University of Copenhagen

Msc in Actuarial Mathematics

Exercises

Author:
Joakim Bilyk

Date:
March 30, 2023

Abstract

This document contain exercises in probability theory and mathematical statistics applied in finance, life insurance and non-life insurance.

Keywords: *probability theory, insurance mathematics, life insurance, non-life insurance, stochastic differential equations.*

Contents

Chapter 1

Introduction

1.1 Abbreviations

Below is given the abbreviations used when referencing to books:

Chapter	Abbreviation	Source
Basic Life Insurance Mathematics		
Stochastic Processes in Life Insurance Mathematics		
Life Insurance Mathematics	Asmussen	*Risk and Insurance*: A Graduate Text by Soren Asmussen and Mogens Steffensen (2020).
	Bladt	Notes from lectures in Liv2.
Topics in Life Insurance Mathematics	Asmussen	*Risk and Insurance*: A Graduate Text by Soren Asmussen and Mogens Steffensen (2020).
Continuous Time Finance	Bjork	*Arbitrage Theory in Continuous Time (Fourth edition)* by Thomas Bjork, Oxford University Press (2019).
Basic Non-Life Insurance Mathematics		
Stochastic Processes in Life Insurance Mathematics		
Topics in Non-Life Insurance Mathematics		
Probabilistic Machine Learning	*None*	Slides from lectures.
Quantative Risk Management		
Measure Theory	Bjork	*Arbitrage Theory in Continuous Time (Fourth edition)* by Thomas Bjork, Oxford University Press (2019).
	Protter	*Probability Essentials (2. edition)* by Jean Jacod and Philip Protter (2004).
Random Variables	Bjork	*Arbitrage Theory in Continuous Time (Fourth edition)* by Thomas Bjork, Oxford University Press (2019).
	Hansen	*Stochastic Processes* (2. edition) by Ernst Hansen (2021).
Discrete Time Stochastic Processes	Hansen	*Stochastic Processes* (2. edition) by Ernst Hansen (2021).

Chapter	Abbreviation	Source
Continuous Time Stochastic Processes	Bjork	*Arbitrage Theory in Continuous Time (Fourth edition)* by Thomas Bjork, Oxford University Press (2019).
Stochastic Calculus	Bjork	*Arbitrage Theory in Continuous Time (Fourth edition)* by Thomas Bjork, Oxford University Press (2019).
	Bladt	Notes from lectures in Liv2.
Linear Algebra	Wiki	Wikipedia

1.2 To-do work

Chapter	Note	Progress
ML	Exercises week 1	

Chapter 2

Continuous Time Finance

2.1 Week 1

Probability exercises

Let $(W(t))_{t\geqslant}$ be a Brownian motion (Bjork, Definition 4.1).

Exercise 1. Show that the following processes also are Brownian motions.

i. $(-W(t))_{t\geqslant 0}$ (symmetry)
ii. For any $s \geqslant 0$, $(W(t+s) - W(s))_{t\geqslant 0}$ (time-homogeneity).
iii. For every $c > 0$, $(cW(t/c^2))_{t\geqslant 0}$ (scaling).

Solution (i).

By assumption W is a Brownian motion and so it follows that

$$-W_0 = -1 \cdot 0 = 0$$

Furthermore, for $r < s \leqslant t < u$ it holds that $W_u - W_t$ and $W_s - W_r$ is independent. By seperate transformations the independence property is preserved and $-(W_u - W_t)$ and $-(W_s - W_r)$ is independent. Next, for a normal distributed random variable $N \sim \mathcal{N}(\mu, \sigma^2)$ it holds, that for a scaler $c \in \mathbb{R}$ we have $cN \sim \mathcal{N}(c\mu, c^2\sigma^2)$. Then obviously;

$$-(W_t) = (-1)W_t \overset{d}{=} \mathcal{N}((-1)\cdot 0, (-1)^2(t-s)) \overset{d}{=} \mathcal{N}(0, t-s).$$

Lastly, let $\omega \in \Omega$ and consider the sample path $s \mapsto (-W_s)(\omega)$. Clearly for two continuous functions f and g it holds that $(g \circ f)$ is continuous. Then with $g(f) = -f$ and $f(t) = W_t(\omega)$"/ > it follows that $(-W_t) = (g \circ W)(t)$ is also continuous.

Solution (ii).

Much like the previous exercise we define a new process and show the properties hold. Let $s \geqslant 0$ be chosen arbitrary. Now define $X_t = W(t+s) - W(s)$.

First, we let $t = 0$ and see

$$X_0 = W(0+s) - W(s) = W(s) - W(s) = 0.$$

Secondly, we have that for $r < u$:

$$X_u - X_r = W(u+s) - W(s) - (W(r+s) - W(s)) = W(u+s) - W(r+s) \sim \mathcal{N}(0, u+s-(r+s)) = \mathcal{N}(0, u-r).$$

and since for $r < u \leqslant k < l$ the translation $r+s < u+s \leqslant k+s < l+s$ still holds and $X_l - X_k = W(l+s) - W(k+s)$ and $X_u - X_r = W(u+s) - W(k+s)$ are independent. Finally since $W_t(\omega)$ is continuous in t hence the translation W_{t+s} is continouos. Adding a constant yields a function that is also continuous, hence X_t is continuous.

Solution (iii).

Let $c > 0$ be given. We show that

$$X_t = cW\left(\frac{t}{c^2}\right)$$

is a Brownian motion. We simply show the four properties. Let $t = 0$ and notice

$$X_0 = cW\left(\frac{0}{c^2}\right) = cW(0) = 0.$$

The second property follows from seperate transformation and that for $r < u \leqslant s < t$ we consider

$$X_u - X_r = c\left(W\left(\frac{u}{c^2}\right) - W\left(\frac{r}{c^2}\right)\right) \qquad \text{and} \qquad X_t - X_s = c\left(W\left(\frac{t}{c^2}\right) - W\left(\frac{s}{c^2}\right)\right)$$

and since $c, r, u, t, s > 0$ we have the same order for the scaled version of r, u, t, s and hence we have two independent RV scaled by c. Then by seperate transformations the variables is still independent. Next for the third property:

$$X_t - X_s = c\left(W\left(\frac{t}{c^2}\right) - W\left(\frac{s}{c^2}\right)\right) \sim \mathcal{N}\left(c \cdot 0, c^2\left(\frac{t}{c^2} - \frac{s}{c^2}\right)\right) = \mathcal{N}(0, t-s).$$

Where we use the properties of scaling a normal distributed random variable i.e. for $c > 0$ and $N \sim \mathcal{N}(\mu, \sigma^2)$ it follows that $cN \sim \mathcal{N}(c\mu, c^2\sigma^2)$. Finally, the forth property follows since $g(f) = cf$ is continuous and $h(t) = t/c^2$ is continuous, then for any continuous function $f(s)$ it follows that $(g \circ f \circ h) = g(f(h(t)))$ is continuous.

Proposition B.37. Let $(\Omega, \mathcal{F}, P)$ be a given probability space, let $\mathcal{G}$ be a sub-sigma-algebra of $\mathcal{F}$, and let X be a square integrable random variable. Consider the problem of minimizing

$$E\left[(X - Z)^2\right]$$

where Z is allowed to vary over the class of all square integrable $\mathcal{G}$ measurable random variables. The optimal solution $\hat{Z}$ is then given by.
$$\hat{Z} = E[X|\mathcal{G}].$$

Exercise 2. *(Bjork, exercise B.11.)* Prove proposition B.37 by going along the following lines.

a. Prove that the "estimation error" $X - E[X|\mathcal{G}]$ is orthogonal to $L^2(\Omega, \mathcal{G}, P)$ in the sence that for any $Z \in L^2(\Omega, \mathcal{G}, P)$ we have
$$E[Z \cdot (X - E[X|\mathcal{G}])] = 0$$
b. Now prove the proposition by writing

$$X - Z = (X - E[X|\mathcal{G}]) + (E[X|\mathcal{G}] - Z)$$

and use the result just proved.

Solution (a).

Let $X \in L^2(\Omega, \mathcal{F}, P)$ be a random variable. Now consider an arbitrary $Z \in L^2(\Omega, \mathcal{G}, P)$. Recall that $\mathcal{G} \subset \mathcal{F}$ and so X is also in $Z \in L^2(\Omega, \mathcal{G}, P)$, as it is bothe square integrable and $\mathcal{G}$-measurable. Then

$$E\left[Z \cdot (X - E[X|\mathcal{G}])\right] = E\left[Z \cdot X\right] - E\left[Z \cdot E[X|\mathcal{G}]\right].$$

Then by using the law of total expectation and secondly that Z is $\mathcal{G}$-measurable we have that

$$E\left[Z \cdot X\right] = E\left[E[Z \cdot X|\mathcal{G}]\right] = E\left[Z \cdot E[X|\mathcal{G}]\right].$$

Combining the two equations gives the desired result.

Solution (b).

Obviously, we have that

$$X - Z = X - Z + E[X|\mathcal{G}] - E[X|\mathcal{G}] = (X - E[X|\mathcal{G}]) + (E[X|\mathcal{G}] - Z).$$

Then squaring the terms gives

$$(X - Z)^2 = (X - E[X|\mathcal{G}])^2 + (E[X|\mathcal{G}] - Z)^2 + 2(X - E[X|\mathcal{G}])(E[X|\mathcal{G}] - Z)$$

Taking expectation on each side and using linearity of the expectation we have that

$$E[(X - Z)^2] = E\left[(X - E[X|\mathcal{G}])^2\right] + E\left[(E[X|\mathcal{G}] - Z)^2\right] + 2E\left[(X - E[X|\mathcal{G}])(E[X|\mathcal{G}] - Z)\right].$$

We can now use that $E[X|\mathcal{G}] - Z$ is $\mathcal{G}$-measurable with the above result on the last term.

$$E[(X - Z)^2] = E\left[(X - E[X|\mathcal{G}])^2\right] + E\left[(E[X|\mathcal{G}] - Z)^2\right].$$

Now since X is given the term $E\left[(X - E[X|\mathcal{G}])^2\right]$ is simply a constant not depending on the choice og Z. The optimal choice of Z is then $E[X|\mathcal{G}]$ since this minimizes the second term. The statement is then proved.

Exercise 3. Discuss the following theory/results of Moment generating functions (Laplace transform).

Let X be a random variable with distribution function $F(x) = P(X \leqslant x)$ and Y be a random variable with distribution function $G(y) = P(Y \leqslant y)$.

Definition. The moment generating function or Laplace transform of X is

$$\psi_X(\lambda) = E\left[e^{\lambda X}\right] = \int_{-\infty}^{\infty} e^{\lambda x} dF(x)$$

provided the expectation is finite for $|\lambda| < h$ for some $h > 0$.

The MGF uniquely determine the distribution of a random variable, due to the following result.

Theorem 1. *(Uniqueness)* If $\psi_X(\lambda) = \psi_Y(\lambda)$ when $|\lambda| < h$ for some $h > 0$, then X and Y has the same distribution, that is, $F = G$.

There is also the following result of independence for Moment generating functions.

Theorem 1. *(Independence)* If

$$E\left[e^{\lambda_1 X + \lambda_2 Y}\right] = \psi_X(\lambda_1)\psi_Y(\lambda_2)$$

for $|\lambda_i| < h$ for $i = 1, 2$ for some $h > 0$, then X and Y are independent random variables.

Example. Recall that the Moment generating function of a normal (Gaussian) distribution is given by

$$\psi_X(\lambda) = E\left[e^{\lambda X}\right] = \exp\left(\lambda\mu + \frac{\lambda^2}{2}\sigma^2\right)$$

where X is normally distributed with mean μ and variance σ^2 and $\lambda \in \mathbb{R}$ is a constant. Since a Brownian motion $W(t)$ is normally distributed with zero mean and variance t, we have that

$$E[\exp(\lambda W(t))] = \exp\left(\frac{\lambda^2}{2}t\right).$$

Discussion.

Exercise 4. *(Bjork, exercise C.8.(a-c))* Let W be a Brownian motion. Notice that for the natural filtration $\mathcal{F}_s = \sigma(W_t | t \leqslant s)$ $W_t - W_s$ is independent of $\mathcal{F}_s$

a. Show that W_t is a martingale.
b. Show that $W_t^2 - t$ is a martingale.
c. Show that $\exp(\lambda W_t - \frac{\lambda^2}{2}t)$ is a martingale.

Solution (a).

We show that for the natural filtration that W_t is a martingale. This include showing integrability and the martingale property. For the first we note that for a normal distributed random variable with mean 0 we have

$$E[|N|] = \int_{-\infty}^{\infty} |x| dF_N(x) = 2\int_0^{\infty} x dF_N(x)$$

since the distribution is symmetric. Substituting the distribution function $\Phi(x) = P(N \leqslant x)$ in we see that

$$E[|N|] = 2\int_0^{\infty} x d\Phi(x) = 2\int_0^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/(2\sigma^2)} dx = (*)$$

by substituting $u = x^2/(2\sigma^2)$ $(x = \sqrt{2\sigma^2 u})$ we have that

$$\frac{dx}{du} = \frac{1}{2}\sqrt{2\sigma^2 u} 2\sigma^2 = (\sigma^2)^{3/2}\sqrt{2}u \iff dx = (\sigma^2)^{3/2}\sqrt{2}u\ du$$

hence

$$(*) = \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^{\infty} \sqrt{2\sigma^2 u} e^{-u} (\sigma^2)^{3/2} \sqrt{2} u\ du = \frac{2\sqrt{2\sigma^2}(\sigma^2)^{3/2}\sqrt{2}}{\sqrt{2\pi\sigma^2}} \int_0^{\infty} \sqrt{u} e^{-u} u\ du.$$

This then simplify to

$$(*) = \frac{(2\sigma^2)^{3/2}}{\sqrt{\pi}} \int_0^{\infty} u^{3/2} e^{-u}\ du = (2\sigma^2)^{1/2} \sqrt{\frac{2\sigma^2}{\pi}} \int_0^{\infty} u^{3/2} e^{-u}\ du = \sqrt{\frac{2\sigma^2}{\pi}} < \infty.$$

(Obviously the above is not derived correctly, but the end expression is valid, source: link) However since

$$W_t = W_t - 0 = W_t - W_0 \sim \mathcal{N}(0, t)$$

we have that $E|W_t| < \infty$ as desired.

Next, we have that

$$E[W_t|\mathcal{F}_s] = E[W_t - W_s|\mathcal{F}_s] + W_s = 0 + W_s = W_s.$$

In the above we used that $W_t - W_s$ is $\mathcal{F}_s$-measurable with mean 0. Then it follows that W_t is a martingale.

Solution (b).

Let $M_t = W_t^2 - t$. First, we observe that two measurable functions composed is still a measurable function. Hence we know that M_t is measurable wrt. the filtration since W_t is measurable and $w \mapsto w^2 + t$ is measurable. Secondly, we have that

$$E[|W_t^2 - t|] \leqslant E|W_t^2| + E|t| = t + t = 2t < \infty$$

where we use the triangle inequality. Thirdly, for the martingale property we have that for $t > s$:

$$E[M_t|\mathcal{F}_s] = E[W_t^2 - t|\mathcal{F}_s] = E[W_t^2 + W_s^2 - 2W_tW_s - W_s^2 + 2W_tW_s - t|\mathcal{F}_s]$$

which by linearity and independence of increments to the filtration gives

$$E[M_t|\mathcal{F}_s] = E[(W_t - W_s)^2 - W_s^2 + 2W_tW_s - t|\mathcal{F}_s] = t - s - t + E[2W_tW_s - W_s^2|\mathcal{F}_s]$$

However since W_s is measurable wrt. the filtration at time s the above is

$$E[M_t|\mathcal{F}_s] = 2W_sE[W_t|\mathcal{F}_s] - W_s^2 - s = 2W_s^2 - W_s^2 - s = W_s^2 - s = M_s.$$

Since from (a) we know that W_t is a martingale. Then we arrive at the desired result.

Solution (c).

