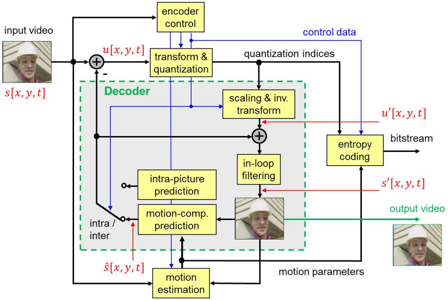
Source Coding and Compression

Heiko Schwarz



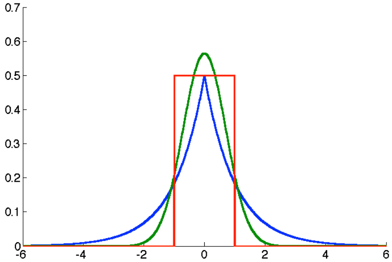
Contact:

Dr.-Ing. Heiko Schwarz

heiko.schwarz@hhi.fraunhofer.de

PartI:
Source Coding Fundamentals

Probability, Random Variables and Random Processes



Outline

Part I: Source Coding Fundamentals

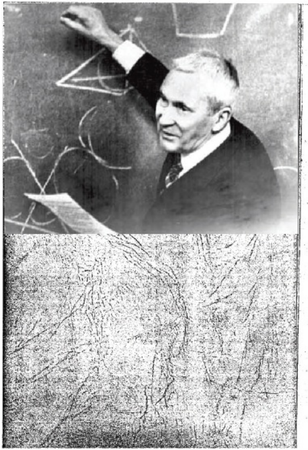
- **Review: Probability, Random Variables and Random Processes**
  - Probability
  - Random Variables
  - Random Processes
- Lossless Source Coding
- Rate-Distortion Theory
- Quantization
- Predictive Coding
- Transform Coding

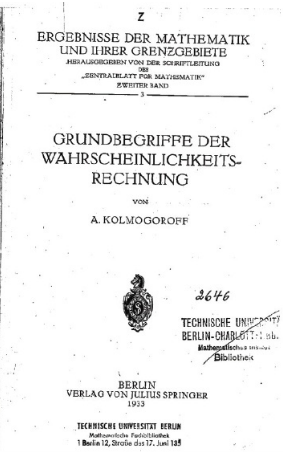
Part II: Application in Image and Video Coding

- Still Image Coding / Intra-Picture Coding
- Hybrid Video Coding (From MPEG-2 Video to H.265/HEVC)

Probability

- Probability theory:
  Branch of mathematics for description and modelling of random events
- Modern probability theory – the axiomatic definition of probability – introduced by KOLMOGOROV





Z

ERGEBNISSE DER MATHEMATIK
UND IHRER GRENZGEBIETE
HERAUSGEGEBEN VON DER SCHRIFTLEITUNG
DES
„ZENTRALBLATT FÜR MATHEMATIK"
ZWEITER BAND
3

GRUNDBEGRIFFE DER
WAHRSCHEINLICHKEITS-
RECHNUNG

VON

A. KOLMOGOROFF

2646

BERLIN
VERLAG VON JULIUS SPRINGER
1933

Definition of Probability

- Experiment with an uncertain outcome: **Random experiment**
- Union of all possible **outcomes** ζ of the random experiment:
  Certain event or **sample space** $\mathcal{O}$ of the random experiment
- **Event**: Subset $\mathcal{A} \subseteq \mathcal{O}$
- **Probability**: Measure $P(\mathcal{A})$ assigned to $\mathcal{A}$ satisfying the following three axioms

