STAT40810 — Stochastic Models

Brendan Murphy

Week 8

Stochastic Processes

Stochastic processes

Definition - stochastic process

A stochastic process (or a random process) is a collection of random variables indexed by time, $\{X_t, t \in \mathcal{T}\}$. The set of all possible values for the X_t is called the state space, $\mathcal{X}$.

A stochastic process can also be seen as a random function $\mathcal{T} \to \mathcal{X}$.

Reminder - discrete/continuous space/time

Discrete / continuous state space

If $\mathcal{X}$ is

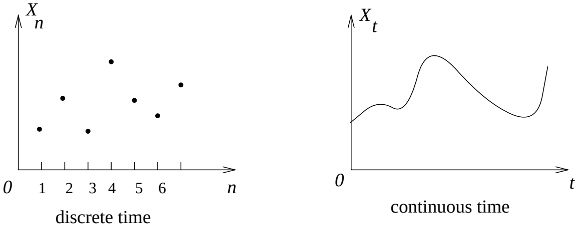
- an interval $[a, b]$, or the set of real numbers $\mathbb{R}$, or $\mathbb{R}^k$: "continuous state space".
- a finite or countable set: "discrete state space".

Discrete / continuous time

If $\mathcal{T}$ is

- an interval $[a, b]$ or the set of real numbers $\mathbb{R}$: "continuous time".
- a finite or countable set: "discrete time".

Illustration



This figure shows a sample path for each of two stochastic processes.

Example 1: Euro/dollar exchange rate

Figure : Euro/dollar exchange rate (ECB data).

- The state space is continuous, $\mathcal{X} = \mathbb{R}_+$.
- As the exchange rate is updated with very high frequency, we can view it as a continuous time process.

Example 2: Annual inflation rate in the USA

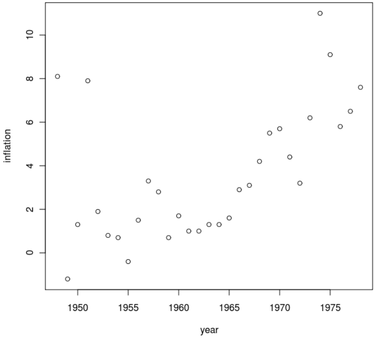


Figure : Inflation (source: Economic Report of the President 2004).

- The state space is continuous, $\mathcal{X} = \mathbb{R}$.
- Discrete time: $t \in \mathcal{T} = \{1948, 1949, \dots, 1978\}$.

Example 3: Annual number of major earthquakes

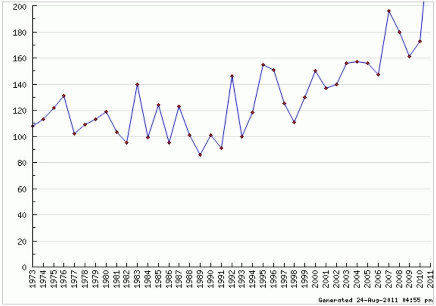


Figure : Magnitude ≥ 6. Source dlindquist.com.

- The state space is discrete, $\mathcal{X} = \mathbb{N}$.
- Discrete time: $t \in \mathcal{T} = \{1973, \ldots, 2011\}$.

Example 4: Clickstream Data

- Heckerman et al. recorded the click patterns of users on MSNBC.
- The category of consecutive webpages visited was recorded:

```
1: frontpage
2: news
3: tech
4: local
5: opinion
6: on-air
7: misc
8: weather
9: msn-news
10: health
11: living
12: business
13: msn-sports
14: sports
15: summary
16: bbs
17: travel
```

- For example (user 1):

```
12 12 12 14 1 12 15 2 15 2 12 14 7 4 3 11 10 6 5 8 5 2 2 2 1 2 2 15 2 12 12 12 12 12 12
12 12 12 12 12 12 12 12 3 3 3 3 3 3 9 13 13 8 13 8 12 12 9 9 13 13 13 13 8 8 3 14
```

- For example (user 2):

```
9 9 13 13 13 13 8 8 3 14 9 13 13 8 3 11 2 2 2 10 9 13 7 9 9 9 9 1 8 2 12 7 4 1 11 5 6 9
9 9 4 4 7 13 12 11 13 13 13 13 13 13 13 14 13 14 14 14 13 7 8
```

