in a "risk neutral world" (see later)

$r$ (constant) market interest rate

In general, (*) is difficult to compute.

However, if $S_t$ is a Markov process, (*) can be calculated in many cases

## 4. Simulation of stoch. proc.

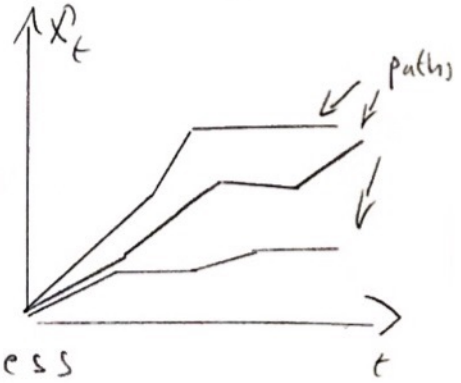
Simulation of a random variable $X$ (e.g. $X = S_T$) means to imitate the outcomes of $X$ by computer power

→ One way: Outcomes of $X$ are generated by means of pseudo rand. numbers (of a uniform distribution) according to an algorithm in the computer (see later)

Problem: How can a stochastic proc. $X_t$ (e.g. $X_t = S_t$) be simulated?

→ In general difficult, since the outcomes of $X_t$ are given by paths ($t \mapsto X_t$) (not single numbers!)

→ infinite dimensional problem




However, if $X_t$ is a Markov process simulation techniques are available

→ Application: The simulated paths of $X_t$ can be e.g. used to approximate option values of the form (*) (if explicit pricing formulas are not available)

## 5. Filtering in engineering:

**Ex.:** Consider a space shuttle with position $X_t$ at time $t$ sending signals $Z_t$ to (monitoring centre) NASA for locating purposes




Transmission code converts $X_t$ into $Z_t$

$$\longrightarrow \quad Z_t = \int_0^t c(X_s)\, ds$$

for a function $c : \mathbb{R}^3 \to \mathbb{R}$

**However:** $Z_t$ is distorted by noise! (on the way to earth)

$\longrightarrow$ NASA receives signals of the form:

$$Z_t = \int_0^t c(X_s)\, ds + B_t$$

**Noise source:** $B_t$ (e.g. Brownian motion)

**or more generally:**

(system) $X_t$ modelled by a Markov process

(observations) $Z_t = \int_0^t c(X_s)\, ds + B_t$

$\longrightarrow$ **Filtering problem:**

What is the **best estimate** $\hat{X}_t$ of the state $X_t$ **based on** the observations $Z_s$, $0 \le s \le t$?

$\longrightarrow$ $\hat{X}_t$ can be computed by exploiting the Markov property of $X_t$

## 6. Branching processes in biology

Branching processes were originally introduced by Galton and Watson (1870s) to study the phenomenon of disappearance of family names in populations

Assumption: Family names are passed from fathers to sons

At time $t=0$ there is 1 male, i.e. $X_0 = 1$
This male dies and is replaced at time $t=1$ by $N_1^1$ sons

$\longrightarrow$ $X_1 := N_1^1$ number of family name bearers in the 1st generation

These sons also die and leave at time $t=2$
$N_1^2, \dots, N_{X_1}^2$ sons

$\longrightarrow$ $X_2 = N_1^2 + \dots + N_{X_1}^2$ number of family name bearers in the 2nd gen.

$\vdots$

More generally, we get

$X_n = N_1^n + \dots + N_{X_{n-1}}^n$ number of fam. name bearers in the $n$-th gen.