1. Probabilities are non-negative real numbers: $P(\mathcal{A}) \geq 0, \quad \forall \mathcal{A} \subseteq \mathcal{O}$
2. Probability of the certain event: $P(\mathcal{O}) = 1$
3. If $\{\mathcal{A}_i : i = 0, 1, \cdots\}$ is a countable set of events such that $\mathcal{A}_i \cap \mathcal{A}_j = \emptyset$ for $i \neq j$, then

$$P\left(\bigcup_i \mathcal{A}_i\right) = \sum_i P(\mathcal{A}_i) \tag{14}$$

Independence and Conditional Probability

- Two events $\mathcal{A}_i$ and $\mathcal{A}_j$ are **independent** if

$$\boxed{P(\mathcal{A}_i \cap \mathcal{A}_j) = P(\mathcal{A}_i)\, P(\mathcal{A}_j)} \tag{15}$$

- The **conditional probability** of an event $\mathcal{A}_i$ given another event $\mathcal{A}_j$, with $P(\mathcal{A}_j) > 0$ is

$$\boxed{P(\mathcal{A}_i|\mathcal{A}_j) = \frac{P(\mathcal{A}_i \cap \mathcal{A}_j)}{P(\mathcal{A}_j)}} \tag{16}$$

- Direct consequence: BAYES' theorem

$$P(\mathcal{A}_i|\mathcal{A}_j) = P(\mathcal{A}_j|\mathcal{A}_i)\,\frac{P(\mathcal{A}_i)}{P(\mathcal{A}_j)} \quad \text{with} \quad P(\mathcal{A}_i),\, P(\mathcal{A}_j) > 0 \tag{17}$$

- Definitions (15) and (16) also imply that,
  if $\mathcal{A}_i$ and $\mathcal{A}_j$ are independent and $P(\mathcal{A}_j) > 0$, then

$$P(\mathcal{A}_i \,|\, \mathcal{A}_j) = P(\mathcal{A}_i) \tag{18}$$

Random Variables

- **Random variable** S:

  Function of the sample space $\mathcal{O}$ that assigns a real value $S(\zeta)$ to each outcome $\zeta \in \mathcal{O}$ of a random experiment

- Define: **Cumulative distribution function** (cdf) of a random variable S:

$$\boxed{F_S(s) = P(S \leq s) = P(\,\{\zeta :\; S(\zeta) \leq s\}\,)} \tag{19}$$

- Properties of cdfs:
  - $F_S(s)$ is non-decreasing
  - $F_S(-\infty) = 0$
  - $F_S(\infty) = 1$

Joint Cumulative Distribution Function

- **Joint cdf** or **joint distribution** of two random variables X and Y

$$F_{XY}(x, y) = P(X \leq x, Y \leq y) \tag{20}$$

- N-dimensional random vector $\boldsymbol{S} = (S_0, \cdots, S_{N-1})^{\mathrm{T}}$:
  Vector of random variables S_0, S_1, $\cdots$, S_{N-1}

- **N-dimensional cdf**, **joint cdf**, or **joint distribution**:

$$\boxed{F_{\boldsymbol{S}}(\boldsymbol{s}) = P(\boldsymbol{S} \leq \boldsymbol{s}) = P(S_0 \leq s_0, \cdots, S_{N-1} \leq s_{N-1})} \tag{21}$$

  with $\boldsymbol{S} = (S_0, \cdots, S_{N-1})^{\mathrm{T}}$ being a random vector

- Joint cdf of two random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$

$$F_{\boldsymbol{X}\boldsymbol{Y}}(\boldsymbol{x}, \boldsymbol{y}) = P(\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Y} \leq \boldsymbol{y}) \tag{22}$$

Conditional Cumulative Distribution Function

- **Conditional cdf** of random variable S given event $\mathcal{B}$ with $P(\mathcal{B}) > 0$

$$\boxed{F_{S|\mathcal{B}}(s \,|\, \mathcal{B}) = P(S \leq s \,|\, \mathcal{B}) = \frac{P(\{S \leq s\} \cap \mathcal{B})}{P(\mathcal{B})}} \tag{23}$$

