Laboration 3: Probability theory and simulation

The goal of this laboration is for you to become more familiar with some important areas within the course:

- Conditional distributions
- Dependence
- Functions of stochastic variables

Stochastic simulation is a very important tool in technical and economical modeling and as an illustration of random variation we will look at the price development where the change in price is modeled both as a log-normal distribution and as a Geometrical Brownian Motion.

1 Preparatory exercises

As a preparation for the laboration you should read Chapter 4.5–4.6, 5.2, 6 and 7:1–2 (Blom: Bok A), (Independent r.v., functions of a random variable, conditional distribution, expectation and variance of a r.v. and of a linear function of a r.v.) and the entire laboration assignment, in particular the appendix about Wiener processes and Geometrical Brownian Motion.

At the start of the laboration you will have brought the solutions to exercises (a)–(d):

(a) Define the following concepts: independent stochastic variables, expectation (mean), variance, covariance, correlation, and conditional density function.

(b) Let $X_1$ and $X_2$ be independent and normally distributed with means 0 and variances 1 and let

$$
Y_1 = X_1,
$$

$$
Y_2 = \rho X_1 + X_2 \sqrt{1 - \rho^2}.
$$

Calculate $\mathbb{V}(Y_1)$, $\mathbb{V}(Y_2)$ and $\mathbb{C}(Y_1, Y_2)$, as well as the correlation coefficient $\rho(Y_1, Y_2)$.

(c) Let $(X, Y)$ follow a two-dimensional normal distribution with $m_X = 1$, $m_Y = 2$, $\sigma_X = 1$, $\sigma_Y = 0.5$, and $\rho = 0.6$. Determine the distribution of $X$ given $Y = 1$, and given $Y = 3$.

(d) Let $W(t) \in N(0, 0.1 \sqrt{t})$ be a Wiener process and let $X(t) = 0.8 \cdot e^{0.3t + W(t)}$ be the corresponding Geometrical Brownian Motion (see appendix). Determine the distribution of $X(t)$.

2 Conditional distributions

First write `addpath 'g:\matstat\ikurs'` to reach the special m-routines or, if this does not work, download them from the home page and save them in your `matlab`-folder.

In this part we will shed some light on the concept of conditional distributions. This is important an important concept since conditional distributions and, in particular, their means and variances are basic tools for all prediction and reconstruction in stochastic systems. The intention is that you should practice on correlation as a measure of dependence between two stochastic variables $X$ and $Y$. We will work with a two-dimensional normal distribution $(X, Y)$. 

The density function of a two-dimensional normal distribution with means \( m_X, m_Y \), standard deviations \( \sigma_X, \sigma_Y \) and correlation coefficient \( \rho = \rho(X, Y) = \frac{C(X, Y)}{\sigma_X \sigma_Y} \) is given by

\[
f_{X,Y}(x,y) = K \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{(x-m_X)^2}{\sigma_X^2} + \frac{(y-m_Y)^2}{\sigma_Y^2} - 2\rho \frac{(x-m_X)}{\sigma_X} \frac{(y-m_Y)}{\sigma_Y} \right] \right\},
\]

where

\[
K = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}}.
\]

What can we say about the dependence between \( X \) and \( Y \) if \( \rho = 0 \)? If \( \rho = 1 \)?

**Answer:** ...

By determining the conditional density function \( f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \) one sees that the conditional distribution of \( X \) given \( Y = y \) is a one-dimensional normal distribution with

- \( \mathbb{E}(X \mid Y = y) = m_X + \rho \frac{\sigma_X}{\sigma_Y} (y - m_Y) \),
- \( \text{Var}(X \mid Y = y) = \sigma_X^2 (1 - \rho^2) \).

Note that the conditional mean equals \( m_X \) plus a correction term depending linearly on \( y \), while the conditional variance only depends on \( \rho \). (Analogous formulae hold for \( n \)-dimensional normal distributions.)

You shall now study graphically how the conditional distribution, mean, and variance of \( X \) change when we fiddle with the different parameters. In other words, how does our information about \( X \) change after we have observed \( Y = y \)?

