Math 632: Introduction to Stochastic Processes

Fall 2012

University of Wisconsin at Madison

Instructor: David F. Anderson

Review of Probability

Probability theory is used to model experiments whose outcome can not be predicted with certainty beforehand.

For any such experiment, there is a triple $(\Omega, \mathcal{F}, P)$, called a *probability space*, where

- ► Ω is the *sample space*,

- ► $\mathcal{F}$ is a collection of *events*,

- ► P is a *probability measure*.

We will consider each in turn.

The sample space Ω

The *sample space* of an experiment is the set of all possible outcomes.

Elements of Ω are called *sample points* and are often denoted by ω. Subsets of Ω are referred to as *events*.

Example
Consider the experiment of rolling a six-sided die. Then the natural sample space is $\Omega = \{1, 2, 3, 4, 5, 6\}$. $\square$

Example
Consider the experiment of tossing a coin three times. Let us write 1 for heads and 0 for tails. Then the sample space consists of all sequences of length three consisting only of zeros and ones. Each of the following representations is valid:

$$
\begin{aligned}
\Omega &= \{0, 1\}^3 \\
&= \{0, 1\} \times \{0, 1\} \times \{0, 1\} \\
&= \{(x_1, x_2, x_3) : x_i \in \{0, 1\} \text{ for } i = 1, 2, 3\} \\
&= \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (1, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0), (1, 1, 1)\}.
\end{aligned}
$$

The sample space Ω

Example

Roll two die. There are

$$6^2 = 36$$

possible outcomes. The sample space could be

$$\Omega = \{(m, n) : 1 \le m, n \le 6\}.$$

The **event** that the "sum is 9" is then

$$B = \{(3, 6), (4, 5), (5, 4), (6, 3)\}.$$

Example

Consider the experiment of waiting for a bacteria to divide. In this case, it is natural to take as our sample space all values greater than or equal to zero. That is,

$$\Omega = \{t \ : \ t \ge 0\},$$

where the units of t are specified as hours, for example. $\qquad\qquad\square$

The sample space Ω

Terminology: A set that is finite or countably infinite is called *discrete*.

The collection of events $\mathcal{F}$

- ▶ Events are simply subsets of the state space Ω.

- ▶ They are often denoted by A, B, C, etc., and they are usually the objects we wish to know the probability of.

- ▶ They can be described in words, or using mathematical notation.

Example 1, continued. Let A be the event that a 2 or a 4 is rolled. That is, $A = \{2, 4\}$. □

Example 2, continued. Let A be the event that the final two tosses of the coin are tails. Thus,

$$A = \{(1, 0, 0), (0, 0, 0)\}.$$

□

Example 4, continued. Let *A* be the event that it took longer than 2 hours for the bacteria to divide. Then,

$$A = \{t \ : \ t > 2\}.$$

$\square$

We will often have need to consider the unions and intersections of events.

1. We write

$$A \cup B$$

for the union of *A* and *B*, and

2. either

$$A \cap B$$

or

$$AB$$

for the intersection.

The collection of events $\mathcal{F}$

For discrete sample spaces, $\mathcal{F}$ will contain all subsets of Ω, and will play very little role. This is the case for nearly all of the models in this course.

When the state space is more complicated, $\mathcal{F}$ is assumed to be a σ-field (or σ-algebra). That is, it satisfies the following three axioms:

1. $\Omega \in \mathcal{F}$.

2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$, where A^c is the complement of A.

3. If $A_1, A_2, \ldots, \in \mathcal{F}$, then

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}.$$

By DeMorgan's law, we also get intersections:

$$\bigcap_i A_i = \left(\bigcup_i A_i^c \right)^c \in \mathcal{F}.$$

The probability measure P

Definition
The real valued function P, with domain $\mathcal{F}$, is a *probability measure* if it satisfies the following three axioms

1. $P(\Omega) = 1$.

2. If $A \in \mathcal{F}$, then $P(A) \geq 0$.

3. If for a sequence of events $A_1, A_2, \ldots,$, we have that $A_i \cap A_j = \emptyset$ for all $i \neq j$ (i.e. the sets are *mutually exclusive* or disjoint) then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

The probability measure P

The following is a listing of some of the basic properties of any probability measure.

Lemma

Let $P(\cdot)$ be a probability measure. Then

1. *If $A_1, \ldots, A_n$ is a finite sequence of mutually exclusive events, then*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

2. $P(A^c) = 1 - P(A)$.

