9831 Probability and Stochastic Processes for Finance, Fall 2016

Lecture 7

Anja Richter

Department of Mathematics

Outline

- Girsanov Theorem
- Martingale representation theorem
- Numeraire

1 Girsanov theorem

1.1 Absolute continuous and equivalent measures, the Radon-Nikodym theorem

Recall:

Definition

Let $\mathbb{P}$ and $\tilde{\mathbb{P}}$ be probability measures on $(\Omega, \mathcal{F})$. $\tilde{\mathbb{P}}$ is said to be **absolutely continuous with respect to** $\mathbb{P}$, $\tilde{\mathbb{P}} \ll \mathbb{P}$, if for all $A \in \mathcal{F}$ with $\mathbb{P}(A) = 0$ we have $\tilde{\mathbb{P}}(A) = 0$.

If $\tilde{\mathbb{P}} \ll \mathbb{P}$ and $\mathbb{P} \ll \tilde{\mathbb{P}}$ then we say that $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are **equivalent** and write $\mathbb{P} \sim \tilde{\mathbb{P}}$.

Radon-Nikodym theorem

Let $\mathbb{P}$ and $\tilde{\mathbb{P}}$ be probability measures on $(\Omega, \mathcal{F})$, and $\tilde{\mathbb{P}} \ll \mathbb{P}$. Then there exists a random variable Z, such that

$$\mathbb{E}(Z) = 1 \quad \text{and} \quad \tilde{\mathbb{P}}(A) = \int_A Z(\omega)\, d\mathbb{P}(\omega) = \mathbb{E}(1_A Z) \ \text{ for all } A \in \mathcal{F}. \tag{1}$$

Moreover, if Z_1 and Z_2 both satisfy (1) then $\tilde{\mathbb{P}}(Z_1 \neq Z_2) = \mathbb{P}(Z_1 \neq Z_2) = 0$.

Proof: See Shreve or Refresher, Lecture 4.

1.2 How to construct new probability measures $\tilde{\mathbb{P}} \ll \mathbb{P}$

Easy converse of the Radon-Nikodym theorem.

Theorem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let Z be any r.v. such that $\mathbb{P}(Z \geq 0) = 1$ and $\mathbb{E}(Z) = 1$.

For each $A \in \mathcal{F}$ define

$$\tilde{\mathbb{P}}(A) = \int_A Z(\omega)\, d\mathbb{P}(\omega). \tag{2}$$

- Then $\tilde{\mathbb{P}}$ is a probability measure on $(\Omega, \mathcal{F})$ and $\tilde{\mathbb{P}} \ll \mathbb{P}$.
- Furthermore, we have for r.v.s X that

$$\tilde{\mathbb{E}}(X) = \mathbb{E}(XZ).$$

- If, in addition $\mathbb{P}(Z > 0) = 1$, then we also have that $\mathbb{P} \ll \tilde{\mathbb{P}}$ (so that $\mathbb{P} \sim \tilde{\mathbb{P}}$), and for all $A \in \mathcal{F}$

$$\mathbb{P}(A) = \int_A \frac{1}{Z(\omega)}\, d\tilde{\mathbb{P}}(\omega),$$

  and for every non-negative random variable Y

$$\mathbb{E}(Y) = \tilde{\mathbb{E}}\left(Y \frac{1}{Z}\right).$$

Definition

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\tilde{\mathbb{P}}$ be a probability measure that is equivalent to $\mathbb{P}$.

Let Z be an a.s. positive r.v. that relates the two measure via (2).

Then Z is called the **Radon-Nikodym derivative of $\tilde{\mathbb{P}}$ wrt.** $\mathbb{P}$, and we write

$$Z = \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}.$$

Assume μ_X is the distribution measure of r.v. X and μ_Y the distribution measure of r.v. Y.

Then absolute continuity $\mu_X \ll \mu_Y$ can be reconciled as $\mathrm{supp} X \subseteq \mathrm{supp} Y$ and the Radon-Nikodym derivative as

$$\frac{d\mu_X}{d\mu_Y}(x) = \lim_{A \downarrow \{x\}} \frac{\mathbb{P}[X \in A]}{\mathbb{P}[Y \in A]}.$$

If X and Y are continuous random variables with density f_X and f_Y, i.e. distribution measures are given as

$$\mu_X(A) = \int_A f_X(x)dx \qquad \text{and} \qquad \mu_Y(A) = \int_A f_Y(y)dy \qquad (A \in \mathcal{F}),$$

then

$$\frac{d\mu_X}{d\mu_Y}(x) = \frac{f_X(x)}{f_Y(x)}.$$

Moreover, we can calculate the expectation related to X by the expectation related to Y but weighted by the Radon-Nikodym derivative as follows:

$$\mathbb{E}[h(X)] = \int h(x) f_X(x) dx = \int h(x) \frac{f_X(x)}{f_Y(x)} f_Y(x) dx = \mathbb{E}\left[h(Y) \frac{f_X(Y)}{f_Y(Y)}\right].$$

Example

Let $X \sim N(m_X, \sigma_X^2)$ and $Y \sim N(m_Y, \sigma_Y^2)$. We have that $\mathrm{supp} X = \mathrm{supp} Y = \mathbb{R}$. Hence, the Radon-Nikodym derivative $\frac{d\mu_X}{d\mu_Y}(x)$ for the distribution measures μ_X and μ_Y of X and Y respectively is given by

$$\frac{d\mu_X}{d\mu_Y}(x) = \frac{f_X(x)}{f_Y(x)} = \frac{\sigma_Y}{\sigma_X} e^{\frac{(x-m_Y)^2}{2\sigma_Y^2} - \frac{(x-m_X)^2}{2\sigma_X^2}}.$$

Moreover, since in this case $\mu_X \sim \mu_Y$,

$$\frac{d\mu_Y}{d\mu_X}(x) = \frac{f_Y(x)}{f_X(x)} = \frac{\sigma_X}{\sigma_Y} e^{\frac{(x-m_X)^2}{2\sigma_X^2} - \frac{(x-m_Y)^2}{2\sigma_Y^2}}.$$

1.3 Illustrative example: Baby Girsanov.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X \sim N(0,1)$ under $\mathbb{P}$.

