Sequences of independent, identically distributed random variables

Thursday, September 04, 2008
12:04 PM

Office hours:
Tuesdays 3-4 PM
Wednesdays 3-4 PM

Homework 1 should be posted Friday, September 5 and due Monday, September 22.

Remark: Often times, one discusses random variables without making explicit reference to some mysterious "probability space" Ω.
When this is done, what you can think of is that the state space S is the same as the probability space Ω.

Examples: Random variables on a finite state space; simplify specify the probability

$$p_x = P(X = x) = P_X(\{x\})$$

for each $x \in S$ with

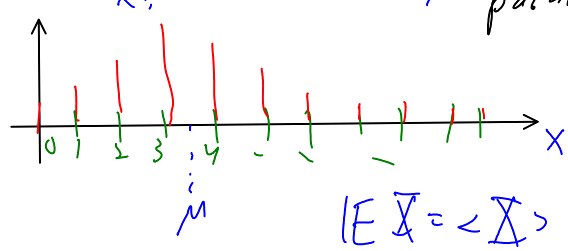
$$\sum_{x \in S} p_x = 1$$

A, C, T, G

Poisson distribution: $S = \mathbb{Z}_{\geq 0} = \{0, 1, 2, \ldots\}$

$$P_X(B) = \sum_{x \in B} p_x \quad \text{(discrete state space)}$$

$$p_x = \frac{\mu^x}{x!} e^{-\mu}$$ where μ is parameter



$$(\mathbb{E} X = \langle X \rangle = \mu)$$

Uniform distribution:

$$X \sim U([a,b])$$

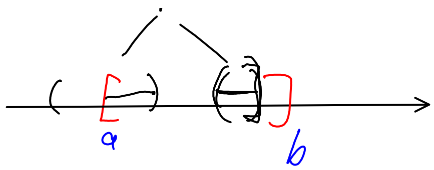
parameters: a, b

$$S = [a,b]$$

$$P_X(B) = \int_{B \cap [a,b]} \frac{1}{b-a}\, dx$$

probability density function (PDF)

B



(pdf)

This is a continuous state space, so probabilities must come from integrating (probability density function), not from summing probabilities for individual values.

$$P_X(\{x\}) = P(X = x) = 0$$

Remark: Could also take $S = \mathbb{R}$
And then just include an indicator function of the interval $[a,b]$
In the probability density function.

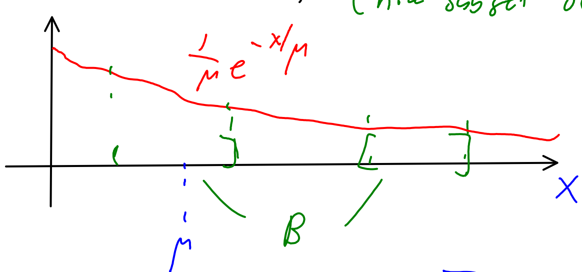
Remark: Can define uniform distribution over any reasonable domain, even in multiple dimensions similarly.

Exponential distribution: $S = \mathbb{R}_{\geq 0} = [0, \infty)$

$$P_X(B) = \int_B \frac{1}{\mu} e^{-x/\mu}\, dx$$

probability density function

for any $B \in \mathcal{B}(S)$ (nice subset of S)



Parameter $\mu = \mathbb{E} X = \langle X \rangle$

Gaussian (normal) distribution: don't think we'll need it in this class...

Now let's consider an elementary stochastic process which is simply a collection of independent, identically distributed random variables (continuing from the last lecture).

$$\{X_n\}_{n=0}^{\infty} \; ; \; T = \{0, 1, 2, \ldots\}$$

$$S = \text{arbitrary}$$

Examples:

Simple model for a DNA sequence, under hypothesis that the particular nucleotide (A,C,T,G) at one location does not have any connection with nucleotides at other locations.

$$X_n \in \{A, C, T, G\} \text{ (or } \{1, 2, 3, 4\})$$

$p_A = .3, \quad p_C = .2, \quad p_T = .23, \quad p_G = .27$

Consider a server which keeps track of the number of service requests in each regular time period (each day)

$$X_n = \# \text{ demands in day } n$$

Consider a sequence of time intervals between important events (decays from a radioactive nucleus, arrival of a new service request, arrival of a bus in a poorly managed transportation system)

$$X_n = \text{times between event } n \text{ and } n+1$$

(also typical to take an exponential distribution for these interevent times…Poisson process)

Consider observations of a microparticle undergoing Brownian motion At regularly spaced time instants $n \Delta t$

$$\vec{X}_n = \vec{X}((n+1)\Delta t) - \vec{X}(n \Delta t)$$

trajectory of Brownian particle in space

$\{\vec{X}_n\}_{n=0}^{\infty}$:

Typically modeled as a sequence of independent, identically distributed Gaussian random variables

The nice thing about sequences of independent random variables is that you can basically compute everything explicitly.

Now we will move to more nontrivial stochastic processes presently, but you will see that typically these stochastic processes are somehow built up from sequences of independent random variables (usually). It's this connection which makes computations possible.

Finite state, discrete time (FSDT) Markov chain

Lawler, Ch.1

$S = \{1, 2, \ldots, M\}$

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

(no problem with $T = \mathbb{Z}$)

There are two fundamental (and equivalent) ways of characterizing and defining a FSDT Markov chain.

