Mathematical Framework for Stochastic Processes

Monday, September 14, 2015
2:02 PM

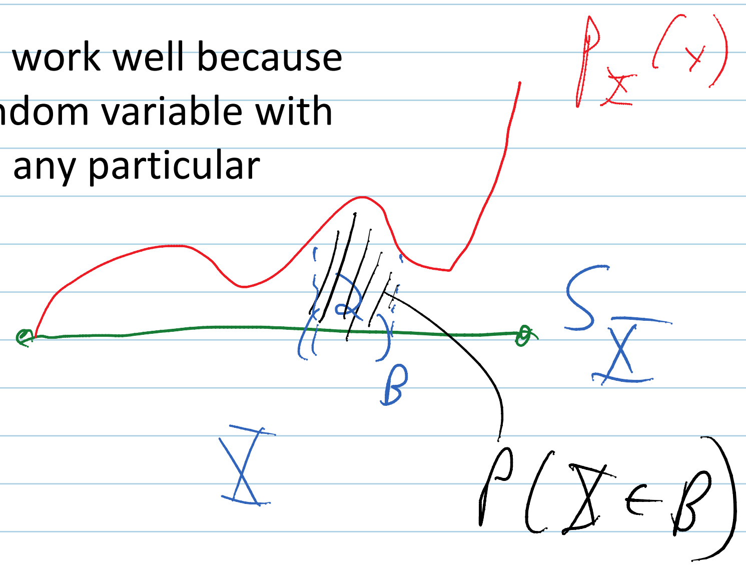
Reading:

- Karlin and Taylor Sec. 1.1-1.3

Homework 1 posted, due Friday, October 2 at 5 PM.

Continuous random variables

Probability mass functions don't work well because typically the probability for a random variable with a continuous state space to take any particular value is just $P(X = x) = 0$.



So we must come up with an alternative way to represent probability distributions for continuous random variables. For an important class of continuous random variables, namely the ``absolutely continuous'' random variables, one can define a probability density function (pdf) $p_X(x)$ which has the property that for any reasonable (Borel) subset $B \subset S_X$:

$$P(X \in B) = \int_B p_X(x)dx$$

PDFs always have the following properties:

- $p_X(x) \geq 0$
- $\int_{S_X} p_X(x)dx = 1$

But PDFs are not probabilities, so they don't have to sum to 1, nor must it be the case that $p_X(x) \leq 1$. No!

Any PDF with the above true properties (and not pathological properties) will give rise to a valid absolutely continuous random variable.

What kind of continuous random variables are not absolutely continuous?

- "singular continuous": rarely occur in practice

If a PDF doesn't describe a probability, then what is it? It describes a density of probabliity just as in continuum mechanics, mass densities describe densities of mass.

That is, if $p_X(x)$ is continuous at a point x_0 (which it doesn't need to be), then one can use the mean value theorem from calculus to show that:

$$\lim_{\epsilon \downarrow 0} \frac{P(|X - x_0| < \epsilon)}{2\epsilon} = \lim_{\epsilon \downarrow 0} \frac{\int_{x_0-\epsilon}^{x_0+\epsilon} p_X(x)dx}{2\,\epsilon} = p_X(x_0)$$

The formulas for expectation also involve integrals:

$$\mathbb{E}\,X = \int_{S_X} p_X(x)\,x\,dx$$

Important examples

1. Uniform distribution

$X \sim U(a,b)$ (uniformly distributed on the interval (a,b))

