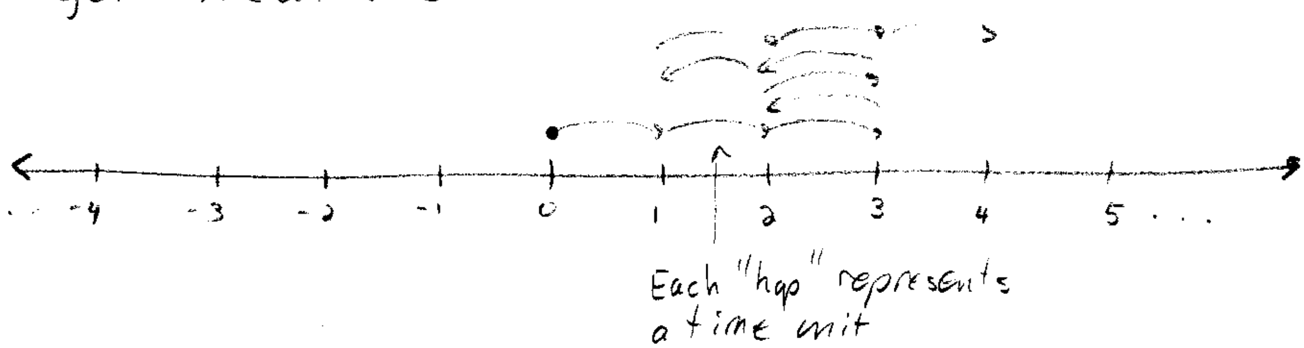
Chapter 1 Introduction and Review of Probability

This is a course about stochastic processes, which are systems that evolve in time according to probabilistic rules. More precisely, a stochastic process is determined by one or more random variables that depend on time. By "time", we usually mean the kind of time we know and love, but it really could be any one dimensional, independent variable that is "time like" in the sense of having a direction.

Example 1.0.1 Random Walk

We consider the motion of a particle that fluctuates randomly. We suppose the particle is restricted to move along the x-axis occupying integer locations



We assume that the particle changes position one unit in space every time unit. From a given position, it may move right with a probability p, $0 \le p \le 1$, and left with probability $1-p$. Here, the stochastic process of interest is the location of the particle at time t. We plot three "realizations" of a random walk on page 3.

Example 1.0.2

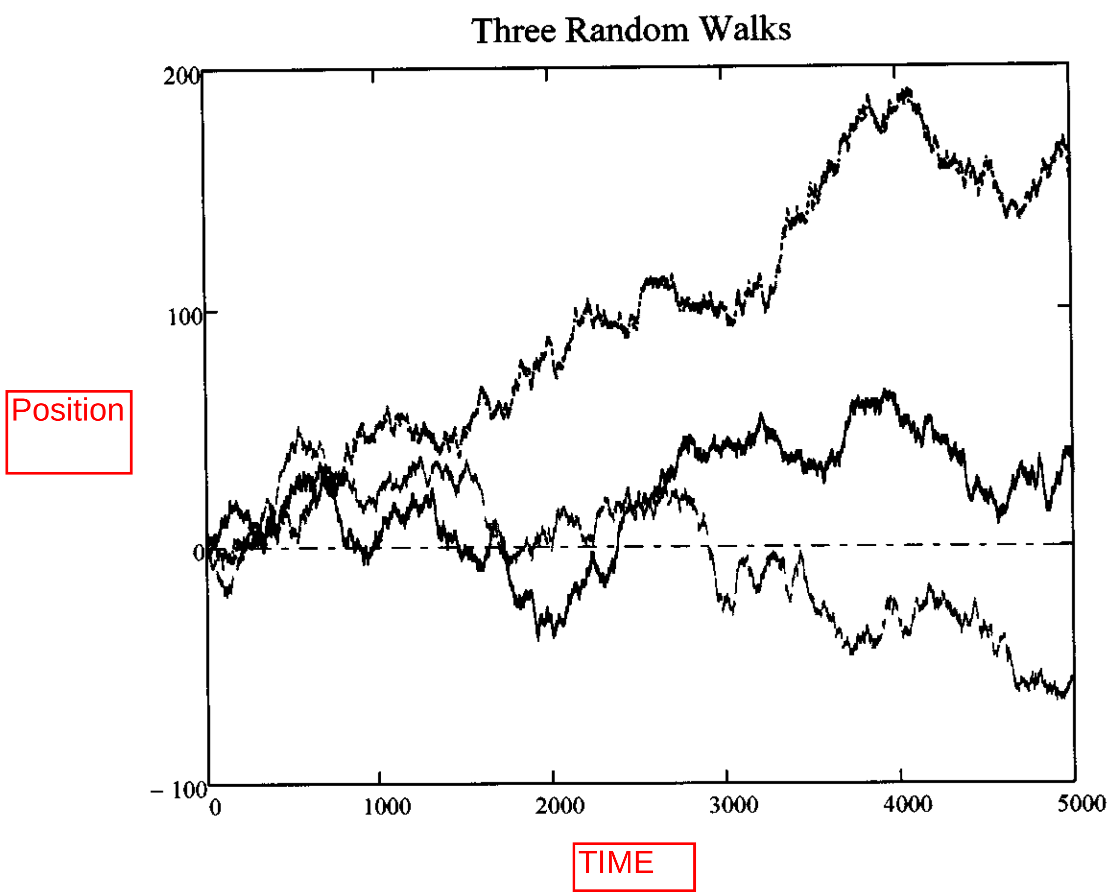
A roulette wheel has 38 numbered slots for the ball to rest. During each spin, we bet \$1 that the ball will end up resting on an odd spot. We start with \$10 and record our fortune after each spin. The chances of winning on each spin is $18/38$ and losing $20/38$. Do we go broke and how long might that take?