- Conditional cdf of a random variable X given another random variable Y

$$\boxed{F_{X|Y}(x|y) = \frac{F_{XY}(x,y)}{F_Y(y)} = \frac{P(X \leq x, Y \leq y)}{P(Y \leq y)}} \tag{24}$$

- Conditional cdf of a random vector $\boldsymbol{X}$ given another random vector $\boldsymbol{Y}$

$$F_{\mathbf{X}|\mathbf{Y}}(\boldsymbol{x}|\boldsymbol{y}) = \frac{F_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y})}{F_{\boldsymbol{Y}}(\boldsymbol{y})} \tag{25}$$

Continuous Random Variables

- A random variables S is called a **continuous random variable**, if and only if its cdf $F_S(s)$ is a continuous function

- Define: **Probability density function** (pdf) for continuous random variables

$$\boxed{f_S(s) = \frac{\mathrm{d}F_S(s)}{\mathrm{d}s} \quad \Longleftrightarrow \quad F_S(s) = \int_{-\infty}^{s} f_S(t)\,\mathrm{d}t} \tag{26}$$

- Properties of pdfs:
  - $f_S(s) \geq 0$, $\forall s$
  - $\int_{-\infty}^{\infty} f_S(t)\,\mathrm{d}t = 1$

Examples for Pdfs

- **Uniform pdf**:

$$f_S(s) = \begin{cases} 1/A & : \quad -A/2 \leq s \leq A/2 \\ 0 & : \quad \text{otherwise} \end{cases}, \qquad A > 0 \tag{27}$$

- **Laplacian pdf**:

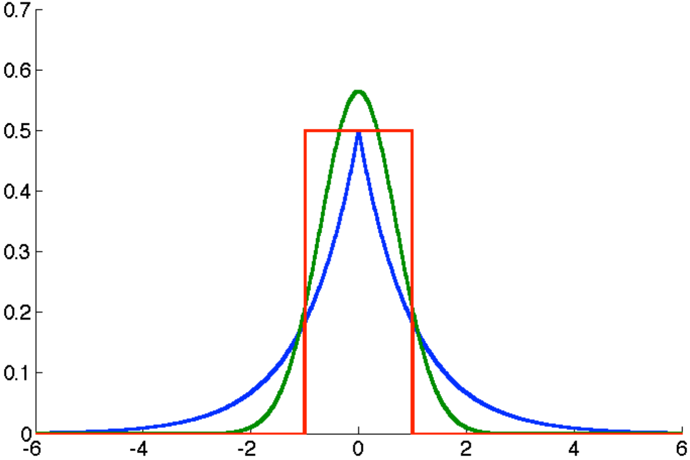
$$f_S(s) = \frac{1}{\sigma_S \sqrt{2}} \, e^{-|s-\mu_S|\sqrt{2}/\sigma_S}, \qquad \sigma_S > 0 \tag{28}$$

- **Gaussian pdf**:

$$f_S(s) = \frac{1}{\sigma_S \sqrt{2\pi}} \, e^{-(s-\mu_S)^2/(2\sigma_S^2)}, \qquad \sigma_S > 0 \tag{29}$$

Generalized Gaussian Distribution

$$f_S(s) = \frac{\beta}{2\,\alpha\,\Gamma(1/\beta)} \cdot e^{-(|x-\mu|/\alpha)^{\beta}} \qquad \Gamma(z) = \int_0^{\infty} e^{-t}\, t^{z-1}\, \mathrm{d}t \qquad (30)$$



Joint and Conditional Pdfs

- **N-dimensional pdf**, **joint pdf**, or **joint density**

$$\boxed{f_{\boldsymbol{S}}(\boldsymbol{s}) = \frac{\partial^N F_{\boldsymbol{S}}(\boldsymbol{s})}{\partial s_0 \cdots \partial s_{N-1}}} \tag{31}$$