Example 5: DNA Sequences

- DNA is a long polymer made from units called nucleotides.
- There are four nucleotides:

```
A: adenine
C: cytosine
G: guanine
T: thymine
```

```
TGGCGCTGGG CGCAATGCGC GCCATTACCG AGTCCGGGCT GCGCGTTGGT GCGGATATCT
CGGTAGTGGG ATACGACGAT ACCGAAGACA GCTCATGTTA TATCCCGCCG TTAACCACCA
TCAAACAGGA TTTTCGCCTG CTGGGGCAAA CCAGCGTGGA CCGCTTGCTG CAACTCTCTC
AGGGCCAGGC GGTGAAGGGC AATCAGCTGT TGCCCGTCTC ACTGGTGAAA AGAAAAACCA
CCCTGGCGCC CAATACGCAA ACCGCCTCTC CCCGCGCGTT GGCCGATTCA TTAATGCAGC
TGGCACGACA GGTTTCCCGA CTGGAAAGCG GGCAGTGAGC GCAACGCAAT TAATGTGAGT
TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC TTCCGGCTCG TATGTTGTGT
GGAATTGTGA GCGGATAACA ATTTCACACA GGAAACAGCT A
```

- DNA can be viewed as a stochastic process where position in the sequence takes the role of time.

Example 1: white noise

Definition - white noise

The stochastic process $\{X_t, t \in \mathcal{T}\}$, with $\mathcal{T} = \mathbb{N}$ or $\mathcal{T} = \mathbb{Z}$, is said to be a white noise if the variable X_t's are iid (independent and identically distributed).

Definition - zero-mean white noise

A white noise with $\mathbb{E}(X_t) = 0$.

Definition - symmetric white noise

A white noise where the distribution of X_t is symmetric.

Example: $X_t \sim \mathcal{N}(0, \sigma^2)$.

Example 2: random walk

Definition - random walk

Let $\{Z_t, t \in \mathcal{T}\}$ be a white noise with $\mathcal{T} = \mathbb{N}$, put

$$X_t \equiv \sum_{i=1}^{t} Z_i = X_{t-1} + Z_t$$

The stochastic process $\{X_t\}$ is called a random walk, and the variables Z_t are called the increments or steps of the walk.

Definition - simple random walk

$Z_t = 1$ with probability p, $Z_t = -1$ with probability $1 - p$.

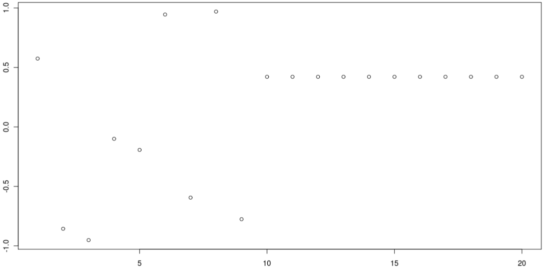
Definition - symmetric random walk

The distribution of Z_t is symmetric.

Notion of stationarity - main idea

Intuitively, a process is said to be stationary if its statistical properties do not change with time.

One could guess that imposing that all the X_t's have the same distribution is a good mathematical interpretation of this idea. However, it is not the case. E.g., X_1, ..., X_{10} are iid $\mathcal{N}(0, \sigma^2)$ and $X_{10} = X_{11} = X_{12} = \ldots$.



Stationarity

This is why the definition of stationarity is a bit more involved. It is, however, one of the most important notions about stochastic processes.

Definition - stationarity

The process $\{X_t\}$ is said to be stationary if for any t_1, ..., t_k and t,

$$(X_{t_1}, \ldots, X_{t_k}) \text{ and } (X_{t_1+t}, \ldots, X_{t_k+t})$$

have the same distribution.

Examples

- It is obvious that any white noise $\{X_t\}$ is stationary.
- One show that a random walk is never stationary (except in the degenerate case where the increments are not random $Z_t = 0$).

Markov property

Definition - Markov property

Let $\{X_t, t \in \mathcal{T}\}$ be a discrete time process with $\mathcal{T} = \mathbb{N}$ or $\mathcal{T} = \mathbb{Z}$. It is said to have the Markov property if for any $n \in \mathcal{T}$,

$$p(X_{n+1}|X_n, X_{n-1}, X_{n-2}, \dots) = p(X_{n+1}|X_n).$$

"If we know the current state of the process, what happened before is irrelevant with respect to the future."

Examples: a white noise and a random walk always satisfy the Markov property.

Markov property: Further result

Theorem

Let $\{X_t\}$ satisty the Markov property then for any $n \in \mathcal{T}$ and $k \in \mathbb{N}$,

$$p(X_{n+k}|X_n, X_{n-1}, X_{n-2}, \dots) = p(X_{n+k}|X_n).$$