Let $M_t = \exp\left(\lambda W_t - \frac{\lambda^2}{2}t\right)$. First, by composition of measurable functions M_t is $\mathcal{F}_t$-measurable. Secondly, we have using the MGF for a normal distributed random variable:

$$E|M_t = E\left(\exp\left(\lambda W_t - \frac{\lambda^2}{2}t\right)\right) \leqslant E\left(\exp\left(\lambda W_t\right)\right) = \exp\left(\frac{\lambda^2}{2}t\right) < \infty.$$

Thirdly, we consider

$$E[M_t|\mathcal{F}_s] = E\left[\left(\exp\left(\lambda W_t - \frac{\lambda^2}{2}t\right)\right)\Big|\mathcal{F}_s\right] = \exp\left(-\frac{\lambda^2}{2}t\right)E\left[\left(\exp\left(\lambda W_t\right)\right)|\,\mathcal{F}_s\right].$$

By adding and subtracting W_s in the exponent we get

$$\begin{aligned} E[M_t|\mathcal{F}_s] &= \exp\left(-\frac{\lambda^2}{2}t\right)E\left[\left(\exp\left(\lambda(W_t - W_s) + \lambda W_s\right)\right)|\,\mathcal{F}_s\right] \\ &= \exp\left(-\frac{\lambda^2}{2}t\right)\exp\left(\frac{\lambda^2}{2}(t-s)\right)E\left[\left(\exp\left(\lambda W_s\right)\right)|\,\mathcal{F}_s\right]. \end{aligned}$$

Using that $E\left[\left(\exp\left(\lambda W_s\right)\right)|\,\mathcal{F}_s\right] = \exp\left(\lambda W_s\right)$ and combining the exponents gives the desired:

$$E[M_t|\mathcal{F}_s] = \exp\left(\lambda W_s - \frac{\lambda^2}{2}s\right) = M_s.$$

2.2 Week 2

Exercise 1 *(Bjork 4.1)* Compute the stochastic differential dZ_t when

a. $Z_t = e^{\alpha t}$.
b. $Z_t = \int_0^t g_s\, dW_s$, where g is an adapted stochastic process.
c. $Z_t = e^{\alpha W_t}$.
d. $Z_t = e^{\alpha X_t}$, where X has stochastic differential $dX_t = \mu\, dt + \sigma\, dW_t$ and μ, σ is constants.
e. $Z_t = X_t^2$, where X has stochastic differential $dX_t = \alpha X_t\, dt + \sigma X_t\, dW_t$.

Solution (a).

Let $Z_t = e^{\alpha t}$, then we see that $f(t,x) = e^{\alpha t}$ and the the following relevant derivatives is

$$\frac{\partial f}{\partial t}(t,x) = \alpha e^{\alpha t}, \quad \frac{\partial f}{\partial x}(t,x) = 0, \quad \frac{\partial f}{\partial x^2}(t,x) = 0.$$

Since Z does not depend on any stochastic process, we will content with $X_t = 0$, that is $\mu_t = \sigma_t = 0$. Then by theorem 4.11 (Ito's formula) we have

$$dZ_t = \left(\alpha e^{\alpha t} + 0 + 0\right)\, dt + 0 = \alpha e^{\alpha t}\, dt,$$

as expected. ▫

Solution (b).

Let $Z_t = \int_0^t g_s\, dW_s$, where g is an adapted stochastic process. We see that if we set $X_t = \int_0^t g_s\, dW_s$ then

$$dX_t = 0\, dt + g_t\, dW_t.$$

Then we have the function $f(t,x) = x$ and the relevant derivatives are:

$$\frac{\partial f}{\partial t}(t,x) = 0, \quad \frac{\partial f}{\partial x}(t,x) = 1, \quad \frac{\partial f}{\partial x^2}(t,x) = 0.$$

This then gives

$$dZ_t = \left(0 + 0 + \frac{1}{2} g_t \cdot 0\right) \, dt + g_t \cdot 1 \, dW_t = g_t \, dW_t,$$

as expected. ▫

Solution (c).

Let $Z_t = e^{\alpha W_t}$. Then we may set $X_t = W_t$ and we then have $\mu_t = 0$ and $\sigma_t = 1$. The function $f(t,x) = e^{\alpha x}$ and the relevant derivatives are:

$$\frac{\partial f}{\partial t}(t,x) = 0, \quad \frac{\partial f}{\partial x}(t,x) = \alpha e^{\alpha x}, \quad \frac{\partial f}{\partial x^2}(t,x) = \alpha^2 e^{\alpha x}.$$

Then the dynamics of Z_t is as follows

$$\begin{aligned} dZ_t &= \left(0 + 0 + \frac{1}{2} 1^2 \alpha^2 e^{\alpha X_t}\right) \, dt + 1 \alpha e^{\alpha X_t} \, dW_t \\ &= \frac{\alpha^2}{2} e^{\alpha X_t} \, dt + \alpha e^{\alpha X_t} \, dW_t \\ &= \frac{\alpha^2}{2} Z_t \, dt + \alpha Z_t \, dW_t. \end{aligned}$$

As desired. ▫.

Solution (d).

Let $Z_t = e^{\alpha X_t}$, where X has stochastic differential $dX_t = \mu \, dt + \sigma \, dW_t$ and μ, σ is constants. Then we have been given the definition of X_t and we set $f(t,x) = e^{\alpha x}$. The relevant derivatives are then:

$$\frac{\partial f}{\partial t}(t,x) = 0, \quad \frac{\partial f}{\partial x}(t,x) = \alpha e^{\alpha x}, \quad \frac{\partial f}{\partial x^2}(t,x) = \alpha^2 e^{\alpha x}.$$

We may now derive the dynamics of Z_t:

$$\begin{aligned} dZ_t &= \left(0 + \mu \alpha e^{\alpha X_t} + \frac{1}{2} \sigma^2 \alpha^2 e^{\alpha X_t}\right) \, dt + \sigma \alpha e^{\alpha X_t} \, dW_t \\ &= \left(\mu + \frac{1}{2} \sigma^2 \alpha\right) \alpha e^{\alpha X_t} \, dt + \sigma \alpha e^{\alpha X_t} \, dW_t \\ &= \left(\mu + \frac{1}{2} \sigma^2 \alpha\right) \alpha Z_t \, dt + \sigma \alpha Z_t \, dW_t. \end{aligned}$$

As desired. ▫.

Solution (e).

Let $Z_t = X_t^2$, where X has stochastic differential $dX_t = \alpha X_t \, dt + \sigma X_t \, dW_t$. Then we set $f(t,x) = x^2$ and the relevant derivatives are:

$$\frac{\partial f}{\partial t}(t,x) = 0, \quad \frac{\partial f}{\partial x}(t,x) = 2x, \quad \frac{\partial f}{\partial x^2}(t,x) = 2.$$

Given this we have the dynamics of Z_t as follows

$$\begin{aligned} dZ_t &= \left(0 + \alpha X_t 2 X_t + \frac{1}{2} (\sigma X_t)^2 2\right) \, dt + \sigma X_t 2 X_t \, dW_t \\ &= \left(2\alpha + \sigma^2\right) X_t^2 \, dt + 2\sigma X_t^2 \, dW_t \\ &= \left(2\alpha + \sigma^2\right) Z_t \, dt + 2\sigma Z_t \, dW_t. \end{aligned}$$

As desired. ▫.

Exercise 2 *(Bjork 4.2)* Compute the stochastic differential for Z when $Z_t = (X_t)^{-1}$ and X has the stochastic differential

$$dX_t = \alpha X_t \; dt + \sigma X_t \; dW_t.$$

Furthermore, by using the definition $Z = X^{-1}$ you can in fact express the right-hand side of dZ entirely in terms of Z itself (rather then in terms of X). Thus Z satisfies a stochastic differential equation. Which one?

Solution.

We see that $f(t,x) = 1/x$ and so the relevant derivatives is

$$\frac{\partial f}{\partial t}(t,x) = 0, \quad \frac{\partial f}{\partial x}(t,x) = -\frac{1}{x^2}, \quad \frac{\partial f}{\partial x^2}(t,x) = \frac{2}{x^3}.$$

Then we by Ito's formula we have

$$\begin{aligned} dZ_t &= \left(0 - \alpha X_t \frac{1}{X_t^2} + \frac{1}{2}\sigma^2 X_t^2 \frac{2}{X_t^3}\right) \; dt - \sigma X_t \frac{1}{X_t^2} \; dW_t \\ &= \left(-\alpha \frac{1}{X_t} + \sigma^2 \frac{1}{X_t}\right) \; dt - \sigma \frac{1}{X_t} \; dW_t \\ &= (\sigma^2 - \alpha) Z_t \; dt - \sigma Z_t \; dW_t. \end{aligned}$$

We also notice that

$$Z_t = \frac{1}{X_t} \Rightarrow dZ_t = d\left(\frac{1}{X_t}\right) = -\left(\frac{1}{X_t}\right)^2 \; dX_t = -Z_t^2(\alpha X_t \; dt + \sigma X_t \; dW_t)$$

Hence we may insert $X_t = Z_t^{-1}$ and optain

$$dZ_t = -Z_t^2 \left(\alpha \frac{1}{Z_t} \; dt + \sigma \frac{1}{Z_t} \; dW_t\right) = -\alpha Z_t \; dt - \sigma Z_t \; dW_t.$$

Which clearly is faulty.. ▫

Exercise 3. *(Bjork 4.3)* Let $\sigma(t)$ be a given deterministic function of time and define the process X by

$$X_t = \int_0^t \sigma(s) \; dW_s.$$

Use the technique discribed in example 4.17 in order to show that the characteristic function of X_t (for a fixed t) is given by

$$E[e^{iuX_t}] = \exp\left\{-\frac{u^2}{2}\int_0^t \sigma^2(s) \; ds\right\}, \quad u \in \mathbb{R},$$

thus showing that X_t is normally distributed with zero mean and a variance given by

$$Var[X_t] = \int_0^t \sigma^2(s) \; ds.$$

Solution.

We follow along the lines of

1. Determine the dynamics of $Z_t = e^{iuX_t}$ (for fixed u).
2. Write the integral form of Z_t.
3. Take expectation.
4. Solve ODE.

"1)" Set $f(t,x) = e^{iuX_t}$ then the relevant derivatives are

$$\frac{\partial f}{\partial t}(t,x) = 0, \quad \frac{\partial f}{\partial x}(t,x) = iue^{iuX_t} = iuZ_t, \quad \frac{\partial f}{\partial x^2}(t,x) = i^2u^2e^{iuX_t} = -u^2Z_t.$$

Recall that $dX_t = \sigma(t)\ dW_t$, then by Ito's formula we have

$$dZ_t = \left(-\sigma(t)^2\frac{1}{2}u^2Z_t\right)\ dt + \sigma(t)iuZ_t\ dW_t. \tag{*}$$

"2)" We can now write (*) on integral form as below

$$Z_t = Z_0 - \frac{u^2}{2}\int_0^t \sigma^2(s)Z_s\ ds + iu\int_0^t \sigma(s)Z_s\ dW_s,$$

where $Z_0 = e^{iuX_0} = 1$.

"3)" Taking expectation now yields

$$E[Z_t] = 1 - \frac{u^2}{2}\int_0^t \sigma^2(s)E[Z_s]\ ds + iuE\left[\int_0^t \sigma(s)Z_s\ dW_s\right] = 1 - \frac{u^2}{2}\int_0^t \sigma^2(s)E[Z_s]\ ds,$$

since any expectaion of an integral wrt. a Brownian motion is 0 (proposition 4.5).

"4)" Now we see that the t-derivative gives

$$dE[Z_t] = -\frac{u^2}{2}\sigma^2(t)E[Z_t]\ dt, \quad E[Z_0] = 1.$$

This is a ordinary differential equation with solution $y(t) = \exp\{-u^2/2\int_0^t \sigma^2(s)\ ds\}$ (check by differentiating) hence

$$E[e^{iuX_t}] = E[Z_t] = \exp\left\{-\frac{u^2}{2}\int_0^t \sigma^2(s)\ ds\right\}.$$

We recognize this as the characteristic function of a normally distributed random variable with variance $\int_0^t \sigma^2(s)\ ds$ as desired. (X_t follows this distributions since characteristic functions determine the distribution)
□

Exercise 4 *(Bjork 4.4)* Suppose that X has the stochastic differential

$$dX_t = \alpha X_t\ dt + \sigma_t\ dW_t,$$

where α is a real number and σ_t is a integrable adapted stochastic process. Use the technique in example 4.17 in order to determine the function $m(t) = E[X_t]$.

Solution.

We follow the same steps as the previous exercise. We have been given the dynamics of X hence we may write it on integral form.

$$X_t = X_0 + \alpha\int_0^t X_s\ ds + \int_0^t \sigma(s)\ dW_s.$$

Then taking expectation now gives

$$E[X_t] = X_0 + \alpha\int_0^t E[X_s]\ ds.$$

Hence $E[X_t]$ follows from the solution to the ODE below

$$dE[X_t] = \alpha E[X_t] \Rightarrow E[X_t] = C\cdot\exp\{\alpha t\}.$$

Then obviously $C = X_0$ and we arrive at the solution $E[X_t] = X_0e^{\alpha t}$, where X_0 is some deterministic value.
□

Exercise 5 *(Bjork 4.5)* Suppose that the process X has a stochastic differential

$$dX_t = \mu_t \, dt + \sigma_t \, dW_t,$$

and that $\mu_t \geqslant 0$ with probability one for all $t \geqslant 0$. Show that this implies that X is a sub-martingale.

Solution.

Note that we are (strictly speaking) supposed to show adaptation and integrability, we will however only fokus on the submartingale property.

"$E[X_t|\mathcal{F}_s] \geqslant X_s$" Intuitively speaking, the statement is obvious since we have with probability one a positive upwards drift with Brownian distortion (i.e. martingale). Formally, we will show the statement by first writing X_t on integral form

$$X_t = x_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s.$$

And so

$$X_t - X_s = \int_s^t \mu_u \, du + \int_s^t \sigma_u \, dW_u.$$

We then have

$$\begin{aligned}
E[X_t \mid \mathcal{F}_s] - X_s &= E[X_t - X_s \mid \mathcal{F}_s] \\
&= E\left[\int_s^t \mu_u \, du + \int_s^t \sigma_u \, dW_u \,\middle|\, \mathcal{F}_s \right] \\
&= E\left[\int_s^t \mu_u \, du \,\middle|\, \mathcal{F}_s \right] + E\left[\int_s^t \sigma_u \, dW_u \,\middle|\, \mathcal{F}_s \right] \\
&= E\left[\int_s^t \mu_u \, du \,\middle|\, \mathcal{F}_s \right] \geqslant 0.
\end{aligned}$$

Then adding X_s to the above inequality yields the result. ▫

Exercise 6 *(Bjork 4.7)* The objective of this exercise is to give an argument for the formal identity

$$dW_1(t) \cdot dW_2(t) = 0,$$

when W_1 and W_2 are independent Brownian motions. Let us therefore fix a time t, and divide the inerval $[0,t]$ into equidistant points $0 = t_0 < t_1 < \cdots < t_n = t$, where $t_i = \frac{i}{n} \cdot t$. We use the notation

$$\Delta W_i(t_k) = W_i(t_k) - W_i(t_{k-1}), \; i = 1, 2.$$

Now define Q_n by

$$Q_n = \sum_{k=1}^{n} \Delta W_1(t_k) \cdot \Delta W_2(t_k).$$

Show that $Q_n \to 0$ in L^2, i.e. show that

$$E[Q_n] = 0, Var[Q_n] \to 0.$$

Solution.

We wish to show the statement

$$E[(Q_n - 0)^2] = E[Q_n^2] \to 0,$$

as $n \to \infty$. Recall that

$$Var[Q_n] = E[Q_n^2] - E[Q_n]^2,$$

hence if Q_n has mean 0, then showing convergence in L^2 is equivalent to showing variance going to 0. Let us start by showing the mean is 0.

