Dependable Systems

2. Chapter

# Recap: Stochastic

Prof. Matthias Werner
Operating Systems Group



## 2.1 Basics

### Motivation

- Faults and load do not behave deterministically ➡ random
- However: randomness can be “tamed” by probability calculus / stochastics
- **Probability calculus**
  - Roots from considerations on gambling
  - Empirical definition by LAPLACE; axiomatic foundation by LAPLACE
- Probability calculus is part of stochastics; other parts are error analysis, statistics, ...





## Basic Terms

**Experiment (with chance):** Any experiment with unpredictable results, e.g., tossing a die

**Event space:** Overall set of possible events for an experiment (also: event field, $\Omega$)

**Event:** Any (in theory) observable result of an experiment

- **Certain event:** event, that will be always result of an experiment (i.e., $\Omega$)
- **Impossible event:** event, that is never a result of an experiment (i.e., $\emptyset$)
- **Random event:** a event, that is neither certain nor impossible ➡ *Events can be constructed from other events by means of set theory*

**Elementary events:** event, that can not constructed by merging of other events
➡ *The impossible event is not a elementary event*

**Random variable:** mapping of the an experiment’s possible outcomes to real numbers





## Probability

- **Observation:** The relative number of occurrences of a selected outcome (event $A$) become stable for a big number of repetitions of experiments
- **Notation:** $\Pr(A)$ or $P(A)$ = probability that event $A$ occurs
- **Definition by LAPLACE:**

$$\Pr(A) = \frac{\text{Number of elementary events supporting } A}{\text{Number of elementary events}}$$

- **Definition by KOLMOGOROW:** Given an event space $\Omega$ and events $A_i$
  1. $\Pr(A_i) \geq 0$
  2. $\Pr(\Omega) = 1$
  3. $\Pr(A_1 \cup A_2 \cup \ldots) = \Pr(A_1) + \Pr(A_2) + \cdots$ if any pair $A_i$, $A_j$ is disjoint (i.e., $A_i \cap A_j = \emptyset$)

## Probability: Basic Properties

A number of properties can be derived from axioms:

- $\Pr(\emptyset) = 0$
- $\Pr(A) = 1 - \Pr(\bar{A})$, where $\bar{A}$ is complement event of $A$
- $\Pr(\bar{A} \cap B) = \Pr(B) - \Pr(A \cap B)$
- $\Pr(A - B) = \Pr(A) - \Pr(A \cap B)$
- $B \subseteq A \Rightarrow \Pr(B) \leq \Pr(A)$
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$





## Independency and Exclusion

Two events $A$ and $B$ are **independent** if

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$$

Two events $A$ and $B$ are **mutual exclusive** if

$$A \cap B = \emptyset$$

**Please note!**

Sometimes, independency and exclusion are confused.
Obviously, for independent events may $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B) \neq 0$, but exclusion requires $\Pr(A \cap B) = \Pr(\emptyset) = 0$.





## Conditional Probabilities

**Conditional probability** $\Pr(A|B)$ of an event $A$ under the condition that $B$ is already given, is

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} \quad \text{with } \Pr(B) \neq 0$$

For independent events $A$ and $B$:

$$\Pr(A|B) = \Pr(A)$$





## Conditional Probabilities (cont.)

Two events $A$ and $B$ are **conditional independent**, if

$$\Pr((A \cap B)|C) = \Pr(A|C)\Pr(B|C)$$

**Note:**

Conditional independence does not imply independence!

From definition of conditional probability one can derive:

**Theorem 2.1 (Multiplication Theorem)**

$\Pr(A \cap B) = \Pr(A|B)\Pr(B) = \Pr(B|A)\Pr(A)$ *(with $\Pr(A), \Pr(B) \neq 0$)*

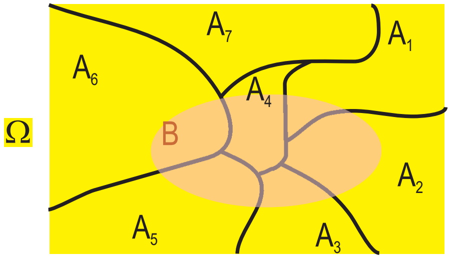
## Total Probability

Let $A_1, A_2, \ldots A_n$ be pair-wise excluding random events, i.e.,

- $\forall i, j, i \neq j, A_i \cap A_j = \emptyset$
- $\bigcup_i A_i = \Omega$

Let $B$ a random event with $\Pr(B) > 0$
Then:

$$\Pr(B) = \sum_{i=1}^{n} \Pr(A_i) \Pr(B|A_i)$$








## Bayes' Theorem

Let $A_1, A_2, \ldots A_n$ random events with mutual exclusion

- $\forall i, j, i \neq j, A_i \cap A_j = \emptyset$
- $\bigcup_i A_i = \Omega$
- all $\Pr(A_i)$ are known (**a priori** probability)

