

Ein endlicher Ergebnisraum ist eine nichtleere Menge $\Omega = \{\xi_1, \xi_2, \dots, \xi_N\}$. Die Elemente $\xi_n \in \Omega$ heißen **Ergebnisse**. Jede Teilmenge $A \subset \Omega$ wird als **Ereignis**, jede einelementige Teilmenge $\{\xi_n\} \subset \Omega$ wird als **Elementarereignis** bezeichnet. Der Ergebnisraum Ω und die leere Menge \emptyset sind stets Ereignisse, Ω heißt das **sichere**, \emptyset das **unmögliche Ereignis**.

Aus einem endlichen Ergebnisraum Ω lassen sich

- 1 unmögliches Ereignis (die leere Menge \emptyset),
- $\binom{N}{1}$ einelementige Ereignisse,
- $\binom{N}{2}$ zweielementige Ereignisse,
- \vdots
- $\binom{N}{N-1}$ $(N-1)$ -elementige Ereignisse und
- 1 N -elementiges Ereignis (der gesamte Ergebnisraum Ω)

konstruieren.

Die Menge aller Ereignisse ist hier also die Potenzmenge $\mathfrak{P}(\Omega)$ des Ergebnisraums Ω . Die Anzahl der Ereignisse ist dann

$$|\mathfrak{P}(\Omega)| = \sum_{n=0}^N \binom{N}{n} = 2^N.$$

Sind A und B Ereignisse und gilt $AB = \emptyset$, so heißen A und B **disjunkt** oder **unvereinbar**.

De Morgan:
 $A \cup B = \overline{\overline{A} \cap \overline{B}}$

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

LAPLACE

Tritt bei N unabhängigen Wiederholungen des durch $\Omega, \mathfrak{P}(\Omega)$ beschriebenen Zufallsexperiments $A \in \mathfrak{P}(\Omega)$ genau $h_N(A)$ -mal ein, heißen $h_N(A)$ die **absolute Häufigkeit** und

$$H_N(A) = \frac{h_N(A)}{N} \quad (2.2-1)$$

die **relative Häufigkeit** von A in N Versuchen.

$$\forall A \in \mathfrak{P}(\Omega): H_N(\overline{A}) = 1 - H_N(A)$$

$$\forall A, B \in \mathfrak{P}(\Omega):$$

$$H_N(A \cup B) = H_N(A) + H_N(B) - H_N(AB)$$

Treten alle Elementarereignisse $\{\xi_n\}; n = 1, 2, \dots, N$; eines endlichen Ergebnisraums $\Omega = \{\xi_1, \xi_2, \dots, \xi_N\}$ gleich häufig auf, ist das zugehörige Zufallsexperiment ein **Laplacesches Zufallsexperiment**.

In einem Laplaceschen Zufallsexperiment ist

$$P(A) = \frac{\text{Anzahl der Elementarereignisse } \{\xi_n\} \subset A}{\text{Gesamtzahl der Elementarereignisse}}$$

die **Wahrscheinlichkeit des Ereignisses A** .

Beispiel: Eine Urne enthält N (bis auf die Farbe) gleiche Kugeln, von denen M rot und $N - M$ weiß sind. Aus der Urne werden zufällig n Kugeln gezogen.

Frage: Wie groß ist die Wahrscheinlichkeit $P(A)$ dafür, daß unter den n gezogenen Kugeln k rote und $n - k$ weiße sind?

$$P_k = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \quad (\text{hypergeometrische Verteilung})$$

KOLMOGOROFF

Ein nichtleeres System \mathfrak{B} von Teilmengen eines Ergebnisraums Ω heißt **σ -Algebra** (über Ω), wenn gilt

$$(i) \quad A \in \mathfrak{B} \Rightarrow \overline{A} \in \mathfrak{B}, \quad (2.3-1)$$

$$(ii) \quad A_n \in \mathfrak{B}; n = 1, 2, \dots; \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{B}. \quad (2.3-2)$$

Ω und \emptyset liegen stets in \mathfrak{B} .