3. $P(\emptyset) = 0$.

4. *If $A \subset B$, then $P(A) \leq P(B)$.*

5. $P(A \cup B) = P(A) + P(B) - P(AB)$.

Conditional probability and independence

Suppose we are interested in the probability that some event *A* took place, though we have some extra information in that we know some other event *B* took place.

For example, suppose

1. that we want to know the probability that a fair die rolled a 4

2. given that we know an even number came up.

Most people would answer this as 1/3, as there are three possibilities for an even number, $\{2, 4, 6\}$, and as the die was fair, each of the options should be equally probable. The following definition generalizes this intuition.

Definition
For two events $A, B \subset \Omega$, the *conditional probability of A given B* is

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

provided that $P(B) > 0$.

Example

The probability that it takes a bacteria over 2 hours to divide is 0.64, and the probability it takes over three hours is 0.51. What is the probability that it will take over three hours to divide, given that two hours have already passed?

Solution: Let

1. A be the event that the bacteria takes over three hours to split and
2. B be the event that it takes over two hours to split.

Then, because $A \subset B$,

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A)}{P(B)} = \frac{.51}{.6} = .85.$$

$\square$

We intuitively think of *A* being independent from *B* if

$$P(A|B) = P(A), \quad \text{and} \quad P(B|A) = P(B).$$

More generally, we have the following definition.

Definition
The events $A, B \in \mathcal{F}$ are called *independent* if

$$P(AB) = P(A)P(B).$$

It is straightforward to check that the definition of independence implies both

$$P(A|B) = P(A), \quad \text{and} \quad P(B|A) = P(B).$$

Independence

By definition

$$P(A|B) = \frac{P(AB)}{P(B)}, \quad \text{and} \quad P(B|A) = \frac{P(AB)}{P(A)}.$$

Rearranging terms yields

$$P(AB) = P(A|B)P(B), \quad \text{and} \quad P(AB) = P(B|A)P(A).$$

We also have:

$$P(AB \mid C) = P(A \mid BC)P(B \mid C).$$

Random variables

Definition
A *random variable X* is a real-valued (measurable) function defined on the sample space Ω. That is, $X : \Omega \to \mathbb{R}$.

- If the range of X is finite or countably infinite, then X is said to be a *discrete random variable*.

- Otherwise X is said to be a *continuous random variable*.

Random variables

Example

Suppose we roll two die and take $\Omega = \{(i, j) \mid i, j \in \{1, \ldots, 6\}\}$. We let

$$X(i, j) = i + j$$

be the discrete random variable giving the sum of the rolls. The range is $\{2, \ldots, 12\}$. $\square$

Example

Consider two bacteria, labeled 1 and 2. Let T_1 and T_2 denote the times they will divide to give birth to daughter cells, respectively. Then, $\Omega = \{(T_1, T_2) \mid T_1, T_2 \geq 0\}$. Let X be the continuous random variable giving the time of the first division: $X(T_1, T_2) = \min\{T_1, T_2\}$. The range of X is $t \in \mathbb{R}_{\geq 0}$. $\square$

Random variables

Definition

If X is a random variable, then the function F_X, or simply F, defined on $(-\infty, \infty)$ by

$$F_X(t) = P\{X \le t\}$$

is called the *distribution function*, or *cumulative distribution function*, of X.

Theorem (Properties of the distribution function)

Let X be a random variable defined on some probability space $(\Omega, \mathcal{F}, P)$, with distribution function F. Then,

1. *F is nondecreasing. Thus, if $s \le t$, then*
   $F(s) = P\{X \le s\} \le P\{X \le t\} = F(t)$.
2. *$\lim_{t \to \infty} F(t) = 1$.*
3. *$\lim_{t \to -\infty} F(t) = 0$.*
4. *F is right continuous. So, $\lim_{h \to 0^+} F(t + h) = F(t)$ for all $t \in \mathbb{R}$.*

Random variables

Definition

Let X be a discrete random variable. Then for $x \in \mathbb{R}$, the function

$$p_X(k) = P\{X = k\}$$

is called the *probability mass function* of X.