$N_K^n \triangleq$ number of male descendants of the $K$-th male in the $(n-1)$-th gen.

reasonable assumption:
For each $n$ $N_K^n$, $K \geq 0$ are independent and identically distributed (see later)

$\longrightarrow$ $X_t, t \in \mathbb{N}_0$ becomes a special case of a Markov process, called branching process

## 2. Prerequisites

We need the following fundamental notions:

**Def 2.1** (Sample space)

A set $\Omega \neq \emptyset$ representing the collection of all possible outcomes of a random experiment is called **sample space**

**Ex.** Throw a die $\longrightarrow$ outcomes $1, 2, \dots, 6$

$\longrightarrow$ reasonable choice: $\Omega = \{1, 2, \dots, 6\}$

In the sequel we shall denote by

$\mathcal{F}$

a family of subsets of $\Omega$ such that

(i) $\emptyset \in \mathcal{F}$

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$ ($A^c := \Omega - A$ (without))

(iii) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i \geq 1} A_i \in \mathcal{F}$

A subset $A \subseteq \Omega$ with $A \in \mathcal{F}$ is called an **event** and $\mathcal{F}$ is referred to as **$\sigma$-algebra**

**Ex.** $\Omega = \{1, \dots, 6\}$. Choose e.g.

$\mathcal{F} := \mathcal{P}(\Omega)$

the class of all subsets of $\Omega$

E.g. $A = \{1, 2, 3\}$ ($\in \mathcal{F}$) the event that one throws 1, 2 or 3

$\longrightarrow$ $A^c = \{1, \dots, 6\} - \{1, 2, 3\} = \{4, 5, 6\}$ is the complementary event to $A$

$A = \Omega$ sure event

$A = \emptyset$ impossible event, that is e.g. to throw 7

Def. $A_1 = \{1, 2, 3\}$, $A_2 = \{3, 4\}$

$\longrightarrow$ $A := A_1 \cup A_2$ the event that $A_1$ or $A_2$ occur, i.e. $\{1, 2, 3, 4\}$

$A := A_1 \cap A_2$ the event that $A_1$ and $A_2$ occur, i.e. $\{3\}$

**Def 2.2** (Probability measure on $\mathcal{F}$)

A function

$$P : \mathcal{F} \to [0,1],$$

that is a function which assigns each event $A \in \mathcal{F}$ to a number in $[0,1]$ is called <u>probability measure</u> on $\mathcal{F}$, if

(i) $P(\emptyset) = 0$, $P(\Omega) = 1$

(ii) $A_1, \dots, A_n, \dots \in \mathcal{F}$ with $A_i \cap A_j = \emptyset$, $i \neq j$ (i.e. disjoint events), Then

$$P\left(\bigcup_{i \geq 1} A_i\right) = \sum_{i \geq 1} P(A_i)$$

<u>Interpretation</u>:

$P(A)$ = "the prob. that the event $A$ occurs"

<u>Rem.</u>: $A := \bigcup_{i \geq 1} A_i$ the event that $A_1$ or $A_2$ or ... occur

<u>Ex.</u>: a) $\Omega = \{1, \dots, 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ class of all subsets of $\Omega$

$P(A) := (\text{number of elements in } A)/6$

e.g. $P(\{1,2,3\}) = \frac{1}{2}$ $\longrightarrow$ prob. that one throws 1, 2 or 3

b) Dirac measure $P = \delta_\omega$:

$$\delta_\omega(A) := \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

<u>Useful properties of $P$</u>:

(i) $A \subseteq B \Rightarrow P(A) \leq P(B)$

(e.g. $\Omega = \{1, \dots, 6\}$, $A = \{1,2\}$, $B = \{1,2,3\}$)

(ii) $P(A^c) = 1 - P(A)$ ($A^c := \Omega - A$)

(iii) $P\left(\bigcup_{i \geq 1} A_i\right) \leq \sum_{i \geq 1} P(A_i)$ for <u>not</u> necessarily disjoint sets

The triple

$$(\Omega, \mathcal{F}, P)$$

is called <u>probability space</u>

<u>Def 2.3</u> (Random variable)

