Probability and Stochastic Processes - Lecture 17

Michael Evans
University of Toronto
https://utstat.utoronto.ca/mikevans/stac62/staC622021.html

2021

Definition III.6.1 Suppose $\{(t, X_t) : t \in T\}$ is a stochastic process such that $E(X_t^2) < \infty$ for all $t \in T$. Then define the *mean function* by $\mu : T \to R^1$ by $\mu(t) = E(X_t)$ and the *autocovariance function* by $\sigma : T \times T \to R^1$ by $\sigma(s, t) = Cov(X_s, X_t)$ provided these expectations exist. The *autocorrelation function* $\rho : T \times T \to R^1$ is defined by $\rho(s, t) = \sigma(s, t)/\sigma^{1/2}(s, s)\sigma^{1/2}(t, t)$ provided $\sigma(t, t) > 0$ for every $t \in T$. ■

Example III.6.1 *i.i.d. process*

- the r.v.'s $\{X_t : t \in T\}$ are mutually statistically independent with each $E(X_t) = m$ and $Var(X_t) = v$

- then $\mu(t) = E(X_t) = m$ and

$$\sigma(s,t) = \begin{cases} v & s = t \\ 0 & s \neq t \end{cases}, \rho(s,t) = \begin{cases} 1 & s = t \\ 0 & s \neq t \end{cases}$$

- for Bernoulli(p) process $m = p, v = p(1-p)$ ■

Example III.6.2 *Gaussian processes*

- recall the r.v.'s $\{X_t : t \in T\}$ are such that for any $\{t_1, \ldots, t_n\} \subset T$ then

$$\begin{aligned}(X_{t_1}, \ldots, X_{t_n}) &\sim N_n\left(\left(\begin{array}{c}\mu(t_1)\\ \vdots \\ \mu(t_n)\end{array}\right), \left(\begin{array}{ccc}\sigma(t_1,t_1) & \cdots & \sigma(t_1,t_n)\\ \vdots & & \vdots \\ \sigma(t_n,t_1) & \cdots & \sigma(t_n,t_n)\end{array}\right)\right)\\ &= N_n\left((\mu(t_i)), (\sigma(t_i,t_j))\right)\end{aligned}$$

- note that a Gaussian process is completely specified by the mean and autocovariance functions ■

- so if we specify $\mu : T \to R^1$ and $\sigma : T \times T \to R^1$ have we correctly defined a Gaussian process?

- there are no restrictions on μ but σ has to have the property that for any $\{t_1, \ldots, t_n\} \subset T$ then the $n \times n$ matrix $(\sigma(t_i, t_j))$ is symmetric and positive semidefinite

- so a function $\sigma : T \times T \rightarrow R^1$ is a valid autocovariance function whenever $\sigma(s,t) = \sigma(t,s)$ and for any $\{t_1, \ldots, t_n\} \subset T$ and $\mathbf{c} = (c_1, \ldots, c_n)' \in R^n$ then

$$\sum_{i=1}^{n} \sum_{i=1}^{n} c_i c_j \sigma(t_i, t_j) = \mathbf{c}'(\sigma(t_i, t_j))\mathbf{c} \geq 0$$

- a Gaussian process exists with given time domain T, mean function $\mu : T \rightarrow R^1$ and autocovariance function $\sigma : T \times T \rightarrow R^1$ since if $\{Z_t : t \in T\}$ is a collection of i.i.d. $N(0,1)$ random variables, then define, for any $\{t_1, \ldots, t_n\} \subset T$,

$$\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_n} \end{pmatrix} = \begin{pmatrix} \mu(t_1) \\ \vdots \\ \mu(t_n) \end{pmatrix} + \begin{pmatrix} \sigma(t_1, t_1) & \cdots & \sigma(t_1, t_n) \\ \vdots & & \vdots \\ \sigma(t_n, t_1) & \cdots & \sigma(t_n, t_n) \end{pmatrix}^{1/2} \begin{pmatrix} Z_{t_1} \\ \vdots \\ Z_{t_n} \end{pmatrix} \tag{1}$$

and this is a valid s.p. by what we have proved about marginalizing the multivariate normal and the Kolmogorov Consistency Theorem