By the axioms of probability, a probability mass function p_X satisfies

$$P\{X \in A\} = \sum_{k \in A} p_X(k).$$

Random variables

Definition

Let X be a continuous random variable with distribution function $F(t) = P\{X \leq t\}$. Suppose that there exists a nonnegative, integrable function $f : \mathbb{R} \to [0, \infty)$, or sometimes f_X, such that

$$F(x) = P\{X \in (-\infty, x]\} = \int_{-\infty}^{x} f(y)dy.$$

Then the function f is called the *probability density function* of X.

We now have that for any $A \subset \mathbb{R}$ (or, more precisely, for any A in "Borel sets" of $\mathbb{R}$, but we are going to ignore this point),

$$P\{X \in A\} = \int_{A} f_X(x)dx.$$

Expectations of random variables

Let X be a random variable. Then, the *expected value* of X is

$$\mathbb{E}[X] = \sum_{x \in \mathcal{R}(X)} x \cdot P\{X = x\}$$

in the case of discrete X, and

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)dx,$$

in the case of continuous X.

The expected value of a random variable is also called its *mean* or *expectation* and is often denoted μ or μ_X.

Example

Consider a random variable taking values in $\{1, \ldots, n\}$ with

$$P\{X = i\} = \frac{1}{n}, \quad i \in \{1, \ldots, n\}.$$

We say that X is distributed uniformly over $\{1, \ldots, n\}$. What is the expectation?

Solution. We have

$$\mathbb{E}[X] = \sum_{k=1}^{n} kP\{X = k\} = \sum_{ik=1}^{n} k\frac{1}{n} = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}.$$

$\square$

Example

Suppose that X is exponentially distributed with density function

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , \quad x \geq 0 \\ 0 & , \quad \text{else} \end{cases} ,$$

where $\lambda > 0$ is a constant. In this case,

$$\mathbb{E}X = \int_0^\infty x \lambda e^{-\lambda x} dx = \frac{1}{\lambda}. \tag{1}$$

$\square$

Expectations of random variables

Theorem

Let X be a random variable and let $g : \mathbb{R} \to \mathbb{R}$ be a function. Then,

$$\mathbb{E}[g(X)] = \sum_{x \in \mathcal{R}(X)} g(x) p_X(x),$$

in the case of discrete X, and

$$\mathbb{E}[g(X)] = \int_{\infty}^{\infty} g(x) f_X(x) dx,$$

in the case of continuous X.

An important property of expectations is that for any random variable X, real numbers $\alpha_1, \ldots, \alpha_n$, and functions $g_1, \ldots, g_n : \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}[\alpha_1 g_1(X) + \cdots + \alpha_n g_n(X)] = \alpha_1 \mathbb{E}[g_1(X)] + \cdots + \alpha_n \mathbb{E}[g_n(X)].$$

Variance of a random variable

The variance gives a measure on the "spread" of a random variable around its mean.

Definition
Let μ denote the mean of a random variable X. The *variance* and *standard deviation* of X are

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$
$$\sigma_X = \sqrt{\mathrm{Var}(X)},$$

respectively.

Also,

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Some common discrete random variables

Bernoulli random variables: X is a *Bernoulli* random variable with parameter $p \in (0, 1)$ if

$$P\{X = 1\} = p,$$
$$P\{X = 0\} = 1 - p.$$

For a Bernoulli random variable with a parameter of p,

$$\mathbb{E}[X] = p \quad \text{and} \quad \text{Var}(X) = p(1 - p).$$

For any event $A \in \mathcal{F}$, we define the indicator function 1_A, or I_A, to be equal to one if A occurs, and zero otherwise. That is,

$$1_A(\omega) = \left\{ \begin{array}{ll} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{array} \right. .$$

1_A is a Bernoulli random variable with parameter $P(A)$.

Binomial random variables: Consider n independent repeated trials of a Bernoulli random variable. Let X be the number of "successes" (i.e. 1's) in the n trials.

The range of X is $\{0, 1, \ldots, n\}$ and the probability mass function is

$$P\{X = k\} = \left\{ \begin{array}{ll} \binom{n}{k} p^k (1-p)^{n-k} & , \quad \text{if } k \in \{0, 1, \ldots, n\} \\ 0 & , \qquad \text{else} \end{array} \right. .$$

For a binomial random variable with parameters n and p,

$$\mathbb{E}[X] = np \quad \text{and} \quad \text{Var}(X) = np(1-p).$$

Some common discrete random variables

Poisson random variables: A random variable with range $\{0, 1, 2, \dots\}$ is a *Poisson random variable* with parameter $\lambda > 0$ if

$$P\{X = k\} = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!} & , \quad k = 0, 1, 2, \dots \\ 0 & , \qquad \text{else} \end{cases} .$$

For a Poisson random variable with a parameter of λ,

$$\mathbb{E}[X] = \lambda \quad \text{and} \quad \text{Var}(X) = \lambda.$$

Later we will see the Poisson process, which will be an essential tool in this class.