- **Conditional pdf** or **conditional density** $f_{S|\mathcal{B}}(s|\mathcal{B})$ of a random variable S given an event $\mathcal{B}$

$$f_{S|\mathcal{B}}(s|\mathcal{B}) = \frac{\mathrm{d}\, F_{S|\mathcal{B}}(s|\mathcal{B})}{\mathrm{d}\, s} \tag{32}$$

- Conditional density of a random variable X given another random variable Y

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} \tag{33}$$

- Conditional density of a random vector $\boldsymbol{X}$ given another random vector $\boldsymbol{Y}$

$$f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y}) = \frac{f_{\boldsymbol{XY}}(\boldsymbol{x},\boldsymbol{y})}{f_{\boldsymbol{Y}}(\boldsymbol{y})} \tag{34}$$

Discrete Random Variables

- A random variable S is called a **discrete random variable**, if and only if its cdf $F_S(s)$ represents a staircase function
- Discrete random variable S takes values of countable set $\mathcal{A} = \{a_0, a_1, \ldots\}$
- Define: **Probability mass function** (pmf) for discrete random variables:

$$\boxed{p_S(a) = P(S = a) = P(\, \{\zeta :\ S(\zeta) = a\} \,)} \tag{35}$$

- Cdf of discrete random variable

$$\boxed{F_S(s) = \sum_{a \le s} p(a)} \tag{36}$$

- Pdf can be constructed using the Dirac delta function δ

$$f_S(s) = \sum_{a \in \mathcal{A}} \delta(s - a)\, p_S(a) \tag{37}$$

Examples for Pmfs

- **Binary pmf**:

$$\mathcal{A} = \{a_0, a_1\} \qquad p_S(a_0) = p, \qquad p_S(a_1) = 1 - p \tag{38}$$

- **Uniform pmf**:

$$\mathcal{A} = \{a_0, a_1, \cdots, a_{M-1}\} \qquad p_S(a_i) = 1/M \qquad \forall\, a_i \in \mathcal{A} \tag{39}$$

- **Geometric pmf**:

$$\mathcal{A} = \{a_0, a_1, \cdots\} \qquad p_S(a_i) = (1-p)\, p^i \qquad \forall\, a_i \in \mathcal{A} \tag{40}$$

Joint and Conditional Pmfs

- **N-dimensional pmf** or **joint pmf** for a random vector $\boldsymbol{S} = (S_0, \cdots, S_{N-1})^{\mathrm{T}}$

$$\boxed{p_{\boldsymbol{S}}(\boldsymbol{a}) = P(\boldsymbol{S} = \boldsymbol{a}) = P(S_0 = a_0, \cdots, S_{N-1} = a_{N-1})} \qquad (41)$$

- Joint pmf of two random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$: $p_{\boldsymbol{XY}}(\boldsymbol{a_x}, \boldsymbol{a_y})$

- **Conditional pmf** $p_{S|\mathcal{B}}(a \,|\, \mathcal{B})$ of a random variable S given an event $\mathcal{B}$, with $P(\mathcal{B}) > 0$

$$p_{S|\mathcal{B}}(a \,|\, \mathcal{B}) = P(S = a \,|\, \mathcal{B}) \qquad (42)$$

- Conditional pmf of a random variable X given another random variable Y

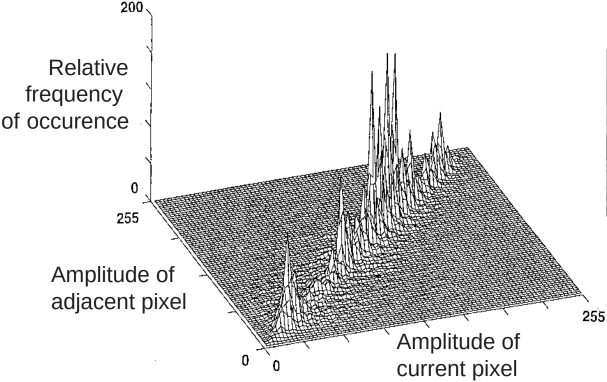
$$p_{X|Y}(a_x|a_y) = \frac{p_{XY}(a_x, a_y)}{p_Y(a_y)} \qquad (43)$$