- note - (1) gives a method for simulating a Gaussian process (not necessarily the best way for large n)

- suppose $T = [0, \infty)$

- since T is a continuous, unbounded set we can't generate a full sample function

- so choose $t_{up} \in T$ and $N \in \mathbb{N}$ and put $t_i = t_{up}(i-1)/2^N$ for $i = 1, \ldots, 2^N + 1$

- then generate the Z_{t_i} i.i.d. $N(0,1)$ and use (*) to get the values of the X_{t_i} plotting the points (t_i, X_{t_i}) to approximate a sample function

- the following is an example of a sample function, with $t_{up} = 2.5, N = 10$, of a Brownian motion $\{(t, B_t) : t \in [0, \infty)\}$ which is a Gaussian process with

$$T = [0, \infty), B_0 = 0, \mu(t) = 0, \sigma(s,t) = \tau^2 \min(s,t)$$

and $\tau^2 > 0$

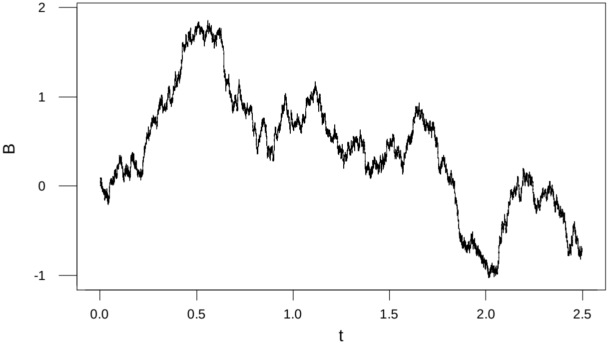

Figure: Simulated Brownian motion.

Definition III.6.2 When $T \subset R^k$ a process with mean function μ and autocovariance function σ is called *weakly stationary* if $\mu(t)$ is constant in t and $\sigma(s,t) = \kappa(s-t)$ for some $\kappa : R^k \to R^1$. ■

note - κ must satisfy $\kappa(0) \geq 0$, $\kappa(t) = \kappa(-t)$ and for all $\{t_1, \ldots, t_n\} \subset T$ and $\mathbf{c} = (c_1, \ldots, c_n)' \in R^n$ then $\sum_{i=1}^n \sum_{j=1}^n c_i c_j \kappa(t_i - t_j) \geq 0$ and such a κ is called a *positive semidefinite function* (positive definite when corresponding matrices are p.d.)

- there are theorems concerning such κ, for example, $\kappa(t) = \exp(-\tau^2 ||t||^2)$ for $\tau^2 > 0$ is positive definite

Example III.6.3 *Random walks*

- suppose the r.v.'s $\{Z_t : t \in \mathbb{N}\}$ are i.i.d. with mean and variance

- then the process $\{(t, X_t) : t \in \mathbb{N}\}$ defined by $X_t = \sum_{i=1}^{t} Z_i$ is called a *random walk* (starting from 0)

- a *simple random walk* arises when $Z_t \sim -1 + 2\text{Bernoulli}(p)$ so $P(Z_t = -1) = 1 - p, P(Z_t = 1) = p$ and so for the random walk

$$\begin{aligned}
\mu(t) &= E(X_t) = \sum_{i=1}^{t} E(Z_t) = tE(Z_1) = t(-(1-p)+p) = (2p-1)t \\
\sigma(s,t) &= Cov(X_s, X_t) = Cov\left(\sum_{i=1}^{s} Z_i, \sum_{j=1}^{t} Z_j\right) = \sum_{i=1}^{s}\sum_{j=1}^{t} Cov(Z_i, Z_j) \\
&= \sum_{i=1}^{\min\{s,t\}} Var(Z_i) = \min\{s,t\} Var(Z_1) = 4p(1-p)\min\{s,t\}
\end{aligned}$$

so not weakly stationary

$$\rho(s,t) = \frac{4p(1-p)\min\{s,t\}}{\sqrt{4p(1-p)s}\sqrt{4p(1-p)t}} = \frac{\min\{s,t\}}{\sqrt{st}}$$