Define a new random variable $\widetilde{X} = X + \mu$.

Obviously $\widetilde{X} \sim N(\mu, 1)$ under $\mathbb{P}$.

We shall define a new probability $\tilde{\mathbb{P}}$ on Ω, which will be equivalent to $\mathbb{P}$, such that $\widetilde{X} \sim N(0,1)$ under this new probability $\tilde{\mathbb{P}}$. In other words, we are able to change the mean of a normal r.v. on Ω without affecting its variance or the r.v. itself by the change of probability on Ω.

First off, notice that the r.v. $e^{-\mu X - \frac{\mu^2}{2}}$

has

- mean 1 (in $\mathbb{P}$) and
- is almost surely (actually, surely) positive,

so it may serve as a Radon-Nikodym derivative.

Define a new probability $\tilde{\mathbb{P}}$ by the Radon-Nikodym derivative

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{-\mu X - \frac{\mu^2}{2}}.$$

Let's see what the distribution for $\widetilde{X}$ is under $\tilde{\mathbb{P}}$. Given a set $A \in \mathcal{F}$, consider

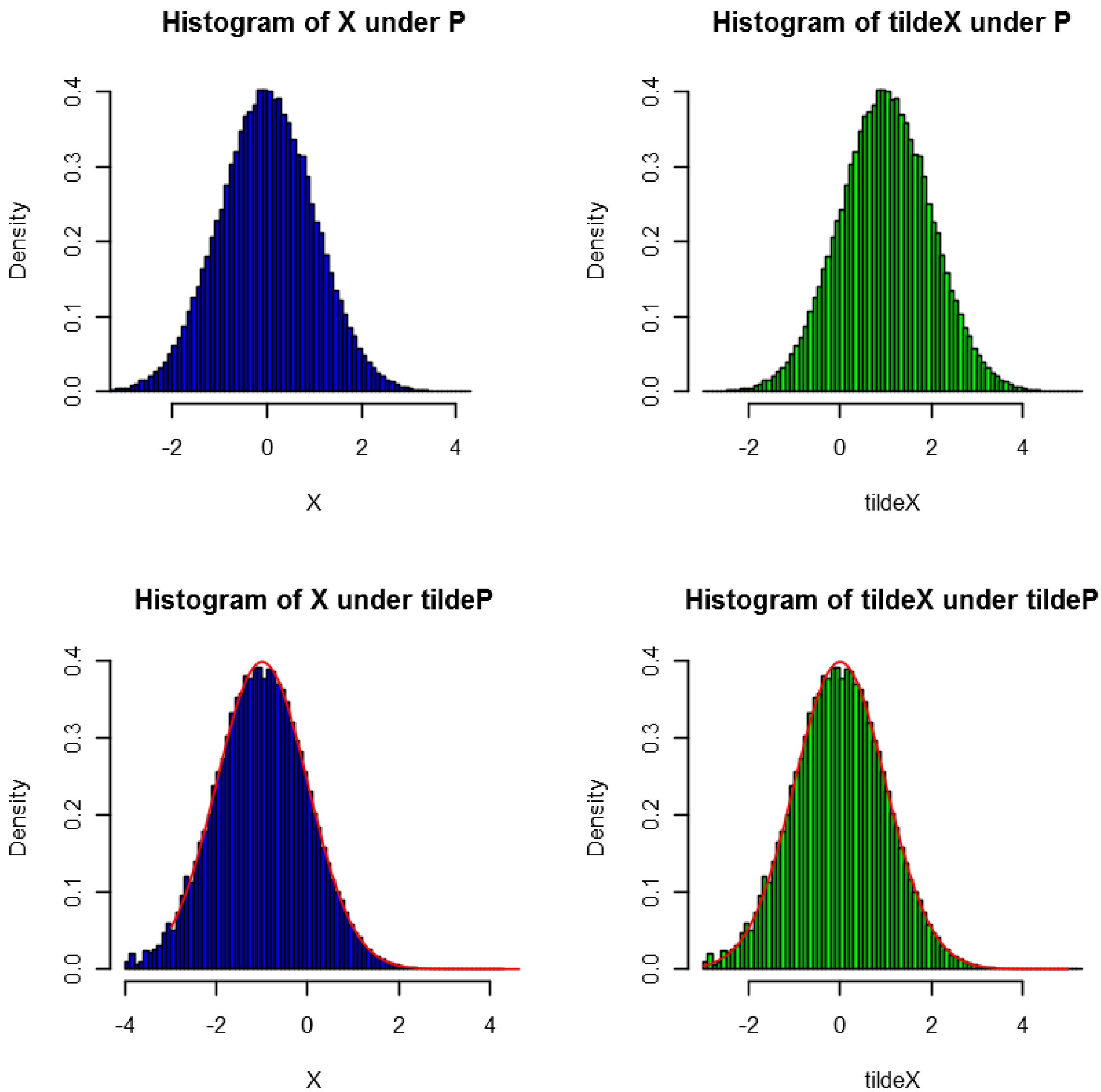
$$\begin{aligned}
\tilde{\mathbb{P}}[\widetilde{X} \in A] &= \int_\Omega 1_A(\widetilde{X}) d\tilde{\mathbb{P}} \\
&= \int_\Omega 1_A(X+\mu) e^{-\mu X - \frac{\mu^2}{2}} d\mathbb{P} \\
&= \int_{-\infty}^{\infty} 1_A(x+\mu) e^{-\mu x - \frac{\mu^2}{2}} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \quad (\text{since } X \sim N(0,1) \text{ under } \mathbb{P}) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 1_A(y) e^{-\mu(y-\mu) - \frac{\mu^2}{2}} e^{-\frac{(y-\mu)^2}{2}} dy \quad (\text{change of variable } y = x + \mu) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 1_A(y) e^{-\frac{y^2}{2}} dy.
\end{aligned}$$