Below we will clarify our vague definition of a stochastic process and study these examples in detail. First, we review some elementary probability.

§1.1 Nonnegative, integer-valued random variables
(see T.K. § I.2.2)

We briefly review some probability fundamentals. You should use these first few lectures to review the material in Chapters I and II in the text.

Let X be a random variable with range $\{0, 1, 2, \ldots \infty\}$.


Three Random Walks
Position
TIME
200
100
0
− 100
0
1000
2000
3000
4000
5000

Example 1.1.1

If $\mathbb{X}$ is the time it takes for a random event to occur, and the event never occurs, we might assign $\mathbb{X} = \infty$.

Definition 1.1.1

We let
$$P(\mathbb{X} = k) = p_k \quad , k = 0, 1, 2, \ldots$$
be the probability that $\mathbb{X}$ has value k. We define the probability mass function to be $p_{\mathbb{X}}(i) = P(\mathbb{X} = i)$.

It follows that

$$P(\mathbb{X} < \infty) = \sum_{k=0}^{\infty} p_k = \lim_{N \to \infty} \sum_{k=0}^{N} p_k$$

Definition 1.1.2

We define

$$p_\infty = P(\mathbb{X} = \infty) = 1 - \sum_{k=0}^{\infty} p_k$$

We are interested in describing statistics of such random variables.

Definition 1.1.3

We define the expected value of $\mathbb{X}$

$$E(X) = \begin{cases} \sum_{k=0}^{\infty} k\,p_k , & P(X=\infty)=0, \\ \infty , & P(X=\infty)>0 \end{cases}$$

In most situations we will encounter, $P(X=\infty)=0$.

Recall from standard probability theory,

<u>Theorem 1.1.1</u>
If f is a nonnegative function,

$$(1.1.1) \qquad E(f(X)) = \sum_{k=0}^{\infty} f(k)\,p_k$$

<u>Definition 1.1.4</u>
We have to treat the case when a function of a random variable can take on positive and negative values. In that case, we set

$$f^{+} = \max\{f, 0\},$$

$$f^{-} = -\min\{f, 0\},$$

so $f(x) = f^{+}(x) - f^{-}(x)$. If at least one of $E(f^{+}(X))$ or $E(f^{-}(X))$ is finite, we set

$$E(f(X)) = E(f^{+}(X)) - E(f^{-}(X)).$$

If both $E(f^{+}(X))$ and $E(f^{-}(X))$ are infinite, then

$E(f(X))$ is not defined.

It follows that if $\sum_{k=0}^{\infty} |f(k)| p_k < \infty$, $E(f(X))$ is defined and is finite.

Definition 1.1.5

If $p_\infty = 0$, then the n^{th} moment of X is $E(X^n)$ and the n^{th} central moment is

$$E\left((X - E(X))^n \right)$$

When $n = 2$, we call this the variance,

$$Var(X) = E\left((X - E(X))^2\right) = E(X^2) - (E(X))^2$$

Example 1.1.2 The Binomial Distribution

This is the distribution of the number of successes in n Bernoulli trials (think flipping a coin) where the probability of success is p.

$$P(X = k) = b(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k} \quad 0 \le k \le n$$

$$E(X) = np$$

$$Var(X) = np(1-p)$$

(In TK, I §3, you can find a review of some of the basic distributions.)