We have that

$$\begin{aligned}
Q_n &= \sum_{k=1}^{n} \Delta W_1(t_k) \cdot \Delta W_2(t_k) \\
&= \sum_{k=1}^{n} (W_1(t_k) - W_1(t_{k-1})) \cdot (W_2(t_k) - W_2(t_{k-1})) \\
&\overset{\mathcal{D}}{=} \sum_{k=1}^{n} XY,
\end{aligned}$$

where $X, Y \sim \mathcal{N}(0, t_k - t_{k-1}) = \mathcal{N}(0, 1/n)$ and independent random variable. This is justified since the increments of the Brownian motion has mean 0 and variance equal to the increment size. Now this implies, that we need to show that $E[XY] = 0$ and that $Var[XY]$ is sufficiently small in terms of n such that it is summable. We see that

$$E[XY] = E[X]E[Y] = 0^2 = 0.$$

Here we use independence. We now know that the mean is

$$E[Q_n] = \sum_{k=1}^{n} E[XY] = 0.$$

We know from basic properties of variance that

$$\begin{aligned}
Var(Q_n) &= \sum_{k=1}^{n} Var(XY) = \sum_{k=1}^{n} E[(XY)^2] \\
&= \sum_{k=1}^{n} \frac{1}{n^2} = \frac{1}{n^2} n \\
&= \frac{1}{n} \to 0, \ n \to \infty.
\end{aligned}$$

And so the result follows. □

Exercise 7 *(Bjork 4.8)* Let X and Y be given as the solutions to the following system of stochastic differential equations.

$$\begin{aligned}
dX_t &= \alpha X_t \, dt - Y_t \, dW_t, & X_0 &= x_0, \\
dY_t &= \alpha Y_t \, dt + X_t \, dW_t, & Y_0 &= y_0.
\end{aligned}$$

Note that the initial values x_0 and y_0 are deterministic constants.

a. Prove that the process R defined by $R_t = X_t^2 + Y_t^2$ is deterministic.
b. Compute $E[X_t]$.

Solution (a).

We see that

$$dR_t = d(X_t^2 + Y_t^2) = d(X_t^2) + d(Y_t^2)$$

Hence we may start by considering de dynamics of the processes X_t^2 and Y_t^2. We see that for the process $Z_t = X_t^2$ we may set $f(t,x) = x^2$ and the relevant derivatives are

$$\frac{\partial f}{\partial t}(t,x) = 0,\ \frac{\partial f}{\partial x}(t,x) = 2x,\ \frac{\partial^2 f}{\partial x^2}(t,x) = 2.$$

By Ito's formula we have

$$d(X_t^2) = \left(\alpha X_t 2X_t + Y_t^2 2\right)\ dt - Y_t 2X_t\ dW_t = 2(\alpha X_t^2 + Y_t^2)\ dt - 2X_tY_t\ dW_t.$$

By the same concept we have

$$d(Y_t^2) = \left(\alpha Y_t 2Y_t + X_t^2 2\right)\ dt + X_t 2Y_t\ dW_t = 2(\alpha Y_t^2 + X_t^2)\ dt + 2X_tY_t\ dW_t.$$

Combining we get the dynamics

$$\begin{aligned} dR_t &= 2(\alpha X_t^2 + Y_t^2)\ dt - 2X_tY_t\ dW_t \\ &\quad + 2(\alpha Y_t^2 + X_t^2)\ dt + 2X_tY_t\ dW_t \\ &= (2\alpha + 1)(X_t^2 + Y_t^2)\ dt \\ &= (2\alpha + 1)R_t\ dt \end{aligned}$$

Hence R_t has deterministic derivative and therefore a deterministic process. In fact, the solution to above is

$$R_t = R_0 \exp\{(2\alpha + 1)t\} = (x_0^2 + y_0^2)e^{(2\alpha+1)t},$$

which is clearly deterministic. □

Solution (b).

We start by acknowledging that the differential form of X may be written on integral form:

$$X_t = x_0 + \alpha \int_0^t X_s\ ds - \int_0^t Y_s\ dW_s.$$

Taking expectation we see that

$$E[X_t] = x_0 + \int_0^t E[X_s]\ ds$$

as the last term has mean 0 according to proposition 4.5. Then the above may be written on the differential form

$$dE[X_t] = E[X_t]\ dt$$

Hence we have that

$$E[X_t] = x_0 e^t.$$

Hence X_t has mean not depending on the tragetory of the sister-process Y_t. □

2.3 Week 3

Exercise 1. *(Bjork 5.1)* Show that the scalar SDE

$$\begin{cases} dX_t = \alpha X_t \; dt + \sigma \; dW_t, \\ \qquad X_0 = x_0, \end{cases}$$

has the solution

$$X(t) = e^{\alpha t} x_0 + \sigma \int_0^t e^{\alpha(t-s)} \; dW_s,$$

by differentiating X as defined by the equation above and showing that X so defined satisfies the SDE.

Solution.

We move forward by rewriting the solution in terms of three processes Z, Y and R as

$$X_t = \underbrace{x_0 e^{\alpha t}}_{:=Y_t} + \underbrace{\sigma e^{\alpha t}}_{:=Z_t} \underbrace{\int_0^t e^{-\alpha s} \; dW_s}_{R_t} = Y_t + Z_t \cdot R_t.$$

We furthermore see easily that the dynamics of the processes individually has dynamics

$$\begin{aligned} dY_t &= \alpha x_0 e^{\alpha t} \; dt = \alpha Y_t \; dt, & Y_0 &= x_0, \\ dZ_t &= \alpha \sigma^{\alpha t} \; dt = \alpha Z_t \; dt, & Z_0 &= \sigma, \\ dR_t &= e^{-\alpha t} \; dW_s, & R_0 &= 0. \end{aligned}$$

We then have the following function

$$f\left(t, y, z, r\right) = y + zr.$$

With the following multi-dimensional process

$$dM_t = \begin{bmatrix} \alpha Y_t \\ \alpha Z_t \\ 0 \end{bmatrix} dt + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & e^{-\alpha t} \end{bmatrix} \begin{bmatrix} dW_t \\ dW_t \\ dW_t \end{bmatrix},$$

with

$$C = \sigma \sigma^\top = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & e^{-\alpha t} \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & e^{-2\alpha t} \end{bmatrix}.$$

That is $X_t = f(t, M_t)$. We can then use the multidimensional version of Ito's formula.

$$\begin{aligned} dX_t &= df(t, M_t) \\ &= \left(\frac{\partial f}{\partial t}(t, M_t) + \sum_{i=1}^{3} \mu_i \frac{\partial f}{\partial x^i}(t, M_t) + \frac{1}{2} \sum_{i,j=1}^{3} C_t^{ij} \frac{\partial^2 f}{\partial x^i \partial x^j}(t, M_t) \right) \; dt + \sum_{i=1}^{3} \frac{\partial f}{\partial x^i}(t, M_t) \sigma_t^i \; dW_t \\ &= (0 + \alpha Y_t + \alpha Z_t R_t) \; dt + Z_t e^{-\alpha t} \; dW_t \\ &= \left(\alpha x_0 e^{\alpha t} + \alpha \sigma e^{\alpha t} \int_0^t e^{-\alpha \sigma} \; dW_s \right) \; dt + \sigma e^{\alpha t} e^{-\alpha t} \; dW_t \\ &= \left(\alpha x_0 e^{\alpha t} + \alpha \sigma \int_0^t e^{(t-s)\alpha} \; dW_s \right) \; dt + \sigma \; dW_t \\ &= \alpha X_t \; dt + \sigma \; dW_t. \end{aligned}$$

Then this solution does in fact satisfies the differential form. We furthermore have that $X_0 = x_0$ and the desired result follows. ▫

Exercise 2. *(Bjork 5.5)* Suppose that X satisfies the SDE

$$dX_t = \alpha X_t \, dt + \sigma X_t \, dW_t.$$

Now define Y by $Y_t = X_t^\beta$, where β is a real number. Then Y is also a GBM process. Compute dY_t and find out which SDE Y satisfies.

Solution.

If we set $f(t,x) = x^\beta$, we have the relevant derivatives as follows

$$\frac{\partial f}{\partial t}(t,x) = 0, \ \frac{\partial f}{\partial x}(t,x) = \beta x^{\beta-1}, \ \frac{\partial^2 f}{\partial x^2}(t,x) = \beta(\beta-1)x^{\beta-2}.$$

Then by applying Ito's formula we have

$$\begin{aligned}
dY_t &= df(t,X_t) \\
&= \left(0 + \beta X_t^{\beta-1}\alpha X_t + \frac{1}{2}\sigma^2 X_t^2 \beta(\beta-1)X_t^{\beta-2}\right) \, dt + \sigma X_t \beta X_t^{\beta-1} \, dW_t \\
&= \left(\alpha\beta + \frac{1}{2}\sigma^2\beta(\beta-1)\right) X_t^\beta \, dt + \sigma\beta X_t^\beta \, dW_t \\
&= \left(\alpha\beta + \frac{1}{2}\sigma^2\beta(\beta-1)\right) Y_t \, dt + \sigma\beta Y_t \, dW_t \\
&= \alpha^Y Y_t \, dt + \sigma^Y Y_t \, dW_t,
\end{aligned}$$

where $\alpha^Y = \left(\alpha\beta + \frac{1}{2}\sigma^2\beta(\beta-1)\right)$ and $\sigma^Y = \sigma\beta$. Futhermore $Y_0 = y_0 = x_0^\beta$. Then by definition of GBM we have that Y_t is a GBM as desired. ▫

Exercise 3. *(Bjork 5.6)* Suppose that X satisfies the SDE

$$dX_t = \alpha X_t \, dt + \sigma X_t \, dW_t,$$

and Y satisfies

$$dY_t = \gamma Y_t \, dt + \delta Y_t \, dV_t,$$

where V is a Brownian motion which is independent of W. Define $Z = X/Y$ and derive an SDE for Z by computing dZ. If X is nominal income and Y describe inflation then Z describes real income.

Solution.

We have that for the function $f(t,x,y) = x/y$ and wish to determine the derivative of the stochastic process $Z_t = f(t,X_t,Y_t)$. We do this by applying Ito's formula in the multidimensional case. That is

$$\begin{aligned}
df(t,X_t,Y_t) &= \frac{\partial f}{\partial t}(t,X_t,Y_t) \, dt + \frac{\partial f}{\partial x}(t,X_t,Y_t) \, dX_t + \frac{\partial f}{\partial y}(t,X_t,Y_t) \, dY_t \\
&+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,X_t,Y_t)(dX_t)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(t,X_t,Y_t)(dY_t)^2 \\
&+ \frac{1}{2}\frac{\partial^2 f}{\partial x \partial y}(t,X_t,Y_t)(dX_t)(dY_t) \\
&= \frac{1}{Y_t}(\alpha X_t \, dt + \sigma X_t \, dW_t) - \frac{X_t}{Y_t^2}(\gamma Y_t \, dt + \delta Y_t \, dV_t) + \frac{1}{2}2\frac{X_t}{Y_t^3}(\gamma Y_t \, dt + \delta Y_t \, dV_t)^2 \\
&- \frac{1}{2}\frac{1}{Y_t^2}(\gamma Y_t \, dt + \delta Y_t \, dV_t)(\alpha X_t \, dt + \sigma X_t \, dW_t)
\end{aligned}$$

Calculating further gives

$$\begin{aligned}
(dY_t)^2 &= \gamma^2 Y_t^2 (dt)^2 + \delta^2 Y_t^2 (dV_t)^2 + 2\gamma Y_t \delta Y_t \; dt \cdot dV_t \\
&= 0 + \delta^2 Y_t^2 dt + 0 = \delta^2 Y_t^2 dt \\
(dX_t)(dY_t) &= (\gamma Y_t \; dt + \delta Y_t \; dV_t)(\alpha X_t \; dt + \sigma X_t \; dW_t) \\
&= \gamma Y_t \alpha X_t \; (dt)^2 + \gamma Y_t \sigma X_t \; dt \cdot dW_t \\
&+ \delta Y_t \alpha X_t \; dt \cdot dV_t + \gamma Y_t \sigma X_t \; (dW_t)(dV_t) \\
&= 0 + 0 + 0 + 0 = 0
\end{aligned}$$

Hence we conclude that

$$\begin{aligned}
df(t, X_t, Y_t) &= \alpha \frac{X_t}{Y_t} \; dt + \sigma \frac{X_t}{Y_t} \; dW_t - \gamma \frac{X_t Y_t}{Y_t^2} \; dt + \delta \frac{X_t Y_t}{Y_t^2} \; dV_t \\
&+ \frac{1}{2} 2 \frac{X_t}{Y_t^3} \delta^2 Y_t^2 dt \\
&= \left(\alpha Z_t - \gamma Z_t + Z_t \delta^2\right) \; dt + \sigma Z_t \; dW_t + \delta Z_t \; dV_t \\
&= \left(\alpha - \gamma + \delta^2\right) Z_t \; dt + \sigma Z_t \; dW_t + \delta Z_t \; dV_t.
\end{aligned}$$

As desired the above is the SDE for the process Z_t. □

Exercise 4. *(Bjork 5.9)* Use a stochastic representation result in order to solve the following boundary value problem in the domain $[0, T] \times \mathbb{R}$.

$$\frac{\partial F}{\partial t} + \mu x \frac{\partial F}{\partial x} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 F}{\partial x^2} = 0,$$

with $F(T, x) = \log(x^2)$. Here μ and σ are assumed to be know constants.

Solution.

We use proposition 5.5 Feymann-Kac with $\mu(t, x) = \mu x$ and $\sigma(t, x) = \sigma x$. We know that, given that the process

$$\sigma X_t \frac{\partial F}{\partial x}(x, X_t) \in \mathcal{L}^2,$$

Then F has stochastic representation

$$F(t, x) = E[\log(X_T^2) \mid X_t = x],$$

with stochastic process X_t satisfying the SDE

$$dX_t = \mu X_t \; dt + \sigma X_t \; dW_t, \qquad X_t = x.$$

Now, since X_t satisfies the above SDE, we see that X_t is a GBM. Then by proposition 5.2 we have

$$X_T = x \exp\left\{ \left(\mu - \frac{1}{2}\sigma^2\right)(T - t) + \sigma(W_T - W_t) \right\}.$$

Inserting this we find that

$$\begin{aligned}
F(t, x) &= E\left[\log(x^2 \exp\left\{\left(2\mu - \sigma^2\right)(T - t) + 2\sigma(W_T - W_t)\right\}) \;\middle|\; X_t = x\right] \\
&= E\left[\log(x^2) + \left(2\mu - \sigma^2\right)(T - t) + 2\sigma(W_T - W_t) \;\middle|\; X_t = x\right] \\
&= 2\log(x) + \left(2\mu - \sigma^2\right)(T - t) + 2\sigma E\left[W_T - W_t \mid X_t = x\right] \\
&= 2\log(x) + \left(2\mu - \sigma^2\right)(T - t).
\end{aligned}$$

Using that the Brownian motion has increments with mean 0. ▫

Exercise 5. *(Bjork 5.13)* Solve the boundary value problem

$$\frac{\partial F}{\partial t}(t,x,y)+\frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial x^2}(t,x,y)+\frac{1}{2}\delta^2\frac{\partial^2 F}{\partial y^2}(t,x,y)=0,$$

with $F(T,x,y)=xy$.

Solution.

We see if this problem fit into the context of Feymann-Kac's multi-dimensional proposition 5.8. Comparing the above PDE with the propositions we see that

$$\mu_i(t,X_t,Y_t)=0$$

for $i=1,2$ representing the assets X_t and Y_t. We furthermore have the matrix C

$$\begin{aligned}C&=\sigma(t,X_t,Y_t)\sigma(t,X_t,Y_t)^\top\\&=\begin{bmatrix}\sigma_{1,1}&\sigma_{1,2}\\\sigma_{2,1}&\sigma_{2,2}\end{bmatrix}\begin{bmatrix}\sigma_{1,1}&\sigma_{2,1}\\\sigma_{1,2}&\sigma_{2,2}\end{bmatrix}\\&=\begin{bmatrix}\sigma_{1,1}^2+\sigma_{1,2}^2&\sigma_{1,1}\sigma_{2,1}+\sigma_{2,2}\sigma_{1,2}\\\sigma_{1,1}\sigma_{2,1}+\sigma_{2,2}\sigma_{1,2}&\sigma_{2,2}^2+\sigma_{2,1}^2\end{bmatrix}\end{aligned}$$

Where we have that the only none-zero entrances is the diagonal with

$$\begin{aligned}C_{1,1}&=\sigma_{1,1}^2+\sigma_{1,2}^2=\sigma^2,\\C_{2,2}&=\sigma_{2,2}^2+\sigma_{2,1}^2=\delta^2.\end{aligned}$$

and obviously $r=0$ and $\Phi(x,y)=xy$. From proposition 5.8 we then have that F has stochastic representation:

$$F(t,x,y)=e^{-r(T-t)}E^Q[\Phi(X_T,Y_T)\mid X_t=x,Y_t=y]=E^Q[X_TY_T\mid X_t=x,Y_t=y]$$

Having given the C matrix we have the following

$$C=\begin{bmatrix}\sigma^2&0\\0&\delta^2\end{bmatrix}=\sigma\sigma^\top\iff\sigma=\begin{bmatrix}\sigma&0\\0&\delta\end{bmatrix}.$$

Hence in the stochastic representation X and Y has dynamics

$$\begin{cases}dX_t=\sigma\ dW_t\\dY_t=\delta\ dV_t.\end{cases}$$

where W and B are independent Brownian motions. In particular we have that X and Y are maringales and independent i.e.

$$F(t,x,y)=E^Q_{X_t=x}[X_T]\cdot E^Q_{Y_t=y}[Y_T]=xy.$$

And so we arrive at the desired result. ▫

Exercise 6. *(Exam 2017/18, problem 1, question (a)-(b))* Let W_t denote a Brownian motion and let

$$\mathcal{F}_t=\mathcal{F}_t^W=\sigma(\{W_s\mid 0\leqslant s\leqslant t\}).$$

Let $T>0$ be a given and fixed time.

