9831 Probability and Stochastic Processes for Finance, Fall 2016

Lecture 2

Anja Richter

Department of Mathematics

Outline

- Multivariate normal r.v.s
- Brownian motion

1 Multivariate normal random variables

Recall: X is said to have a **(one-dimensional) normal distribution with mean** $\mu \in \mathbb{R}$ **and variance** $\sigma^2 > 0$ if X has density

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

To include the degenerate case $\sigma = 0$, we shall say that X has a degenerate normal distribution if $X = \mu$ a.s..

The degenerate distribution, clearly, has no density, its distribution μ_X is δ_μ.

Definition 1

A random vector $X = (X_1, X_2, \ldots, X_n)$ is said to be **Gaussian (equivalently, normal)** if for every $a \in \mathbb{R}^n$ the random variable

$$\langle a, X \rangle := \sum_{i=1}^{n} a_i X_i$$

has a (one-dimensional, possibly degenerate) normal distribution.

Theorem 2

An n-dimensional Gaussian vector X has a density iff C is non-degenerate.

In this case

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det C}}\, e^{-\frac{1}{2}\langle (x-\mu, C^{-1}(x-\mu)\rangle}.$$

The proof can be found in Jacod/Protter, Probability essentials, Chapter 16, and is based on the following useful fact:

Lemma 3

Let $X = (X_1, X_2, \ldots, X_n)$ be a r.v. with density $f_X(x)$.

Then for every invertible $n \times n$ matrix A and $\mu \in \mathbb{R}^n$ the random vector $Y = \mu + AX$ has density

$$f_Y(y) = \frac{1}{|\det A|}\, f_X(A^{-1}(y - \mu)).$$

Proof: Let $B \in \mathcal{B}^n$, $g(x) = \mu + Ax$, and $D = g^{-1}(B)$.

Then by the change of variable formula we have

$$P(Y \in B) = P(X \in D) = \int_D f_X(x)\,dx = \int_{g^{-1}(B)} f_X(x)\,dx = \int_B f_X(A^{-1}(y - \mu))|\det A^{-1}|\,dy.$$

This implies that Y has density

$$f_Y(y) = f_X(A^{-1}(y - \mu))|\det A^{-1}| = \frac{1}{|\det A|}\, f_X(A^{-1}(y - \mu)).$$

Definition 4

- The **characteristic function of a r.v.** $X = (X_1, X_2, \dots, X_n)$ is

$$\phi_X(t) = E(e^{i\langle t, X\rangle}), \quad t \in \mathbb{R}^n.$$

Characteristic functions uniquely characterize the distribution of a r.v., i.e. if two r.v. have the same characteristic function then they have the same distribution.

Example 5

- $Z \sim N(0,1)$:

$$\begin{aligned}
\phi_Z(t) &= E(e^{itZ}) \\
&= \int_{-\infty}^{\infty} e^{itz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= \int_{0}^{\infty} e^{itz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \int_{-\infty}^{0} e^{itz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= \int_{0}^{\infty} e^{itz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz + \int_{0}^{\infty} e^{-itz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= 2\int_{0}^{\infty} \cos(tz) \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz
\end{aligned}$$

  Compute the derivative and integrate by parts:

$$\begin{aligned}
\phi_Z'(t) &= -\frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{z^2}{2}} z \sin(tz) dz \\
&= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \sin(tz) d\left(e^{-\frac{z^2}{2}}\right) \\
&= \frac{2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sin(tz)\Big|_0^{\infty} - \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{z^2}{2}} t \cos(tz) dz \\
&= -t\phi_Z(t)
\end{aligned}$$

  The ODE $\phi_Z(t) = -t\phi_Z(t)$, has the solution $\phi_Z(t) = ce^{-t^2/2}$. As $\phi_Z(0) = 1$, the integration constant is $c = 1$. Hence we have

$$\phi_Z(t) = e^{-t^2/2}.$$

- $X \sim N(\mu, \sigma^2)$, then

$$\phi_X(t) = E(e^{itX}) = e^{it\mu - \sigma^2 t^2/2}.$$

  Indeed, write $X = \mu + \sigma Z$ where $Z \sim N(0,1)$. Then

$$\phi_X(t) = E(e^{it(\mu+\sigma Z)}) = e^{it\mu} E(e^{it\sigma Z}) = e^{it\mu} e^{-\sigma^2 t^2/2}$$