Some common continuous random variables

Uniform random variables. Consider an interval (a, b), where we will often have $a = 0$ and $b = 1$. The random variable is said to be uniformly distributed over (a, b) if

$$F(t) = \begin{cases} 0 & t < 1 \\ (t - a)/(b - a) & a \leq t < b \\ 1 & t \geq b \end{cases},$$

$$f(t) = F'(t) = \begin{cases} \frac{1}{b-a} & a < t < b \\ 0 & else \end{cases}.$$

For a uniform random variable over the interval (a, b),

$$\mathbb{E}[X] = \frac{a + b}{2} \quad \text{and} \quad \text{Var}(X) = \frac{(b - a)^2}{12}.$$

Normal random variables. Also called *Gaussian* random variables,

1. play a central role in the theory of probability due to their connection to the central limit theorem and Brownian motions.

A random variable X is called a *normal* with mean μ and variance σ^2, and we write $X \sim N(\mu, \sigma^2)$, if its density is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

A *standard normal* random variable is a normal random variable with $\mu = 0$ and $\sigma = 1$. For a normal random variable with parameters μ and σ^2,

$$\mathbb{E}[X] = \mu \quad \text{and} \quad \text{Var}(X) = \sigma^2.$$

Some common continuous random variables

Exponential random variables. Many simulation methods will consists of generating a sequence of correctly chosen exponential random variables.

A random variable X has an *exponential distribution* with parameter $\lambda > 0$ if it has a probability density function

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , \quad x \geq 0 \\ 0 & , \quad \text{else} \end{cases}.$$

For an exponential random variable with a parameter of $\lambda > 0$,

$$\mathbb{E}[X] = \frac{1}{\lambda} \quad \text{and} \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

Theorem

Let U be uniformly distributed on the interval $(0, 1)$. Suppose that $p_k \geq 0$ for each $k \in \{0, 1, \ldots, \}$, and that $\sum_k p_k = 1$. Define

$$q_k = P\{X \leq k\} = \sum_{i=0}^{k} p_i.$$

Let

$$X = \min\{k \mid q_k \geq U\}.$$

Then,

$$P\{X = k\} = p_k.$$

More than one random variable

To discuss more than one random variable defined on the same probability space $(\Omega, \mathcal{F}, P)$, we need joint distributions.

Definition
Let $X_1, \ldots, X_n$ be discrete random variables with domain Ω. Then

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = P\{X_1 = x_1, \ldots, X_n = x_n\}$$

is called the *joint probability mass function* of $X_1, \ldots, X_n$.

Definition
We say that $X_1, \ldots, X_n$ are *jointly continuous* if there exists a function $f(x_1, \ldots, x_n)$, defined for all reals, such that for all $A \subset \mathbb{R}^n$

$$P\{(X_1, \ldots, X_n) \in A\} = \int \cdots \int_{(x_1,\ldots,x_n \in A)} f(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

The function $f(x_1, \ldots, x_n)$ is called the *join probability density function*.

More than one random variable

Expectations are found in the obvious way.

Theorem
If $h : \mathbb{R}^n \to \mathbb{R}$ then

$$\mathbb{E}[h(X_1, \ldots, X_n)] = \sum_{x_1 \in \mathcal{R}(X_1)} \cdots \sum_{x_n \in \mathcal{R}(X_n)} h(x_1, \ldots, x_n) p_{X_1, \ldots, X_N}(x_1, \ldots, x_n).$$

Corollary
For random variables X and Y on the same probability space

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

More than one random variable

Definition
The random variables X and Y are independent if for any sets of real numbers A and B

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}.$$

This implies that X and Y are independent if and only if

$$p(x, y) = p_X(x)p_Y(y)$$
$$f(x, y) = f_X(x)f_Y(y),$$

for discrete and continuous random variables, respectively.