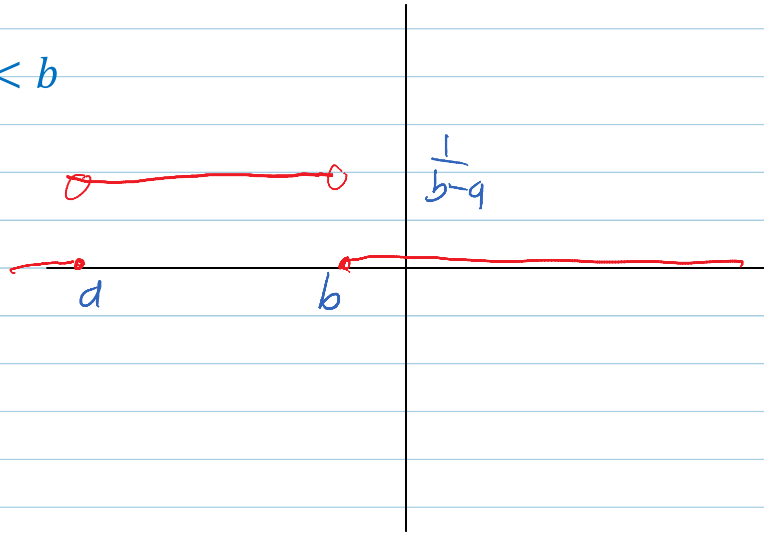
We can take $S_X = \mathbb{R}$ or any subset of $\mathbb{R}$ that contains the interval

(a, b). We'll make the first choice.

The PDF:

$p_X(x) = \frac{1}{b-a}$ for $a < x < b$

0 else



One can calculate that:

$$\mathbb{E}X = \int_{S_X} x\, p_X(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}$$

The primary use of the uniform distribution in this class is that random number generators are typically built up from an algorithm that generates a $U(0,1)$ random variable.

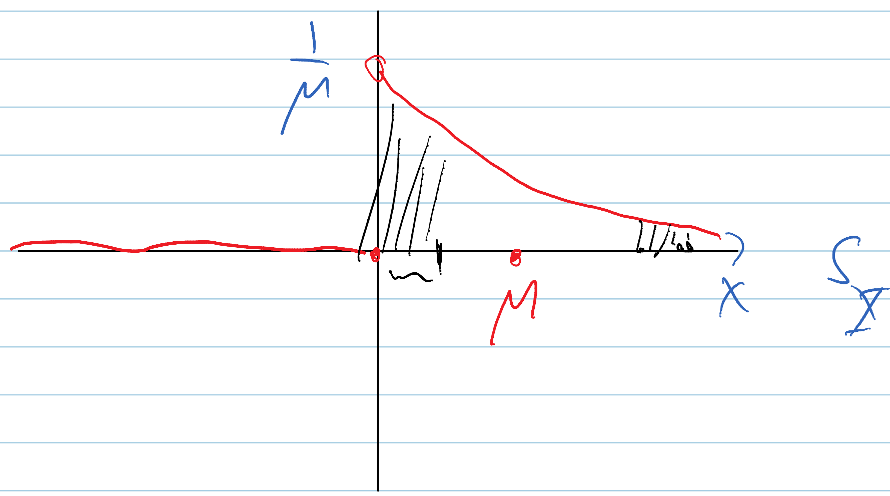
2. Exponential distribution

$X \sim \mathrm{Exp}(\mu)$

$S_X = \mathbb{R}$ (or $\mathbb{R}_+$)

$p_X(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}$ for $x > 0$

0 else



One parameter μ which turns out to be equal to the mean of the random variable:

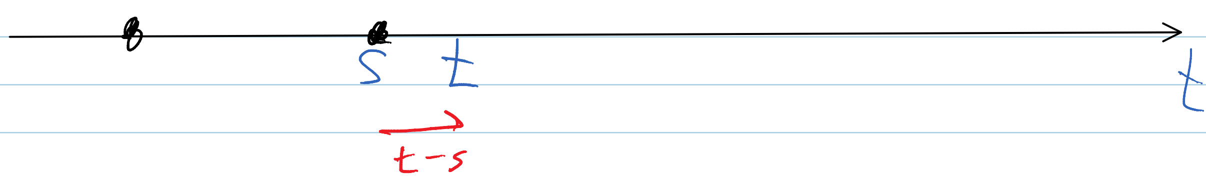
$$\mathbb{E}\,X = \int_{S_X} x\, p_X(x) dx = \int_0^\infty x \frac{1}{\mu} e^{-\frac{x}{\mu}} dx = \mu$$

The exponential distribution has the important property of being memoryless when interpreted as a model for a time delay.