Somit läßt sich aus jedem Teilsystem $M \subset \mathfrak{P}(\Omega)$ eindeutig eine σ -Algebra $\mathfrak{B} = \mathfrak{B}(M)$, nämlich die **von M erzeugte σ -Algebra**, konstruieren, für die gilt:

$$\mathfrak{B}(M) \supset M$$

Ist $\mathfrak{B}' \supset M$ eine σ -Algebra

$$\Rightarrow \mathfrak{B}' \supset \mathfrak{B}(M),$$

d.h. $\mathfrak{B} = \mathfrak{B}(M)$ ist die **kleinste σ -Algebra**, die M enthält.

Ein wichtiges Beispiel einer σ -Algebra über dem überabzählbaren Ergebnisraum $\Omega = \mathbb{R}$ ist die aus der Menge der halboffenen Intervalle $O = \{(a, b] \subset \mathbb{R}\}$ erzeugte σ -Algebra $\mathfrak{B}(O)$. Sie wird auch **Borelsche σ -Algebra** genannt.

Ein höchstens abzählbares System $\{A_n \in \mathfrak{B} : A_k A_n = \emptyset, k \neq n\}$ heißt **vollständige Ereignisdisjunktion** (im engeren Sinne), wenn gilt $\bigcup_{n=1}^{\infty} A_n = \Omega$.

Kolmogoroffsche Axiome:

Gegeben seien ein Ergebnisraum Ω und eine geeignete σ -Algebra \mathfrak{B} über Ω . Die Elemente von \mathfrak{B} sind also die Ereignisse des Zufallsexperiments. Eine Funktion P , die jedem Ereignis $A \in \mathfrak{B}$ eine reelle Zahl zuordnet, erfülle

$$\text{Axiom 1:} \quad \forall A \in \mathfrak{B}: P(A) \geq 0 \quad (2.3-4)$$

$$\text{Axiom 2:} \quad P(\Omega) = 1 \quad (2.3-5)$$

Axiom 3: Für paarweise disjunkte Ereignisse $A_n \in \mathfrak{B}; n = 1, 2, \dots$; gilt

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \quad (2.3-6)$$

$P(A)$ heißt **Wahrscheinlichkeit des Ereignisses A** .

Damit kann für jedes Zufallsexperiment mit der Ergebnismenge Ω , einer geeigneten σ -Algebra \mathfrak{B} über Ω und der Wahrscheinlichkeit aus Definition 2.3-3 ein **Wahrscheinlichkeitsraum** $(\Omega, \mathfrak{B}, P)$ zur Modellierung des Zufallsexperimentes gefunden werden.

$$(i) \quad P(\emptyset) = 0$$

$$(ii) \quad P(\overline{A}) = 1 - P(A)$$

$$(iii) \quad P(A \cup B) = P(A) + P(B) - P(AB)$$

Für eine vollständige Ereignisdisjunktion $\{A_n \in \mathfrak{B} : A_k A_n = \emptyset, k \neq n\}$ folgt

$$P\left(\sum_{n=1}^{\infty} A_n\right) = P(\Omega) = 1. \quad (2.3-14)$$

$$A, B \in \mathfrak{B} \text{ mit } B \subset A \Rightarrow P(B) \leq P(A)$$

Kombinatorik

	Permutation	Kombination	Variation
Auswahl von K aus N Elementen	Anordnungen in bestimmter Reihenfolge	Auswahl ohne Beachtung der Reihenfolge	Auswahl mit Beachtung der Reihenfolge
ohne Wiederholung	$N!, n=k$	$\binom{N}{K}$	$K! \binom{N}{K}$
mit Wiederholung	$\frac{N!}{K!}$ bzw. $\frac{N!}{A! B! \dots}$	$\binom{N+K-1}{K}$	N^K

Bedingte Wahrscheinlichkeiten

$A, B \in \mathfrak{B}$ und $P(B) > 0$. Dann heißt

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A_n)P(B|A_n)}{P(B)} \quad (3.1-1)$$

bedingte Wahrscheinlichkeit von A unter der Bedingung B .