Hence $\widetilde{X}$ is standard normal under $\tilde{\mathbb{P}}$.

On the other hand, one can show that, under $\tilde{\mathbb{P}}$, $X \sim N(-\mu, 1)$.

Simulation of Baby Girsanov

In [2]:
```
# The code demonstrates baby Girsanov by simulation
# requires a library called weight
# If necessary, install the package by using the following line of code
# install.packages("weights")

library('weights');

NSim <- 1e5
mu <- 1

X <- rnorm(NSim)
tildeX <- X + mu  # y normally distributed with mean mu

par(mfrow=c(2,2))
hist(X,prob=T,breaks=100,col='blue',xlim=c(-3,5),main='Histogram of X under
P')
hist(tildeX,prob=T,breaks=100,col='green',add=F,main='Histogram of tildeX unde
r P')

# Radon-Nikodym derivative
RN.derivative <- function(x, mu) exp(-mu*x - mu^2/2)

# Now weigh the samples of X by the Radon-Nikodym derivative
wgt <- RN.derivative(X,mu=mu)

wtd.hist(X,weight=wgt,prob=T,breaks=100,col='blue',main='Histogram of X under
 tildeP')
# superimpose normal density
curve(dnorm(x, mean = -mu),from=-3,to=5,col='red',add=T)

# Now weigh the samples of tildeX by the Radon-Nikodym derivative
wgt <- RN.derivative(tildeX,mu=mu)

wtd.hist(tildeX,weight=wgt,prob=T,breaks=100,col='green',main='Histogram of ti
ldeX under tildeP')
# superimpose normal density
curve(dnorm(x, mean = 0),from=-3,to=5,col='red',add=T)
```


Histogram of X under P
Density
X
Histogram of tildeX under P
Density
tildeX
Histogram of X under tildeP
Density
X
Histogram of tildeX under tildeP
Density
tildeX

1.4 Girsanov theorem

Let μ_t be an adapted process and B_t a standard Brownian motion. Let

$$\tilde{B}_t = B_t + \int_0^t \mu_s ds$$

be a Brownian motion with drift μ_t. Define the process Z_t as

$$Z_t = e^{-\int_0^t \mu_s dB_s - \frac{1}{2}\int_0^t \mu_s^2 ds}$$

and assume that $E[\int_0^T \mu_s^2 Z_s^2 ds] < \infty$.

Then $\mathbb{E}[Z_T] = 1$ and under the new probability measure $\tilde{\mathbb{P}}$

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = Z_T = e^{-\int_0^T \mu_t dB_t - \frac{1}{2}\int_0^T \mu_t^2 dt},$$

the process $\tilde{B}$ is a Brownian motion.

On the other hand, under the probability measure $\tilde{\mathbb{P}}$, B is a Brownian motion with drift $-\mu_t$.

Remarks

- Girsanov theorem allows us to change the drift of Brownian motions.
- With deterministic drift, the result was discovered earlier on, nowadays known as the Cameron-Martin theorem:

  Let $h(t)$ be a deterministic $L^2[0,T]$ function.

  Let $\tilde{B}_t = B_t + \int_0^t h(s)ds$ be the Brownian motion with (deterministic) drift h.

  Define the probability $\tilde{\mathbb{P}}$ on Ω by

  $$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{-\int_0^T h(t)dB_t - \frac{1}{2}\int_0^T h^2(t)dt}.$$

  Then the process $\tilde{B}_t$ under $\tilde{\mathbb{P}}$ is a Brownian motion, whereas B_t under $\tilde{\mathbb{P}}$ becomes a Brownian motion with drift $-h$. In this case, since h is deterministic, the Radon-Nikodym derivative is lognormally distributed under $\mathbb{P}$.

- Note that, by applying Ito's formula, one may show that (under $\mathbb{P}$) Z_t is indeed the solution to the SDE

  $$dZ_t = -\mu_t Z_t dB_t, \quad Z_0 = 1.$$

  Integrating this, we have

  $$Z_T = 1 - \int_0^T \mu_t Z_t dB_t. \qquad (*)$$

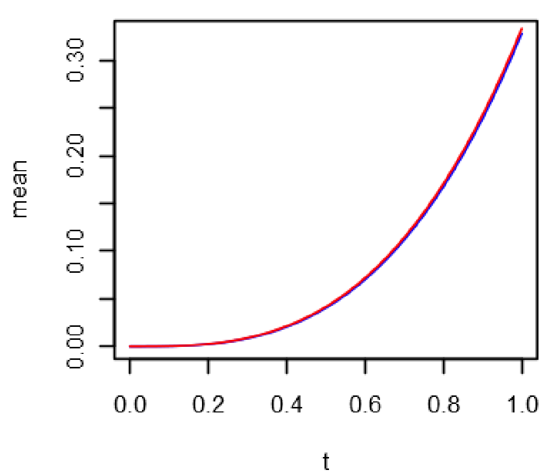
- The technicality and difficulty in applying Girsanov theorem is the integrability condition $E[\int_0^T \mu_s^2 Z_s^2 ds] < \infty$ which ensures that the stochastic integral in $(*)$ is defined and is a martingale.