Let $f(t)$ be a bounded deterministic continuous function. Define the two processes

$$\begin{cases} X_t = \int_0^t f(u)\ dW_u, \\ M_t^{(\lambda)} = \exp\left\{\lambda X_t - \frac{\lambda^2}{2}\int_0^t f^2(u)\ du\right\}, \end{cases}$$

where $\lambda \in \mathbb{R}$ is a constant.

a. Show that $M^{(\lambda)}$ is a martingale with $E[M_t^{(\lambda)}] = 1$.

Let $0 < s < t$ and $\lambda_1, \lambda_2 \in \mathbb{R}$ be given and fixed.

b. i. Show that

$$\begin{aligned} M_s^{(\lambda_1)} &= E\left[\frac{M_s^{(\lambda_1)} M_t^{(\lambda_2)}}{M_s^{(\lambda_2)}} \;\middle|\; \mathcal{F}_s\right] \\ &= E\left[\exp\left\{\lambda_1 X_s + \lambda_2(X_t - X_s) - \frac{1}{2}\lambda_1^2 \int_0^s f^2(u)\ du - \frac{1}{2}\lambda_2^2 \int_s^t f^2(u)\ du\right\} \;\middle|\; \mathcal{F}_s\right] \end{aligned}$$

ii. Show that X_s and $X_t - X_s$ are normally distributed and independent.

Solution (a).

First, we see that since X_t is on integral form we know that

$$\begin{cases} dX_t = f(t)\ dW_t \\ X_0 = 0. \end{cases}$$

Hence we may represent M as $M_t^{(\lambda)} = g(t, X_t, Y_t)$ given by

$$g(t, x, y) = \exp\left\{\lambda x - \frac{\lambda^2}{2} y\right\},$$

where $Y_t = \int_0^t f^2(u)\ du$ with dynamics

$$\begin{cases} dY_t = f^2(t)\ dt \\ Y_0 = 0. \end{cases}$$

Hence by the multidimensional Ito's formula we have the dynamics of M given by

$$\begin{aligned} dM_t^{(\lambda)} &= g_t\ dt + g_x\ dX_t + g_y\ dY_t + \frac{1}{2} g_{yy}\ (dY_t)^2 + \frac{1}{2} g_{xx}\ (dX_t)^2 + f_{xy}(dX_t)(dY_t) \\ &= 0 + \lambda g\ dX_t - \frac{\lambda^2}{2} g\ dY_t + 0 + \frac{1}{2}\lambda^2 g\ (dX_t)^2 + 0 \\ &= \lambda M_t^{(\lambda)} f(t)\ dW_t - \frac{1}{2}\lambda^2 M_t^{(\lambda)} f^2(t)\ dt + \frac{1}{2}\lambda M_t^{(\lambda)} f^2(t)\ dt \\ &= \lambda f(t) M_t^{(\lambda)}\ dW_t, \end{aligned}$$

And so we see that M is a martingale as it only has dynamics wrt. the Brownian motion W (assuming $\lambda f_t M_t^{(\lambda)} \in \mathcal{L}^2$). Furthermore we have that

$$M_0^{(\lambda)} = g(0, X_0, Y_0) = \exp\left\{\lambda X_0 - \frac{1}{2}\lambda^2 Y_0\right\} = e^0 = 1$$

and so we have $E[M_t^{(\lambda)}] = M_0^{(\lambda)} = 1$ as desired. □

Solution (b).

"(i)" We have from the previous exercise

$$
\begin{aligned}
&\frac{M_s^{(\lambda_1)} M_t^{(\lambda_2)}}{M_s^{(\lambda_2)}} \\
&= \exp\left\{\lambda_1 X_s - \frac{1}{2}\lambda_1^2 \int_0^s f^2(u)\ du\right\} \exp\left\{\lambda_2 X_t - \frac{1}{2}\lambda_2^2 \int_0^t f^2(u)\ du\right\} \exp\left\{\frac{1}{2}\lambda_2^2 \int_0^s f^2(u)\ du - \lambda_2 X_s\right\} \\
&= \exp\left\{\lambda_1 X_s - \frac{1}{2}\lambda_1^2 \int_0^s f^2(u)\ du + \lambda_2 X_t - \frac{1}{2}\lambda_2^2 \int_0^t f^2(u)\ du + \frac{1}{2}\lambda_2^2 \int_0^s f^2(u)\ du - \lambda_2 X_s\right\} \\
&= \exp\left\{\lambda_1 X_s + \lambda_2 (X_t - X_s) - \frac{1}{2}\lambda_1^2 \int_0^s f^2(u)\ du - \frac{1}{2}\lambda_2^2 \int_s^t f^2(u)\ du\right\}
\end{aligned}
$$

and so the conclusion follows. □

"(ii)" We have that from lemma 4.18 that

$$X_s = \int_0^s f(u)\ dW_u \sim \mathcal{N}\left(0, \int_0^s f^2(u)\ dW_u\right)$$

furthermore we have that

$$X_t - X_s = \int_s^t f(u)\ dW_u \sim \mathcal{N}\left(0, \int_s^t f^2(u)\ dW_u\right).$$

In regard to the independence claim we could check identity below

$$E[e^{t_1 X} e^{t_2 Y}] = E[e^{t_1 X}] E[e^{t_2 Y}]$$

where X, Y are independent random variables. The above identity holds if and only if X and Y are independent. From above we have that

$$M_s^{(\lambda_1)} = E[e^{\lambda_1 X_s} e^{\lambda_2 (X_t - X_s)} \mid \mathcal{F}_s] e^{-\frac{1}{2}\lambda_1^2 \int_0^s f^2(u)\ du - \frac{1}{2}\lambda_2^2 \int_s^t f^2(u)\ du}$$

and so taking expectation we have

$$1 = E[e^{\lambda_1 X_s} e^{\lambda_2 (X_t - X_s)}] e^{-\frac{1}{2}\lambda_1^2 \int_0^s f^2(u)\ du - \frac{1}{2}\lambda_2^2 \int_s^t f^2(u)\ du}$$

Which the gives

$$E[e^{\lambda_1 X_s} e^{\lambda_2 (X_t - X_s)}] = e^{\frac{1}{2}\lambda_1^2 \int_0^s f^2(u)\ du + \frac{1}{2}\lambda_2^2 \int_s^t f^2(u)\ du} = E[e^{\lambda_1 X_s}] E[e^{\lambda_2 (X_t - X_s)}]$$

and so the conclusion is that X_s and $X_t - X_s$ are independent. □

Exercise 7. *(Exam 2019/20, problem 1, question (a))*

Solution.

Exercise 8. *(Exam 2020/21, problem 1, question (a)-(b))*

Solution.

Extra-Exercise 1. *(Bjork 5.7)*

Solution.

Extra-Exercise 2. *(Bjork 5.8)*

Solution.

Extra-Exercise 3. *(Bjork 5.10)*

Solution.

Extra-Exercise 4. *(Bjork 5.11)*

Solution.

Extra-Exercise 5. *(Bjork 5.12)*

Solution.

2.4 Week 4

Exercise 1. *(Bjork 7.1)* Consider the standard Black-Scholes model and a T-claim $\mathcal{X}$ of the form $\mathcal{X} = \Phi(S_t)$. Denote the corresponding arbitrage free price processes by Π_t.

a. Show that, under the martingale measure Q, Π_t has a local rate of return equal to the short rate r. In other words shot that Π_t has a differential of the form

$$d\Pi_t = r\Pi_t \ dt + g_t \ dW_t^Q.$$

Hint: Use the Q-dynamics of S together with the fact that F satisfies the pricing PDE.

b. Show that, under the martingale measure Q, the process $Z_t = \frac{\Pi_t}{B_t}$ is a **martingale**. More precisely show that the stochastic differential for Z has zero drift term, i.e. is of the form

$$dZ_t = Z_t \sigma_t^Z \ dW_t^Q.$$

Determine also the diffusion process Σ_t^Z (in terms of the pricing function F and its derivatives).

Solution (a).

First we have that the dynamics of S and B are given by

$$\begin{aligned} dS_t &= \mu S_t \ dt + \sigma S_t \ dW_t, \\ dB_t &= rB_t \ dt. \end{aligned}$$

We know that $\Pi_t[X] = F(t, S_t)$ for some smooth function F hence by Ito's formula we have

$$\begin{aligned} d\Pi_t &= \frac{\partial F}{\partial t}(t, S_t) \ dt + \frac{\partial F}{\partial s}(t, S_t) \ dS_t + \frac{1}{2}\frac{\partial^2 F}{\partial s^2}(t, S_t) \ (dS_t)^2 \\ &= F_t \ dt + F_s(rS_t \ dt + \sigma S_t \ dW_t^Q) + \frac{1}{2}F_{ss}(rS_t \ dt + \sigma S_t \ dW_t^Q)^2 \\ &= (F_t + rS_tF_s) \ dt + F_s\sigma S_t \ dW_t^Q + \frac{1}{2}F_{ss}\sigma^2 S_t^2 \ dt \\ &= (F_t + rS_tF_s + \frac{1}{2}F_{ss}\sigma^2 S_t^2) \ dt + F_s\sigma S_t \ dW_t^Q. \end{aligned}$$

By setting $g_t = F_s\sigma S_t$ and restating the Black-Scholes equation we have

$$rF = F_t + rsF_s + \frac{1}{2}F_{ss}\sigma^2 s^2$$

hence

$$d\Pi_t = rF \ dt + g_t \ dW_t^Q$$

as desired. ▫

Solution (b).

Consider the stochastic process $Z_t = F(t, \Pi_t, B_t)$ given by $F(t, \pi, b) = \frac{\pi}{b}$. Then by Ito's formula we have the dynamics of Z as

$$
\begin{aligned}
dZ_t &= F_t\ dt + F_\pi\ d\Pi_t + F_b\ dB_t + \frac{1}{2}F_{\pi\pi}\ (d\Pi_t)^2 + \frac{1}{2}F_{bb}\ (dB_t)^2 + F_{b\pi}\ (d\Pi_t)(dB_t) \\
&= 0 + \frac{1}{B_t}\ d\Pi_t - \frac{\Pi_t}{B_t^2}\ dB_t + \frac{1}{2}0g_t\ dt + \frac{\Pi_t}{B_t^3}0 - \frac{1}{B_t^2}0 \\
&= \frac{1}{B_t}(r\Pi_t\ dt + g_t\ dW_t^Q) - \frac{\Pi_t}{B_t^2}(rB_t\ dt) \\
&= \frac{1}{B_t}(r\Pi_t\ dt - r\Pi_t\ dt + g_t\ dW_t^Q) \\
&= \frac{1}{B_t}g_t\ dW_t^Q.
\end{aligned}
$$

hence we have that Z_t is a Q-martingale since it only has dynamics in terms of the Brownian motion W^Q. Additionally we may represent the process as

$$
Z_t = \Pi_0 + \int_0^t \frac{g_s}{B_s}\ dW_s^Q.
$$

From this it is clear that Z_t is a Q-martingale. ▫

Exercise 2. *(Bjork 7.2)* Consider the standard Black-Scholes model. An innovative company, F&H Inc., has produced the derivative "the Golden Logarith", henceforth abbreviated as the GL. The holder of a GL with maturity time T, denotet as GL_t, will, at time T, obtain the sum $\Phi(S_T) = \log\ S_T$. Note that if $S_T < 1$ this means that the holder has to pay a positive amount to F&H Inc. Determine the arbitrage free price process for the GL_t.

Solution.

We know that in the BS model the simple derivative has to have the smooth pricing function $F(t,s)$ that is the solution to the boundary value problem.

$$
\begin{cases}
F_t(t,s) + rsF_s(t,s) + \frac{1}{2}\sigma^2 s^2 F_{ss}(t,s) - rF(t,s) = 0 \\
F(T,s) = \Phi(s).
\end{cases}
$$

Which has the stochastic representation (proposition 7.11) given by the risk neutral valuation formula:

$$
F(t,s) = e^{-r(T-t)}E_{t,s}^Q[\Phi(S_T)] = e^{-r(T-t)}E_{t,s}^Q[\log(S_T)],
$$

with S_t having dynamics

$$
dS_t = rS_t\ dt + \sigma S_t\ dW_t^Q.
$$

Then by Ito's formula on the function $f(t,s) = \log(s)$ we have

$$
d(\log(S_t)) = df(t,S_t) = f_t\ dt + f_s\ dS_t + \frac{1}{2}f_{ss}(dS_t)^2.
$$

Since $f_t = 0$ and $f_s = 1/s$ and $f_{ss} = -1/s^2$ we have

$$
\begin{aligned}
d(\log(S_t)) &= \frac{1}{S_t}(rS_t\ dt + \sigma S_t\ dW_t^Q) - \frac{1}{2}\sigma^2 S_t^2 \frac{1}{S_t^2}\ dt \\
&= \left(r - \frac{1}{2}\sigma^2\right)\ dt + \sigma\ dW_t^Q.
\end{aligned}
$$

Then given that $S_t = s$ we have

$$\begin{aligned}\log(S_u) &= \log s + \int_t^u r - \frac{1}{2}\sigma^2 \, dv + \sigma \int_t^u dW_v^Q \\ &= \log s + \left(r - \frac{1}{2}\sigma^2\right)(u-t) + \sigma(W_u^Q - W_t^Q)\end{aligned}$$

Taking expectation we then have

$$E^Q[\log(S_T) \mid S_t = s] = \log s + \left(r - \frac{1}{2}\sigma^2\right)(T-t),$$

given that

$$\Pi_t(\Phi(S_t)) = e^{-r(T-t)} \log s + e^{-r(T-t)}\left(r - \frac{1}{2}\sigma^2\right)(T-t).$$

The arbitrage free price is the the above. $\square$

Exercise 3. *(Bjork 7.4)* Consider the standard Black-Scholes model. Derive the arbitrage free price process for the T-claim X where X is given by $X = S_T^\beta$. Here β is a known constant.

Solution.