Theorem 6

$X = (X_1, X_2, \ldots, X_n)$ is a Gaussian vector iff its characteristic function has the form

$$\phi_X(t) = \exp\left(i\langle t, \mu\rangle - \frac{1}{2}\langle t, Ct\rangle \right), \; t \in \mathbb{R}^n, \tag{1}$$

where $\mu \in \mathbb{R}^n$ and C is $n \times n$ symmetric non-negative semi-definite matrix.

C is then the covariance matrix of X and μ is the mean of X.

Proof:

- Suppose (1) holds. Let $a \in \mathbb{R}^n$ and consider $Y = \langle a, X\rangle$. Then for $t \in \mathbb{R}$

  $$\phi_Y(t) = \phi_X(at) = e^{it\langle \mu, a\rangle - t^2\langle a, Ca\rangle/2},$$

  which means that $\phi_Y(t)$ is the characteristic function of a normal r.v. with mean $\langle \mu, a\rangle$ and variance $\langle a, Ca\rangle$.

  Therefore, Y is normal. Since a was arbitrary, we conclude that X is Gaussian.

- Now suppose that X is Gaussian.

  Then we know that for every $a \in \mathbb{R}$ the random variable $Y = \langle a, X\rangle$ is (one-dimensional) normal. We can compute the mean and variance of Y:

  $$\begin{aligned} E(Y) &= \langle a, \mu\rangle \\ \mathrm{Var}(Y) &= E((\langle a, X\rangle - \langle a, \mu\rangle)^2) \\ &= E((a^T(X-\mu))^2) \\ &= E((a^T(X-\mu)(X-\mu)^T a)) \\ &= a^T C a \\ &= \langle a, Ca\rangle, \end{aligned}$$

  where $C = E((X-\mu)(X-\mu)^T)$.

  Since Y is normal,

  $$\phi_Y(t) = E(e^{it\langle a, \mu\rangle - t^2\langle a, Ca\rangle/2}).$$

  Then

  $$\phi_X(a) = E(e^{i\langle a, X\rangle}) = \phi_Y(1) = e^{i\langle a, \mu\rangle - \langle a, Ca\rangle/2},$$

  and we conclude that the characteristic function of X has the claimed form (replace a with t).

Theorem 7

Let $X = (X_1, X_2, \ldots, X_n)$ be a Gaussian vector. The components $X_1, X_2, \ldots, X_n$ are independent iff they are uncorrelated (i.e. C is diagonal).

The proof is based on the fact that r.v. $X_1,\ X_2, \ldots,\ X_n$ are independent if the characteristic function of $(X_1, X_2, \ldots, X_n)$ splits into the product of characteristic functions of $X_1,\ X_2, \ldots,\ X_n$. We omit the details.

Theorem 8

Let X be an n-dimensional Gaussian vector with mean vector μ.

Then there exist independent normal r.v.s $Y_1, Y_2, \ldots, Y_n$ with $Y_i \sim N(0, \sigma_i^2),\ \sigma_i \geq 0$, and an orthogonal $n \times n$ matrix A such that $X = \mu + AY$, where $Y = (Y_1, Y_2, \ldots, Y_n)$.

Proof: Let C be the covariance matrix of X. Since C is symmetric, non-negative semi-definite, there is an orthogonal matrix A and a diagonal matrix D with non-negative diagonal elements such that $C = ADA^{-1}$.

Set $Y = A^T(X - \mu)$. Since X is Gaussian, Y (as an affine transformation of X) is also Gaussian.

Moreover, $E(Y) = 0$ and

$$E(YY^T) = E(A^T(X-\mu)(X-\mu)^T A) = A^T C A = A^T(ADA^T)A = D,$$

since $AA^T = A^T A = I_n$.

By Theorem 7, $Y_1, Y_2, \ldots, Y_n$ are independent and $\mathrm{Var}(Y_i) = D_{ii} =: \sigma_i^2 \geq 0$.