- a *Gaussian random walk* when $\{Z_t : t \in \mathbb{N}\}$ are i.i.d. $N(m, \tau^2)$

$$\mu(t) = E(X_t) = \sum_{i=1}^{t} E(Z_t) = tE(Z_1) = mt$$

$$\sigma(s,t) = Cov(X_s, X_t) \overset{\text{as above}}{=} \min\{s,t\} Var(Z_1) = \tau^2 \min\{s,t\}$$

$$\rho(s,t) = \frac{\tau^2 \min\{s,t\}}{\sqrt{\tau^2 s}\sqrt{\tau^2 t}} = \frac{\min\{s,t\}}{\sqrt{st}}$$

- in general

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_t \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & & \vdots & 0 \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_t \end{pmatrix} = A\mathbf{Z}_t$$

and the finite joint distributions of $\{X_t : t \in \mathbb{N}\}$ are defined consistently and so by KCT this defines a s.p. and it is a Gaussian process ■

Example III.6.4

- suppose the r.v.'s $\{Z_t : t \in \mathbb{Z}\}$ are i.i.d. $N(0, \tau^2)$ and define $\{X_t : t \in \mathbb{Z}\}$ by $X_t = Z_t + \theta Z_{t-1}$ for some constant $\theta \in R^1$ so $\mu(t) = E(X_t) = \mu(t) = E(Z_t) + \theta E(Z_{t-1}) = 0$ and

$$\begin{aligned}
\sigma(s,t) &= Cov(X_s, X_t) = E(X_s X_t) - E(X_s)E(X_t) \\
&= E((Z_s + \theta Z_{s-1})(Z_t + \theta Z_{t-1})) \\
&= E(Z_s Z_t) + \theta[E(Z_s Z_{t-1}) + E(Z_{s-1} Z_t)] + \theta^2 E(Z_{s-1} Z_{t-1}) \\
&= \begin{cases} 0 & s < t-1 \\ \tau^2\theta & s = t-1 \\ \tau^2 + \tau^2\theta^2 & s = t \\ \tau^2\theta & s = t+1 \\ 0 & s > t+1 \end{cases}
\end{aligned}$$

$$\begin{pmatrix} X_t \\ X_{t+1} \\ \vdots \\ X_{t+n} \end{pmatrix} = \begin{pmatrix} \theta & 1 & 0 & \cdots & & 0 \\ 0 & \theta & 1 & \cdots & & 0 \\ \vdots & & & \vdots & & 0 \\ 0 & 0 & & \cdots & \theta & 1 \end{pmatrix} \begin{pmatrix} Z_{t-1} \\ Z_t \\ \vdots \\ Z_{t+n} \end{pmatrix} = A\mathbf{Z}_{t-1,t+n}$$

and so $\{(t, X_t) : t \in \mathbb{Z}\}$ is a Gaussian process

- **note** - $\sigma(s,t) = \kappa(s-t)$ where

$$\kappa(t) = \begin{cases} 0 & t < -1 \\ \tau^2\theta & t = -1 \\ \tau^2 + \tau^2\theta^2 & t = 0 \\ \tau^2\theta & t = 1 \\ 0 & t > 1 \end{cases}$$

and so this is a weakly stationary Gaussian process ■

Exercise III.6.1 If r.v.'s $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ all have finite second moments, then for constants $a_0, a_1, , \ldots, a_m, b_0, b_1, , \ldots, b_n$ prove that.

$$Cov\left(a_0 + \sum_{i=1}^{m} a_i X_i, b_0 + \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j Cov(X_i, Y_j).$$

Exercise III.6.2 If r.v.'s $X_1, \ldots, X_m$ all have finite second moments then for constants $a_0, a_1, , \ldots, a_m$ prove that

$$Var\left(a_0 + \sum_{i=1}^{m} a_i X_i\right) = \sum_{i=1}^{m} a_i^2 Var(X_i) + 2\sum_{i<j} a_i a_j Cov(X_i, X_j).$$

Specialize this result to the case where $X_1, \ldots, X_m$ are mutually statistically independent.

Exercise III.6.3 In Examples III.6.3 and III.6.4 determine the joint distribution of $(X_1, \ldots, X_t)'$ in the Gaussian case.