Mathematically, it has the property that, for $0 < s < t$:

$P(T > t | T > s) = P(T > t - s)$ if $T \sim$Exp (μ)

Interpretation: Think of T as the time required to wait until some incident occurs.



Then if we have waited a time s without incident, then we know $T > s$. Then the probability that the incident still has not occurred at a later time t is the same as just recomputing probabilities by starting over from the last observed time s.

Exponentially distributed random variables will arise naturally in stochastic processes where the time to wait for an incident does not depend on the past.

It turns out that, up to weird pathological examples,

- exponential distribution is the only continuous random variable that has the memoryless property
- geometric distribution is the only discrete random

variable that has the memoryless property

$$P(T>t \mid T>s) = \frac{Pr((T>t)\cap(T>s))}{Pr(T>s)}$$

$$= \frac{Pr(T>t)}{Pr(T>s)}$$

$$= \frac{\frac{1}{\mu}\exp(-t/\mu)}{\frac{1}{\mu}\exp(-s/\mu)}$$

$$= \exp(-(t-s)/\mu)$$

$$= P(T>t-s)$$

$$= \int_{t-s}^{\infty} \underbrace{p_T(t')}_{=\frac{1}{\mu}e^{-t'/\mu}}\, dt'$$

Stochastic Process Theory

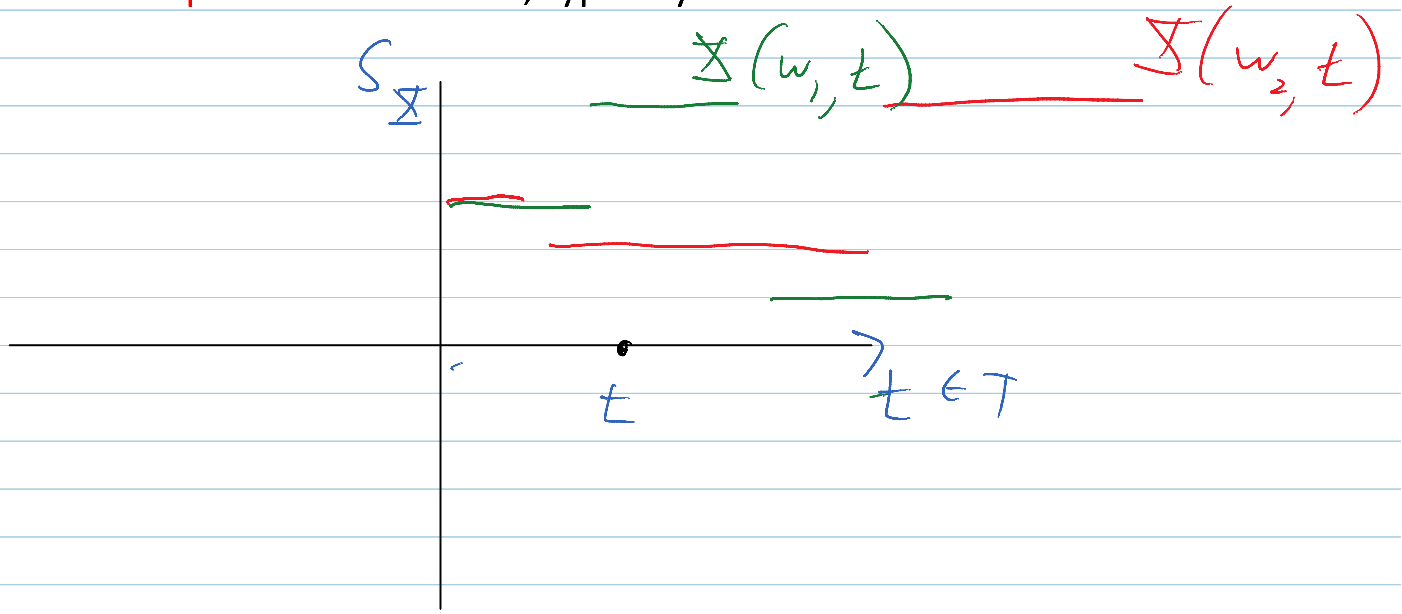
A stochastic process X is a mapping:

$X: \Omega \times T \rightarrow S_X$

where as before:

- Ω is the sample space (of uncertain possibiliites)

- S_X is the state space
- T is the parameter domain, typically time



In practice, there are a few ways to organize the thinking about a stochastic process X:

1. Joint map from sample space and parameter domain into state space:

   $X = X(\omega, t)$ where $\omega \in \Omega$ is a particular realization and $t \in T$ is a particular ``time''

   This is a natural way to think about stochastic or Monte Carlo simulations, where one will typically loop over both the time variable t and the realization ω, the latter by using different random number seeds.
2. Think of X by first viewing it for arbitrary fixed value of $t_1 \in T$, then $X(\omega, t_1)$ is just a random variable. So we can think about a stochastic process as a family of (not necessarily independent) random variables $X = \{X(\cdot, t)\}_{t \in T}$
   This can sometimes work fine for discrete parameter domains T, but is awkward for continuous parameter domains due to the uncountably infinite collection of random variables. Still possible to proceed with this framework; need to impose niceness conditions (separability). When we have separability then the family of random variables $X = \{X(\cdot, t)\}_{t \in T}$ is completely

described by simply defining its behavior for any finite-dimensional distribution $\{X(\cdot,t)\}_{t\in K}$ where K is a finite subset of T. This point of view is mostly valuable for defining Gaussian random fields which generalize multivariate Gaussian random variables to Gaussian random functions. Outside the scope of this class.

3. Think about X for a fixed realization $\omega_1 \in \Omega$; then $X(\cdot,\omega_1)$ is just a trajectory. Then we can think about $X = \{X(\cdot,\omega)\}_{\omega\in\Omega}$ as a random trajectory, i.e., with a state space being a space of trajectories or paths. This is sometimes appealing to mathematicians and theoretical physicists (path integral…).
4. A less frequently used perspective is the ``weak'' perspective, which is to say that any reasonable functional of the stochastic process is a random variable:

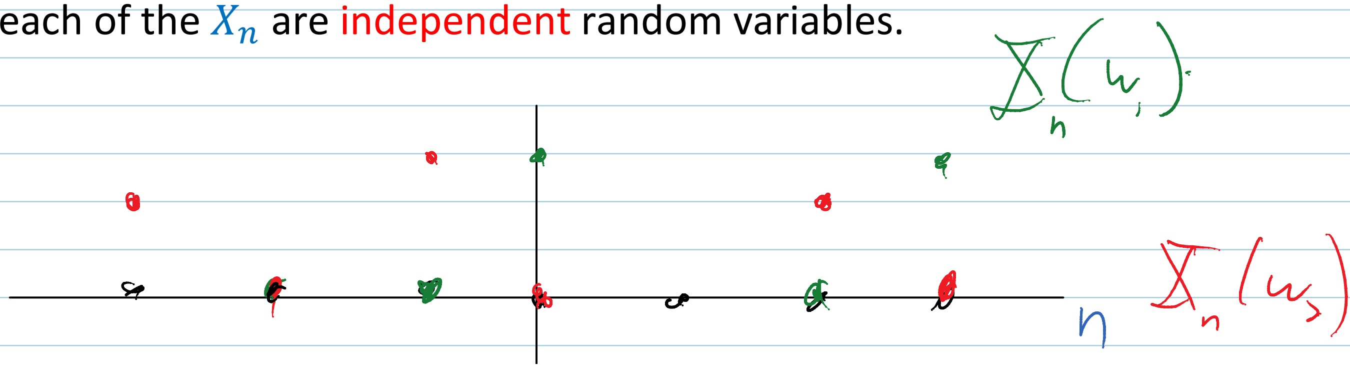
$$X(3), \quad \frac{dX}{dt}(2) + X(3), \quad \int_0^\infty X(t)\,dt$$