Theorem 9

Let X be an n-dimensional Gaussian vector and Y be an m-dimensional Gaussian vector.

If X and Y are independent then (X, Y) is an $n+m$-dimensional Gaussian vector.

Proof: Let $Z = (X, Y)$. By independence of X and Y,

$$\phi_Z(t) = \phi_X(u)\phi_Y(v), \quad t = (u, v),\ u \in \mathbb{R}^n,\ v \in \mathbb{R}^m.$$

Therefore,

$$\phi_Z(t) = e^{i\langle u, \mu_x\rangle - \frac{1}{2}\langle u, C_X u\rangle} e^{i\langle v, \mu_y\rangle - \frac{1}{2}\langle v, C_Y v\rangle} = e^{-\langle (u,v), (\mu_X, \mu_Y)\rangle - \frac{1}{2}\langle t, Ct\rangle},$$

where

$$C = \begin{bmatrix} C_X & 0 \\ 0 & C_Y \end{bmatrix}$$

By Theorem 6, we conclude that Z is $n+m$-dimensional Gaussian vector.

2 Brownian motion

2.1 Definition

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Suppose that for each $\omega \in \Omega$ there is a continuous function B_t, $t \geq 0$, such that $B_0 = 0$.

Then $B := (B_t)_{t \geq 0}$ is called a **(standard) Brownian motion** if

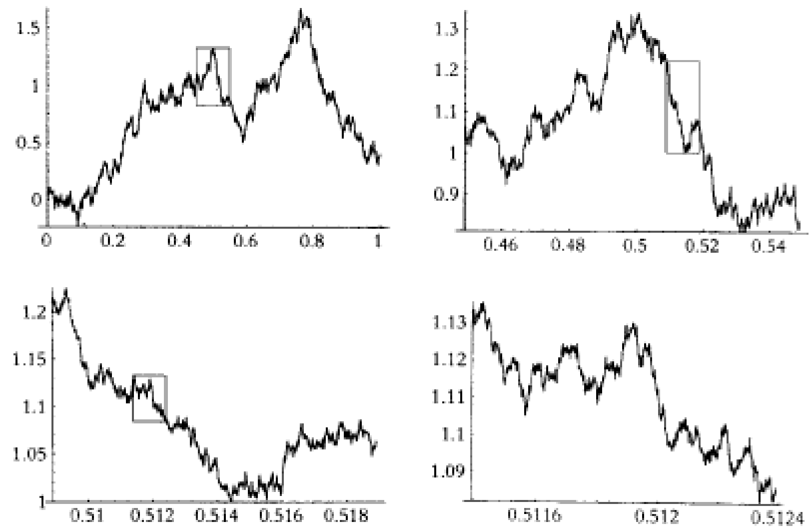
1. for each $m \in \mathbb{N}$ and $t_0 = 0 < t_1 < \cdots < t_n$ the r.v.s
$$B_{t_1},\ B_{t_2} - B_{t_1},\ \cdots,\ B_{t_n} - B_{t_{n-1}}$$
are independent,
2. for all $s > 0,\ t \geq 0$ the r.v. $B_t - B_s$ has a normal distribution with mean 0 and variance $t - s$.

Remark: The standard Brownian motion starts at 0. A Brownian motion starting at $x \neq 0$ is obtain by shifting $x + B_t$.

2.2 First properties of Brownian motion

The following properties hold for Brownian motion B.

- *Symmetry* The process $-B_t$, $t \geq 0$, is a Brownian motion.
- *Time-homogeneity* For any $s > 0$, the process $B_{t+s} - B_s$, $t \geq 0$, is also a Brownian motion.
- *Self-similarity* For every $c > 0$, the process cB_{t/c^2}, $t \geq 0$, is a Brownian motion.



Zooming in on Brownian motion.

- *Time inversion* The process X defined by $X_0 = 0$, $X_t = tB_{1/t}$ for $t > 0$, is a Brownian motion.