A function

$$X : \Omega \to \mathbb{R}$$

is called <u>random variable</u> (r.v.) on our probability space, if the set

$$A := \{\omega \in \Omega : X(\omega) < t\}$$

is an event (i.e. $A \in \mathcal{F}$) for all $t$

<u>Rem.:</u> $\{X < t\} := \{\omega \in \Omega : X(\omega) < t\}$

<u>Ex.:</u> $\Omega = \{1, \dots, 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ the class of all subsets of $\Omega$

$\longrightarrow$ <u>any</u> function $X : \Omega \to \mathbb{R}$ is a r.v.

e.g. $X(\omega) = \omega$ if $\omega = 1, \dots, 6$, which models the outcomes of a rolling die

<u>Ex.:</u> As an application to finance we may think of $\mathcal{F}$ to be a model for the <u>entirety of market information</u> (i.e. market events) ~~up to time $T$~~.

$\longrightarrow$ It is "clear" that $\{S_T < t\}$ (stock price at time $T$)

is a (possible) market event for all $t$ (#)

Def. 2.3 $\Longrightarrow$ $X := S_T$ is a r.v.

**Def. 2.4** (Random vector)

A function

$$X : \Omega \to \mathbb{R}^n$$

with

$$X(\omega) = (X_1(\omega), \dots, X_n(\omega))$$

is called <u>random vector</u>, if the components $X_1, X_2, \dots, X_n$ are r.v.'s

**Def. 2.5** (Distribution function)

Let $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ be a random vector

Then the function

$$F_X : \mathbb{R}^n \to [0,1]$$

given by

$$F_X(x_1, \dots, x_n) := P(X_1 \le x_1, \dots, X_n \le x_n)$$

for $x_1, \dots, x_n \in \mathbb{R}$

is called <u>distribution function</u> of $X$

We say two rand. vectors $X, Y \in \mathbb{R}^n$ have the <u>same distribution</u> (i.e. $X \sim Y$), if

$$F_X(x_1, \dots, x_n) = F_Y(x_1, \dots, x_n)$$

for all $x_1, \dots, x_n \in \mathbb{R}$

**Rem. 2.6**: Let $X$ be a <u>discrete</u> r.v., that is $X$ takes values in a countable set $J$ (i.e. $J = \{x_1, x_2, x_3, \dots\}$, $x_i \neq x_j$, $i \neq j$)

Then

$$F_X(x) = P(X \le x) = P\Big(\bigcup_{i : x_i \le x} \{X = x_i\}\Big) \overset{\text{disjoint}}{\underset{\text{Def. 2.2(ii)}}{=}} \sum_{i : x_i \le x} P(X = x_i)$$

$\Rightarrow$ $F_X$ determined by the probabilities

$$p_i := P(X = x_i), \quad i = 1, 2, \dots$$

$\longrightarrow$ Examples of the distr. of discrete r.v.'s:

## Ex. 2.7 (Binomial distr.)

Consider a particular event $A$ of a random experiment (e.g. to throw 2 or 3).
Suppose the experiment is repeated $n$ times.
Def. the r.v. $X$ which counts the occurence of $A$
$\Rightarrow X \in \{0, 1, \dots, n\}$
Denote by $p \in [0,1]$ the probability that $A$ occurs.
Then one shows that

$$P(X = k) = b(k, n, p),$$

where

$$b(k, n, p) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \dots, n$$

$\longrightarrow$ $X$ binomially distributed with parameters $n, p$ (i.e. $X \sim bin(n, p)$)

$\binom{n}{k} := \frac{n!}{(n-k)!\,k!}$ ; $m! := m(m-1)\cdot(m-2)\cdot\ldots\cdot 1$, $m \in \mathbb{N}$
$0! := 1$

## Ex. 2.8 (Poisson distr.)

A r.v. $X \in \{0, 1, 2, \dots\}$ is Poisson distributed with param. $\lambda > 0$ (i.e. $Pois(\lambda)$), if