We may solve the pricing problem by evaluating the risk neautral valuation formula 7.11 given by

$$F(t,s) = e^{-r(T-t)} E_{t,s}^Q[\Phi(S_T)] = e^{-r(T-t)} E_{t,s}^Q[S_T^\beta].$$

We know that S_t is a Geometric Brownian motion with drift $(r - \sigma^2/2)$ and diffusion σ i.e.

$$S_T = s \cdot \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma(W_T^Q - W_t^Q)\right\}$$

wrt. the martingale-measure Q. Here we assume that $S_t = s$. Then we know that S_T^β is likewise a GBM given as

$$S_T^\beta = s^\beta \cdot \exp\left\{\beta\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \beta\sigma(W_T^Q - W_t^Q)\right\}.$$

Evaluating the price process then becomes

$$\begin{aligned}\Pi_t &= e^{-r(T-t)} E^Q\left[s^\beta \cdot \exp\left\{\beta\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \beta\sigma(W_T^Q - W_t^Q)\right\}\right] \\ &= e^{-r(T-t)} s^\beta e^{\beta\left(r - \frac{1}{2}\sigma^2\right)(T-t)} E^Q\left[e^{\beta\sigma(W_T^Q - W_t^Q)}\right].\end{aligned}$$

Evaluating the expectation may be done by calculating the MGF of a $\mathcal{N}(0, T-t)$ variable i.e.

$$E^Q\left[e^{\beta\sigma(W_T^Q - W_t^Q)}\right] = e^{0 \cdot 1 + (T-t)(\beta\sigma)^2/2} = e^{\beta^2\sigma^2\frac{T-t}{2}}.$$

Combining these two equations yields

$$\log \Pi_t = -r(T-t) + \beta \log s + \beta\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \beta^2\sigma^2\frac{T-t}{2} = -r(T-t) + \beta \log s + \left(\beta r + \frac{1}{2}\sigma^2\beta(\beta-1)\right)(T-t).$$

We arrive at the arrived result. $\square$

Exercise 4. *(Bjork 7.5)* A so-called **binary option** is a claim which pays a certain amount if the stock prices at a certain date falls within some pre-specified interval. Otherwise nothing will be paid out. Consider a binary option which pays K dollars to the holder at date T if the stock price at time T is in the interval $[\mu, \beta]$. Determine the arbitrage free price. The pricing formula will involve the standard Gaussian cumulative distrbution function N.

Solution.

First we see that the claim may be written on the form

$$\Phi(S_T) = 1_{S_T \in [a,b]} K,$$

using the values $a < b$ for the interval endpoints (instead of μ, β). We then by the risk neutral valuation formula must have

$$\Pi_t = e^{-r(T-t)} E^Q_{t,s}[1_{S_T \in [a,b]} K].$$

Under closer inspection we see that we must evaluating the expectation of the indicator under the measure Q. Then we have

$$E^Q_{t,s}[1_{S_T \in [a,b]}] = Q(S_T \in [a,b] \mid S_t = s) = (*).$$

When assuming $S_t = s$ we must have that under Q that

$$S_T = s \cdot \exp\left\{ \left(r - \frac{1}{2}\sigma^2 \right)(T-t) + \sigma \left(W^Q_T - W^Q_t \right) \right\}.$$

Then

$$\begin{aligned}
(*) &= Q(S_T \leqslant b \mid S_t = s) - Q(S_T < a \mid S_t = s) \\
&= Q(S_T \leqslant b \mid S_t = s) - Q(S_T \leqslant a \mid S_t = s) \\
&= Q\left(s \cdot \exp\left\{ \left(r - \frac{1}{2}\sigma^2 \right)(T-t) + \sigma \left(W^Q_T - W^Q_t \right) \right\} \leqslant b \right) \\
&- Q\left(s \cdot \exp\left\{ \left(r - \frac{1}{2}\sigma^2 \right)(T-t) + \sigma \left(W^Q_T - W^Q_t \right) \right\} \leqslant a \right) \\
&= Q\left(\frac{1}{\sqrt{T-t}} \left(W^Q_T - W^Q_t \right) \leqslant \frac{1}{\sigma\sqrt{T-t}} \left\{ \log b - \log s + \left(\frac{1}{2}\sigma^2 - r \right)(T-t) \right\} \right) \\
&- Q\left(\frac{1}{\sqrt{T-t}} \left(W^Q_T - W^Q_t \right) \leqslant \frac{1}{\sigma\sqrt{T-t}} \left\{ \log a - \log s + \left(\frac{1}{2}\sigma^2 - r \right)(T-t) \right\} \right) \\
&= N\left(\frac{1}{\sigma\sqrt{T-t}} \left\{ \log b - \log s + \left(\frac{1}{2}\sigma^2 - r \right)(T-t) \right\} \right) \\
&- N\left(\frac{1}{\sigma\sqrt{T-t}} \left\{ \log a - \log s + \left(\frac{1}{2}\sigma^2 - r \right)(T-t) \right\} \right) \\
&= N(d_b) - N(d_a),
\end{aligned}$$

as $W^Q_T - W^Q_t$ is $\mathcal{N}(0, T-t)$ distributed under the Q-measure. The function d_c is defined as

$$d_c = \frac{1}{\sigma\sqrt{T-t}} \left\{ \log c - \log s + \left(\frac{1}{2}\sigma^2 - r \right)(T-t) \right\},$$

as expected. We then arrive at the price

$$\Pi_t = e^{-r(T-t)} K(N(d_b) - N(d_a)),$$

as desired. □

Exercise 5. *(Bjork 7.6)* Consider the standard Black-Scholes model. Derive the arbitrage free price process for the T-claim X where X is given by $X = \frac{S_{T_1}}{S_{T_0}}$. The times T_0 and T_1 are given and the claim is paid out at time T_1.

Solution.

Firstly, we may use that S is a GBM. Setting $S_t = s$ we have that

$$\begin{aligned}
X = \frac{S_{T_1}}{S_{T_0}} &= \frac{s \cdot \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T_1 - t) + \sigma\left(W_{T_1}^Q - W_t^Q\right)\right\}}{s \cdot \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T_0 - t) + \sigma\left(W_{T_0}^Q - W_t^Q\right)\right\}} \\
&= \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T_1 - t - T_0 + t) + \sigma\left(W_{T_1}^Q - W_t^Q - W_{T_0}^Q + W_t^Q\right)\right\} \\
&= \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T_1 - T_0) + \sigma\left(W_{T_1}^Q - W_{T_0}^Q\right)\right\}.
\end{aligned}$$

Then it follows from proposition 7.11 that the price process is given by

$$\begin{aligned}
\Pi_t &= e^{-r(T_1-t)} E_{t,s}^Q[X] \\
&= e^{-r(T_1-t)} E^Q\left[\exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T_1 - T_0) + \sigma\left(W_{T_1}^Q - W_{T_0}^Q\right)\right\}\right] \\
&= \exp\left\{-r(T_1 - t) + \left(r - \frac{1}{2}\sigma^2\right)(T_1 - T_0)\right\} E^Q\left[\exp\left\{\sigma\left(W_{T_1}^Q - W_{T_0}^Q\right)\right\}\right] \\
&= \exp\left\{-r(T_1 - t) + \left(r - \frac{1}{2}\sigma^2\right)(T_1 - T_0) + \frac{1}{2}\sigma^2(T_1 - T_0)\right\} \\
&= \exp\left\{-r(T_1 - t) + r(T_1 - T_0)\right\} \\
&= \exp\left\{-r(T_1 + T_0 - T_1 - t)\right\} = \exp\left\{-r(T_0 - t)\right\}
\end{aligned}$$

as desired. □

Exercise 6. *(Exam 2017/18, problem 2, question (c)-(d))* Consider a standard Black-Scholes model, that is, a model consisting of a bank account B_t with P-dynamics given by

$$dB_t = rB_t\,dt,\ B_0 = 1$$

and a stock S_t with P-dynamics given by

$$dS_t = \alpha S_t\,dt + \sigma S_t\,d\overline{W}_t,\ S_0 = s > 0$$

where $r, \alpha \in \mathbb{R}$ and $\sigma > 0$ are constants and $\overline{W}_t$ is a P-Brownian motion. Let $T > 0$ be a given and fixed date.

Consider the derivative that at time T pays

$$X = \max\left\{\min\left\{S_T, K_2\right\}, K_1\right\},$$

where $0 < K_1 < K_2$ are constants.

a. Determine the arbitrage free price of derivative X at time $t < T$.

Consider a new derivative that at time T pays

$$Y = (S_T^2 - K^2)^+ - (K^2 - S_T^2)^+.$$

b. i. Determine the arbitrage free price of derivative Y at time $t < T$.
   ii. Find a hedging portfolio for derivative Y.

Let $h(t) = (h_0(t), h_1(t))$ be a portfolio where

$$h_0(t) = -e^{r(T-2t)+\sigma^2(T-t)}S^2(t)$$

is the number of units in the bank account at time t and

$$h_1(t) = 2e^{(r+\sigma^2)(T-t)}S(t)$$

is the number of shares in the stock at time t. Let $V^h(t)$ denote the associated value process.

c. Determine whether the portfolio h is self-financing or not.
d. Compute $V^h(T)$.

Solution (b).

(i): We start by seeing that the derivative pays out

$$Y = \begin{cases} S_T^2 - K^2 & \text{if } S_T^2 \geqslant K^2, \\ -(K^2 - S_T^2) & \text{if } S_T^2 < K^2. \end{cases}$$

hence the payout is $Y = S_T^2 - K^2 = \Phi(S_T)$ where $\Phi(s) = s^2 - K^2$. That is Y is in fact a simple claim. We have from the risk neutral valueation formula 7.11 that

$$\begin{aligned} \Pi_t[Y] &= e^{-r(T-t)}E^Q_{t,s}[S_T^2 - K^2] \\ &= e^{-r(T-t)}E^Q_{t,s}[S_T^2] - e^{-r(T-t)}K^2. \end{aligned}$$

Recall that under the martingale measure Q we have that S_t is a GBM hence

$$S_t = s \cdot \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\left(W_T^Q - W_t^Q\right)\right\}$$

then

$$S_T^2 = s^2 \cdot \exp\left\{2\left(r - \frac{1}{2}\sigma^2\right)(T-t) + 2\sigma\left(W_T^Q - W_t^Q\right)\right\}.$$

Inserting this into the risk neutral valuation formula we get

$$\begin{aligned} \Pi_t[Y] &= e^{-r(T-t)}E^Q_{t,s}[S_T^2] - e^{-r(T-t)}K^2 \\ &= e^{-r(T-t)}s^2e^{2\left(r-\frac{1}{2}\sigma^2\right)(T-t)}E^Q\left[\exp\left\{2\sigma\left(W_T^Q - W_t^Q\right)\right\}\right] - e^{-r(T-t)}K^2 \\ &= e^{-r(T-t)}s^2e^{2\left(r-\frac{1}{2}\sigma^2\right)(T-t)}e^{\frac{1}{2}4\sigma^2(T-t)} - e^{-r(T-t)}K^2 \\ &= e^{-r(T-t)}\left(s^2e^{(2r-\sigma^2)(T-t)+\frac{1}{2}4\sigma^2(T-t)} - K^2\right) \\ &= e^{-r(T-t)}\left(s^2e^{(2r+\sigma^2)(T-t)} - K^2\right). \end{aligned}$$

The arbitrage free price of the derivative is then given above. ▫

(ii): From theorem 8.5 we can determine a hedging portfolio with weightings

$$
\begin{aligned}
w_t^B &= \frac{\Pi_t - S_t \frac{\partial \Pi}{\partial s}}{\Pi_t} \\
&= 1 - \frac{S_t 2 S_t e^{-r(T-t)} e^{(2r+\sigma^2)(T-t)}}{e^{-r(T-t)}\left(S_t^2 e^{(2r+\sigma^2)(T-t)} - K^2\right)} \\
&= 1 - \frac{2S_t^2 e^{(2r+\sigma^2)(T-t)}}{S_t^2 e^{(2r+\sigma^2)(T-t)} - K^2} \\
&= 1 - \frac{2}{1 - K^2 S_t^{-2} e^{(2r+\sigma^2)(t-T)}} \\
w_t^S &= \frac{2}{1 - K^2 S_t^{-2} e^{(2r+\sigma^2)(t-T)}}.
\end{aligned}
$$

In absolute terms we will hold the portfolio

$$
\begin{aligned}
h_t^S &= 2S_t e^{-r(T-t)} e^{(2r+\sigma^2)(T-t)} \\
h_t^B &= \frac{e^{-r(T-t)}\left(s^2 e^{(2r+\sigma^2)(T-t)} - K^2\right) - S_t h_t^S}{B_t} \\
&= \frac{e^{-r(T-t)}\left(s^2 e^{(2r+\sigma^2)(T-t)} - K^2\right) - S_t h_t^S}{e^{rt}} \\
&= e^{-rT} s^2 e^{(2r+\sigma^2)(T-t)} - e^{-rT} K^2 - e^{-rt} S_t h_t^S.
\end{aligned}
$$

The portfolio above will hedge Y with probability one. □

Solution (c).

We assume no dividends and no consumption that is $c_t = 0$ and $dD_t^i = 0$ for $i = 0, 1$. Then the portfolio is self-financing if and only if the value process has dynamics.

$$h_0(t)\ dB_t + h_1(t)\ dS_t = 0$$

This is given in lemma 6.12.

THE BELOW IS IN WORKS AND NOT CORRECT!

Now we have that the value process is given by

$$V_t^h = h_0(t) B_t + h_1(t) S_t.$$

Using the representation $V_t^h = f(h_0(t), B_t) + f(h_1(t), S_t)$ given by $f(x, y) = xy$ we have

$$dV_t^h = df(h_0(t), B_t) + df(h_1(t), S_t).$$

Using Ito's formula on each term we have

$$
\begin{aligned}
df(h_0(t), B_t) &= B_t\ dh_0(t) + h_0(t)\ dB_t + (dB_t)(dh_0(t)), \\
df(h_1(t), S_t) &= S_t\ dh_1(t) + h_1(t)\ dS_t + (dS_t)(dh_1(t)),
\end{aligned}
$$

since of cause $f_{xx} = f_{yy} = 0$. We can the determine the dynamics of the portfolio by

$$
\begin{aligned}
dh_0(t) &= -(-2t-\sigma^2)S_t^2 e^{r(T-2t)+\sigma^2(T-t)}\ dt \\
&- 2S_t e^{r(T-2t)+\sigma^2(T-t)}\ dS_t \\
&- \frac{1}{2}2e^{r(T-2t)+\sigma^2(T-t)}\ (dS_t)^2 \\
&= (-2t-\sigma^2)h_0(t)\ dt + \frac{2}{S_t}h_0(t)\ (\mu S_t\ dt + \sigma S_t\ dW_t) + \frac{1}{S_t^2}h_0(t)\sigma^2 S_t^2\ dt \\
&= (\mu-1)2h_0(t)\ dt + 2\sigma h_0(t)\ dW_t
\end{aligned}
$$

and

$$
\begin{aligned}
dh_1(t) &= (-r-\sigma^2)2e^{(r+\sigma^2)(T-t)}S_t\ dt \\
&+ 2e^{(r+\sigma^2)(T-t)}\ dS_t + 0 \\
&= (-r-\sigma^2)h_1(t)\ dt + \frac{1}{S_t}h_1(t)(\mu S_t\ dt + \sigma S_t\ dW_t) \\
&= (-r-\sigma^2+\mu)h_1(t)\ dt + h_1(t)\sigma\ dW_t
\end{aligned}
$$

And so in total

$$
\begin{aligned}
dV_t^h(t) &= df(h_0(t), B_t) + df(h_1(t), S_t) \\
&= B_t\ dh_0(t) + h_0(t)\ dB_t + (dB_t)(dh_0(t)) \\
&+ S_t\ dh_1(t) + h_1(t)\ dS_t + (dS_t)(dh_1(t)) \\
&= B_t\ ((\mu-1)2h_0(t)\ dt + 2\sigma h_0(t)\ dW_t) + h_0(t)\ rB_t\ dt + 0 \\
&+ S_t\ ((-r-\sigma^2+\mu)h_1(t)\ dt + h_1(t)\sigma\ dW_t) + h_1(t)\ (\mu S_t\ dt + \sigma S_t\ dW_t) + \sigma^2 S_t h_1(t)\ dt \\
&= \left[B_t(\mu-1)2h_0(t) + h_0(t)rB_t + S_t(-r-\sigma^2+\mu)h_1(t) + h_1\mu S_t + \sigma^2 S_t h_1(t)\right]\ dt \\
&+ \left[B_t 2\sigma h_0(t) + S_t h_1\sigma + h_1\sigma S_t\right]\ dW_t \\
&= \left[(2\mu-2+r)B_t h_0(t) + (-r+2\mu)S_t h_1(t)\right]\ dt \\
&+ \left[B_t h_0(t) + h_1 S_t\right] 2\sigma\ dW_t \\
&= V_t^h 2\mu\ dt + V_t^h\ dW_t
\end{aligned}
$$

Solution (d).