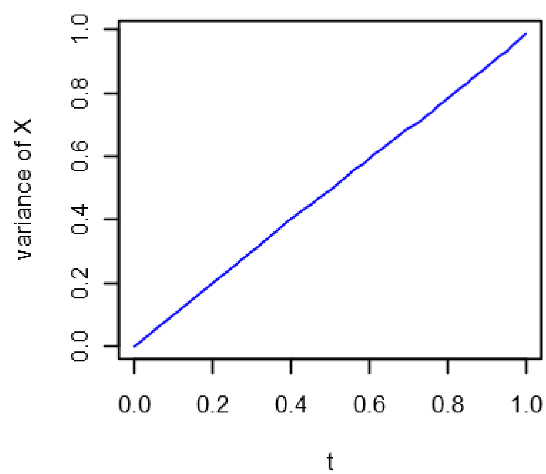
  This condition is usually difficult to check. A sufficient condition is given by the **Novikov condition**:
  $$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T \mu_t^2 dt}\right] < \infty.$$

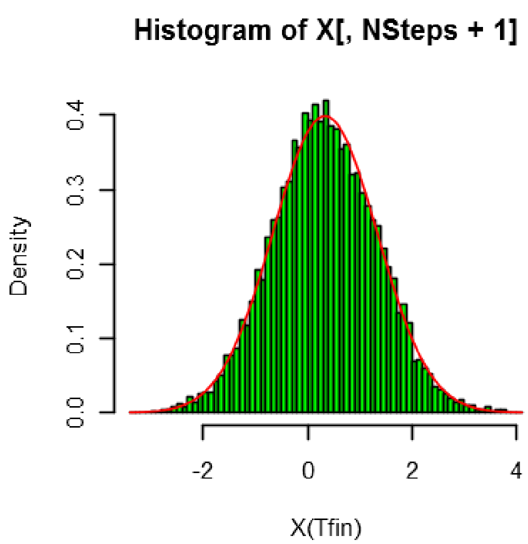
Simulation of Girsanov

In [8]:

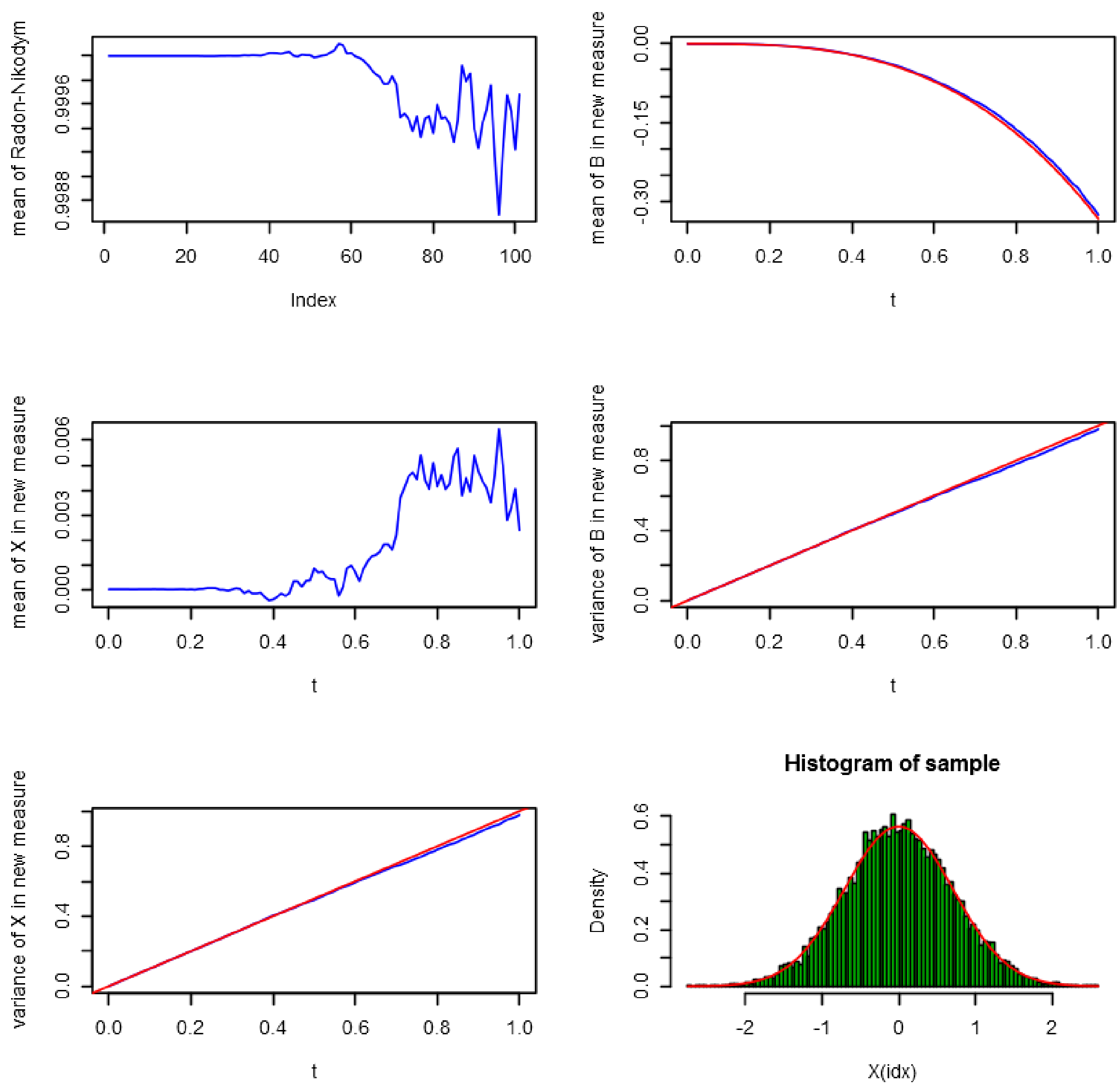
```
# Demonstrate Girsanov theorem by simulation

NSim <- 1e4   # number of samples at each step
NSteps <- 100   # number of steps
Tfin <- 1    # terminal time
dt <- Tfin/NSteps

B <- matrix(0,NSim,NSteps+1)  # initialize the Brownians
dB <- matrix(0,NSim,NSteps+1)

# Simulate Browian paths
for (i in 1:NSteps){
  db <- rnorm(NSim)
  db <- db - mean(db)  # now db has mean 0
  db <- db/sd(db)  # now db has variance 1
  dB[,i] <- sqrt(dt)*db
  B[,i+1] <- B[,i] + dB[,i]
}

# Brownian motion with drift h
h <- function(t) t^2
t <- (0:NSteps)*dt
# matrixize t
# the command works like repmat in Matlab
tM <- matrix(t,NSim,length(t),byrow=T)
hdt <- h(tM)*dt
H <- t(apply(hdt,1,FUN=cumsum))  # H is the integral of h, the drift
H <- cbind(0,H[,-dim(H)[2]]) # remove the last column of H and add 0 as the fi
rst column
X <- B + H

# plots
par(mfrow=c(2,2))

# evolution of mean
plot(t,colMeans(X),type='l',col='blue',ylab='mean')
curve(1/3*x^3,col='red',add=T)

# evolution of variance
plot(t,apply(X,2,FUN=var),type='l',col='blue',ylab='variance of X')

# histogram at terminal time
hist(X[,NSteps+1],prob=T,col='green',breaks=100,xlab='X(Tfin)')
curve(dnorm(x,mean=1/3*Tfin^3,sd=sqrt(Tfin)),col='red',add=T)

# So it does look like a Brownian motion with drift h
```