Let $B \subseteq \Omega$ a random event with $\Pr(B) > 0$
Then:

**Theorem 2.2 (Bayes' Theorem)**

$$\Pr(A_i|B) = \frac{\Pr(A_i \cap B)}{\Pr(B)} = \frac{\Pr(A_i) \Pr(B|A_i)}{\sum_j \Pr(A_j) \Pr(B|A_j)}$$

$\Pr(A_i|B)$ is called **a posteriori** probability





## 2.2 Random Variables and Distributions

- $\Pr(X \leq t)$ denotes the probability that a **random variable** $X$ has a value that is smaller or equal to $t$
- $t$ can be viewed as a function parameter
- The function $F_X(t) = \Pr(X \leq t)$ ist called **distribution function** of $X$
  ($t, F_X(t) \in \mathbb{R}$)
- It can be used to get the probability that the value of $X$ is in interval $(a, b]$:

$$\Pr(a < X \leq b) = F_X(b) - F_X(a)$$





## Random Variables and Distributions (cont.)

- Continuous distributions are frequently described by the corresponding **density function** $f_X(t)$

$$F_X(t) = \int_{-\infty}^{t} f_X(\tau) d\tau$$

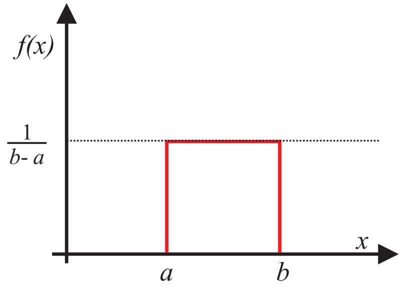
- Then:

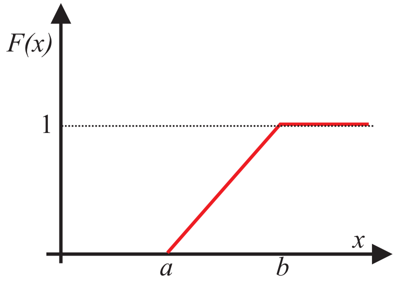
$$\Pr(a < X \leq b) = \int_{a}^{b} f_X(t) dt$$

# Popular Distributions: Rectangular Distribution

For the rectangular distribution, any value within an interval $I = [a, b]$ has the same probability.

$$f(t) = \begin{cases} \frac{1}{b-a} & , \text{ if } a \le t \le b \\ 0 & , \text{ otherwise} \end{cases} \quad F(t) = \begin{cases} 0 & , \text{ if } t < a \\ \frac{t-a}{b-a} & , \text{ if } a \le t \le b \\ 1 & , \text{ if } t > b \end{cases}$$



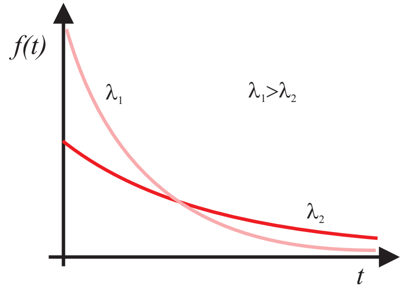


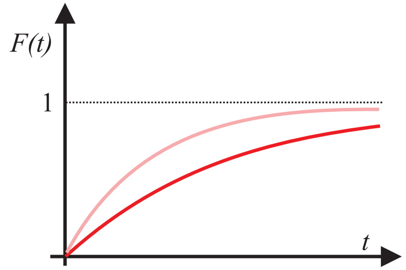



# Popular Distributions: Exponential Distribution

In processes of disintegration, the value of a random variabel depends frequently on the residual.
Then, we get an **exponential distribution**.

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, \text{ if } t \ge 0 \\ 0 \quad\quad , \text{ otherwise} \end{cases} \qquad F(t) = \begin{cases} 0 \quad\quad\quad , \text{ if } t > 0 \\ 1 - e^{-\lambda t}, \text{ if } t \ge 0 \end{cases}$$