We compute V_T^h easily by inserting h_0 and h_1 below

$$
\begin{aligned}
V_T^h &= B_T h_0(T) + S_T h_1(T) \\
&= B_T\left(-e^{r(T-2T)+\sigma^2(T-T)}S_T^2\right) + S_T\left(2e^{(r+\sigma^2)(T-T)}S_T\right) \\
&= -S_T^2 + 2S_T^2 = S_T^2.
\end{aligned}
$$

and so h hedge the payout $\Phi(S_T) = S_T^2$. □

Solution.

Exercise 7. *(Exam 2018/19, problem 1)*

Solution.

Exercise 8. *(Exam 2018/19, problem 2, question (a))*

Solution.

Extra-Exercise 1. *(Bjork 7.7)*

Solution.

2.5 Week 5

Exercise 1. *(Bjork 10.1)* Consider the standard Black-Scholes model. Fix the time of maturity T and consider the following T-claim X:

$$X = \begin{cases} K & \text{if } S_T \leqslant A, \\ K + A - S_T & \text{if } A < S_T < K + A, \\ 0 & \text{if } S_T > K + A. \end{cases}$$

This contract can be replicated using a portfoliom consisting solely of bonds, stock, and European call options, which is constant over time. Determine this portfolio as well as the arbitrage free price of the contract.

Solution.

We see that the put option with strike $K + A$ gives the payout

$$P_{K+A}(S_T) = \begin{cases} K + A - S_T & \text{if } S_T \leqslant A, \\ K + A - S_T & \text{if } A < S_T < K + A, \\ 0 & \text{if } S_T > K + A. \end{cases}$$

Hence this asset will *"almost"* gives the wanted payout except for the event $(S_T \leqslant A)$. If we find an asset giving the payout $A - S_T$ if and only if the event $(S_T \leqslant A)$ occurs, we may short this asset. It happens that the put with strike A has the wanted payout that is

$$P_A(S_T) = \begin{cases} A - S_T & \text{if } S_T \leqslant A, \\ 0 & \text{if } A < S_T < K + A, \\ 0 & \text{if } S_T > K + A. \end{cases}$$

Then making the portfolio of one long position in the put P_{K+A} and a short position in the put P_A will replicate X. We know from the put-call parity that this exact portfolio may be replicated by $K + A - A = K$ long zero-coupon bonds, long call with strike $K + A$ and short call with strike A. Notice no position is taking on the underlying stock since we both go long and short on a put. Let us check if this portfolio give the wantet payout:

$$V_T^h = \begin{cases} K + 0 - 0 = K & \text{if } S_T \leqslant A, \\ K + 0 - (S_T - A) = K + A - S_T & \text{if } A < S_T < K + A, \\ K + (S_T - K - A) - (S_T - A) = 0 & \text{if } S_T > K + A. \end{cases}$$

Then the portfolio give the desired payout. Given the price process for the call option and the zero-coupon bond we have that the value process is given by

$$\begin{aligned} V_t^h &= Ke^{-r(T-t)} + c(K + A; t, T) - c(A; t, T) \\ &= e^{-r(T-t)} \{K + (K + A)N(d_2(K + A; t, S_t) - AN(d_2(A; t, S_t))\} \\ &+ S_t \{N(d_1(K + A; t, S_t)) - N(d_1(A; t, S_t))\} \end{aligned}$$

with d_1 and d_2 as given in the Black-Scholes formula. The price of the portfolio is given by the above value process. □

Exercise 2. *(Bjork 10.2)* The setup is the same as the previous exercise. Here the contract is a so-called **straddle**, defined by

$$X = \begin{cases} K - S_T & \text{if } 0 < S_T \leqslant K, \\ S_T - K & \text{if } S_T > K. \end{cases}$$

Determine the constant replicating portfolio as well as the arbitrage free price of the contract.

Solution.

We search for portfolio paying the payout above. Recall that a call option with strike K has payout $S_T - K$ if $S_T \geqslant K$ and a put option with strike K has payout $K - S_T$ if $S_T \leqslant K$. Hence by longing one call with strike K and long one put with strike K will give the desired payout.

We know that we kan replicate this portfolio by buying K zero-coupon bonds, long two call options with strike K and shorting the underlying stock. Lets see what this yields

$$V_T^h = \begin{cases} K + 2 \cdot 0 - S_T = K - S_T & \text{if } 0 < S_T \leqslant K, \\ K + 2 \cdot (S_T - K) - S_T = S_T - K & \text{if } S_T > K. \end{cases}$$

As desired. The price is the then value process with

$$\Pi_t = V_t^h = Ke^{-r(T-t)} + 2c(t,T) - S_t.$$

Giving the desired result. □

Exercise 3. *(Bjork 10.3)* The setup is the same as the previous exercise. We will now study a so-called bull spread (see Fig. 10.7). With this contract we can, to a limited extent, take advantage of an increase in the market price while being protected from a decrease. The contract is defined by

$$X = \begin{cases} B & \text{if } S_T > B, \\ S_T & \text{if } A \leqslant S_T \leqslant B, \\ A & \text{if } S_T < A. \end{cases}$$

We have of course the relation $A < B$. Determine the constant replicating portfolio as well as the arbitrage free price of the contract.

Solution.

Again we search for a replicating portfolio. If we long one stock we get the payout S_T only and so we want to recieve an additional payout on the events S_T falls outside the interval $[A, B]$. We want to have an asset paying $-(S_T - A)$ on the event $S_T < A$ and an asset paying $-(S_T - B)$ on the event $S_T > B$. The two assets are: one put with strike A giving the payout $A - S_T$ and a short position call option with strke B. Hence we may replicate the put with A bonds, one call and a short position. In total we hold A bonds, one call option with strike A and a short on a call with strike B. We will then recieve the payout:

$$V_T^h = \begin{cases} A + S_T - A - (S_T - B) = B & \text{if } S_T > B, \\ A + S_T - A - 0 = S_T & \text{if } A \leqslant S_T \leqslant B, \\ A + 0 + 0 = A & \text{if } S_T < A. \end{cases}$$

As desired. The value process then give the portfolio price:

$$\Pi_t = V_t^h = Ae^{-r(T-t)} + c(A; t, T) - c(B; t, T)$$

as desired. □

Exercise 4. *(Exam 2017/18, problem 2, question (a)-(b))* Consider a standard Black-Scholes model, that is, a model consisting of a bank account B_t with P-dynamics given by

$$dB_t = rB_t\ dt,\ B_0 = 1$$

and a stock S_t with P-dynamics given by

$$dS_t = \alpha S_t\ dt + \sigma S_t\ d\overline{W}_t,\ S_0 = s > 0$$

where $r, \alpha \in \mathbb{R}$ and $\sigma > 0$ are constants and $\overline{W}_t$ is a P-Brownian motion. Let $T > 0$ be a given and fixed date.

Consider the derivative that at time T pays

$$X = \max\left\{\min\left\{S_T, K_2\right\}, K_1\right\},$$

where $0 < K_1 < K_2$ are constants.

a. Determine the arbitrage free price of derivative X at time $t < T$.

Consider a new derivative that at time T pays

$$Y = (S_T^2 - K^2)^+ - (K^2 - S_T^2)^+.$$

b. i. Determine the arbitrage free price of derivative Y at time $t < T$.
   ii. Find a hedging portfolio for derivative Y.

Let $h(t) = (h_0(t), h_1(t))$ be a portfolio where

$$h_0(t) = -e^{r(T-2t)+\sigma^2(T-t)}S^2(t)$$

is the number of units in the bank account at time t and

$$h_1(t) = 2e^{(r+\sigma^2)(T-t)}S(t)$$

is the number of shares in the stock at time t. Let $V^h(t)$ denote the associated value process.

c. Determine whether the portfolio h is self-financing or not.
d. Compute $V^h(T)$.

Solution (a).

We see that the derivative is the bull spread given by the payout function

$$X = \begin{cases} K_2 & \text{if } S_T > K_2, \\ S_T & \text{if } K_1 \leqslant S_T \leqslant K_2, \\ K_1 & \text{if } S_T < K_1. \end{cases}$$

We know from exercise 10.3 that this can be replicated by holding K_1 bonds, one call option with strike K_1 and a short on a call with strike K_2. The arbitrage free price of the derivative is then the value process of the mentioned portfolio i.e.

$$\Pi_t[X] = K_1 e^{-r(T-t)} + c(K_1; t, T) - c(K_2; t, T),$$

where c denotes the pricing function for a European call option (non-instructive parameters supressed). □

Solution (b).

(i): We start by seeing that the derivative pays out

$$Y = \begin{cases} S_T^2 - K^2 & \text{if } S_T^2 \geqslant K^2, \\ -(K^2 - S_T^2) & \text{if } S_T^2 < K^2. \end{cases}$$

hence the payout is $Y = S_T^2 - K^2 = \Phi(S_T)$ where $\Phi(s) = s^2 - K^2$. That is Y is in fact a simple claim. We have from the risk neutral valueation formula 7.11 that

$$
\begin{aligned}
\Pi_t[Y] &= e^{-r(T-t)} E^Q_{t,s}[S_T^2 - K^2] \\
&= e^{-r(T-t)} E^Q_{t,s}[S_T^2] - e^{-r(T-t)} K^2.
\end{aligned}
$$

Recall that under the martingale measure Q we have that S_t is a GBM hence

$$
S_t = s \cdot \exp\left\{ \left(r - \frac{1}{2}\sigma^2 \right)(T-t) + \sigma \left(W_T^Q - W_t^Q \right) \right\}
$$

then

$$
S_T^2 = s^2 \cdot \exp\left\{ 2\left(r - \frac{1}{2}\sigma^2 \right)(T-t) + 2\sigma \left(W_T^Q - W_t^Q \right) \right\}.
$$

Inserting this into the risk neutral valuation formula we get

$$
\begin{aligned}
\Pi_t[Y] &= e^{-r(T-t)} E^Q_{t,s}[S_T^2] - e^{-r(T-t)} K^2 \\
&= e^{-r(T-t)} s^2 e^{2\left(r - \frac{1}{2}\sigma^2\right)(T-t)} E^Q\left[\exp\left\{ 2\sigma \left(W_T^Q - W_t^Q \right) \right\} \right] - e^{-r(T-t)} K^2 \\
&= e^{-r(T-t)} s^2 e^{2\left(r - \frac{1}{2}\sigma^2\right)(T-t)} e^{\frac{1}{2}4\sigma^2(T-t)} - e^{-r(T-t)} K^2 \\
&= e^{-r(T-t)} \left(s^2 e^{(2r-\sigma^2)(T-t) + \frac{1}{2}4\sigma^2(T-t)} - K^2 \right) \\
&= e^{-r(T-t)} \left(s^2 e^{(2r+\sigma^2)(T-t)} - K^2 \right).
\end{aligned}
$$

The arbitrage free price of the derivative is then given above. ▫

(ii): From theorem 8.5 we can determine a hedging portfolio with weightings

$$
\begin{aligned}
w_t^B &= \frac{\Pi_t - S_t \frac{\partial \Pi}{\partial s}}{\Pi_t} \\
&= 1 - \frac{S_t 2 S_t e^{-r(T-t)} e^{(2r+\sigma^2)(T-t)}}{e^{-r(T-t)} \left(S_t^2 e^{(2r+\sigma^2)(T-t)} - K^2 \right)} \\
&= 1 - \frac{2 S_t^2 e^{(2r+\sigma^2)(T-t)}}{S_t^2 e^{(2r+\sigma^2)(T-t)} - K^2} \\
&= 1 - \frac{2}{1 - K^2 S_t^{-2} e^{(2r+\sigma^2)(t-T)}} \\
w_t^S &= \frac{2}{1 - K^2 S_t^{-2} e^{(2r+\sigma^2)(t-T)}}.
\end{aligned}
$$

In absolute terms we will hold the portfolio

$$
\begin{aligned}
h_t^S &= 2 S_t e^{-r(T-t)} e^{(2r+\sigma^2)(T-t)} \\
h_t^B &= \frac{e^{-r(T-t)} \left(s^2 e^{(2r+\sigma^2)(T-t)} - K^2 \right) - S_t h_t^S}{B_t} \\
&= \frac{e^{-r(T-t)} \left(s^2 e^{(2r+\sigma^2)(T-t)} - K^2 \right) - S_t h_t^S}{e^{rt}} \\
&= e^{-rT} s^2 e^{(2r+\sigma^2)(T-t)} - e^{-rT} K^2 - e^{-rt} S_t h_t^S.
\end{aligned}
$$

The portfolio above will hedge Y with probability one. ▫

Exercise 5. *(Exam 2019/20, problem 2)*

Solution.

Exercise 6. *(Exam 2020/21, problem 2, question (a)-(b))*

Solution.

Exercise 7. *(Exam 2020/21, problem 3, question (b))*

Solution.

Extra-Exercise 1. *(Bjork 10.4)*

Solution.

2.6 Week 6

Exercise 1. *(Exam 2017/18, problem 1, question (c))* Let W_t denote a Brownian motion and let

$$\mathcal{F}_t = \mathcal{F}_t^W = \sigma(\{W_s \mid 0 \leqslant s \leqslant t\}).$$

Let $T > 0$ be a given and fixed time.

Let $f(t)$ be a bounded deterministic continuous function. Define the two processes

$$\begin{cases} X_t = \int_0^t f(u)\ dW_u, \\ M_t^{(\lambda)} = \exp\left\{\lambda X_t - \frac{\lambda^2}{2}\int_0^t f^2(u)\ du\right\}, \end{cases}$$

where $\lambda \in \mathbb{R}$ is a constant.

c. Compute the mean value of $M_T^{(\lambda)} \log(M_T^{(\lambda)})$.

Solution (c).

We recall the definition of $M_t^{(\lambda)}$ and observe that

$$\log M_t^{(\lambda)} = \lambda X_t - \frac{1}{2}\lambda^2 \int_0^t f^2(u)\ du.$$

Furthermore we have the dynamics of $M^{(\lambda)}$ given by the differential form

$$dM_t^{(\lambda)} = \lambda f(t) M_t^{(\lambda)}\ dW_t.$$

with $M_0^{(\lambda)} = 1$. Since we know that $M_t^{(\lambda)}$ is a martingale we have

$$E^P[M_T^{(\lambda)}] = E^P[M_0^{(\lambda)}] = 1,$$

and so we may define a new probability measure as

$$d\tilde{P} = M_T^{(\lambda)}\ dP$$

on $\mathcal{F}_T$. We then have a new Brownian motion $\tilde{W}$ such that

$$dW_t = \lambda f(t)\ dt + d\tilde{W}_t.$$

We can then see

$$
\begin{aligned}
E^P[M_T^{(\lambda)} \log M_T^{(\lambda)}] &= \int M_T^{(\lambda)} \log M_T^{(\lambda)}\ dP = \int M_T^{(\lambda)} \log M_T^{(\lambda)} \frac{1}{M_T^{(\lambda)}}\ d\tilde{P} \\
&= \int \log M_T^{(\lambda)}\ d\tilde{P} = E^{\tilde{P}}[\log M_T^{(\lambda)}].
\end{aligned}
$$

Then we can evaluate the mean value by seeing the X has representation wrt. $\tilde{P}$ by

$$
X_t = \int_0^t f(u)\ (\lambda f(u)\ du + d\tilde{W}_u) = \lambda \int_0^t f^2(u)\ du + \int_0^t f(u)\ d\tilde{W}_u.
$$

Giving that

$$
\begin{aligned}
E^P[M_T^{(\lambda)} \log M_T^{(\lambda)}] &= E^{\tilde{P}}[\log M_T^{(\lambda)}] \\
&= E^{\tilde{P}}\left[\lambda X_T - \frac{1}{2}\lambda^2 \int_0^T f^2(u)\ du\right] \\
&= E^{\tilde{P}}\left[\lambda^2 \int_0^T f^2(u)\ du + \lambda \int_0^T f(u)\ d\tilde{W}_u - \frac{1}{2}\lambda^2 \int_0^T f^2(u)\ du\right] \\
&= \lambda E^{\tilde{P}}\left[\frac{1}{2}\lambda \int_0^T f^2(u)\ du + \int_0^T f(u)\ d\tilde{W}_u\right] \\
&= \frac{1}{2}\lambda^2 \int_0^T f^2(u)\ du + \lambda E^{\tilde{P}}\left[\int_0^T f(u)\ d\tilde{W}_u\right] \\
&= \frac{1}{2}\lambda^2 \int_0^T f^2(u)\ du
\end{aligned}
$$

Since

$$
\tilde{X}_T = \int_0^T f(u)\ d\tilde{W}_u,
$$

is a $\tilde{P}$-martingale. □

Exercise 2. *(Exam 2018/19, problem 2, question (b)i and (c)-(d))*

Solution.