- Conditional pmf of a random vector $\boldsymbol{X}$ given another random vector $\boldsymbol{Y}$

$$p_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{a_x}|\boldsymbol{a_y}) = \frac{p_{\boldsymbol{XY}}(\boldsymbol{a_x}, \boldsymbol{a_y})}{p_{\boldsymbol{Y}}(\boldsymbol{a_y})} \qquad (44)$$

Example for a Joint Pmf

- For example, samples in picture and video signals typically show strong statistical dependencies
- Below: Histogram of two horizontally adjacent sampels for the picture 'Lena'



Expectation

- **Expectation value** or **expected value**
  of a continuous random variable S

$$\boxed{E\{g(S)\} = \int_{-\infty}^{\infty} g(s)\; f_S(s)\; \mathrm{d}s} \qquad (45)$$

  of a discrete random variable S

$$\boxed{E\{g(S)\} = \sum_{a \in \mathcal{A}} g(a)\; p_S(a)} \qquad (46)$$

- Important expectation values are **mean** μ_S and **variance** σ_S^2

$$\boxed{\mu_S = E\{S\}} \qquad \text{and} \qquad \boxed{\sigma_S^2 = E\big\{(S - \mu_s)^2\big\}} \qquad (47)$$

- Expectation value of a function $g(\boldsymbol{S})$ of a set of N random variables $\boldsymbol{S} = \{S_0, \cdots, S_{N-1}\}$

$$\boxed{E\{g(\boldsymbol{S})\} = \int_{\mathcal{R}^N} g(\boldsymbol{s})\, f_{\boldsymbol{S}}(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s}} \qquad (48)$$

Conditional Expectation

- **Conditional expectation value** of function $g(S)$ given an event $\mathcal{B}$, with $P(\mathcal{B}) > 0$

$$E\{g(S)\,|\,\mathcal{B}\} = \int_{-\infty}^{\infty} g(s)\, f_{S|\mathcal{B}}(s\,|\,\mathcal{B})\, \mathrm{d}s \qquad (49)$$

- Conditional expectation value of function $g(X)$ given a particular value y for another random variable Y

$$E\{g(X)\,|\,y\} = E\{g(X)\,|\,Y = y\} = \int_{-\infty}^{\infty} g(x)\, f_{X|Y}(x,y)\, \mathrm{d}x \qquad (50)$$

- Note: $E\{g(X)\,|\,y\}$ is a deterministic function of y

- Conditional expectation value of function $g(X)$ given a random variable Y,

$$E\{g(X)\,|\,Y\} = \int_{-\infty}^{\infty} g(x)\, f_{X|Y}(x,Y)\, \mathrm{d}x, \qquad (51)$$

is another random variable

Iterative Expectation Rule

- Expectation value $E\{Z\}$ of a random variable $Z = E\{g(X)|Y\}$

$$
\begin{aligned}
E\{E\{g(X)|Y\}\} &= \int_{-\infty}^{\infty} E\{g(X)|y\}\; f_Y(y)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} g(x)\, f_{X|Y}(x,y)\,\mathrm{d}x \right) f_Y(y)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} g(x) \left(\int_{-\infty}^{\infty} f_{X|Y}(x,y)\, f_Y(y)\,\mathrm{d}y \right) \mathrm{d}x \\
&= \int_{-\infty}^{\infty} g(x)\, f_X(x)\,\mathrm{d}x \\
&= E\{g(X)\} \qquad (52)
\end{aligned}
$$

$\Longrightarrow$ $E\{E\{g(X)|Y\}\}$ does not depend on the statistical properties of the random variable Y, but only on those of X