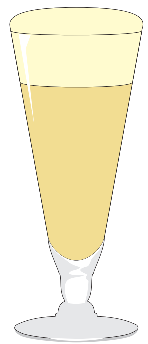



# Why Exponential Function

**Assumption:** relative amount of disintegration per time unit is constant

With other words: number of disintegrating elements is proportional to number of (still) existing elements (e.g., froth)

$$\frac{d}{dt} x(t) = -\lambda \cdot x(t) \qquad (\star)$$







# Why Exponential Function (cont.)

$(\star)$ is a differential equation. It has the general solution

$$x(t) = x_0 \cdot e^{-\lambda \cdot t}$$

For distribution functions, there is a side condition
$\Pr(\Omega) = 1$, i.e., in this case $\int_0^\infty x(t) dt = 1$

$$\int_0^\infty x_0 e^{-\lambda t} dt = 1 \qquad x_0 \int_0^\infty e^{-\lambda t} dt = 1 \qquad x_0 \left[ -\frac{1}{\lambda} e^{-\lambda t} \right]_0^\infty = 1$$

$$-x_0 \frac{[e^{-\lambda\infty} - e^{-\lambda 0}]}{\lambda} = 1 \qquad x_0 \cdot \frac{1}{\lambda} = 1 \qquad x_0 = \lambda$$

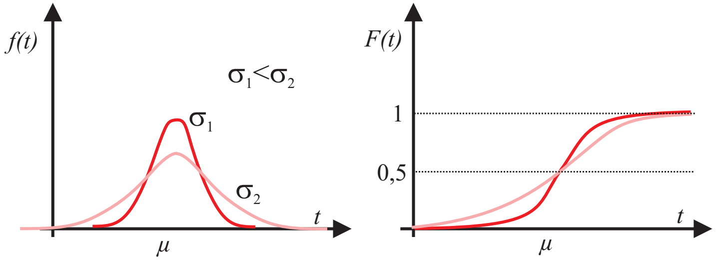
# Popular Distributions: Normal Distribution

Probably, the best-known distribution **Normal distribution** or **GAUSSIAN** distribution.

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}} \qquad (\sigma > 0)$$

The distribution function $F(t) = \int_{-\infty}^{t} f(\tau)d\tau$ can not be given in a closed form.

However, one can use distribution tables. ($\mu = 0, \sigma^2 = 1$).








# Expectation

Distributions are completely characterized by the distribution function.
But: often, we want a more “compact” description

- **Expectation value:** Mean value of the random variable for a big number of experiments

$$\mathrm{E}[X] = \int_{-\infty}^{+\infty} t \cdot f_X(t)dt$$

- Rules:
  - $\mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b$
  - $\mathrm{E}\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} \mathrm{E}[X_i]$ (if all $X_i$ are independent)





# Variance

- Expectation value is a kind of “mass center”
- “Closeness” to this center:

$$\mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2]$$

- $\mathrm{Var}[X]$ is called **variance** (also a bit imprecise **dispersion**).

$$\mathrm{Var}[X] = \int_{-\infty}^{\infty} (t - \mathrm{E}[X])^2 f_X(t)dt$$
$$= \int_{-\infty}^{\infty} t^2 f_X(t)dt - \left(\int_{-\infty}^{\infty} t \cdot f_X(t)dt\right)^2$$





# Moments of Popular Distributions

| **distribution** | **density** $f(x)$ | $\mathrm{E}[X]$ | $\mathrm{Var}[X]$ |
|---|---|---|---|
| Rectangular | $\frac{1}{b-a}$ | $\frac{a+b}{2}$ | $\frac{(b-a)^2}{12}$ |
| Exponential | $\lambda \cdot e^{-\lambda \cdot t}$ | $\frac{1}{\lambda}$ | $\frac{1}{\lambda^2}$ |
| Normal | $\frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}}$ | $\mu$ | $\sigma^2$ |

# 2.3 Multivariate Random Variables

- Experiments can depend on more than one random variable
- Combination of random variables $X_1, X_2, \ldots, X_n$ by a function $Z = g(X_1, X_2, ..., X_n)$ results in a random variable, agian

**General for two distributions:**

$$F_Z(t) = \Pr(Z \leq t) = \iint\limits_{g(x,y)\leq t} f_{XY}(x,y)dx\ dy$$

Here, $f_{XY}(x,y)$ is the density of the common distribution $F_{XY}(x,y) = \Pr(X \leq x \wedge Y \leq y)$.





# Multivariate Distributions

- Common distribution $f_{XY}(x,y)$ describes dependency between $X$ and $Y$
- For independent $X$ and $Y$:

$$F_Z(t) = \Pr(Z \leq t) = \iint\limits_{g(x,y)\leq t} f_X(x)f_Y(y)dx\ dy$$





# Special Cases

**Multiplication:** ($Z = X \cdot Y, X, Y \geq 0$)

$$f_Z(t) = \int\limits_{-\infty}^{\infty} f_X(\tau)f_Y(\frac{t}{\tau})\frac{1}{|\tau|}d\tau$$

**Addition:** ($Z = X + Y, X, Y > 0$)

$$f_Z(t) = \int\limits_{-\infty}^{\infty} f_X(\tau)f_Y(t-\tau)d\tau$$

- The operation $\int\limits_{-\infty}^{\infty} f_1(\tau)f_2(t-\tau)d\tau$ is called **convolution** of $f_1$ and $f_2$ and denoted by $f_1(t) * f_2(t)$.