More than one random variable

Theorem
Let X and Y be independent random variables. Then for all real valued functions g and h,
$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)].$$

One important application of the above theorem is the relation

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

if X and Y are independent.

More generally, if $X_1, X_2, \ldots, X_n$ are independent random variables, then

$$\mathbb{E}[X_1 \cdots X_n] = \mathbb{E}[X_1] \cdots \mathbb{E}[X_n].$$

Variance of linear combinations.

Suppose that

$$X = X_1 + X_2 + \cdots + X_n.$$

We already know that for any X_i defined on the same probability space

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i].$$

For the variance of a linear combination, a direct calculation shows that for $a_i \in \mathbb{R}$,

$$\text{Var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \text{Var}(X_i) + 2 \sum \sum_{i<j} a_i a_j Cov(X_i, X_j),$$

Therefore, if the X_i are pairwise independent,

$$\text{Var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \text{Var}(X_i).$$

Variance of linear combinations.

Example

Let X be a binomial random variable with parameters n and p. Since X is the number of successes in n independent trials, we can write

$$X = X_1 + \cdots + X_n,$$

where X_i is 1 if *ith* trial was success, and zero otherwise.

Since the X_i's are independent Bernoulli random variables and $\mathbb{E}[X_i] = P\{X_i = 1\} = p$:

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} p = np.$$

Because each of the X_i's are independent

$$\text{Var}(X) = \sum_{i=1}^{n} \text{Var}(X_i) = \sum_{i=1}^{n} p(1-p) = np(1-p).$$

$\square$

Variance of linear combinations.

Proposition

Let $X_1, \ldots, X_n$ be n independent random variables with mean μ and variance σ^2. Let $\bar{X} = (1/n)(X_1 + \cdots X_n)$ be the average of the sample. Then

$$\mathbb{E}(\bar{X}) = \mu, \quad Var(\bar{X}) = \frac{\sigma^2}{n}.$$

Proof.

Calculating shows

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{1}{n}n\mu = \mu$$

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{1}{n}(X_1 + \cdots + X_n)\right) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}.$$

$\square$

Weak Law of Large Numbers

Let $X_1, X_2, X_3, \ldots$ be a sequence of independent and identically distributed random variables with $\mu = \mathbb{E}[X_i]$ and $\sigma^2 = \mathrm{Var}(X_i) < \infty$, $i = 1, 2, \ldots$. Then for all $\epsilon > 0$

$$\lim_{n \to \infty} P \left\{ \left| \frac{X_1 + \cdots + X_n}{n} - \mu \right| > \epsilon \right\} = 0.$$

Strong law of large numbers

Let $X_1, X_2, X_3, \ldots$ be a sequence of independent and identically distributed random variables with mean μ. then

$$P \left\{ \lim_{n \to \infty} \frac{X_1 + \cdots + X_n}{n} = \mu \right\} = 1.$$

So

$$\bar{X} = (X_1 + \cdots X_n)/n$$

converges to μ almost surely, or with a probability of one.

Central limit theorem

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each with expectation μ and variance σ^2. Then the distribution of

$$Z_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

converges to the distribution of a standard normal random variable. That is, for any $t \in (-\infty, \infty)$

$$\lim_{n \to \infty} P\{Z_n \leq t\} = \lim_{n \to \infty} P\left\{ \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq t \right\}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx.$$

Stochastic versus deterministic models

Before proceeding too far, it is important to understand the basic terms "stochastic" and "deterministic."

A process is *deterministic* if its future is completely determined by its present and past. Examples of deterministic processes include solutions to differential and difference equations.

Example

The initial value problem

$$\dot{x}(t) = 3x(t) \qquad x(0) = 2,$$

has the solution $x(t) = 2e^{3t}$. $\qquad\qquad$ □

Example

Consider the difference equation

$$F_1 = F_2 = 1$$
$$F_n = F_{n-1} + F_{n-2}, \quad \text{for } n > 2.$$

Then $\{F_n\}_{n=1}^{\infty}$ is the well known Fibonacci sequence: $\{1, 1, 2, 3, 5, 8, \dots\}$. □

Stochastic versus deterministic models

On the other hand, a *stochastic process* is a random process evolving in time.

Informally, this means that even if you have full knowledge of the state of the system (and it's entire past), you can not be sure of it's value at future times.

More formally, a stochastic process is a collection of random variables, $X(t)$ or X_t, indexed by time.

