

MATH 468 / 568 Spring 2010

klin@math.arizona.edu

Lecture 1

Introduction

In Nature we frequently encounter quantities that fluctuate in a seemingly random manner. The “static” or “noise” that you hear in car radio, minute-to-minute fluctuations in stock prices, daily fluctuations in temperature are all examples. Stochastic processes are mathematical objects used to model systems that exhibit random variations across time (or space). In this course we will develop the basic theory of stochastic processes, using mainly the tools and language of probability.

Historically, much of the theory of stochastic processes was developed to understand (1) the random movement of atoms and molecules at the molecular scale; (2) financial problems; and (3) filtering out noise from measurements (an important problem during World War II). It has since blossomed into a branch of mathematics with a large number of applications, to physics, engineering, biology, chemistry, economics,

I’ll begin with a very brief review of probability theory. Since you are expected to have learned all this material in MATH 464, these notes only sketch out what you will need to know for this course. See Prof. Watkins’s notes from MATH 363 for details (you can find a link on the on-line course syllabus). You might also consult a standard text book like *A First Course in Probability* by Sheldon Ross.

REVIEW OF PROBABILITY

Sample Spaces, Events, and Probability

The fundamental object in probability is the *sample space*, which is the set of all possible *outcomes* of a “chance experiment.” An *event* is a set of outcomes, i.e., a subset of the sample space.

For example, if our experiment consists of tossing a fair coin 8 times, then the sample space consists of all possible sequences of 8 H’s (heads) and T’s (tails). The cardinality (size) of the sample space is $2^8 = 256$. The event that “the first toss is an H” is represented by the set of all sequences that begin with an H; this event has cardinality 128. If we toss the coin an infinite number of times, the sample space is the set of all infinite sequences of H’s and T’s. In the latter case, the sample space is infinite (in fact, it’s *uncountable*).

While in theory we can define from scratch a sample space and the relevant events for each experiment we might study, in practice we often construct sample spaces and events from simpler ones. For example, if our chance experiment is “flip a coin once,” the sample space is just $\{H, T\}$. If we now flip the coin 20

times, the sample space is then $\{H, T\} \times \dots \times \{H, T\} = \{H, T\}^{20}$, the Cartesian product of 20 copies of the single-coin-flip sample space $\{H, T\}$. (The Cartesian product of two sets A and B is the set $A \times B = \{(x, y) : x \in A, y \in B\}$.) In general, if there are n chance experiments we can perform and the i th experiment has sample space Ω_i , then the “compound experiment” in which we perform the n experiments in order, one after another, has sample space $\Omega_1 \times \dots \times \Omega_n$.

Event algebra. Similarly, we can express complicated events in terms of simpler events; this is an important method for simplifying probability problems. The basic operations correspond to the basic logical operations *AND*, *OR*, and *NOT*:

$$\begin{array}{lll} \text{A and B both occurred} & \Leftrightarrow & A \cap B \text{ (intersection of A and B)} \\ \text{A or B (or both) occurred} & \Leftrightarrow & A \cup B \text{ (union of A and B)} \\ \text{A did not occur} & \Leftrightarrow & A^c = \Omega \setminus A \text{ (complement of A in } \Omega) \end{array}$$

With the operations above we can build a large set of sample spaces and events.

As an example, let A denote the event (the first 5 cards dealt out of a well-shuffled deck of 52 form a straight flush). (A straight flush is a set of 5 cards of the same suit with consecutive numbers.) Directly calculating the probability of the event A is a somewhat tedious problem. However, if we let A_i denote the event (the first 5 cards form a straight flush and the lowest card is i), then we can write $A = A_1 \cup \dots \cup A_{10}$. (We think of an ace as a 1; 10-J-Q-K-A is considered a straight, i.e., ace can be high or low (but not both).) Since the events are “pairwise-disjoint” (see below), this means the probability of A is the sum of the probabilities of $P(A_i)$, which are easier to compute.

Probability. A *probability function* (more precisely a *probability measure*) is a function P that assigns to each *event* a number and satisfying the following axioms:

- i. For all events E , $0 \leq P(E) \leq 1$.
- ii. $P(\Omega) = 1$.
- iii. If A_1, A_2, \dots is a sequence of *pairwise-disjoint events* (i.e., $A_i \cap A_j = \emptyset$ whenever $i \neq j$), then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

From these axioms we can derive many useful identities, e.g., $P(A^c) = 1 - P(A)$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, $P(\emptyset) = 0$, etc.

As you all know from a previous course like MATH 464, we assign probabilities to events (and not to individual outcomes) because sample spaces may be uncountably infinite, and in such cases it is not possible to assign probabilities to outcomes in a meaningful way.

I’d like to emphasize again how useful it is the event algebra is: it is frequently useful for transform a complicated problem into a much simpler one. As a familiar example, consider the event E = (out of k people, at least 2 have the

same birthday). Computing $P(E)$ directly (assuming all days of the year are equally likely and that there are no leap years) is not as easy as computing $P(E^c)$, the probability that no 2 people have the same birthday: the latter is just $1 - \frac{365 \cdot 364 \cdot \dots \cdot (365 - k + 1)}{365^k}$.

Conditional probability. Let A be an event with $P(A) > 0$. If the event A is known to have occurred, it reduces the number of possible outcomes and thus changes the sample space. This, in turn, means we may need to change the probability we assign to other events. For example, if we roll a die without looking at the outcome but are told that the result is odd, the probability that we rolled a 5 is $1/3$, not $1/6$.

To capture this idea, we define the *probability of B given A* (or the *probability of B conditioned on A*) by

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

We think of the event A as the new sample space, and for every event B we can define a new probability $P'(B) := P(B|A)$. It is easy to check that the function P' satisfies the axioms of probability.

Two events are *independent* if $P(B|A) = P(B)$, i.e., knowing that A has occurred doesn't affect the probability of B occurring. It is easy to show that A and B are independent if and only if $P(A \cap B) = P(A) \cdot P(B)$. If we have n events A_1, A_2, \dots, A_n , they are independent if

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \dots \cdot P(A_{i_k})$$

for all subsets $1 \leq i_1 < i_2 < \dots < i_k \leq n$. (Why is it not enough to ask that the events are "pairwise-independent," i.e., $P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$ for all pairs i, j ? Think about an experiment with 3 fair coin tosses: let $A =$ (1 st and 2 nd tosses are the same), $B =$ (2 nd and 3 rd tosses are the same), and $C =$ (1 st and 3 rd tosses are the same). Verify that any two of these events are independent, but $P(A \cap B \cap C) \neq P(A)P(B)P(C)$.)

Two consequences of the definition of conditional probability are:

- 1) For any collection of events A_1, \dots, A_k such that $P(A_i) > 0$ for all i , we have $P(A_1 \cap \dots \cap A_k) = P(A_k | A_1 \cap \dots \cap A_{k-1}) P(A_{k-1} | A_1 \cap \dots \cap A_{k-2}) \dots P(A_2 | A_1) P(A_1)$
- 2) Suppose we have a collection of events A_1, \dots, A_k such that $P(A_i \cap A_j) = 0$ whenever $i \neq j$ (this says these events are "mutually exclusive," i.e., no two ever occur at the same time), and $A_1 \cup \dots \cup A_k = \Omega$. Then for any event B , one can show that

$$P(B) = \sum_{i=1}^k P(B|A_i)P(A_i).$$

(Thanks to Misha Stepanov for catching an egregious error in an earlier version of these notes.)