Random Processes

- Series of random experiments at time instants t_n, with $n = 0, 1, 2, \ldots$
- Outcome of experiment: Random variable $S_n = S(t_n)$
- **Discrete-time random process**: Series of random variables $\boldsymbol{S} = \{S_n\}$
- Statistical properties of discrete-time random process $\boldsymbol{S}$: N-th order joint cdf

$$F_{\boldsymbol{S}_k}(\boldsymbol{s}) = P(\boldsymbol{S}_k^{(N)} \leq \boldsymbol{s}) = P(S_k \leq s_0, \cdots, S_{k+N-1} \leq s_{N-1}) \tag{53}$$

- **Continuous random process**

$$f_{\boldsymbol{S}_k}(\boldsymbol{s}) = \frac{\partial^N}{\partial s_0 \, \cdots \, \partial s_{N-1}} F_{\boldsymbol{S}_k}(\boldsymbol{s}) \tag{54}$$

- **Discrete random process**

$$F_{\boldsymbol{S}_k}(\boldsymbol{s}) = \sum_{\boldsymbol{a} \in \mathcal{A}^N} p_{\boldsymbol{S}_k}(\boldsymbol{a}) \tag{55}$$

$\mathcal{A}^N$ product space of the alphabets $\mathcal{A}_n$ and

$$p_{\boldsymbol{S}_k}(\boldsymbol{a}) = P(S_k = a_0, \cdots, S_{k+N-1} = a_{N-1}) \tag{56}$$

Autocovariance and Autocorrelation Matrix

- N-th order **autocovariance matrix**

$$\boxed{\boldsymbol{C}_N(t_k) = E\left\{\left(\boldsymbol{S}_k^{(N)} - \boldsymbol{\mu}_N(t_k)\right)\left(\boldsymbol{S}_k^{(N)} - \boldsymbol{\mu}_N(t_k)\right)^{\mathrm{T}}\right\}} \tag{57}$$

- N-th order **autocorrelation matrix**

$$\boxed{\boldsymbol{R}_N(t_k) = E\left\{\left(\boldsymbol{S}_k^{(N)}\right)\left(\boldsymbol{S}_k^{(N)}\right)^{\mathrm{T}}\right\}} \tag{58}$$

- Note the following relationship

$$\begin{aligned}
\boldsymbol{C}_N(t_k) &= E\left\{\left(\boldsymbol{S}_k^{(N)} - \boldsymbol{\mu}_N(t_k)\right)\left(\boldsymbol{S}_k^{(N)} - \boldsymbol{\mu}_N(t_k)\right)^{\mathrm{T}}\right\} \\
&= E\left\{\left(\boldsymbol{S}_k^{(N)}\right)\left(\boldsymbol{S}_k^{(N)}\right)^{\mathrm{T}}\right\} - E\left\{\boldsymbol{S}_k^{(N)}\right\}\boldsymbol{\mu}_N(t_k)^{\mathrm{T}} \\
&\quad -\boldsymbol{\mu}_N(t_k)\,E\left\{\boldsymbol{S}_k^{(N)}\right\}^{\mathrm{T}} + \boldsymbol{\mu}_N(t_k)\,\boldsymbol{\mu}_N(t_k)^{\mathrm{T}} \\
&= \boldsymbol{R}_N(t_k) - \boldsymbol{\mu}_N(t_k)\,\boldsymbol{\mu}_N(t_k)^{\mathrm{T}}
\end{aligned} \tag{59}$$

Stationary Random Process

- **Stationary random process**:
  Statistical properties are invariant to a shift in time
  $\Longrightarrow$ $F_{\boldsymbol{S}_k}(\boldsymbol{s})$, $f_{\boldsymbol{S}_k}(\boldsymbol{s})$ and $p_{\boldsymbol{S}_k}(\boldsymbol{a})$ are independent of t_k
  and are denoted by $F_{\boldsymbol{S}}(\boldsymbol{s})$, $f_{\boldsymbol{S}}(\boldsymbol{s})$ and $p_{\boldsymbol{S}}(\boldsymbol{a})$, respectively
  $\Longrightarrow$ $\boldsymbol{\mu}_N(t_k)$, $\boldsymbol{C}_N(t_k)$ and $\boldsymbol{R}_N(t_k)$ are independent of t_k
  and are denoted by $\boldsymbol{\mu_N}$, $\boldsymbol{C}_N$ and $\boldsymbol{R}_N$, respectively