Stochastic update rule: (Resnick Sec. 2.1)

Specify functions $f_n(i, z)$ such that

$$f_n : \{1, 2, \ldots, M\} \times S_Z \to \{1, 2, \ldots, M\}$$

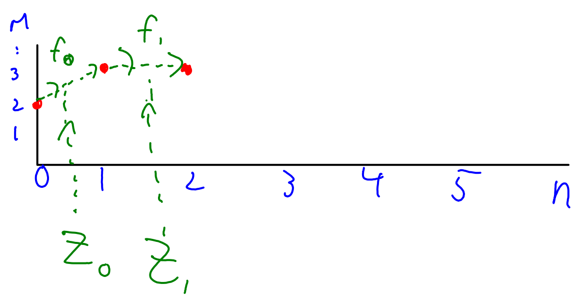
And a sequence of independent, identically distributed random variables $\{Z_n\}_{n=0}^{\infty}$ with $Z_n \in S_Z$

$$X_{n+1} = f_n(X_n, Z_n)$$

Initial condition X_0 is a random variable with specified probability distribution

$$P_{X_0}(\{i\}) = \phi_i \quad \text{for } i \in \{1, \ldots, M\}$$

Intuitively, the stochastic update rule says that given the history of the stochastic process (Markov chain) up to time step (epoch) n, the next value of the Markov chain at epoch n+1 is obtained by applying a certain updating function f_n to the current state (at epoch n) and some independent noise source.



Usually, we will be interested in the special case of time-homogenous Markov chains where the dynamics do not depend explicitly on the epoch ($f_n = f$).

(Dumb) example: $X_{n+1} = (X_n^2 + 2 - Z_n) \pmod{10}$

where $S_Z = \{-1, 0, 2\}$

$$P_{Z_n}(\{-1\}) = \frac{1}{4}, \quad P_{Z_n}(\{0\}) = \frac{1}{3} \quad P_{Z_n}(\{2\}) = \frac{1}{4}$$

and $$X_0 = \begin{cases} 2 & \text{w/ prob. } 1/3 \\ 5 & \text{" } 1/6 \\ 7 & \text{" } 1/2 \end{cases}$$

$$S = \{0, 1, 2, \ldots, 9\}$$

$$\vec{\phi} = \left(\underset{\substack{\uparrow \\ \phi_0}}{0}, 0, \underset{\substack{\uparrow \\ \phi_2}}{1/3}, 0, 0, \underset{\substack{\uparrow \\ \phi_5}}{1/6}, 0, \underset{\substack{\uparrow \\ \phi_7}}{1/2}, 0, 0\right)$$

A sequence of independent, identically distributed random variables on a finite state space is a special case of a Markov chain with a simplified stochastic update rule:

$$X_{n+1} = f(Z_n)$$

Now let's consider Markov chains from the more classical point of view.

Finite-state, discrete time Markov chains are defined to be those sequences of random variables with finite state space $S = \{1, \ldots, M\}$
On time domain $T = \mathbb{Z}_{\geq 0}$ or $\mathbb{Z}$
Such that the Markov property holds

$$P(X_{n+1} = j \mid \overbrace{X_0 = i_0, X_1 = i_1, X_2 = i_2, \ldots,}^{\text{irrelevant for future}} X_n = i_n)$$

$$= P(X_{n+1} = j \mid X_n = i)$$

Intuitively, one formulation of the Markov property says that given the present state of the system, further information about the past is irrelevant for predicting the future.

Another formulation which can be proven to be equivalent:

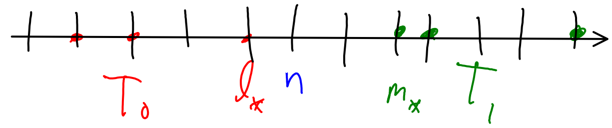
Let T_0 be a subset of "past times" and T_1 a subset of "future times"

$$T_0, T_1 \subseteq T$$

$$l_x = \max\{l : l \in T_0\}$$

$$m_x = \min\{m : m \in T_1\}$$

$$l_x < n < m_x$$



$$P\left(X_m = j_m \text{ for } m \in T_1 \text{ and } X_l = j_l \text{ for } l \in T_0 \,\middle|\, X_n = i_n\right)$$
$$= P\left(X_m = j_m \text{ for } m \in T_1 \,\middle|\, X_n = i_n\right)$$
$$\times\, P\left(X_l = j_l \text{ for } l \in T_0 \,\middle|\, X_n = i_n\right)$$

Given the present state of the system, the future is (conditionally) independent of the past.

This formulation of the Markov property helps to show that there really isn't a direction of time inherent in a Markov chain. It turns out for example, that for a Markov chain (i.e. stochastic property satisfying the Markov property), one can just as well define a stochastic update rule that runs backward in time. It's not the same stochastic update rule as the forward stochastic update rule. For more discussion on this look up "reversible Markov chains." The only thing that really sets a direction of time is the imposition of an initial condition.

The Markov property point of view suggests another way to define a Markov chain model as an (equivalent) alternative to the stochastic update rule. We will focus on time-homogenous Markov chains. Then to define a Markov chain model, one just needs to specify two things:

a. Probability transition matrix P $(M \times M)$

$$P_{ij} = P\left(X_{n+1} = j \,\middle|\, X_n = i\right) \quad \text{for } i,j \in \{1,\ldots,M\}$$

would also depend on n if not time-homogenous

b. Initial probability distribution ϕ (M elements)

$$\phi_j = P(X_0 = j) \quad \text{for } j \in \{1,\ldots,M\}$$

The restrictions in defining these objects:

$$P_{ij} \geq 0$$
$$\sum_{j=1}^{M} P_{ij} = 1$$
$$\phi_j \geq 0$$
$$\sum_{j=1}^{M} \phi_j = 1$$

Any arbitrary choice of initial probability distribution and probability transition matrix satisfying these conditions defines a legitimate time-homogenous FSDT Markov chain.