Example

Consider rolling a fair, six-sided die many times, and for $k \in \{1, 2, \dots\}$, let Z_k be the outcome of the kth roll. Let s

$$X_n = \sum_{k=1}^{n} Z_k.$$

• Thus, X_n is the accumulated total of the first n rolls.

• Knowing $X_1 = 3$ only tells you that $X_2 \in \{4, \dots, 9\}$, with equal probability.

• Note that time, indexed here by n, is discrete in that we only update the system after each roll of the die. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Stochastic versus deterministic models

Example

Consider a frog sitting in a pond with *k* lily pads, labeled 1 through *k*.

1. The frog starts the day by sitting on a randomly chosen pad (for example, they could be chosen with equal probability).

2. However, after a random amount of time, the frog will jump to another pad, also randomly chosen.

3. Letting $t = 0$ denote the start of the day, we let $X(t) \in \{1, \ldots, k\}$ denote the lily pad occupied by the frog at time t. In this example, time is naturally continuous.

However, if we are only interested in which lily pad the frog is on after a given number of jumps,

1. then we may let Z_n denote the lily pad occupied by the frog *after the nth jump*,

2. with Z_0 defined to be the starting lily pad.

3. The process Z_n is discrete in time.

4. The processes Z_n and $X(t)$ are clearly related, and Z_n is usually called the *embedded discrete process associated with* $X(t)$.

Example: Bacterial Growth

Let's consider two oversimplified models for bacterial growth (by *growth* here, I mean the growth of the size of the colony, not of an individual bacterium):

- one deterministic

- one stochastic.

We suppose
- there are 10 bacteria at time zero.
- each bacteria divides at an "average" rate of once per three hours.

Deterministic model: a "reasonable" model would be

$$\frac{d}{dt}x(t) = \frac{1}{3}x(t) \qquad x(0) = 10, \tag{2}$$

with solution

$$x(t) = 10e^{t/3},$$

where the units of t are hours.

Example: Bacterial Growth

Stochastic Model: Without going into the finer details, assume

1. Each bacteria divides after a random (independent, exponential) amount of time with an average wait of 3 hours.

Similar to equation (2) for the deterministic model, it is possible to write down systems of equations describing the time evolution of model

1. Evolution of individual sample paths – instance of experiment (like the ODE model)

2. Evolution of the distribution (probability of being in certain states)

Example: Bacterial Growth - evolution of sample paths

- Below is a plot of the solution of the deterministic system versus three different realizations of the stochastic system.
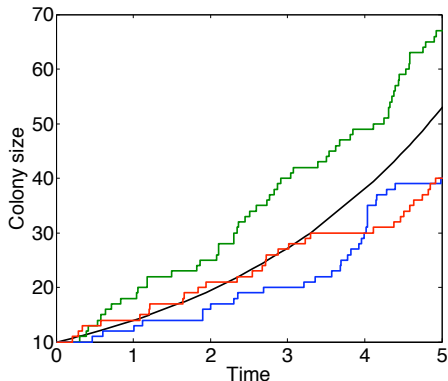
- Stochastic realizations/experiments appear to follow the deterministic system in a "noisy" way.

- It is clear that the behavior of a single realization or experiment of the stochastic system can not be predicted with absolute accuracy.

Example: population growth - evolution of distribution
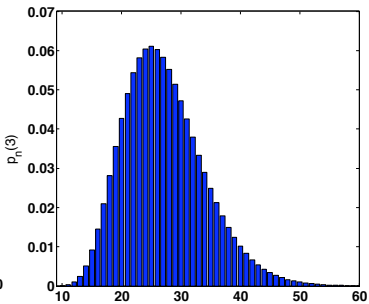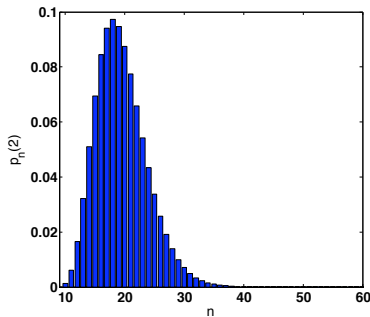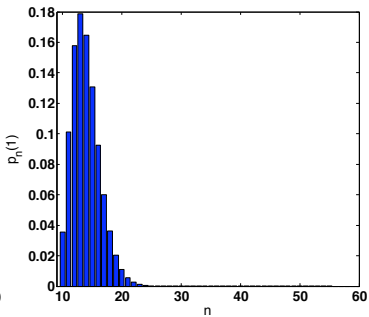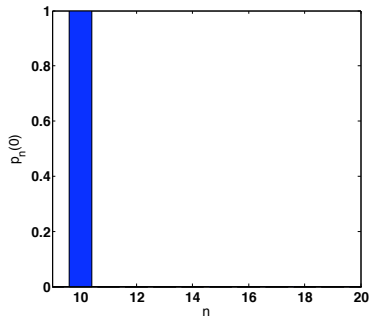
Example: Bacterial Growth and Death