- N-th order autocovariance matrix

$$\boldsymbol{C}_N = E\Big\{ (\boldsymbol{S}^{(N)} - \boldsymbol{\mu}_N)(\boldsymbol{S}^{(N)} - \boldsymbol{\mu}_N)^{\mathrm{T}} \Big\} \tag{60}$$

is a symmetric **Toeplitz matrix**

$$\boldsymbol{C}_N = \sigma_S^2 \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{N-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{N-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{N-1} & \rho_{N-2} & \rho_{N-3} & \cdots & 1 \end{pmatrix} \tag{61}$$

with

$$\rho_k = \frac{1}{\sigma_S^2} E\Big\{ (S_\ell - \mu_S)\,(S_{\ell+k} - \mu_S) \Big\} \tag{62}$$

Memoryless and IID Random Processes

- **Memoryless random process**:
  Random process $\boldsymbol{S} = \{S_n\}$ for which the random variables S_n are independent
- **Independent and identical distributed** (iid) random process:
  Stationary and memoryless random process
- N-th order cdf $F_{\boldsymbol{S}}(\boldsymbol{s})$, pdf $f_{\boldsymbol{S}}(\boldsymbol{s})$, and pmf $p_{\boldsymbol{S}}(\boldsymbol{a})$ for iid processes, with $\boldsymbol{s} = (s_0, \cdots, s_{N-1})^T$ and $\boldsymbol{a} = (a_0, \cdots, a_{N-1})^T$

$$F_{\boldsymbol{S}}(\boldsymbol{s}) = \prod_{k=0}^{N-1} F_S(s_k) \tag{63}$$

$$f_{\boldsymbol{S}}(\boldsymbol{s}) = \prod_{k=0}^{N-1} f_S(s_k) \tag{64}$$

$$p_{\boldsymbol{S}}(\boldsymbol{a}) = \prod_{k=0}^{N-1} p_S(a_k) \tag{65}$$

  $F_S(s)$, $f_S(s)$, and $p_S(a)$ are the marginal cdf, pdf, and pmf, respectively

Markov Processes

- **Markov process**: Future outcomes do not depend on past outcomes, but only on the present outcome,

$$P(S_n \leq s_n \mid S_{n-1} = s_{n-1}, \cdots) = P(S_n \leq s_n \mid S_{n-1} = s_{n-1}) \tag{66}$$

- Discrete Markov processes

$$p_{S_n}(a_n \mid a_{n-1}, \cdots) = p_{S_n}(a_n \mid a_{n-1}) \tag{67}$$

- Example for a discrete Markov process

a	a_0	a_1	a_2
$p(a\|a_0)$	0.90	0.05	0.05
$p(a\|a_1)$	0.15	0.80	0.05
$p(a\|a_2)$	0.25	0.15	0.60
$p(a)$			

Continuous Markov Processes

- Continuous Markov processes

$$f_{S_n}(s_n \mid s_{n-1}, \cdots) = f_{S_n}(s_n \mid s_{n-1}) \tag{68}$$

- Construction of continuous stationary Markov process $\boldsymbol{S} = \{S_n\}$ with mean μ_S, given a zero-mean iid process $\boldsymbol{Z} = \{Z_n\}$

$$S_n = Z_n + \rho\,(S_{n-1} - \mu_S) + \mu_S, \quad \text{with} \quad |\rho| < 1 \tag{69}$$