Gegenereignis: $P(\bar{A}|B) = 1 - P(A|B)$

Im allgemeinen ist $P(A|B) \neq P(B|A)$. Es gilt die Beziehung

$$P(A|B)P(B) = P(B|A)P(A). \quad (3.1-2)$$

Die bedingten Wahrscheinlichkeiten $P(A|B)$ erfüllen für festes $B \in \mathfrak{B}$ die Kolmogoroffschen Axiome. Bedingte Wahrscheinlichkeiten über einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{B}, P)$ sind dort also auch Wahrscheinlichkeiten [Kol33].

Multiplikationsregel für Wahrscheinlichkeiten:

Anm: Die Multiplikationsregel kann man sich aus der Definition (s.o.) bedingten Wahrscheinlichkeit durch Umformen herleiten \rightarrow nichts neues!

$$P(AB) = P(B)P(A|B) \quad P(AB) = P(A)P(B|A).$$

Die wiederholte Anwendung der Multiplikationsregel auf den Durchschnitt N zufälliger Ereignisse liefert:

$$\begin{aligned} P\left(\bigcap_{n=1}^N A_n\right) &= P(A_1)P\left(\bigcap_{n=2}^N A_n|A_1\right) \\ &= P(A_1)P(A_2|A_1)P\left(\bigcap_{n=3}^N A_n|A_1A_2\right) \\ &\vdots \\ &= P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots \\ &\quad \cdots P\left(A_N|\bigcap_{n=1}^{N-1} A_n\right) \end{aligned} \quad (3.1-4)$$

Die Ereignisse A_n ($1 \leq n \leq N$) seien eine vollständige Ereignisdisjunktion und es gelte $P(A_n) > 0 \forall n$. Dann folgt für jedes $B \in \mathfrak{B}$ die **Formel von der totalen Wahrscheinlichkeit**

$$P(B) = \sum_{n=1}^N P(B|A_n)P(A_n) \quad (3.1-5)$$

und, wenn $P(B) > 0$ ist, die **Formel von Bayes**

$$P(A_n|B) = \frac{P(B|A_n)P(A_n)}{\sum_{n=1}^N P(B|A_n)P(A_n)}. \quad (3.1-6)$$

Die Wahrscheinlichkeiten $P(A_n|B)$ werden **a posteriori Wahrscheinlichkeiten** genannt, da sie die Wahrscheinlichkeiten der A_n nach Eintreten von B angeben. Im Gegensatz dazu sind die $P(A_n)$ die **a priori Wahrscheinlichkeiten** für das Auftreten der A_n .

Gilt für $A, B \in \mathfrak{B}$

$$P(A|B) = P(A)$$

heißt A **unabhängig von B** .

D.h. wenn A von B unabhängig ist, ist auch B von A unabhängig. Man sagt A und B seien voneinander unabhängig.

Mit Gleichung (3.2-2) wird oft die Unabhängigkeit von A und B definiert. Diese Definition ist äquivalent zu (3.2-1), wenn $P(B) > 0$ ist.

Weil $AB \subset B$ und $AB \subset A$, gilt (3.2-2) auch, wenn $P(A) = 0$ oder $P(B) = 0$ ist.

Die Ereignisse $A_n \in \mathfrak{B}$ ($n = 1, 2, \dots, N$) heißen **vollständig unabhängig**, wenn für jedes $K \in \{2, 3, \dots, N\}$ und beliebige natürliche Zahlen $1 \leq i_1 < i_2 < \dots < i_K \leq N$

$$P\left(\bigcap_{k=1}^K A_{i_k}\right) = \prod_{k=1}^K P(A_{i_k}) \quad \text{gilt.} \quad (3.2-4)$$

Disjunkte Ereignisse sind in höchstem Maße abhängig!

