Chapter 7: Sums of Random Variables

Sums of Random Variables

- Mean and variance
- PDF of sums of independent RVs
- Laws of large numbers
- Central limit theorems

Mean and Variance

let $X_1, X_2, \ldots, X_n$ be a sequence of RVs

regardless of statistical dependence, we have

$$\mathbf{E}[X_1 + X_2 + \cdots + X_n] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \cdots + \mathbf{E}[X_n]$$

the variance of a sum of RVs is, however, NOT equal to the sum of variances

$$\mathbf{var}(X_1 + X_2 + \cdots + X_n) = \sum_{k=1}^{n} \mathbf{var}(X_k) + \sum_{j=1}^{n} \sum_{k=1}^{n} \mathbf{cov}(X_j, X_k)$$

If $X_1, X_2, \ldots, X_n$ are *uncorrelated*, then

$$\mathbf{var}(X_1 + X_2 + \cdots + X_n) = \mathbf{var}(X_1) + \mathbf{var}(X_2) + \cdots + \mathbf{var}(X_n)$$

PDF of sums of independent RVs

consider the sum of n independent RVs

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

the characteristic function of S_n is

$$\begin{aligned} \Phi_S(\omega) &= \mathbf{E}[e^{j\omega S_n}] = \mathbf{E}[e^{j\omega(X_1+X_2+\cdots+X_n)}] \\ &= \mathbf{E}[e^{j\omega X_1}]\cdots\mathbf{E}[e^{j\omega X_n}] \\ &= \Phi_{X_1}(\omega)\cdots\Phi_{X_n}(\omega) \end{aligned}$$

thus the pdf of S_n is found by finding the inverse Fourier of $\Phi_S(\omega)$:

$$f_S(X) = \mathcal{F}^{-1}[\Phi_{X_1}(\omega)\cdots\Phi_{X_n}(\omega)]$$

Example

find the pdf of a sum of n independent exponential RVs

all exponential variables have parameter α

the characteristic function of a single exponential RV is

$$\Phi_X(\omega) = \frac{\alpha}{\alpha - \mathrm{j}\omega}$$

the characteristic function of the sum is

$$\Phi_S(\omega) = \left(\frac{\alpha}{\alpha - \mathrm{j}\omega}\right)^n$$

we see that S_n is an n-Erlang RV

Sample Mean

let X be an RV with $\mathbf{E}[X] = \mu$ (unknown)

$X_1, X_2, \ldots, X_n$ denote n independent, repeated measurements of X

X_j's are *independent, identically distributed* (i.i.d.) RVs

the **sample mean** of the sequences is used to estimate $\mathbf{E}[X]$:

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

two statistical quantities for characterizing the sample mean's properties:

- $\mathbf{E}[M_n]$: we say M_n is unbiased if $\mathbf{E}[M_n] = \mu$
- $\mathbf{var}(M_n)$: we examine this value when n is large

Sample Mean

the sample mean is an **unbiased estimator** for μ:

$$\mathbf{E}[M_n] = \mathbf{E}\left[\frac{1}{n}\sum_{j=1}^{n} X_j\right] = \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}[X_j] = \mu$$

suppose $\mathbf{var}(X) = \sigma^2$ (true variance)

since X_j's are i.i.d, the variance of M_n is

$$\mathbf{var}(M_n) = \frac{1}{n^2}\sum_{j=1}^{n}\mathbf{var}(X_j) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

hence, the variance of the sample mean approaches zero as the number of samples increases