Exercise 3. *(Exam 2019/20, problem 1, question (b)-(c))*

Solution.

Exercise 4. *(Exam 2019/20, problem 3, question (b))*

Solution.

Exercise 5. *(Exam 2020/21, problem 1, question (c))*

Solution.

Exercise 6. *(Exam 2020/21, problem 2, question (c)-(d))*

Solution.

2.7 Week 7

Exercise 1. *(Exam 2018/19, problem 2, question (b).ii)*

Solution.

Exercise 2. *(Exam 2017/18, problem 3)*

Solution.

Exercise 3. *(Exam 2018/19, problem 3)*

Solution.

Exercise 4. *(Exam 2019/20, problem 3, question (a) and (c)-(e))*

Solution.

Exercise 5. *(Exam 2020/21, problem 3, question (a) and (c))*

Solution.

Chapter 3

Probabilistic Machine Learning

3.1 Week 1

Exercise 1. Let $\tilde{m} = \arg\min_{m \in \mathcal{G}} r(m)$. Proof that

$$r(\hat{m}_n) - r(\tilde{m}) \leqslant 2 \sup_{m \in \mathcal{G}} \left| \hat{R}_n(m) - r(m) \right|.$$

Solution.

We recall that per definition

$$R(\hat{m}_n) = \mathbb{E}[L(Y, \hat{m}_n(X)) \mid \mathcal{D}_n] \quad \text{and} \quad r(\hat{m}_n) = \mathbb{E}[R(\hat{m}_n)]$$

for some estimator $\hat{m}_n$. Assume now that $m(\mathcal{D}_n) = \hat{m}_n \in \mathcal{G}$ is the estimator given the data $\mathcal{D}_n$ from the precedure m. We led $\tilde{m}$ be the Bayes estimator on $\mathcal{G}$. We therefore have

$$\begin{aligned}
r(\hat{m}_n) - r(\tilde{m}) &= r(\hat{m}_n) - r(\tilde{m}) + \hat{R}_n(\hat{m}_n) - \hat{R}_n(\hat{m}_n) + \hat{R}_n(\tilde{m}) - \hat{R}_n(\tilde{m}) \\
&= r(\hat{m}_n) - \hat{R}_n(\hat{m}_n) + \hat{R}_n(\tilde{m}) - r(\tilde{m}) + \underbrace{\hat{R}_n(\hat{m}_n) - \hat{R}_n(\tilde{m})}_{\leqslant 0} \\
&\leqslant \left(r(\hat{m}_n) - \hat{R}_n(\hat{m}_n) \right) + \left(\hat{R}_n(\tilde{m}) - r(\tilde{m}) \right) \\
&\leqslant \left| \hat{R}_n(\hat{m}_n) - r(\hat{m}_n) \right| + \left| \hat{R}_n(\tilde{m}) - r(\tilde{m}) \right| \\
&\leqslant 2 \sup_{m \in \mathcal{G}} \left| \hat{R}_n(m) - r(m) \right|
\end{aligned}$$

as desired. ▫

We consider the Poisson deviance loss

$$L_{\text{pois:dev}}(y_1, y_2) = 2 \left(y_1 \log \frac{y_1}{y_2} - y_1 + y_2 \right).$$

1. Show that $\arg\min r(m) = m^*(x) = E[Y \mid X = x]$.
2. In the squared loss case $L_2(y_1, y_2) = (y_1 - y_2)^2$, we used its Hilbert space property to derive that for any estimator $\hat{m}_n(x)$ $\quad r_2(\hat{m}_n(x)) - r_2(m^*(x)) = \mathbb{E}[L_2(\hat{m}_n(x), m^*(x))]$
   Show that an anaolgue result for the he Poisson deviance loss is not true, i.e.,

$$\begin{aligned}
r_{\text{pois:dev}}(\hat{m}_n(x)) - r_{\text{pois:dev}}(m^*(x)) &= \mathbb{E}\left[2 \left(Y \log \frac{m^*(x)}{\hat{m}_n(x)} - m^*(x) + \hat{m}_n(x) \right) \right] \\
&\neq \mathbb{E}[L_{\text{pois:dev}}(\hat{m}_n(x), m^*(x))]
\end{aligned}$$

Solution (1).

We have that

$$\begin{aligned} r(m) &= \mathbb{E}\left[2\left(Y\log\frac{Y}{m(X)} - Y + m(X)\right)\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[2\left(Y\log\frac{Y}{m(x)} - Y + m(x)\right) \;\middle|\; X = x\right]\right] \end{aligned}$$

We see that the integrand is differentiable in $m = m(x)$ and so we have

$$\begin{aligned} \frac{\partial}{\partial m} r(m) &= \frac{\partial}{\partial m}\mathbb{E}\left[\mathbb{E}\left[2\left(Y\log\frac{Y}{m} - Y + m\right) \;\middle|\; X = x\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{\partial}{\partial m}2\left(Y\log\frac{Y}{m} - Y + m\right) \;\middle|\; X = x\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[2\left(-Y\frac{m}{Y}\frac{Y}{m^2} + 1\right) \;\middle|\; X = x\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[2\left(-\frac{Y}{m} + 1\right) \;\middle|\; X = x\right]\right] \\ &= -2\mathbb{E}\left[\mathbb{E}\left[\frac{Y}{m} \;\middle|\; X = x\right]\right] + 2 \end{aligned}$$

Setting this equal to zero gives

$$1 = \mathbb{E}\left[\mathbb{E}\left[\frac{Y}{m} \;\middle|\; X = x\right]\right] = \frac{1}{m}\mathbb{E}\left[\mathbb{E}\left[Y \mid X = x\right]\right]$$

Giving that the expectation is minimized for

$$m = E\left[Y \mid X = x\right]$$

as desired. □

Solution (2).

Take any estimator $\hat{m}_n$ and consider the risk associated with the estimator:

$$\begin{aligned} r_{\text{pois:dev}}(\hat{m}_n(X)) - r_{\text{pois:dev}}(m^*(x)) &= \mathbb{E}\left[2\left(Y\log\frac{Y}{\hat{m}_n(X)} - Y + \hat{m}_n(X)\right)\right] - \mathbb{E}\left[2\left(Y\log\frac{Y}{m^*(X)} - Y + m^*(X)\right)\right] \\ &= \mathbb{E}\left[2\left(Y\log\frac{Y}{\hat{m}_n(X)} - Y + \hat{m}_n(X) - Y\log\frac{Y}{m^*(X)} + Y - m^*(X)\right)\right] \\ &= \mathbb{E}\left[2\left(Y\log\frac{Ym^*(X)}{\hat{m}_n(X)Y} + \hat{m}_n(X) - m^*(X)\right)\right] \\ &= \mathbb{E}\left[2\left(Y\log\frac{m^*(X)}{\hat{m}_n(X)} + \hat{m}_n(X) - m^*(X)\right)\right] \end{aligned}$$

Yielding the desired result. □

Exercise 3. We consider a simple regression case with no explanatory variables. We denote by $\hat{m}_{n,1}$ the sample mean and by $\hat{m}_{n,2}$ the sample mean. Furthermore, $\hat{R}_{n,1}$ and $\hat{R}_{n,2}$ denote the empirical risk with respect to the L_1 loss and the squared loss respectively.

1. Generate 10,000 iid observations $y_1, ..., y_{10000}$ from a standard normal distribution. Compare

$$\hat{R}_{n,1}(\hat{m}_{n,1}), \hat{R}_{n,1}(\hat{m}_{n,2}), \hat{R}_{n,2}(\hat{m}_{n,1}), \hat{R}_{n,2}(\hat{m}_{n,2}).$$

2. Generate 10,000 iid observations $y_1, ..., y_{10000}$ from a t-distribution with one degree of freedom. Compare
$$\hat{R}_{n,1}(\hat{m}_{n,1}), \hat{R}_{n,1}(\hat{m}_{n,2}), \hat{R}_{n,2}(\hat{m}_{n,1}), \hat{R}_{n,2}(\hat{m}_{n,2}).$$
3. What conclusion can you draw from the two exercises?

Solution (1).

We set the seed to 1 `set.seed(1)` and generate the $n = 10000$ samples.

We may now compute the sample mean and median.

One may recall that median$(X) = \mathbb{E}[X] = 0$ for $X \sim \mathcal{N}(0, 1)$ and so we would expect $\hat{m}_{n,1} \approx \hat{m}_{n,2}$. We can compute the empirical risk wrt. the L_1 and L_2 loss.

Although these are empirical risk we have not computed an estimate of the risk as we should generate more samples of the risk. As such we draw using the above method $S = 1000$ samples of the risks and compute the means.

We may compare the results in the table below.

Measure	Value
$\hat{R}_{n,1}(\hat{m}_{n,1})$	0.7979859
$\hat{R}_{n,1}(\hat{m}_{n,2})$	0.7980077
$\hat{R}_{n,2}(\hat{m}_{n,1})$	1.0004189
$\hat{R}_{n,2}(\hat{m}_{n,2})$	1.0003643

We can see that the following holds:

1. For the L_1 loss the empirical median does better than the empirical mean,
2. For the squared loss the empirical median does worse than the empirical mean.

Solution (2).

We set the seed to 1 `set.seed(1)` and generate the $n = 10000$ samples.

We may now compute the sample mean and median.

One may recall that median$(X) = \mathbb{E}[X] = 0$ for $X \sim \sqcup(1)$ and so we would expect $\hat{m}_{n,1} \approx \hat{m}_{n,2}$. We can compute the empiracal risk wrt. the L_1 and L_2 loss.

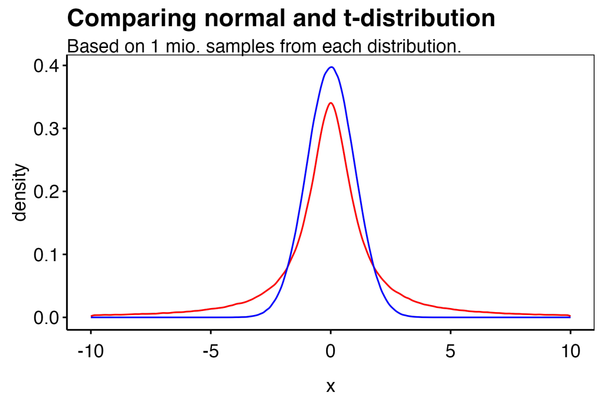
We may compare the results in the table below.

Measure	Value
$\hat{R}_{n,1}(\hat{m}_{n,1})$	11.7586085
$\hat{R}_{n,1}(\hat{m}_{n,2})$	16.7838177
$\hat{R}_{n,2}(\hat{m}_{n,1})$	1.8456626×10^7
$\hat{R}_{n,2}(\hat{m}_{n,2})$	1.84548×10^7

We see that for the t-distribution with one degree of freedom we see the same relation as in the normal case. The median does better in the L_1 loss but worse under the squared loss.

Solution (3).

We can see that the risk is far greater than the normal case. Comparing the density functions of a standard normal distribution and the t-distribution with one degree of freedom we see that the mean is the same but the variance is far greater than the normal case. In fact, with degrees of freedom below 2 we have that the variance is infinite.



This gives that for a given sample size the risk will tend to infinity as n grows to infinity. This in turn explains the large values we see in the risk. We see that for the t-distribution with one degree of freedom we see the same relation as in the normal case. The median does better in the L_1 loss but worse under the squared loss.

Exercise 4. We want to practise model tuning with the mlr3 package. Go through the following steps:

1. Install and load the relevant ml3 packages: mlr3, mlr3learners, mlr3tuning, mlr3mbo.
2. Create a task

- Load the `mtcars` data (write: *data(mtcars)*)
- Use the `as_task_regr` to create a task with `mpg` as target

3. Set `regr.xgboost` as learner with corresponding search space; e.g.,

$$\begin{aligned} eta &= to\_tune(0,1) \\ nrounds &= to\_tune(10,5000) \\ max\_depth &= to\_tune(1,20) \\ colsample\_bytree &= to\_tune(0.1,1) \\ subsample &= to\_tune(0.1,1) \end{aligned}$$

4. Tune your learner on you task using the `tune` function with

- Resampling method: 5-fold cross validation
- Measure: squared loss
- Method: `mbo` or `random search`
- Terminator: 10 evaluations

5. Fit your learner on the task using the optimal hyper parameters calculated

Solution (1).

We install the required packages.

Solution (2).

We start by loading the data.

We now transform the data into a task.

We may now split out data set into random partions of a training dataset and a testing dataset.

We see that the testing dataset has $N_1 = 11$ and the training dataset has $N_2 = 21$. We will be training the model on the training dataset. Therefore we start a task on the subset.

Solution (3).

We now initiate a learner.

We can now look at the current configuration of the leaner by looking at the parameter space of the hyper-parameter.

As the text says we set som of the parameters to as specific section of the parameter space.

Solution (4).

We tune a regression model using `mbo` search.

We can consider the the estimates from each subset and the fitted parameters.

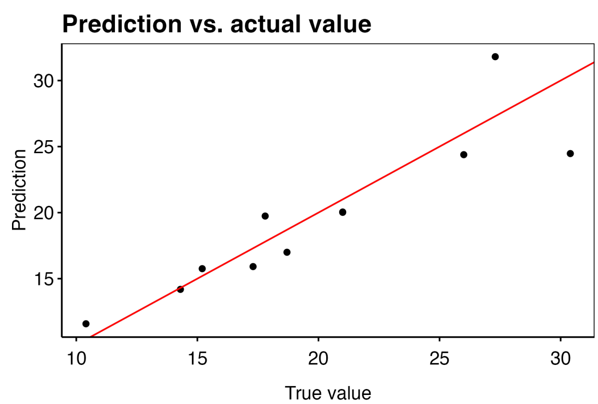
Solution (5).

We can now use optimal paramteres in `instance$result_learner_param_vals` to create a learner that we may fit to the training data.

We now fit the data.

We can now predict onto the testing data.

Using a diagram we can see. That the model does pretty well. In fact the empirical risk is 6.38.



3.2 Week 2

From now on, given sample size n, a dimension p, and a correlation parameter $0 \leqslant$ `rho` < 1 we will generate data X via the following code.

The resulting features will have approximately support on $[1, 1]$ and will have approximately pairwise correlation of size `rho`. Given the distribution of ε and a regression function m, we then generate data $(X_i, Y_i)_{i=1,\ldots,n}$ via

$$Y_i = m(X_i) + \varepsilon_i,$$

where ε_i are iid copies of ε and X_i is the ith row of X.

Exercise 1. We want to try out ordinary least squares regression, lasso, ridge regression and elastic net on some different data generating settings (Model 1–4).

	n	p	s	rho	m	ε
Model 1	1000	100	5	0.3	$\sum_{j=1}^{s} x_j$	$\mathcal{N}(0,1)$
Model 2	1000	100	100	0.3	$\sum_{j=1}^{s} x_j$	$\mathcal{N}(0,1)$
Model 3	1000	100	5	0.3	$\sum_{j=1}^{s} 0.1x_j$	$\mathcal{N}(0,1)$
Model 4	1000	100	100	0.3	$\sum_{j=1}^{s} 0.1x_j$	$\mathcal{N}(0,1)$

a. Given one set of training data for each model tune (i.e. estimate optimal hyperprameter) via 5-fold cross-validation using a search method and number of evals of your choice. You should now have an estimated optimal hyperparameter for every combination of model and method.
b. Generate 100 test sets each of size n for every model and calculate empirical mean and standard deviation of the test error for each method and model using the hyperparameters calculated in (a). Create a table of your results (each missing entry should show hyperparameter: mean(sd)):

	Least Squares	Ridge	Lasso	Elastic Net
Model 1	.	.	.	.
Model 2	.	.	.	.
Model 3	.	.	.	.
Model 4	.	.	.	.

c. Discuss why test error alone might often not be the (only) quantity of interest.

Solution (a).