mean
0.00
0.10
0.20
0.30
0.0
0.2
0.4
0.6
0.8
1.0
t


variance of X
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


Histogram of X[, NSteps + 1]
Density
0.0
0.1
0.2
0.3
0.4
-2
0
2
4
X(Tfin)

In [9]:
```
# Now weigh the Brownian paths as suggested by Girsanov theorem

# calculate the Radon-Nikodym derivative
# dB <- cbind(dB,0)
dExponent <- -h(tM)*dB - 1/2*h(tM)^2*dt
Exponent <- t(apply(dExponent,1,FUN=cumsum))
Exponent <- cbind(0,Exponent[,-dim(Exponent)[2]])  # Add 0 to the first column
 and delete the last
RN.derivative <- exp(Exponent)

# plots
par(mfrow=c(3,2))

plot(colMeans(RN.derivative),type='l',col='blue',ylab='mean of Radon-Nikodym')

# plot the evolution of means for B and X in the new measure
# note that we did not normalize the weights
plot(t,colMeans(RN.derivative*B),type='l',col='blue',ylab='mean of B in new me
asure')
curve(-1/3*x^3,add=T,col='red') # superimpose with H
plot(t,colMeans(RN.derivative*X),type='l',col='blue',ylab='mean of X in new me
asure')

# plot the evolution of variances for B and X in the new measure
varsB <- colMeans(RN.derivative*B^2) - colMeans(RN.derivative*B)^2
plot(t,varsB,type='l',col='blue',ylab='variance of B in new measure')
abline(a=0,b=1,col='red') # superimpose with the straight line y = x
varsX <- colMeans(RN.derivative*X^2) - colMeans(RN.derivative*X)^2
plot(t,varsX,type='l',col='blue',ylab='variance of X in new measure')
abline(a=0,b=1,col='red')


# plot histogram of X in new measure at time index idx
idx <- 51
sample <- X[,idx]
wgt <- RN.derivative[,idx]
wtd.hist(sample,weight=wgt,prob=T,breaks=100,col='green',xlab='X(idx)')
curve(dnorm(x,sd=sqrt(t[idx])),add=T,col='red')
```



mean of Radon-Nikodym
Index
mean of B in new measure
t
mean of X in new measure
t
variance of B in new measure
t
variance of X in new measure
t
Histogram of sample
Density
X(idx)

2 Proof of Girsanov's theorem

- Note that Z_T is indeed a Radon-Nikodym derivative: It is obviously positive and since Z_t is a martingale with $E(Z_0) = 1$, we have $E(Z_T) = 1$. Hence $\tilde{\mathbb{P}}$ is indeed a probability measure and we have $\tilde{\mathbb{P}} \sim \mathbb{P}$.

The proof of Girsanov theorem is based on Lévy's characterization of Brownian.

Recall that, under $\mathbb{P}$, $\tilde{B}_t = B_t + \int_0^t \mu_s ds$ is a Brownian motion with drift μ_t.