Definition

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

A stochastic process $(X_t)_{t \geq 0}$ is called **Gaussian process**, if for each $m \in \mathbb{N}$ and $t_0 = 0 < t_1 < \ldots < t_m$ the random variables $X_{t_1}, X_{t_2}, \ldots, X_{t_m}$ are jointly normally distributed.

Remark:

This means, all its finite dimensional distributions are multivariate normally. Thus, a Gaussian process is fully characterized by its

- mean function $\mu(t) = \mathbb{E}[X_t]$ and
- covariance function $\gamma(t, s) = \mathrm{Cov}(X_t, X_s)$.

Fact: Brownian motion is a Gaussian process, it is fully characterized by the mean and the covariance functions.

- $\mathbb{E}\left[B_t\right] = 0$ for all t
- $\operatorname{Cov}(B_t, B_s) = \min\{s, t\}$

To calculate the covariance, without loss of generality, we assume $s < t$.

$$\begin{aligned} \operatorname{cov}(B_t, B_s) &= \mathbb{E}\left[B_t\, B_s\right] \\ &= \mathbb{E}\left[(B_t - B_s + B_s)B_s\right] \\ &= \mathbb{E}\left[(B_t - B_s)B_s\right] + s \qquad (B_s \sim N(0, s)) \\ &= \mathbb{E}\left[B_t - B_s\right]\, \mathbb{E}\left[B_s\right] + s \quad (\text{ independent increments}) \\ &= s \\ &= \min\{s, t\}. \end{aligned}$$

2.3 Constructing Brownian motion - Donsker's invariance principle

Suppose $\{X_i\}_{i=1}^{\infty}$ is an iid sequence of random variables with mean 0 and and variance 1.

Let $S_n = \sum_{i=1}^{n} X_i$, i.e. S is a symmetric random walk.

Define the function $\mathfrak{S}_n$ of t by

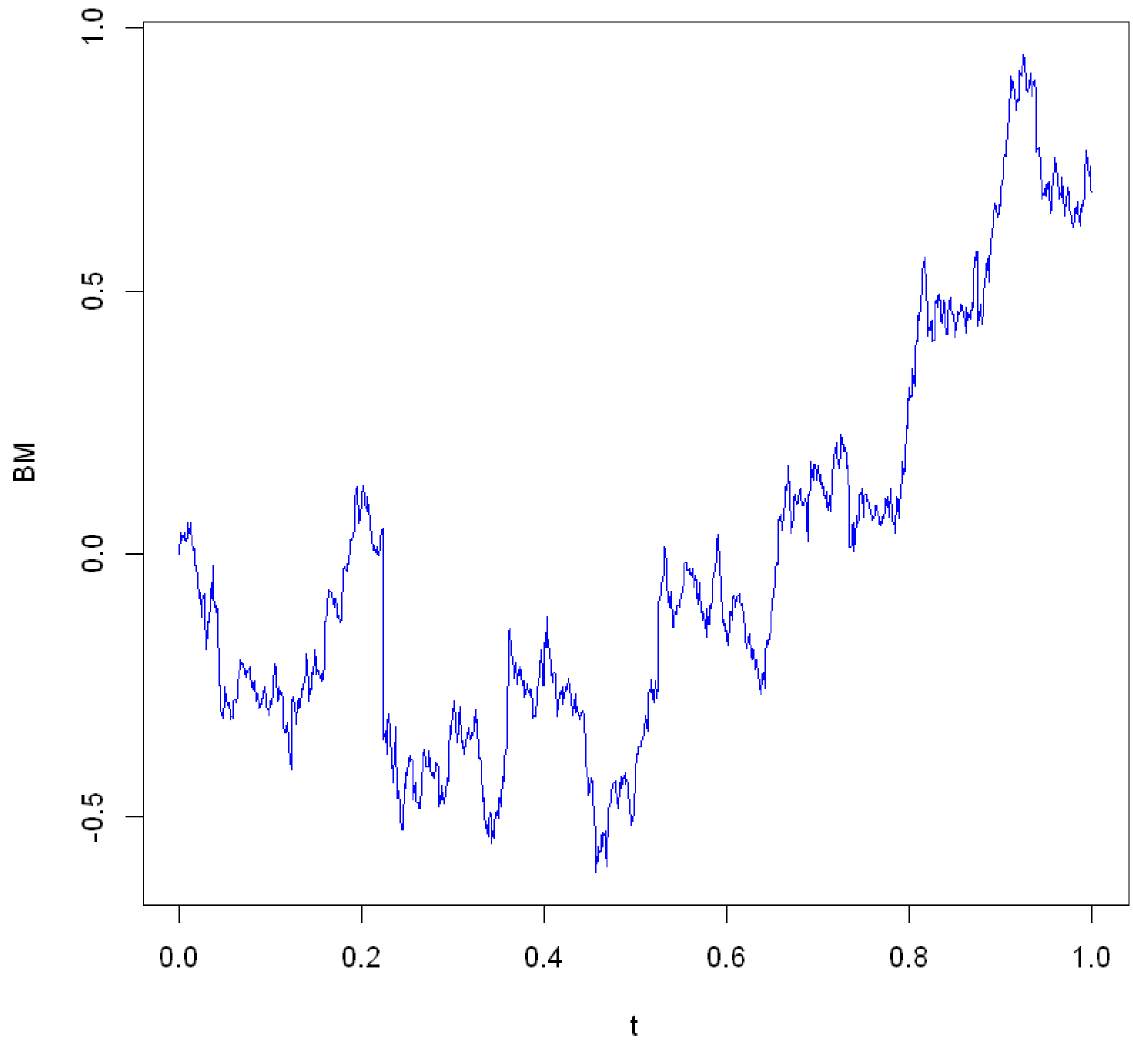
$$\mathfrak{S}_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[S_{i-1} + n\left(t - \frac{i-1}{n}\right) X_i\right] 1_{\left(\frac{i-1}{n}, \frac{i}{n}\right]}(t).$$