$\Longrightarrow$ Variance σ_S^2 of stationary Markov process $\boldsymbol{S}$

$$\sigma_S^2 = E\big\{(S_n - \mu_S)^2\big\} = E\big\{(Z_n + \rho\,(S_{n-1} - \mu_S)\,)^2\big\} = \frac{\sigma_Z^2}{1 - \rho^2} \tag{70}$$

$\Longrightarrow$ Autocovariance function of stationary Markov process $\boldsymbol{S}$

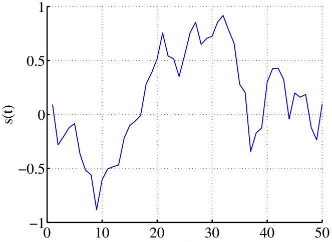
$$\phi_{k,\ell} = \phi_{|k-\ell|} = E\{(S_k - \mu_S)(S_\ell - \mu_S)\} = \sigma_S^2\,\rho^{|k-\ell|} \tag{71}$$

Gaussian Processes

- **Gaussian process**: Continuous process $\boldsymbol{S} = \{S_n\}$ with the property that all finite collections of random variables S_n represent Gaussian random vectors
- N-th order pdf of stationary Gaussian process with N-th order autocorrelation matrix $\boldsymbol{C}_N$ and mean μ_S

$$f_{\boldsymbol{S}}(\boldsymbol{s}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{C}_N|}} e^{-\frac{1}{2}(\boldsymbol{s}-\boldsymbol{\mu}_S)^{\mathrm{T}} \boldsymbol{C}_N^{-1}(\boldsymbol{s}-\boldsymbol{\mu}_S)} \quad \text{with} \quad \boldsymbol{\mu}_S = \begin{bmatrix} \mu_s \\ \vdots \\ \mu_S \end{bmatrix} \qquad (72)$$

- **Stationary Gauss-Markov process**:
  Stationary process that is a Gaussian process and a Markov process
  - IID process $\boldsymbol{Z} = \{Z_n\}$ in (69) has a Gaussian pdf
  - Statistical properties are completely determined by
    - mean μ_S
    - variance σ_S^2
    - correlation factor ρ



Chapter Summary

Random variables

- Discrete and continuous random variables
- Cumulative distribution function (cdf)
- Probability density function (pdf)
- Probability mass function (pmf)
- Joint and conditional cdfs, pdfs, pmfs
- Expectation values and conditional expectation values

Random processes

- Stationary processes
- Memoryless processes
- IID processes
- Markov processes
- Gaussian processes
- Gauss-Markov processes

Exercise 1

Given is a stationary discrete Markov process with the alphabet $\mathcal{A} = \{a_0, a_1, a_2\}$ and the conditional pmfs listed in the table below

a	a_0	a_1	a_2
$p(a\|a_0)$	0.90	0.05	0.05
$p(a\|a_1)$	0.15	0.80	0.05
$p(a\|a_2)$	0.25	0.15	0.60
$p(a)$			

Determine the marginal pmf $p(a)$.

Exercise 2

Investigate the relationship between independence and correlation.

(a) Two random variables X and Y are said to be *correlated* if and only if their covariance C_{XY} is not equal to 0.
Can two independent random variables X and Y be correlated?

(b) Let X be a continuous random variable with a variance $\sigma_X^2 > 0$ and a pdf $f_X(x)$. The pdf shall be non-zero for all real numbers, $f_X(x) > 0$, $\forall x \in \mathbb{R}$.
Furthermore, the pdf $f_X(x)$ shall be symmetric around zero, $f_X(x) = f_X(-x)$, $\forall x \in \mathbb{R}$. Let Y be a random variable given by $Y = a\,X^2 + b\,X + c$ with $a, b, c \in \mathbb{R}$.
For which values of a, b, and c are X and Y uncorrelated?
For which values of a, b, and c are X and Y independent?

(c) Which of the following statements for two random variables X and Y are true?

- If X and Y are uncorrelated, they are also independent.
- If X and Y are independent, $E\{XY\} = 0$.
- If X and Y are correlated, they are also dependent.