There are two specially written m-files, `normal2d` and `condnormal` which give pictures of the density functions involved.

The command `normal2d(m_X, m_Y, \sigma_X, \sigma_Y, \rho)` draws a picture of the two-dimensional density function, its level curves, and the marginal density functions for \( X \) and \( Y \).

The function `condnormal()` draws pictures of the conditional density functions. The command `condnormal(m_X, m_Y, \sigma_X, \sigma_Y, \rho, y, y')`, e.g., generates a picture over the conditional density function of \( X \) given \( Y = y \).

Plot some different distributions and examine how the conditional mean and variance are affected by small or large values of \( \rho \), \( \sigma_X \), and \( \sigma_Y \). What happens if \( \rho = 0 \) or 0.99?

**Answer:** ...

Use, e.g., `condnormal` and `hold on` to study how the density changes with \( \rho \) and \( \sigma_Y \). What happens when you change \( \rho \) and \( \sigma_Y \)?

**Answer:** ...

### 3 Functions of stochastic variables

#### 3.1 Constant price development over time

A certain type of electronical component has, due to perfected production techniques, been able decline in price by a certain percentage each year. If the relative change in price is constant then the price, \( P(t) \), at time \( t \), can be described by the relationship

\[
P(t) = P(0) \cdot r^t
\]
where $P(0)$ is the initial price and $r$ is the yearly price change. Now, suppose that $r = 0.8$, i.e., the price decreases by 20% per year, and that $P(0) = 100$ kr. Plot the price development during the following 10 years:

```matlab
>> r=0.8;
>> P0=100;
>> t=[0:0.5:10]’;
>> Pt=P0*r.^t;
>> plot(t,Pt)
```

The half-life, i.e. the time, $T_{0.5}$, it takes until the price has been halved, i.e., when $P(T_{0.5}) = \frac{P(0)}{2}$, can be calculated as

$$T_{0.5} = \frac{\ln 0.5}{\ln r}.$$ 

As is seen the half-life does not depend on the initial price. In this case

$$T_{0.5} = \frac{\ln 0.5}{\ln 0.8} \approx 3.1 \text{ years.}$$

In reality the decrease in price differs between producers, e.g., due to exchange rates, personell politics, etc. It might not be unreasonable to suppose that the price change, $R$, for a randomly chosen producer follows a log-normal distribution, such that $\ln R \in N(\ln 0.8, \sigma)$. We start by looking at the price development for 10 different producers when $\sigma = 0.05$:

```matlab
>> sigma=0.05;
>> r=lognrnd(log(0.8),sigma,10,1);
>> T50=log(0.5)/log(r)
>> for k=1:10
>> plot(t,P0*r(k).^t)
>> hold on
>> end
>> plot(T50,0,’*’)
>> hold off
```

Does there seem to be a large variation on $P(t)$? On $T_{0.5}$?

**Answer:** …

What does the density function of $T_{0.5}$ look like? What will the expected $T_{0.5}$ be? How large is the variation in $T_{0.5}$? Answer these questions by simulating $T_{0.5}$ 1000 times, plot a histogram (with `hist`) and estimate $E(T_{0.5})$ and $D(T_{0.5})$ using the functions `mean` and `std` respectively.

**Answer:** …

Repeat the simulations with a smaller variation on $\ln R$, e.g. $\sigma = 0.01$. How do $E(T_{0.5})$ and $D(T_{0.5})$ change?

**Answer:** …
3.2 Geometrical Brownian Motion

Previously we assumed that the change in price was constant over time. In reality it is often useful to imagine the price change \( r(t) \) as a stochastic function of the time \( t \). An often used model for stock rates is the Geometrical Brownian Motion, see appendix. In this case the change in \( r(t) \) at time \( t \) can be described by the stochastic differential equation

\[
\begin{align*}
  r(0) &= r_0, \\
  dr(t) &= \alpha \cdot r(t) \, dt + r(t) \, dW(t),
\end{align*}
\]

where \( W(t) \) is a Wiener process. Loosely speaking this means that \( dW(t) \) are independent and \( N(0, \sigma) \)-distributed and that \( W(t) \) is \( N(0, \sigma \sqrt{t}) \)-distributed. On the other hand the \( W(t) \)s at different times are not independent. In other words, the change in \( r(t) \) depends on the size of \( r(t) \) (the term \( \alpha r(t) \, dt \)) as well as on chance (\( dW(t) \)) and the larger \( r(t) \) the larger the jumps (the factor \( r(t) \) in front of \( dW(t) \)).