In fact, $\mathfrak{S}_n$ is simply the linear interpolation of the scaled random walk $\left\{\frac{S_1}{\sqrt{n}}, \frac{S_2}{\sqrt{n}}, \cdots, \frac{S_n}{\sqrt{n}}\right\}$.

Then, $\mathfrak{S}_n \to B$ in distribution as $n \to \infty$, where B denotes a Brownian motion.

In other words, as $n \to \infty$, the linearly interpolated scaled random walk $\mathfrak{S}_n$ converges in distribution to a Brownian motion.

In [25]:
```
# Simulate BM by using the Donsker's invariance principle
N <- 1000
nu <- 3
X <- rt(N,df=nu)/sqrt(nu/(nu-2))
BM <- c(0,cumsum(X)/sqrt(N))
t <- c(0,(1:N)/N)
plot(t,BM,type='l',col='blue')
```



2.4 Brownian motion as martingale

Definition

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

- A **filtration** on $(\Omega, \mathcal{F}, P)$ is a collection of σ-algebras $(\mathcal{F}_t)_{t\geq 0}$ satisfying the following property: $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $0\le s
- Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration, and $(X_t)_{t\geq 0}$ be a stochastic process on $(\Omega, \mathcal{F}, P)$.

  $(X_t)_{t\geq 0}$ is said to be **adapted** to the filtration $(\mathcal{F}(t))_{t\geq 0}$ if X_t is $\mathcal{F}_t$-measurable for each $t \geq 0$.
- Let $(B_t)_{t\geq 0}$ be a Brownian motion on $(\Omega, \mathcal{F}, P)$. A **filtration for the Brownian motion** is a filtration $(\mathcal{F}_t)_{t\geq 0}$ such that
  - $(B_t)_{t\geq 0}$ is adapted to it and
  - $B_t - B_s$ is independent of $\mathcal{F}_s$ for all $0 \leq s \leq t$.
- The **natural filtration** $(\mathcal{F}_t)_{t\geq 0}$ of a Brownian motion is the filtration such that each $\mathcal{F}_t$ is the smallest σ-algebra which contains all sets of the form $\{B_{t_1} \leq x_1, B_{t_2} \leq x_2, \ldots, B_{t_m} \leq x_m\}$ for all $m \geq 1$, $0 \leq t_1 < t_2 < \cdots < t_m \leq t$, $x_1, x_2, \ldots, x_m \in \mathbb{R}$.