Now suppose that we change the model "slightly" in that:

1. we allow bacteria to die as well as divide.

2. we suppose we begin with only two bacteria.

We suppose that they die after about five hours.

Our new deterministic model could be

$$\dot{x}(t) = \frac{1}{3}x(t) - \frac{1}{5}x(t) = \frac{2}{15}x(t), \qquad x(0) = 2,$$
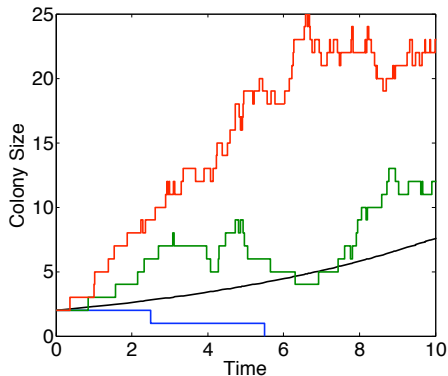
with solution

$$x(t) = 2e^{2t/15}.$$

Example: Bacterial Growth and Death

For the stochastic model, we now model the two possible changes to the size of the colony separately. That is, the next event is *either*

1. a growth event (via a division) or

2. a decrease event (via a death).

Example: Bacterial Growth and Death

▶ Deterministic vs. three realizations/experiments of stochastic system.

▶ The models now behave *qualitatively* differently:

one of the realizations of the stochastic model (i.e. one of the colonies under observation) has been completely wiped out, *something not possible in the deterministic modeling context*.

Example: Lotka-Volterra

Think of A as a prey and B as a predator.

$$A \overset{\kappa_1}{\to} 2A, \qquad A + B \overset{\kappa_2}{\to} 2B, \qquad B \overset{\kappa_3}{\to} \emptyset,$$
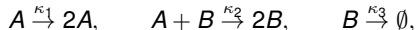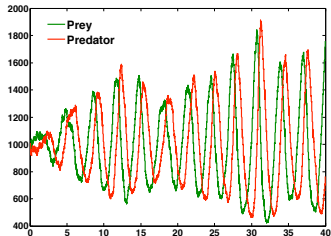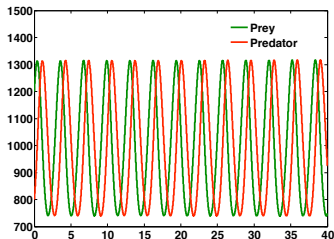
with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.
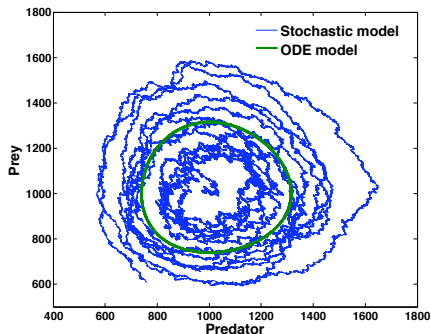
Behavior is qualitatively different:

1. Deterministic always periodic while stochastic oscillates in random way.
2. Predator and/or prey will end up extinct in stochastic model.

Some questions

We start to see some of the different types of questions that become interesting in the stochastic context as opposed to the deterministic:

1. For a given birth and death rate in bacteria example, what is the probability that the colony will eventually die out?

2. For models in which extinction is eventually guaranteed: what is the expected amount of time before extinction?

3. If we know a stochastic processes X_t neither dies out, nor goes to infinity, and if $a < b$ are real numbers, then what is the probability that the value of the process is between a and b for very large t? That is, what is

$$\lim_{t \to \infty} \text{Prob}\{a \le X_t \le b\}?$$

4. How did I make those plots?