- We have $\tilde{B}_0 = 0$ as $B_0 = 0$.
- The process $\tilde{B}_t$ is continuous since both B_t and $\int_0^t \mu_s ds$ are.
- Quadratic variation of $\tilde{B}_t$ is the same as B_t since $\int_0^t \mu_s ds$ will not contribute anything to the quadratic varition of $\tilde{B}_t$ because it is of finite variation. So the quadratic variation of $\tilde{B}_t$ is t, i.e.

$$d[\tilde{B}]_t = d[B]_t = t \quad \mathbb{P}\text{-a.s.}$$

  As $\mathbb{P} \sim \tilde{\mathbb{P}}$, we immediately obtain $d[\tilde{B}]_t = t \quad \tilde{\mathbb{P}}$-a.s.
- Martingality of $\tilde{B}_t$ under $\tilde{\mathbb{P}}$ is the technical part. To this end, we shall need two general lemmas which we state below.

Lemma 1

Let $(\Omega, \mathcal{F}_t, \mathbb{P})$ be a probability space and Z be a positive random variable with $\mathbb{E}[Z] = 1$.

Define $Z_t = \mathbb{E}[Z|\mathcal{F}_t]$, $t \in [0, T]$, hence Z_t is a martingale under $\mathbb{P}$.

Define a new probability $\tilde{\mathbb{P}}$ on Ω by

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = Z.$$

Then we have

(a) if Y is $\mathcal{F}_t$-measurable, then $\tilde{\mathbb{E}}[Y] = \mathbb{E}[YZ_t]$.

(b) if Y is $\mathcal{F}_t$-measurable and $s < t$, then $\tilde{\mathbb{E}}[Y|\mathcal{F}_s] = \dfrac{1}{Z_s}\mathbb{E}[YZ_t|\mathcal{F}_s]$.

Proof

(a) By the definition of $\tilde{\mathbb{P}}$ and properties of conditional expectation, it is straightforward that

$$\tilde{\mathbb{E}}[Y] = \mathbb{E}[YZ] = \mathbb{E}[\mathbb{E}[YZ|\mathcal{F}_t]] = \mathbb{E}[Y\mathbb{E}[Z|\mathcal{F}_t]] = \mathbb{E}[YZ_t].$$

(b) Fix $s \in [0, t]$ and and define

$$X = \frac{1}{Z_s} E(YZ_t\,|\mathcal{F}_s).$$

We want to show that $\tilde{\mathbb{E}}[Y|\mathcal{F}_s] = X$ a.s. For that we need to check whether X is $\mathcal{F}_s$-measurable and the averaging property

$$\int_A Y d\tilde{\mathbb{P}} = \int_A X d\tilde{\mathbb{P}},$$

for $A \in \mathcal{F}_s$.

Obviously X is $\mathcal{F}_s$-measurable. For $A \in \mathcal{F}_s$, we start with the right-hand side and transform it into the left-hand side:

$$\begin{aligned}
\int_A \frac{1}{Z_s}\mathbb{E}[YZ_t|\mathcal{F}_s]d\tilde{\mathbb{P}} &= \tilde{\mathbb{E}}\big[1_A \frac{1}{Z_s}\mathbb{E}[YZ_t|\mathcal{F}_s]\big] \\
&= \mathbb{E}\big[1_A \mathbb{E}[YZ_t|\mathcal{F}_s]\big] \quad \text{(use (a))} \\
&= \int_A \mathbb{E}[YZ_t|\mathcal{F}_s]d\mathbb{P} \\
&= \int_A \mathbb{E}[\mathbb{E}[YZ|\mathcal{F}_t]|\mathcal{F}_s]d\mathbb{P} \quad (\text{definition of } Z_t, Y \ \in \mathcal{F}_t) \\
&= \int_A \mathbb{E}[YZ|\mathcal{F}_s]d\mathbb{P} \quad \text{(tower property)} \\
&= \int_A YZd\mathbb{P} \quad \text{(definition of conditional expectation)} \\
&= \int_A Yd\tilde{\mathbb{P}} \quad \left(\text{since } \frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = Z\right)
\end{aligned}$$

Lemma 2

Under the same assumptions as in Lemma 1 we have that M_t is a martingale under $\tilde{\mathbb{P}}$ if and only if $M_t Z_t$ is a martingale under $\mathbb{P}$.

Proof

We check only the conditional expectation property below.

- "$\Longleftarrow$": For $s<t$,
$$\begin{aligned}\tilde{\mathbb{E}}[M_t|\mathcal{F}_s] &= \frac{1}{Z_s}\mathbb{E}[M_tZ_t|\mathcal{F}_s] \quad \text{(by (b) in Lemma 1)}\\ &= \frac{1}{Z_s}M_sZ_s \quad (\text{since } M_tZ_t \text{ is a martingale under } \mathbb{P}).\\ &= M_s\end{aligned}$$

- "$\Longrightarrow$": For $s<t$,
$$\begin{aligned}\mathbb{E}[M_tZ_t|\mathcal{F}_s] &= Z_s\frac{1}{Z_s}\mathbb{E}[M_tZ_t|\mathcal{F}_s]\\ &= Z_s\tilde{\mathbb{E}}[M_t|\mathcal{F}_s] \quad \text{(by (b) in Lemma 1)}\\ &= M_sZ_s \quad (\text{since } M_t \text{ is a martingale under } \tilde{\mathbb{P}}).\end{aligned}$$

Now we are ready to show that $\tilde{B}_t$ is a $\tilde{\mathbb{P}}$-martingale. According to Lemma 2, it suffices to show that $\tilde{B}_tZ_t$ is a $\mathbb{P}$-martingale. Recall that $\tilde{B}_t$ and Z_t satisfy (under $\mathbb{P}$) respectively the SDEs
$$\begin{aligned}d\tilde{B}_t &= dB_t+\mu_tdt,\\ dZ_t &= -\mu_tZ_tdB_t.\end{aligned}$$

Apply Ito's formula to $\tilde{B}_tZ_t$ we have
$$\begin{aligned}d(\tilde{B}_tZ_t) &= \tilde{B}_tdZ_t+Z_td\tilde{B}_t+d\tilde{B}_tdZ_t\\ &= (-\mu_t\tilde{B}_t+1)Z_tdB_t,\end{aligned}$$

which, since B_t is a $\mathbb{P}$-Brownian motion, has zero drift and therefore is a martingale under $\mathbb{P}$.