In general, $(\mathcal{F}_t)_{t\geq 0}$ can be generated not only by the Brownian motion but by the Brownian motion together with some other process or processes. In such case the filtration $(\mathcal{F}_t)_{t\geq 0}$ can be strictly larger than the natural filtration.

Definition

A stochastic process $(M_t)_{t\geq 0}$, on $(\Omega, \mathcal{F}, P)$ is said to be a **martingale** with respect to filtration $(\mathcal{F}_t)_{t\geq 0}$ if

- (i) $E(|M_t|) < \infty$ for all $t \geq 0$;
- (ii) M_t is $\mathcal{F}_t$-measurable for all $t \geq 0$;
- (iii) $E(M_t \mid \mathcal{F}_s) = M_s$ a.s. for all $t \geq s \geq 0$.

Theorem

Let $(B_t)_{t\geq 0}$ be a Brownian motion and $(\mathcal{F}(t))_{t\geq 0}$ be a filtration for this Brownian motion.

The following processes are martingales relative to $(\mathcal{F}(t))_{t\geq 0}$:

- $(B_t)_{t\geq 0}$;
- $(B_t^2 - t)_{t\geq 0}$;
- $\left(e^{\sigma B_t - \sigma^2 t/2}\right)_{t\geq 0}$ (for each $\sigma \in \mathbb{R}$).

Proof:

- $(B_t)_{t\geq 0}$ is a martingale. Indeed, (i) and (ii) hold by construction.

  Furthermore, we have for all $0 \leq s \leq t$

$$\begin{aligned} E(B_t|\mathcal{F}_s) &= E(B_t - B_s|\mathcal{F}_s) + E(B_s|\mathcal{F}_s) \\ &= E(B_t - B_s) + B_s \\ &= 0 + B_s \\ &= B_s. \end{aligned}$$

- The other two parts are exercises in HW02.

2.5 Extensions of Brownian motion

Brownian motion with drift

Let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space and B_t a Brownian motion on $(\Omega, \mathcal{F}_t, \mathbb{P})$. A stochastic process X of the form

$$X_t = x + B_t + \int_0^t \mu_s ds$$

is called a **Brownian motion with drift** μ_t, where μ_t is adapted to the filtration $\mathcal{F}_t$.

Remark:

- X_t is a Gaussian process if μ_t is deterministic. Then, the mean function is $\mathbb{E}\left[X_t\right] = x + \int_0^t \mu_s ds$ and the covariance function $\gamma(t, s) = \mathrm{Cov}(X_t, X_s) = \min\{t, s\}$.

- We can always transform a Brownian motion with drift into a standard Brownian motion by change of the underlying probability measure so long as the drift μ_t satisfies certain conditions, say, bounded (see future lectures).

Brownian bridge

Let B_t be a Brownian Motion in $[0,1]$. Define the process

$$X_t := B_t - tB_1, \quad t \in [0,1],$$

then X_t is called the (standard) **Brownian bridge**.

- Note that $X_0 = X_1 = 0$.
- X is a Gaussian process.
- The mean and autovariance function of Brownian bridge X_t are $\mathbb{E}\left[X_t\right] = 0$ and $\mathrm{Cov}(X_s, X_t) = s(1-t)$ respectively, for $0 \le s \le t \le 1$. In particular, $\mathrm{Var}(X_t) = t(1-t)$.
- The Brownian bridge appears as the limit process of the normalized empirical distribution function of a sample of iid uniform U(0,1) random variables.