# Application of LAPLACE Transformation

Convolution is not trivial ➡ frequently, LAPLACE Transformation is used

$$\begin{tikzcd} z = f(t) \arrow[r, "\mathcal{L}"] \arrow[d] & Z = F(s) \arrow[d, "\text{solving in frequency domain}"] \\ z(t) & Z(s) \arrow[l, "\mathcal{L}^{-1}"'] \end{tikzcd}$$

Here, it is true: $\mathcal{L}\left(f_1(t) * f_2(t)\right) = \mathcal{L}\left(f_1(t)\right) \cdot \mathcal{L}\left(f_2(t)\right)$

# Appendix A: LAPLACE-Transformation

- Definition
  - transformation time $\rightarrow$ frequency

$$F(s) = \int_{-\infty}^{\infty} f(t) e^{-s \cdot t} dt$$

  - transformation frequency $\rightarrow$ time

$$f(t) = \frac{1}{2\pi j} \int_{\delta - j\infty}^{\delta + j\infty} F(s) e^{st} ds$$

  ($j$ is the imaginary unit)

- Notation:

$$f(t) \circ\!\!-\!\!\bullet F(s) \qquad F(s) \bullet\!\!-\!\!\circ f(t)$$





## Calculation

- ... the hard way: calculate the integral
- ...the long way: decomposition and use of transformation tables:

| $F(s)$ | $f(t) = \mathcal{L}^{-1}\{F(s)\}$ | $F(s)$ | $f(t) = \mathcal{L}^{-1}\{F(s)\}$ |
|---|---|---|---|
| $1$ | $\delta(t)$ | $\frac{1}{s^n}$ | $\frac{t^{n-1}}{(n-1)!}$ |
| $\frac{1}{s}$ | $\mathbf{1}(t)$ | $\frac{1}{s+a}$ | $e^{-at}$ |
| $\frac{1}{s^2}$ | $t$ | $\frac{1}{(s+a)^2}$ | $te^{-at}$ |





## Calculation (cont.)

| $F(s)$ | $f(t)$ | $F(s)$ | $t(t)$ |
|---|---|---|---|
| $\frac{a}{s^2+a^2}$ | $\sin at$ | $\frac{1}{1+sT}$ | $\frac{1}{T} e^{-\frac{t}{T}}$ |
| $\frac{s}{s^2+a^2}$ | $\cos at$ | $\frac{1}{s(1+sT_1)} \cdot \frac{1}{(1+sT_2)}$ | $1 - \frac{T_1}{T_1 - T_2} e^{-\frac{t}{T_1}} + \frac{T_2}{T_1 - T_2} e^{-\frac{t}{T_2}}$ |





## Rules

- **Linearity:** $a_1 f_1(t) + a_2 f_2(t) \circ\!\!-\!\!\bullet a_1 F_1(s) + a_2 F_2(s)$
- **Time scaling:** $f(at) \circ\!\!-\!\!\bullet \frac{1}{a} F\left(\frac{s}{a}\right), a \neq 0$
- **Time shifting:** $f(t-T) \circ\!\!-\!\!\bullet e^{-sT} F(s)$
- **Frequency shifting** $e^{at} f(t) \circ\!\!-\!\!\bullet F(s-a)$
- **Derivation:** $\frac{d}{dt} f(t) \circ\!\!-\!\!\bullet sF(s) - f(-0)$

  ... and for higher derivatives:

$$\frac{d^k}{dt^k} f(t) \circ\!\!-\!\!\bullet s^k F(s) - s^{k-1} f(-0) - s^{k-2} \dot{f}(-0) - \cdots - f^{(k-1)}(-0)$$

# Rules (cont.)

- **Integration:** $\int\limits_0^t f(\tau)d\tau \circ\!\!-\!\!\bullet \frac{1}{s}F(s)$
- **Frequency derivation:** $t^k f(t) \circ\!\!-\!\!\bullet (-1)^k \frac{d^k}{ds^k}F(s)$
- **Convolution:** $f_1(t) * f_2(t) \circ\!\!-\!\!\bullet F_1(s)F_2(s)$
- **Initial value theorem:** $f(+0) = \lim\limits_{t\to+0} f(t) = \lim\limits_{s\to\infty} s\ F(s)$
- **Final value theorem:** $\lim\limits_{t\to\infty} f(t) = \lim\limits_{s\to 0} s\ F(s)$