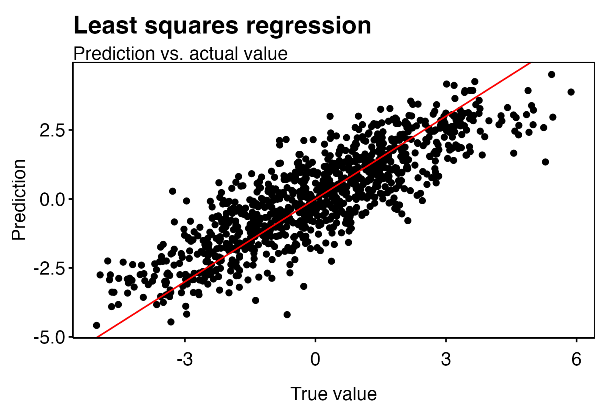
We start by simulating the data with the seed 1 for the training data and 2 for testing data.

Model 1

We may now estimate under model 1 by setting the parameters and generating the data.

We can estimate the optimal parameters under least squares estimation with the `lm` function.

We can use `predict` to see the model prediction vs. the true values.



We can also use `mlr3` to tune optimal hyperparameters for a lasso, ridge and elastic net estimator. We start as usual by slitting the data into training and testing datasets and starting a task.

We can now initiate a leaner. We start by fitting the hyperparameter for the lasso learner.

We use the `glmnet` as it is an algorithm that solves the minimizing problem

$$\hat{\beta}_{\lambda}^{\text{glmnet}} = \arg\min_{\beta \in \mathbb{R}^p} \left\{ \hat{R}_n(\beta) + J_{\lambda}(\beta) \right\}, \qquad J_{\lambda}(\beta) = \lambda \left[\alpha \sum_{j=1}^{p} |\beta_j| + \frac{1-\alpha}{2} \sum_{j=1}^{p} \beta_j^2 \right],$$

One may recall that the lasso estimator is given by

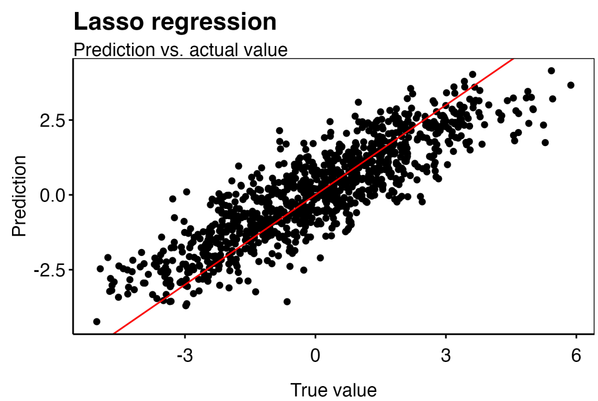
$$\hat{\beta}_{\lambda}^{\text{lasso}} = \arg\min_{\beta \in \mathbb{R}^p} \left\{ \hat{R}_n(\beta) + J_{\lambda}(\beta) \right\}, \qquad J_{\lambda}(\beta) = \lambda \sum_{j=1}^{p} |\beta_j|,$$

and so the lasso estimator is a specialcase of the `glmnet` with $\alpha = 1$. From the above we see that the minimizing algorithm punishes large beta's. The hyperparameter to tune is then $\lambda \geqslant 0$ (`s` is the syntax for λ). Looking at the parameters in the `learner_lasso` one can see the entire hyperparameter space.

Let us tune a lasso learner by setting `alpha = 1` and tuning `s = totune(0,1)` i.e. searching for $\lambda \in [0, 1]$.

Now that we have tuned the algorithm we can fit the model with the above s (0.0618).

Lets quickly look at the predictions.



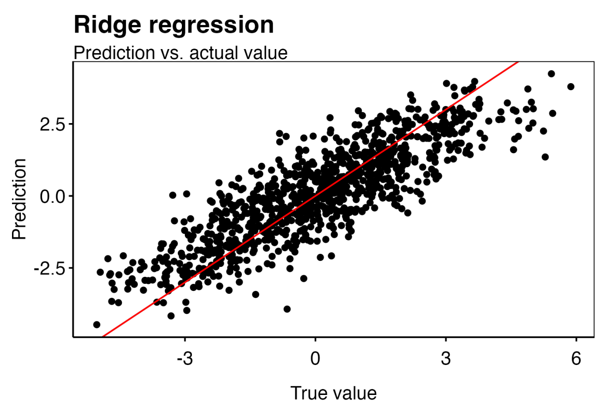
One can get the coefficients of β by predicting the data frame `I = diag(1,nrow = 100)` being the identity.

One can see that the values is close to the true value $\beta = (1, 1, 1, 1, 1, 0, ..., 0)$.

We can now estimate the parameters in ridge regression by setting $\alpha = 0$ and then getting a scaled version of λ.

We fit the model with the above s (0.22).

Lets quickly look at the predictions.

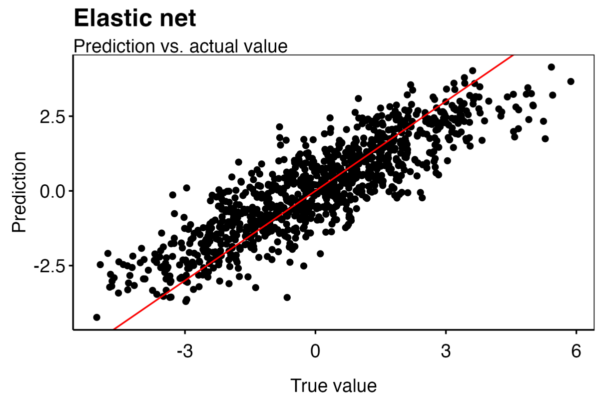


For good measures we calculate the beta's as in lasso.

Finally, let us tune an elastic net.

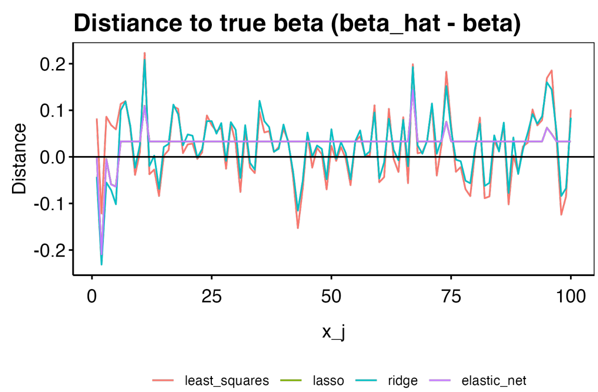
We fit the model with the above s (0.0623).

Lets quickly look at the predictions.



For good measures we calculate the beta's as in lasso.

Let's lastly look at how the different estimators do with respect to the true beta.



We may now estimate the remaining three models.

Model 2-4. We generate the data and estimate under the model. We define a function that compute the estimators.

Now we simply gather the results.

Solution (b).

We simply generate $N = 100$ test sets and compute the mean squared error for each run, then take the empirical mean and standard deviation of the 100 samples of the empirical mean squared error.

Solution (c).

In general, it is better to have as small of a test error as possible. It it however also important not to overcomplicate the model if not all variables is relevant. Ridge regression does well in the models with fewer explanatory variable s as it punishes under L^2 distance rather than L^1. This means that ridge does a better job in setting $\beta_j \approx 0$ for $j > s$.

3.3 Week 3

Exercise 1. Random forests are known for their computational speed and can be used in settings with thousands of features. Here we discuss how it is usually decided where to split a node. Assume we have n observations at a node, and want to find the optimal split. The `mtry` provides us with a list of possible features to split. Fix one feature and assume that observations in that node are sorted according to the value of that feature. In the current node the squared loss is

$$Q = \sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n} x_i^2 - n\overline{x}^2.$$

Define

$$Q_L(k) = \sum_{i=1}^{k}(x_i - \overline{x}_k^+)^2, \qquad Q_R(k) = \sum_{i=k+1}^{n}(x_i - \overline{x}_k^-)^2$$

where

$$\overline{x}_k^+ = \frac{1}{k}\sum_{i=1}^{k} x_i \quad \text{and} \quad \overline{x}_k^- = \frac{1}{n-k}\sum_{i=k+1}^{n} x_i.$$

The node is split at that k that minimizes $Q_L(k) + Q_R(k)$. Discuss how the optimal k can be found more efficiently than calculating for every k $Q_L(k) + Q_R(k)$ from scratch.

Solution.

Exercise 2. We extend the framework from previous week and now also consider classification tasks. When considering a classification task, the only change in the data generation compared to the regression setting is that we generate the response Y via

$$Y = \frac{\operatorname{sign}(m(X_i) + \varepsilon_i) + 1}{2}.$$

Try out least squares regression and classification with least squares and binary loss. As learners, we will work with linear regression/logistic regression (with e.g. elastic net penalization), generalized additive models from the `mgcv` package and random forest via the `ranger` package. Part of the task is to figure out how to optimize the (penalty) parameters to be successful in the sparse settings as given below. We will consider the following models.

	n	p	s	rho	m	ε
Model 1	1000	100	5	0.3	$\sum_{j=1}^{s} x_j$	$\mathcal{N}(0,1)$
Model 3	1000	100	5	0.3	$\sum_{j=1}^{s} 0.1x_j$	$\mathcal{N}(0,1)$
Model 5	1000	100	5	0.3	$\sum_{j=1}^{s} m_j(x_j)$	$\mathcal{N}(0,1)$
Model 6	1000	100	5	0.3	$\sum_{j=1}^{s} m_j(x_j)+$ $\sum_{j=1}^{s-1} m_j(x_j x_{j+1})$	$\mathcal{N}(0,1)$

Here $m_j(x_j) = (-1)^j 2\sin(\pi x_j)$

a. For each of the three settings (regression: squared loss, classification: squared loss, classification: binary loss), given one set of training data for each model tune (i.e. estimate optimal hyperprameter) via 5-fold cross-validation using a search method and number of evals of your choice. You should now have an estimated optimal hyperparameter for every combination of setting, model and learner.
b. Generate 100 test sets each of size n for every setting, model and learner and calculate empirical mean and standard deviation of the test error for each setting, model and learner using the hyperparameters calculated in (a). Create three tables ofor your results (each missing entry should show hyperparameter: mean(sd)):

	Linear model	glm	Random forest
Model 1	·	·	·
Model 3	·	·	·

	Linear model	glm	Random forest
Model 5	·	·	·
Model 6	·	·	·

Solution (a).

Solution (b).

Chapter 4

Topics in Life Insurance

4.1 Computing moments of reserve in timehomogeneous case

In this exercise we consider a time-homogeneous Markov jump process X on the state-space $E = \{1, ..., 5\}$ defined as such

1. Active,
2. Inactive,
3. Unemployed,
4. Active after unemployment,
5. Dead.

The associated intensity matrix $\mathbf{\Lambda}$ is

$$\mathbf{\Lambda}(x) = \mathbf{\Lambda} = \begin{bmatrix} -0.7 & 0.1 & 0.1 & & 0.5 \\ & -0.5 & & & 0.5 \\ & 0.1 & -0.7 & 0.1 & 0.5 \\ & 0.1 & & -0.6 & 0.5 \\ & & & & \end{bmatrix}.$$

Where all zero entries are left blank. We consider the insurance contract that,

- Insured pays ρ as long as he is alive,
- Insured recieve 1 while unemployed or inactive and
- Insured receive 2 if he becomes inactive. If he is active just before the jump this only happends with probability 1/2.

The contract expires at time $T = 1$ and we assume $X(0) = 1$. We want to compute the equivalence premium for a constant interest rate $r = 0$.

a. Determine $\mathbf{B}(t)$ and b.
b. Decompose $\mathbf{\Lambda}$ into $\mathbf{\Lambda} = \mathbf{\Lambda}^0 + \mathbf{\Lambda}^1$ and compute the rewards $\mathbf{R}$.
c. Compute the reserve at time $t = 0$ for $\rho = -1$.
d. Compute the equivalence premium $\hat{\rho}$ and check that $\mathbf{V}(0; X(0) = 0) = 0$.
e. Compute the contributions $\mathbf{C}_r^{(2)}$ and use this to compute the second moment $\mathbf{V}^{(2)}(0; X(0) = 0)$.

Solution (a).

The matrix $\mathbf{B} = \{b_{ij}\}_{i,j \in E}$ is given by $b_{ij} = 2$ if $j = 2$ and $i \in \{1, 3, 4\}$ i.e.

$$\mathbf{B} = \begin{bmatrix} 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The continuous payment rate is

$$b = \begin{pmatrix} \rho \\ 1 \\ 1 \\ \rho \\ 0 \end{pmatrix}.$$

We implement this i R using.

Solution (b).

The $\boldsymbol{\Lambda}$ is decomposed by setting $(\lambda^1_{12}, \lambda^1_{32}, \lambda^1_{42}) = (\lambda_{12}/2, \lambda_{32}, \lambda_{42}/2)$ and the rest to 0 since the payment when transitioning from 1 or 4 only happens with probabiltiy 1/2. The reward matrix is simply given by $\mathbf{R} = \boldsymbol{\Lambda}^1 \bullet \mathbf{B} + \Delta b$ i.e.

$$\mathbf{R} = \begin{bmatrix} \rho & 2\lambda^1_{12} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 2\lambda^1_{32} & 1 & 0 & 0 \\ 0 & 2\lambda^1_{42} & 0 & \rho & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We implement in R.

Solution (c).

We compute the reserve by using the following

$$\prod_t^T \left(\mathbf{I} + \begin{pmatrix} \boldsymbol{\Lambda} - \Delta(r) & \mathbf{R} \\ \mathbf{0} & \boldsymbol{\Lambda} \end{pmatrix} \, dx \right) = \begin{pmatrix} e^{-\int_t^T r \, du} \mathbf{P}(t,T) & \mathbf{V}(t,T) \\ \mathbf{0} & \mathbf{P}(t,T) \end{pmatrix}.$$

We then compute

$$V(t,T) = e_1^\top \mathbf{V}(t,T) e.$$

Solution (d).

The equivalence premium is computed by

$$\hat{\rho} = - \frac{e_1^\top \mathbf{V}(t,T) e}{e_1^\top \mathbf{V}'(t,T) e} \bigg|_{\rho=0},$$

with $\mathbf{V}'(t,T)$ defined by

$$\prod_t^T \left(\mathbf{I} + \begin{pmatrix} \boldsymbol{\Lambda} - \Delta(r) & \frac{\partial}{\partial \rho} \mathbf{R}(\rho) \\ \mathbf{0} & \boldsymbol{\Lambda} \end{pmatrix} \, dx \right) = \begin{pmatrix} e^{-\int_t^T r \, du} \mathbf{P}(t,T) & \mathbf{V}'(t,T) \\ \mathbf{0} & \mathbf{P}(t,T) \end{pmatrix}.$$

We compute

Let us see if the reserve is zero as expected.

Solution (e).

We have that

$$\mathbf{C}_r^{(k)} = \frac{1}{k!} \boldsymbol{\Lambda}^1 \bullet \mathbf{B}^{\bullet k}$$

and

$$\prod_t^T \left(\mathbf{I} + \begin{pmatrix} \mathbf{\Lambda} - k\Delta(r) & \mathbf{R} & \mathbf{C}_r^{(2)} & \cdots & \mathbf{C}_r^{(k)} \\ \mathbf{0} & \mathbf{\Lambda} - (k-1)\Delta(r) & \mathbf{R} & \cdots & \mathbf{C}_r^{(k-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Lambda} \end{pmatrix} dx \right)$$

$$= \begin{pmatrix} * & * & * & \cdots & \mathbf{V}_r^{(k)}(t,T) \\ * & * & * & \cdots & \mathbf{V}_r^{(k-1)}(t,T) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & \mathbf{P}(t,T) \end{pmatrix}.$$

Where

$$\mathbf{V}_r^{(k)}(t,T) = \frac{1}{k!}\mathbf{V}^{(k)}(t,T).$$

We therefore implement a Toeplizs block matrix function that for matrices $\mathbf{M}_1, ..., \mathbf{M}_k$ generates the matrix

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_2 & \mathbf{M}_3 & \cdots & \mathbf{M}_k \\ \mathbf{0} & \mathbf{M}_1 & \mathbf{M}_2 & \cdots & \mathbf{M}_{k-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_1 & \cdots & \mathbf{M}_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{M}_1 \end{pmatrix}$$

This is done in the `bToeplitz`

We can then compute the second moment.

We can then see that the variance is 0.27.