Example 1.1.3 The Poisson Distribution

For $\lambda > 0$, $k = 0, 1, 2, \ldots$

$$\left.\begin{aligned} P(X=k) &= p(k,\lambda) = e^{-\lambda}\lambda^k / k! \\ E(X) &= \lambda \\ Var(X) &= \lambda \end{aligned}\right\} \text{we write } X \sim Poi(\lambda)$$

This distribution is observed empirically in many applications when we count the number of occurences of an unpredictable event over a period of time, where the event is rare on a short time scale, but more frequent on a long time scale. The has to be sufficient randomness. So, the number of tropical cyclones over two decades fits a Poisson distribution very well, but the number of buses running on a regular schedule that arrive at a given stop does not.

The next result says that we can compute the expected value of a random variable by summing the tail probabilities.

<u>Theorem 1.1.2</u>

If X is a nonnegative, integer valued random variable, then

$$(1.1.2) \qquad E(X) = \sum_{k=0}^{\infty} P(X > k)$$

<u>Proof</u>

$$\sum_{k=0}^{\infty} P(X > k) = \sum_{k=0}^{\infty} \sum_{j=k+1}^{\infty} P_j$$

$$= \sum_{j=1}^{\infty} \Big(\underbrace{\sum_{k=0}^{j-1} 1}_{\text{number of repeated values}} \Big) P_j$$

$$= \sum_{j=1}^{\infty} j P_j = E(X).$$

---

§1.2 Independence

Recall in probability that two events A, B are called independent if

$$P(A \cap B) = P(A) P(B)$$

<u>Example 1.1.2</u>

We choose a card at random from a deck of 52 cards. The events $\{card = king\}$ and $\{card = spades\}$ are independent.

Definition 1.2.1

Let $\{X_1, \dots, X_k\}$ be a collection of nonnegative, integer valued random variables. They are (mutually) independent if and only if

$$(1.2.1) \quad P(X_1 = i_1, \dots, X_k = i_k) = P(X_1 = i_1) \cdot \cdots \cdot P(X_k = i_k)$$

for all integers $i_1, \dots, i_k \geq 0$. If we define

$$p_{i_1, \dots, i_k} = P(X_1 = i_1, \dots, X_k = i_k)$$

then

$$p_{i_1, \dots, i_k} = p_{i_1} \cdot \cdots \cdot p_{i_k}.$$

Notes:

(1) We should write $P(\{X_1 = i_1, \dots, X_k = i_k\})$ since we are talking about the probability of the event $\{X_1 = i_1, \dots, X_k = i_k\}$, but this is cumbersome.

(2) Recall that it does not suffice to check (1.2.1) pairwise only.

Theorem 1.2.1

The random variables $\{X_1, \dots, X_k\}$ are independent if and only if

$$E(f_1(X_1) \cdots f_k(X_k)) = E(f_1(X_1)) \cdot \cdots E(f_k(X_k)) \tag{1.2.2}$$

for all bounded functions $f_1, \dots, f_k$.

Proof

($\Rightarrow$)

$$E(f_1(X_1) \cdots f_k(X_k)) = \sum_{i_1, \dots, i_k} f_1(i_1) \cdots f(i_k)\, p_{i_1, \dots, i_k}$$

$$= \sum_{i_1} \sum_{i_2} \cdots \sum_{i_k} f_1(i_1) \cdots f(i_k)\, p_{i_1} \cdots p_{i_k}$$

$$= E(f_1(X_1)) \cdots E(f_k(X_k))$$

($\Leftarrow$)

Set

$$f_j(x) = \begin{cases} 1 & \text{if } x = i_j \\ 0 & \text{otherwise} \end{cases}$$

Then

$$E(f_1(X_1) \cdots f_k(X_k)) = p_{i_1, \dots, i_k} = p_{i_1} \cdots p_{i_k} \quad \text{by assumption}$$

<u>Corollary</u>

If $X_1, \ldots, X_k$ are independent, then

$$(1.2.3) \qquad E(X_1 \cdots X_k) = E(X_1) \cdots E(X_k)$$

and

$$(1.2.4) \qquad Var(a_1 X_1 + \cdots + a_k X_k) = a_1^2 Var(X_1) + \cdots + a_k^2 Var(X_k)$$

for numbers $a_1, \ldots, a_k$, provided $Var(X_i) < \infty$, $1 \le i \le k$.

§1.3 Conditional Probability

We will be greatly concerned with evaluating statements like "What is the probability of event A given event B has occurred?"

Suppose we conduct an experiment N times and we observe the occurrences of two events A, B. We are only interested in outcomes for which B occurs and disregard others. In this smaller set, the proportion of times A occurs is

$$\frac{N_{A \cap B}}{N_B}$$

since B occurs in both. We write

$$\frac{N_{A \cap B}}{N_B} = \frac{N_{A \cap B}/N}{N_B/N}.$$