If we solve the above stochastic differential equation we get

\[
  r(t) = r_0 \exp\left(\alpha - \frac{\sigma^2}{2}\right) t + W(t)
\]

whose distribution at time \( t \) you have calculated in preparatory exercise (d). The specially written routine \( \text{gbr} \) simulates a Geometrical Brownian Motion. The command \( \text{gbr}(\alpha, \sigma, x_0, T, n) \) gives \( n \) different realizations of a GBR with parameters \( \alpha \) and \( \sigma \) with initial value \( x_0 \) at the time-points 0, 1, ..., \( T \).

```matlab
>> help gbr
>> [r,t]=gbr(0.005,0.1,1,100,10);
>> plot(t,r)
```

plots 10 simulations of \( r(t) = e^{W(t)} \) where \( t = 0, \ldots, 100 \) and \( W(t) \in N(0, 0.1 \sqrt{t}) \). Experiment with different parameter values and consider what should happen when \( \alpha > \sigma^2/2, \alpha < \sigma^2/2, \) and \( \alpha = \sigma^2/2 \).

\( \text{Answer: . . .} \)

Also simulate many short series and check that the distribution of \( r(t) \) coincides with the theoretical result from the preparatory exercise. The last row in the \( r \)-matrix, e.g. when \( t = T \), can be extracted with \( r(\text{end},:) \). The density function for a log-normal distribution can be calculated with \( \text{lognpdf} \). Check Laboration 1 to see how to get a density function and a histogram with the same scale in the same plot.

\( \text{Answer: . . .} \)

Calculating the distribution of the half-life now is not easy and is left to some course in extreme value theory.

4 Appendix

4.1 Wiener processes

A Wiener process \( W(t) \) is a sequence of random numbers with the following four properties:

i) \( W(0) = 0 \), i.e. it starts in 0 at time \( t = 0 \).

ii) The increase in a time interval is independent of the increase in all other non-overlapping time intervals, i.e. \( W(t_0) - W(t_0) \) and \( W(t_1) - W(t_1) \) are independent when \( t_0 < t_0 < t_1 < t_1 \).

iii) \( W(t) - W(s) \in N(0, \sigma \sqrt{t-s}) \), i.e. the increase in the interval \( (s, t] \) in normally distributed and the variance depends only on the length of the interval, not on, e.g., its location.
iv) $W(t)$ is continuous.

This means, among other things, that while a Wiener process is continuous everywhere it is so wiggly that it does not have a continuous derivative anywhere. Despite this it is very often used as model in many practical situations.

Figur 1: A–C: Successive enlargements of a Wiener process with $\sigma = 0.1$. D: Corresponding Geometrical Brownian Motion: $X(t) = 0.8 \cdot e^{0.3t + W(t)}$.

### 4.2 Geometrical Brownian Motion

A financial application of the Wiener process is as a model for stock values. It turns out that these often can be described by the following stochastic differential equation:

\[
\begin{align*}
X(0) &= x_0, \\
dX(t) &= \alpha X(t) \, dt + X(t) \, dW(t)
\end{align*}
\]

where $dW(t) \in N(0, \sigma)$ is the change in the Wiener process at time $t$ and $\alpha$ describes the drift in the process. In the stock context $\sigma$ is often called volatility. If one solves the stochastic differential equation ones finds that

\[X(t) = x_0 e^{(\alpha - \sigma^2/2) t + \sigma W(t)}.\]

It can be shown that if $\alpha > \sigma^2/2$ the process grow out of control: $X(t) \to \infty$ when $t \to \infty$. If, on the other hand, $\alpha < \sigma^2/2$ the process will eventually die out: $X(t) \to 0$ when $t \to \infty$. In the case $\alpha = \sigma^2/2$ the process will vary between arbitrarily large and small values.