We have checked all conditions for Levy's characterization of Brownian motions and hence have that $\tilde{B}_t$ is a $\tilde{\mathbb{P}}$-Brownian motion.

3 Application of Girsanov's theorem

3.1 Pricing a binary barrier option in the Black-Scholes model

Recall that the payoff function of a knock-in binary put struck at K and knock-in level L is given by

$$1_{\{\max_{0\le t\le T} S_t \ge L, S_T \le K\}}$$

and its price/premium is given by the expectation under risk neutral probability

$$\mathbb{E}\left[1_{\{\max_{0\le t\le T} S_t \ge L, S_T \le K\}}\right] = \mathbb{P}\left[\max_{0\le t\le T} S_t \ge L, S_T \le K\right].$$

The right hand side is calculated, in the Bachelier model, by applying reflection principle (see Lecture 3).

In the Black-Scholes model, we calculate the probability as follows. Assume the underlying S_t follows the Black-Scholes model:

$$\frac{dS_t}{S_t} = \sigma dB_t \quad \leftrightsquigarrow \quad S_t = S_0 e^{\sigma B_t - \frac{\sigma^2}{2} t}.$$

Hence,

$$\begin{aligned}
\mathbb{P}\left[\max_{0\le u\le t} S_u \ge L, S_T \le K\right] &= \mathbb{P}\left[\max_{0\le t\le T} S_0 e^{\sigma B_t - \frac{\sigma^2}{2}t} \ge L, S_0 e^{\sigma B_T - \frac{\sigma^2}{2}T} \le K\right] \\
&= \mathbb{P}\left[\max_{0\le t\le T} \left\{\sigma B_t - \frac{\sigma^2}{2}t\right\} \ge \log\left(\frac{L}{S_0}\right), \sigma B_T - \frac{\sigma^2}{2}T \le \log\left(\frac{K}{S_0}\right)\right] \\
&= \mathbb{P}\left[\max_{0\le t\le T} \tilde{B}_t \ge \frac{1}{\sigma}\log\left(\frac{L}{S_0}\right), \tilde{B}_T \le \frac{1}{\sigma}\log\left(\frac{K}{S_0}\right)\right],
\end{aligned}$$

where $\tilde{B}_t = B_t - \frac{\sigma}{2}t$ is a Brownian motion with drift $-\frac{\sigma}{2}$.

Unfortunately, reflection principle is not directly applicable because $\tilde{B}_t$ is a not a Brownian motion in $\mathbb{P}$-measure.

However, we can apply Girsanov theorem and define a new measure $\tilde{\mathbb{P}}$ so that $\tilde{B}_t$ is a Brownian motion in $\tilde{\mathbb{P}}$, henceforth reflection principle is applicable.

The price we pay is that, while calculating the probability, the weight of each reflected path has to be adjusted accordingly. Detailed calculations are left as a homework.

3.2 De-drifting an Ito process

Let X_t be an Ito process given by

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s \quad \leftrightsquigarrow \quad dX_t = \mu_t dt + \sigma_t dB_t$$

We can de-drift X_t by applying the Girsanov theorem as follows. Define a new measure $\tilde{\mathbb{P}}$ via the Radon-Nikodym derivative

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{-\int_0^T \frac{\mu_t}{\sigma_t} dB_t - \frac{1}{2}\int_0^T \left(\frac{\mu_t}{\sigma_t}\right)^2 dt}.$$

Then, by Girsanov theorem, under the new measure $\tilde{\mathbb{P}}$, B_t is a Brownian motion with drift $-\frac{\mu_t}{\sigma_t}$, i.e., $dB_t = d\tilde{B}_t - \frac{\mu_t}{\sigma_t} dt$, where $\tilde{B}_t$ is a Brownian motion under $\tilde{\mathbb{P}}$. Thus, the SDE for X_t under $\tilde{\mathbb{P}}$ reads

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t dB_t \\ &= \mu_t dt + \sigma_t \left(d\tilde{B}_t - \frac{\mu_t}{\sigma_t} dt \right) \\ &= \sigma_t d\tilde{B}_t. \end{aligned}$$

In other words, X_t is driftless under $\tilde{\mathbb{P}}$.

Remarks

- The same technique will be applied to characterize the risk neutral probability and the equivalent martingale measure in derivative pricing.
- Of course, a mathematical technicality is to make sure that $e^{-\int_0^T \frac{\mu_t}{\sigma_t} dB_t - \frac{1}{2}\int_0^T \left(\frac{\mu_t}{\sigma_t}\right)^2 dt}$ does define a change of measure, i.e., $\mathbb{E}\left[e^{-\int_0^T \frac{\mu_t}{\sigma_t} dB_t - \frac{1}{2}\int_0^T \left(\frac{\mu_t}{\sigma_t}\right)^2 dt}\right] = 1.$
- Girsanov theorem transforms only the drift, the diffusion part is never changed.

