Chapter 1
Elements of Stochastic Processes

1 Review of Basic Terminology and Properties of Random Variables and Distribution Functions

A random variable X is *discrete* if there exists a finite or numerable set of distinct values λ_i for $i \in \mathbb{Z}$ s.t $Pr\{X = \lambda_i\} = a_i > 0$ and $\sum_i a_i = 1$. If on the other hand $Pr\{X = \lambda\} = 0$ for any λ, the variable is called *continuous*.

Distribution function F is what we normally called the cumulative distribution function (CDF), and it is defined as

$$F(\lambda) = \int_{-\infty}^{\lambda} p(t)dt \tag{1}$$

where $p(t)$ is the *probability density*. p may or may not exist for continuous random variables. For discrete variables this boils down to

$$F(\lambda) = Pr\{X \leq \lambda\}. \tag{2}$$

Moments for random variables are given by

$$E[X^m] = \int_{-\infty}^{\infty} x^m p(x)dx \tag{3}$$

for a moment of order m. This integral turns into a sum for discrete variables. The first moment is the *mean*, m_X. The second central moment (the 2nd moment of random variable $X - m_X$ is the *variance*, σ_X^2.

Functions of random variables are random variables. Expectations of functions of random variables can be easily computed.

1.1 Joint Distribution Functions

The joint distribution of a pair of discrete random variables X, Y is

$$F(\lambda_1, \lambda_2) = Pr\{X \leq \lambda_1, Y \leq \lambda_2\}. \tag{4}$$

The marginal distribution of X is then $F(\lambda, \infty) = \lim_{\lambda_2 \to \infty} F(\lambda, \lambda_2)$. The marginal distribution of Y is defined similarly. The corresponding probability density, if it exists, for a joint distribution would then be

$$F(\lambda_1, \lambda_2) = \int_{-\infty}^{\lambda_1} \int_{-\infty}^{\lambda_2} p(s,t) ds dt. \tag{5}$$

If X, Y have respective means m_X, m_Y, their covariance is

$$\sigma_{XY} = E[(X - m_X)(Y - m_Y)]. \tag{6}$$

If X_1, X_2 are independent with respective distributions F_1, F_2, the distribution F of their sum is

$$F(x) = \int F_1(x - y) dF_2(y) = \int F_2(x - y) dF_1(y). \tag{7}$$

The same convolution goes for the density functions if they exist.

1.2 Characteristic Functions

Chapter 2
Elements of Stochastic Processes

2 Markov Chains

The definition of a Markov chain is quite clear from Section 1. Of usual interest are *discrete time* Markov chains, where the state space is countable and finite. These states are normally labelled by nonnegative integers, since there can be found a one-to-one mapping between the states and a subset of integers. The time at which the chain is evaluated is also labelled by a nonnegative integer. So, a chain that is in state i at time n is represented by

$$X_n = i \tag{8}$$

The probability of $X_{n+1} = j$ given that $X_n = i$, known as the transition probability, is then

$$P_{ij}^{n,n+1} = Pr\{X_{n+1} = j | X_n = i\}. \tag{9}$$

Most chains of interest are *stationery*, indicating that the transition probabilities are independent of n. So, the transition probabilities are represented simply as $P_{ij}^{n,n+1}$.

A Markov chain is fully defined by specifying all the transition probabilities. This can be done by defining a *transition probability matrix*, $\mathbf{P} = \|P_{ij}\|$.

2.1 Example Markov Chains

Refer to Elementary Problem 1a. for a full problem description. In brief, we are to define $\mathbf{P} = \|P_{jk}\|$ for a sequence of coin tosses (with probability of heads p) where the state of the process is the number of heads minus the number of tails at time n. So, it would suffice to find

$$P_{jk} = Pr\{H^{n+1} - T^{n+1} = k | H^n - T^n = j\}. \tag{10}$$

We notice that $H^{n+1} = H^n + h$ and $T^{n+1} = T^n + t$, where $h = 1, t = 0$ if we toss heads at $n+1$ time otherwise $h = 0, t = 1$. So,

$$P_j k = Pr\{H^n + h - T^n - t = k\} = Pr\{h - t = k - j\} \tag{11}$$

When $k - j = 1$ we'd have tossed heads; we get that with probability p. When $k - j = -1$ we'd have tossed tails; we get that with probability $1 - p$. So

$$P_{jk} = \begin{cases} p, & k = j + 1 \\ 1 - p, & k = j - 1 \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

Refer to Elementary Problem 1b. for a full problem description. In brief, we are to find a **P** for a system of two urns that at $n = 0$ contain N black and white balls each, respectively. At each time step, balls are randomly picked from each urn and swapped between the urns. The state of this system is the number of white balls in the first urn. We label these urns β, ω respectively. Now

$$P_{jk} = Pr\{W_\beta^{n+1} = k | W_\beta^n = j\} \tag{13}$$

where W_β^n is the variable for the number of white balls in the β urn at time n. Given that $W_\beta^n = j$, we know that $W_\omega^n = N - j$. For $k = j + 1$ we need to pick a white ball from ω and a black ball from β. For $k = j$ we need to pick either both black balls or white balls. For $k = j - 1$ we need to pick a black ball from ω and a white ball from β. These cases clearly give

$$P_{jk} = \begin{cases} \frac{(N-j)^2}{N^2}, & k = j + 1 \\ \frac{i^2}{N^2}, & k = j - 1 \\ \frac{2j(N-j)}{N^2}, & k = j \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

We can convince ourselves by checking that these transition probabilities sum up to 1.