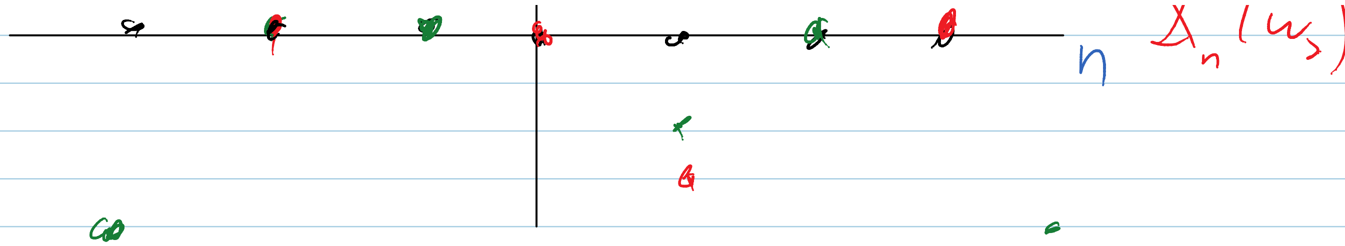
As a first example of a stochastic process, we take a discrete parameter domain:

$T = \mathbb{Z}_{\geq 0} = \{0,1,2,\dots\}$

(though sometimes we also will like to refer to negative times, $T = \mathbb{Z} = \{\dots,-2,-1,0,1,2,\dots\}$

And let us take $\{X_n\}_{n\in T}$ as a stochastic process with the property that each of the X_n are independent random variables.



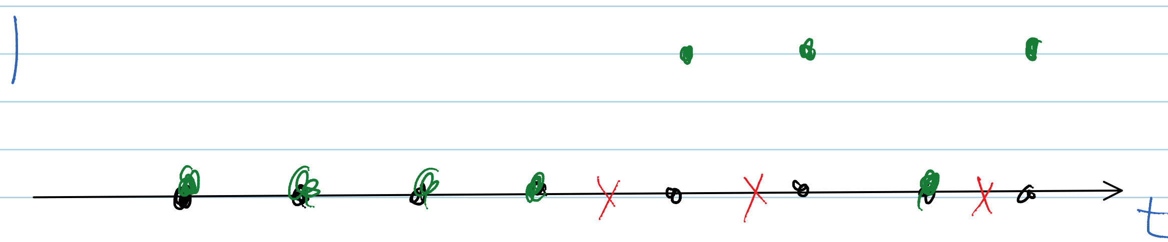


A particular example of such a stochastic process is the Bernoulli process:

- $S_X = \{0,1\}$
- At each $n \in T$, we take X_n to be an independent identically distributed (iid), with pmf:
  - $p_1 = p, p_0 = 1 - p$
    - $P(X_n = 1) = p, P(X_n = 0) = 1 - p$, for all n.
      - if not identically distributed, replace p by p_n

Let's interpret this stochastic model as follows:
Imagine we divide a continuous time interval into equal time segments Δt, and we interpret:

- $X_n = 1$ indicates that an incident occurs during the time interval $[(n-1)\Delta t, n\Delta t)$, while $X_n = 0$ indicates that no incident occurs during that time interval.



Iid random variables in this discrete-time stochastic process reflect a model in which the incidents are occurring independently over different time intervals (no memory!) and are occurring at a constant, homogenous rate.

Two basic questions:

- How many incidents have occurred up to time n?
  - Binomial distribution with n trials and success probability p
- Starting from a given time n, how long do I have to wait until the next incident (in discrete time) happens?
  - Geometric distribution with parameter p describes number of epochs without incident, before the next incident occurs