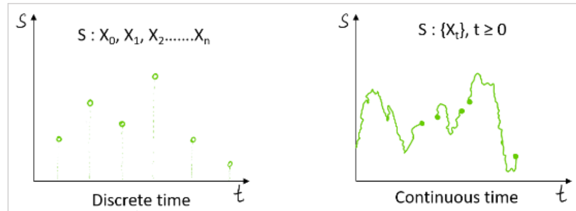
Deepthi Rajashekar
Statistical Learning Study Group

Definition

Stochastic processes (SP) refer to the study of probabilistic systems that are evolving over time. In other words, stochastic processes are a collection of random variables indexed by time. Unlike deterministic models, they model the dynamics of a system's behaviour and hence deemed more realistic. Mathematically, they are a probability distribution over a space of paths.



Use

- Most suited for to model uncertain and varying factors of a system [population growth]
- Analyzing the long-term behaviour of a system [stock price in 6 months]
- Determining the extreme events of a system [market crash]

Classification

The set of all possible values assumed by the stochastic process is called the state space. The set of all possible indexing values (say time) is called as the parameter space. All SPs can be categorized by the nature of their state/parameter spaces as shown alongside.

PARAMETER SPACE

STATE SPACE	Discrete	Continuous
Discrete	Monthly consumer preference	# jobs waiting at *any time* of day
Continuous	*exact* amount of demand each day of the month	waiting time of an arriving job before it is serviced

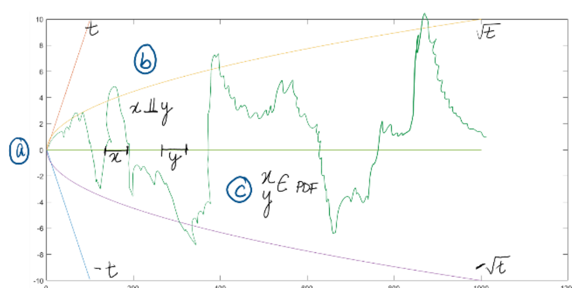
Discrete time stochastic processes

Simple random walk

Consider an independent and identically distributed (i.i.d.) discrete random variable $Y_i = \begin{cases} +1, & with\ prob\ 0.5 \\ -1, & with\ prob\ 0.5 \end{cases}$

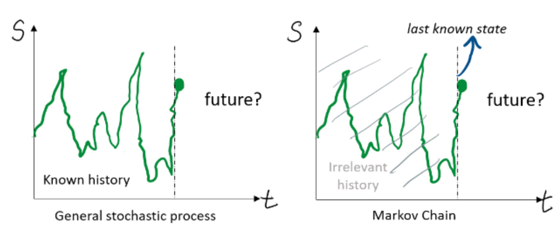
For each t, $X_t = \sum_{i=1}^{t} Y_i$; then S: {X_0, X_1, X_2, ……} is a simple random walk. For very long periods of time, (by the central limit theorem), the following properties hold true:

a. $\mathbb{E}(X_k)$= 0 (with variance = t)
b. Independent increments (paths in different non-overlapping intervals are independent)
c. Stationary (starting point is irrelevant to the distribution)



Markov chain

A stochastic process is called a Markov chain when the state known for *any* specific time is enough to predict the behaviour of the system beyond that point. That is, for each t, X_t only depends on the most recent known state. Formally, for all n:



$$P(X_{t+1} = j | X_1 = i_1, X_2 = i_2, \dots \dots X_t = i_t) = P(X_{t+1} = j | X_t = i_t)$$

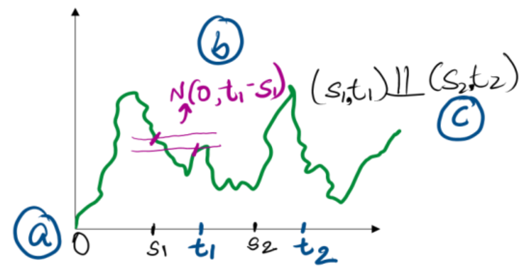
Markov chains with *finite* state spaces can be represented with the transition probability matrix, A = $\begin{pmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{m1} & \cdots & P_{mn} \end{pmatrix}$. Using this matrix, the long-term behaviour of the system (eigen-vector), and all n^{th}-order dependencies (multiplication) can be modelled. A simple random walk is a Markov chain, without a transition probability matrix (because, infinite state space).

Continuous time stochastic processes

Brownian process

Brownian process is the limit of a simple random walk.
Formally, a probability distribution over continuous functions B : $\mathbb{R} \geq 0 \rightarrow \mathbb{R}$, such that:

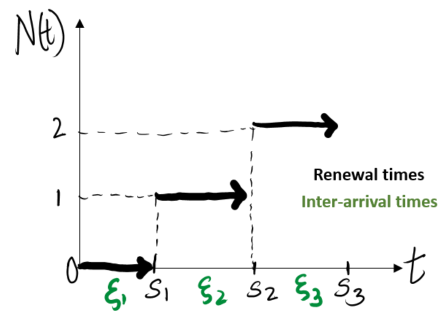
a. $P(B_0) = 1$ (always starts at 0)
b. Independent increments
c. Stationary (difference in intervals is normal); is called Brownian Motion *aka* Weiner Process.
d. Continuous but not differentiable anywhere (needs Ito's calculus)



Poisson process

Poisson process is a simple counting process in which the inter-arrival times is a Poisson distributed random variable.
Formally, if N(t) := the number of events in time (0,t] for all t ≥ 0 then,

$$P(N_t = k) = e^{-\lambda t} . \frac{\lambda t^k}{k!}, k = 0,1,2, \dots.$$



Gaussian process

The central idea of this continuous state and continuous parameter stochastic process is as follows:

a. If the priors are Gaussian, the average prediction from every possible model in the universe will yield a gaussian.
b. For each time, predictions are made without having to learn *a singular* model.
c. The strength of each prediction is influenced by the P(model parameters|Data).

Formally, $P(y|x, D) = \int P(y|x, w)P(w|D)dw$; which simplifies to:$p(y|x, D) \sim N(\mu, \Sigma)$, where ∑ is the covariance matrix. This covariance matrix constitutes the *kernel* (such as radial basis functions) to make the predictions possible!

References:

1. Bhat, U. Narayan, and Gregory K. Miller. *Elements of applied stochastic processes*. Vol. 3. Hoboken^ eN. JNJ: Wiley-Interscience, 2002.
2. Rasmussen, Carl Edward. "Gaussian processes in machine learning." *Summer School on Machine Learning*. Springer, Berlin, Heidelberg, 2003.
3. MIT open courseware (18.S096 Fall 2013); Cornell (Cornell CS4780 SP17)