4 Girsanov theorem for multidimensional Brownian motions

4.1 Theorem

Let B_t be an n dimensional Brownian motion and

$$\tilde{B}_t = B_t + \int_0^t h_s ds$$

be a Brownian motion with drift h_t under $\mathbb{P}$. Note that the Brownian motion B_t and the drift h_t are both vector-valued.

Define a new measure $\tilde{\mathbb{P}}$ by the Radon-Nikodym derivative

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{-\int_0^T h_t \cdot dB_t - \frac{1}{2}\int_0^T |h_t|^2 dt}.$$

Assume $\mathbb{E}\left[e^{-\int_0^T h_t \cdot dB_t - \frac{1}{2}\int_0^T |h_t|^2 dt}\right] = 1.$

Then $\tilde{\mathbb{P}}$ is a probability measure. Moreover, under $\tilde{\mathbb{P}}$

- $\tilde{B}_t$ is a Brownian motion
- B_t becomes a Brownian motion with drift $-h_t$.

Remarks

- Define $Z_t = e^{-\int_0^t h_s \cdot dB_s - \frac{1}{2}\int_0^t |h_s|^2 ds}$. Then Z_t satisfies the SDE driven by the n dimensional Brownian motion

$$dZ_t = -Z_t h_t \cdot dB_t, \quad Z_0 = 1.$$

- Again, we can change the drift of an n dimensional Brownian motion by changing the measure.

4.2 De-drifting multidimensional Ito processes

Consider the n dimensional Ito process X_t driven by an m dimensional Brownian motion B_t

$$dX_t = \mu_t dt + \sigma_t dB_t$$

where μ_t is the drift n-vector and σ_t is the $n \times m$ diffusion matrix.

Define a new measure $\tilde{\mathbb{P}}$ by the Radon-Nikodym derivative

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = e^{\int_0^T \theta_t \cdot dB_t - \frac{1}{2}\int_0^T |\theta_t|^2 dt},$$

where θ_t is a yet to be determined m-vector of L^2 processes.

By Girsanov theorem, B_t becomes a Brownian motion with drift θ in $\tilde{\mathbb{P}}$, i.e., $dB_t = d\tilde{B}_t + \theta_t dt$, where $d\tilde{B}_t$ is a Brownian motion in $\tilde{\mathbb{P}}$.

Therefore, under $\tilde{\mathbb{P}}$, the SDE for X_t becomes

$$dX_t = \mu_t dt + \sigma_t dB_t = \mu_t dt + \sigma_t(d\tilde{B}_t + \theta_t dt) = (\mu_t + \sigma_t \theta_t)dt + \sigma_t d\tilde{B}_t.$$

Hence, X_t is driftless (and thus a martingale) in $\tilde{\mathbb{P}}$ if θ_t is a solution to the linear system

$$\sigma_t \theta_t + \mu_t = 0 \quad \rightsquigarrow \quad \theta_t = -\sigma_t^{-1} \mu_t \ \text{ if } \sigma_t \text{ invertible.}$$

In other words, we can de-drift X_t by a change of measure if the above linear system has a solution for every $t \in [0, T]$ (and of course the Radon-Nikodym derivative corresponding to θ_t defines a probability measure).

Remark

The argument above is the core of the existence of risk neutral prbability and equivalent martingale measure in derivative pricing theory.

5 Martingale representation theorem (MRT)

Bottom line: we know that stochastic integrals with respect to the Brownian motion are martingales.

The MRT states that (under conditions) the converse is also true.

5.1 MRT in dimension $d = 1$

Theorem

Let W_t be a Brownian motion on $(\Omega, \mathcal{F}_t, \mathbb{P})$, where $\mathcal{F}_t$ is the filtration generated by W_t (the **natural filtration**).

Let M_t be a square integrable martingale adapted to $\mathcal{F}_t$.

Then there exists an adapted process θ_t with $\mathbb{E}\left[\int_0^T \theta_t^2 dt\right] < \infty$ such that

$$M_t = M_0 + \int_0^t \theta_s dW_s.$$

Proof: omitted

Remarks

- The adaptedness to the natural filtration of Brownian motion is the key point in the martingale representation theorem. It basically says that if M_t does not contain more information than Brownian motion, then it can be recovered by a stochastic integral with respect to Brownian motion. The theorem does not hold if M_t contains information other than the information in Brownian motion.
- In general, it is hard, if not impossible, to characterize the process θ_t.

Example

Let $M_t = \mathbb{E}\left[W_T^2 \middle| \mathcal{F}_t\right]$.

Then M_t is a martingale adapted to the filtration generated by the Brownian motion W_t and $M_0 = \mathbb{E}\left[W_T^2\right] = T$. We determine an adapted process θ_t such that $M_t = M_0 + \int_0^t \theta_s dW_s$ as follows.

Apply Ito's formula we obtain

$$dW_t^2 = 2W_t dW_t + dt \quad \leftrightsquigarrow \quad W_T^2 = 2\int_0^T W_s dW_s + T.$$

Hence, we obtain the martingale representation for M_t as

$$M_t = \mathbb{E}\left[W_T^2 \middle| \mathcal{F}_t\right] = T + \int_0^t 2\, W_s dW_s.$$

5.2 MRT in dimension $d \geq 1$

Assume that $(\mathcal{F})_{t\in[0,T]}$ is the filtration generated by the standard d-dimensional Brownian motion $(B_t)_{t\in[0,T]}$.

Let $(M_t)_{t\in[0,T]}$ be a (one-dimensional) $\mathcal{F}_t$-martingale with respect to P.

Then there is an adapted d-dimensional process $(\theta_t)_{t\in[0,T]}$, such that

$$M_t = M(0) + \int_0^t \theta_u \cdot dB_u, \quad t \in [0,T].$$