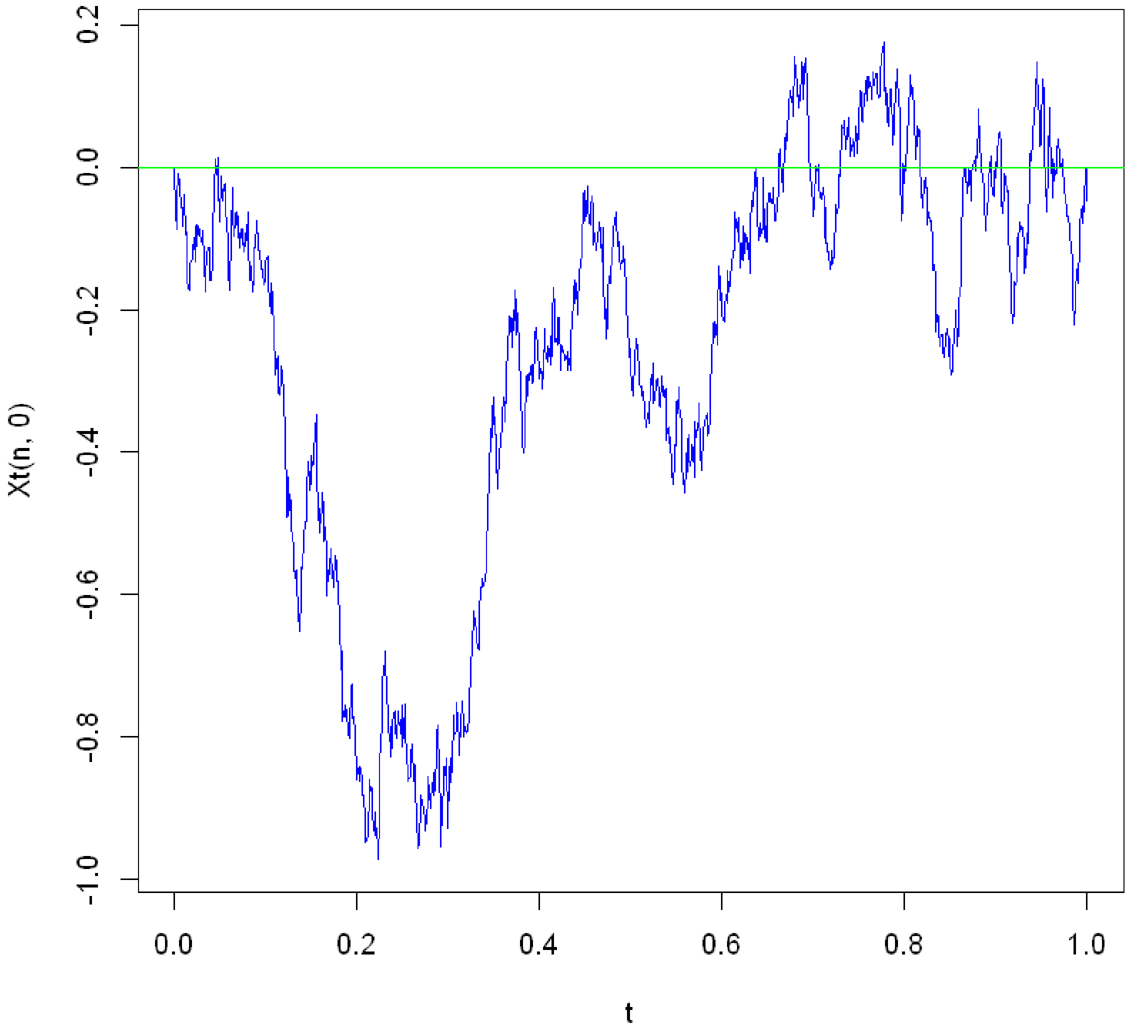
Brownian bridge with inital and terminal points other than 0 is obtained by adding a linear function to the standard Brownian bridge. For example, Brownian bridge X with $X_0 = a$ and $X_1 = b$ is given by