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$\longrightarrow$ a limit case of Ex. 2.7, since:
Let $p = p_n$ in Ex. 2.7 depend on $n$ such that

$$p_n \xrightarrow[n\to\infty]{} 0$$

Choose e.g. $p_n$ s.t.

$$n \cdot p_n \xrightarrow[n\to\infty]{} \lambda$$

for a $\lambda > 0$
Then

$$b(k, n, p_n) \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!} e^{-\lambda}$$

$X \sim Pois(\lambda)$ can be e.g. used to model the number of claims (e.g. fire losses) arriving by a fixed time $t$ (see App. 2d in Sect. 1)

**Def 2.9** (Probability density)

Let $X: \Omega \to \mathbb{R}^n$ be a random vector.

Assume for $x_1, \dots, x_n \in \mathbb{R}$

$$F_X(x_1, \dots, x_n) = \int_{-\infty}^{x_n} \int_{-\infty}^{x_{n-1}} \cdots \int_{-\infty}^{x_1} f_X(y_1, y_2, \dots, y_n)\, dy_1 \cdots dy_n$$

for a function $f_X: \mathbb{R}^n \to [0, \infty)$

Then $f_X$ is called probability density function of $X$

**Ex. 2.10** (Uniform distr.)

Consider e.g. the future life time

$$T = T(x)$$

of a life aged $x$ years

$\longrightarrow$ $x+T$ time of death

Define (Gauss bracket)

$$S := T - [T] \quad (\in [0,1])$$

$\longrightarrow$ remaining fraction of the year of death

e.g. $[0.9] = 0$, $[1.1] = 1$, $[2.7] = 2$, ...

$\longrightarrow$ common assumption:

$S$ uniformly distr. on $[0,1]$ (i.e. $S \sim U(0,1)$), i.e.

$$F_S(y) \overset{\text{def.}}{=} P(S \le y) = y$$

for all $0 \le y \le 1$

$$\Rightarrow F_S(y) = \int_{-\infty}^{y} 1_{[0,1]}(u)\, du$$

$$\Rightarrow f_S(u) = 1_{[0,1]}(u) = \begin{cases} 1, & 0 \le u \le 1 \\ 0 & \text{else} \end{cases}$$

**Ex. 2.11** (Exponential distr.)

Let $T_1 \le T_2 \le T_3 \le \dots$ be the arrival times of claims (e.g. fire losses). See App(.2 in Sect.1

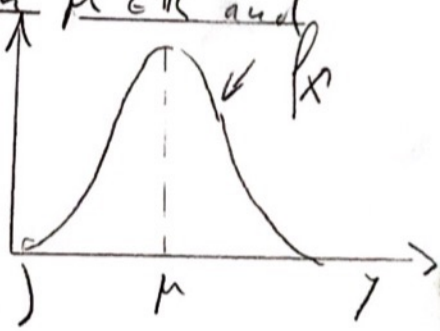
$\longrightarrow$ $W_i := T_i - T_{i-1}$ inter-arrival times

$\to$ time span betw. two claims

$\longrightarrow$ reasonable model: $W_i$ exponentially distr. with param. $\lambda > 0$ (i.e. $W_i \sim \text{Exp}(\lambda)$), that is

(15)
$$F_{W_i}(x) = \int_{-\infty}^{x} f_{W_i}(y)\,dy ,$$
where
$$f_{W_i}(y) = \lambda \cdot e^{-\lambda \cdot y} \cdot 1_{[0,\infty)}(y)$$

Ex. 2.12 (Normal distr.)
A r.v. $X$ is normally distr. with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ (i.e. $X \sim N(\mu, \sigma^2)$), if
$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy ,$$
where
$$f_X(y) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \exp(-(y-\mu)^2/2\cdot\sigma^2)$$




