convergence does not imply a.s. convergence and vice versa.