$$X_t = B_t - tB_1 + a + (b-a)t,$$

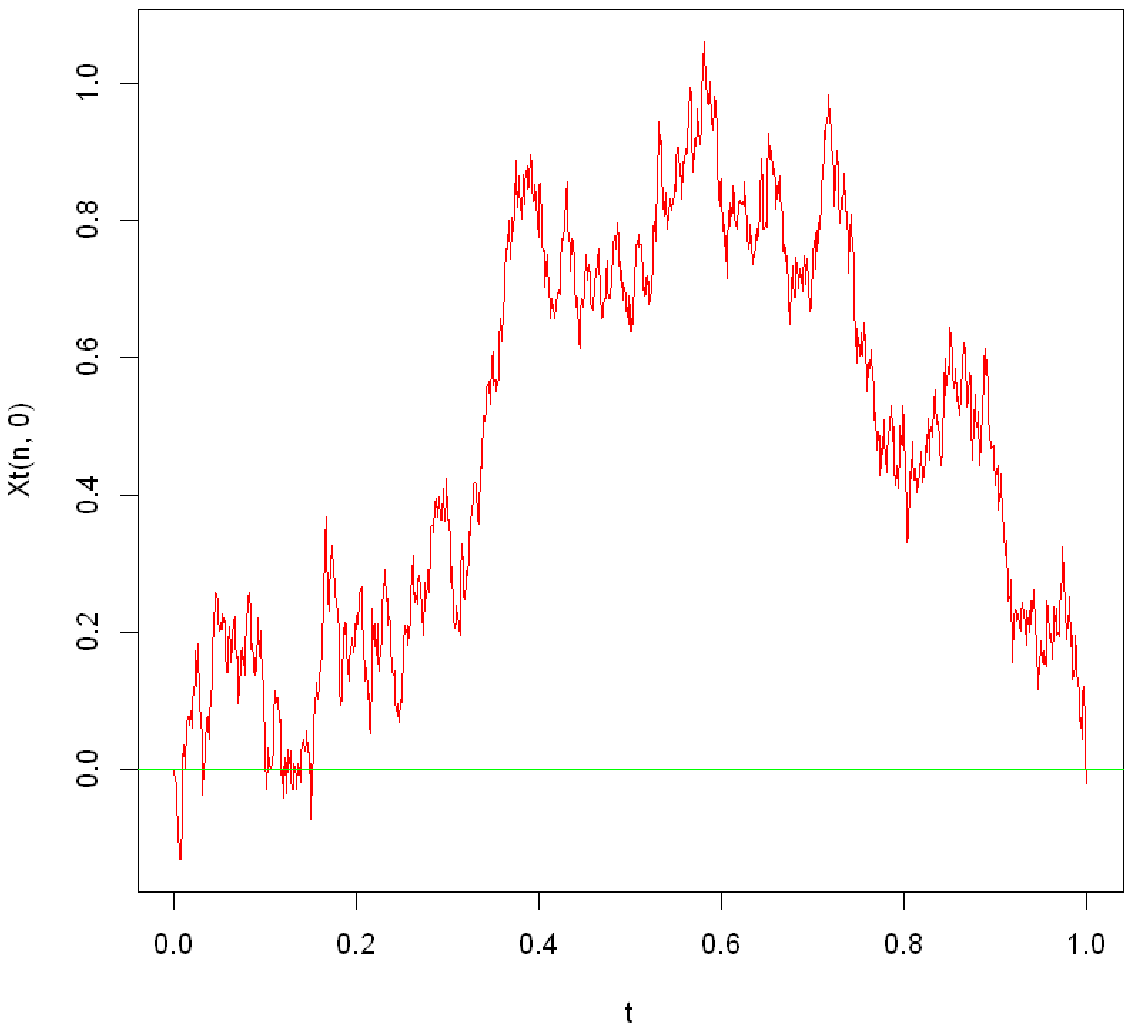
where again B_t is a standard Brownian motion.

In [21]:

```r
# Simulate Brownian bridge X_t = B_t - t*B_1

Bt <- function(n){cumsum(rnorm(n,sd=1/sqrt(n)))}

# Generates Brownian bridge with n points
Xt <- function(n, final){
  tmp <- Bt(n)
  c(0,tmp + (1:n)/n * (final-tmp[n]))
  }

n <- 1000
t <- c(0,(1:n)/n)
plot(t,Xt(n,0),type='l',col='blue')
abline(h=0,col='green')
plot(t,Xt(n,0),type='l',col='red')
abline(h=0,col='green')
```



Xt(n, 0)
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
t



Xt(n, 0)
0.0
0.2
0.4
0.6
0.8
1.0
t
0.0
0.2
0.4
0.6
0.8
1.0