→ Application: modeling of the position of electrons

Def. 2.13 (Independence)
(i) The events $A_1, \dots, A_n$ are independent $\iff$
$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_n)$$
(ii) The r.v.'s $X_1, \dots, X_n$ are independent $\iff$
$$F_Y(y_1, \dots, y_n) = F_{X_1}(y_1) \cdot F_{X_2}(y_2) \cdot \dots \cdot F_{X_n}(y_n)$$
for all $y_1, \dots, y_n \in \mathbb{R}$,
where $Y = (X_1, \dots, X_n)$

Rem. 2.14: (i) Note: If $A$ event, then $X$ given by
$$X(\omega) := 1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0 & \text{else} \end{cases}$$ is a r.v.
$\Rightarrow$ $A_1, \dots, A_n$ indep. events $\iff$
$X_1 := 1_{A_1}, \dots, X_n := 1_{A_n}$ indep. r.v.'s
(ii) A collection of r.v.'s $\{X_i\}_{i \in I}$, $I \neq \emptyset$ is independent, if $\{X_i\}_{i \in J}$ is indep. for all finite subsets $J \subseteq I$
(iii) A sequence of r.v.'s $X_n$, $n \geq 0$, is independent and identically distr. (i.i.d), if $\{X_n\}_{n \geq 1}$ is indep. and $F_{X_i} = F_{X_j}$ for all $i, j \geq 1$

Ex.: Modeling of $W_i$, $i \geq 1$ in Ex 2.11 by i.i.d r.v. with common exponential distr. $Exp(\lambda)$

**Def. 2.15** (stochastic process)
Let $\mathcal{I} \neq \emptyset$ be an index set (e.g. $\mathcal{I} = [0,1]$)
A collection
$$\{X_i\}_{i \in \mathcal{I}}$$
of r.v.'s (or rand. vectors) is called stochastic process.
$\mathcal{I}$ is the parameter space

Ex.: $\mathcal{I} = [0,\infty)$, $S_t$ price of a stock at time $t$
$\longrightarrow$ $\{S_t\}_{0 \leq t < \infty}$ stoch. proc.

**Ex. 2.16** (Gaussian process)
A stoch. proc. $\{X_i\}_{i \in \mathcal{I}}$ is called Gaussian if each finite linear combination
$$X := \sum_{i=1}^{n} \alpha_i X_{t_i}, \quad \alpha_i \in \mathbb{R}, t_i \in \mathcal{I}, n \in \mathbb{N}$$
is a normally distr. r.v. (see Ex. 2.12)
$\longrightarrow$ Ex: Brownian motion (see later)

**Def. 2.17** (Expectation of a discrete r.v.)
Let $X \in \{x_1, x_2, x_3, \dots\}$ be a discrete r.v.
Define
$$p_i := P(\{X = x_i\})$$
Then the expected value of $X$ denoted by $E[X]$ is defined as
$$E[X] := \sum_{i \geq 1} x_i \cdot p_i,$$
if $\sum_{i \geq 1} |x_i| \cdot p_i < \infty$.

Ex.: $X \sim \text{Pois}(\lambda)$. Then $E[X] = \lambda$ (see Exerc. 1, Prob 2)

**Def. 2.18** (Expectation of a r.v. with density)
Let $f_X$ be the prob. density of a r.v. $X$. Then the expected value of $X$ denoted by $E[X]$ is defined as
$$E[X] := \int_{-\infty}^{\infty} y \cdot f_X(y)\, dy,$$
if the right hand side converges

Ex.: $X \sim \text{Exp}(\lambda)$ Then $E[X] = \frac{1}{\lambda}$ (see Exerc. 1, Prob. 2)

**Prop. 2.19**: Let $X = (X_1, \dots, X_n)$ be a rand. vector and $g: \mathbb{R}^n \to \mathbb{R}$ a function.
Denote by

(17) the r.v. $Y := g(X) = g(X_1, \dots, X_n)$, $f_X$ prob. dens. of $X$

Then

$$E[Y] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(y_1, \dots, y_n) \cdot f_X(y_1, \dots, y_n) \, dy_1 \dots dy_n ,$$

if the right hand side converges