Zufallsvariablen

Eine Funktion $X = X(\xi) : \Omega \rightarrow \mathbb{R}$.

die jedem Ergebnis $\xi \in \Omega$ eine reelle Zahl zuordnet, heißt **Zufallsvariable**, wenn das Urbild eines jeden Intervalls $(-\infty, a] \subset \mathbb{R}$ ein Ereignis aus \mathfrak{B} ist: $X^{-1}((-\infty, a]) \in \mathfrak{B}, \forall a \in \mathbb{R}$

Die Funktion $F(x) := P(X \leq x)$ heißt Verteilungsfunktion der Zufallsvariablen X (dabei ist x reell).

Aus Bild 4.1-1 lassen sich folgende **Eigenschaften** von $F(x)$, die allgemein für **Verteilungsfunktionen** gelten, ablesen:

- $F : \mathbb{R} \rightarrow [0, 1]$
- $\lim_{x \rightarrow -\infty} F(x) = 0$
- $F(x)$ ist monoton nichtfallend:
 $x_1 \leq x_2 \Rightarrow F(x_1) \leq F(x_2)$
- $F(x)$ ist rechtsseitig stetig:
 $\lim_{x \rightarrow \infty} F(x) = 1$ $F(x+0) = \lim_{h \rightarrow 0} F(x+h) = F(x) \forall x \in \mathbb{R}$

Die Zufallsvariable X heißt **diskret**, wenn ihr Wertebereich eine endliche oder höchstens abzählbar unendliche Menge ist.

Eine Zufallsvariable X heißt **stetig**, wenn eine integrierbare Funktion

$$f(x) \geq 0 \forall x \in \mathbb{R} \quad (4.1-5)$$

existiert, so daß sich die Verteilungsfunktion $F(x) \forall x \in \mathbb{R}$ in der Form

$$F(x) = \int_{-\infty}^x f(u) du \quad (4.1-6)$$

schreiben läßt. $f(x)$ heißt **Dichte** von X .

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

Die Dichte

$$f(x) = \begin{cases} \frac{1}{2\pi} & \text{für } x \in [-\pi, \pi] \\ 0 & \text{sonst} \end{cases} \quad (4.1-7)$$

beschreibt eine Gleichverteilung über dem Intervall $[-\pi, \pi]$, siehe Bild 4.1-3.

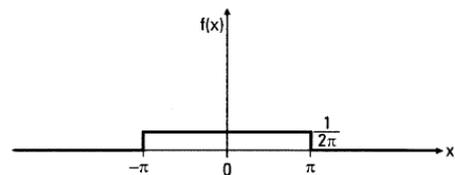


Bild 4.1-3: Gleichverteilungsdichte über $[-\pi, \pi]$

(ii) Durch

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (4.1-8)$$

ist die Dichte der **Standardnormalverteilung** gegeben (Bild 4.1-4).

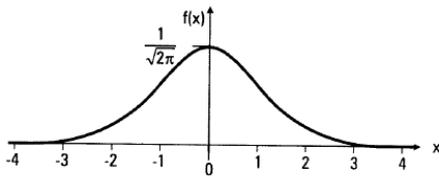


Bild 4.1-4: Dichte der Standardnormalverteilung $\mathcal{N}(0; 1)$

Die Verteilungsfunktion $F(x)$ ist gemäß (4.1-6) eine Stammfunktion der Dichte $f(x)$. Nach dem Hauptsatz der Differential- und Integralrechnung ergibt sich daraus

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x) dx. \quad (4.1-9)$$

Die Wahrscheinlichkeit dafür, daß X einen Wert aus dem Intervall $(x_1, x_2]$ annimmt, ist daher durch die Fläche unter der Dichte über diesem Intervall gegeben. Insbesondere folgt bei stetigen Zufallsvariablen $\forall x$:

$$\begin{aligned} P(X = x) &= \lim_{\Delta x \rightarrow 0} P(x < X \leq x + \Delta x) \\ &= \lim_{\Delta x \rightarrow 0} \int_x^{x+\Delta x} f(x) dx \\ &= 0 \end{aligned} \quad (4.1-10)$$

Folgerung:

Für eine stetige Zufallsvariable X verschwinden alle Wahrscheinlichkeiten der Form $P(X = x)$, obwohl die Ereignisse $\{\xi; X = x\}$ **nicht** mit dem unmöglichen Ereignis zusammenfallen müssen. Bild 4.1-5 zeigt die geometrische Deutung der Wahrscheinlichkeiten und des durch

$$P(x < X \leq x + dx) = dF(x) = f(x) dx \quad (4.1-11)$$

definierten **Wahrscheinlichkeitselements**

Mathematisch (wegen der fehlenden Stetigkeit/Integrierbarkeit der Dichte) **nicht** einwandfrei, aber für Anwendungen oft nützlich, schreibt man für die „Dichte“ einer diskreten Zufallsvariablen, deren Einzelwahrscheinlichkeiten durch (4.1-4) gegeben sind, auch

$$f(x) = \sum_{n=1}^{\infty} P(X = x_n) \delta(x - x_n) \quad (4.1-12)$$

$$\frac{dF(x)}{dx} = f(x)$$

Funktionen von Zufallsvariablen

Beispiel:

$$Y = aX + b, \quad a > 0$$

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right) \quad F_Y(y) = F_X\left(\frac{y-b}{a}\right)$$

Mit $Y = \sigma X + \mu$, $\sigma > 0$, ergibt sich aus der standardnormalverteilten Zufallsvariablen X die allgemeine $\mathcal{N}(\mu; \sigma^2)$ -verteilte Zufallsvariable Y mit der Dichte

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right). \quad (4.2-4)$$

Allgemein gilt: Wenn x_1, x_2, \dots, x_N die reellen Wurzeln der Gleichung $y = g(x)$ sind, kann die Dichte der Zufallsvariablen $Y = g(X)$ in der Form

$$f_Y(y) = \sum_{n=1}^N \frac{f_X(x_n)}{|g'(x_n)|} \quad (4.2-9)$$

geschrieben werden, worin dann natürlich die x_n ; $n = 1, 2, \dots, N$; Funktionen von y sind. Dabei ist selbstverständlich der Definitionsbereich der Funktion $g(x)$ zu beachten.

Kennwerte von Zufallsvariablen

Im Falle seiner Existenz heißt

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \quad (5.1-1)$$

Erwartungswert oder auch **Mittelwert** der Zufallsvariablen X .

Für diskrete ZV berechnet man

$$E(X) = \sum_{n=1}^{\infty} x_n p_n$$

Cauchy-Dichte:

$$f(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x - \mu)^2} \quad (\lambda > 0) \quad (\text{hat keinen Erwartungswert!})$$

X sei eine $\mathcal{N}(\mu; \sigma^2)$ -verteilte Zufallsvariable, d.h. ihre Dichte ist nach (4.2-4)

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \quad E(X) = \mu$$

Im Falle seiner Existenz heißt der Erwartungswert

$$E(X^k) = \int_{-\infty}^{\infty} x^k f(x) dx \quad (5.1-3)$$

das **k-te Moment** der Zufallsvariablen X . Der Erwartungswert

$$E([X - E(X)]^k) = \int_{-\infty}^{\infty} [x - E(X)]^k f(x) dx \quad (5.1-4)$$

ist das **k-te zentrale Moment** der Zufallsvariablen X .

Von besonderer Bedeutung ist das durch

$$D^2(X) = E([X - E(X)]^2) = \text{var}(X) = E(X^2) - (E(X))^2 \quad (5.1-5)$$

gegebene zweite zentrale Moment. Es heißt **Varianz** (oder auch **Dispersion**) der Zufallsvariablen X .

$$D(X) = \sqrt{E([X - E(X)]^2)}$$

ist die **Standardabweichung** der Zufallsvariablen X .

Varianz einer $\mathcal{N}(\mu; \sigma^2)$ -verteilten Zufallsvariablen: $= \sigma^2$

Eine $\mathcal{N}(\mu; \sigma^2)$ -verteilte Zufallsvariable hat den Erwartungswert μ und die Varianz σ^2 . Ihr k -tes zentrales Moment ist allgemein ([Pro95], S.41):

$$E[(X - \mu)^k] = \begin{cases} 1 \cdot 3 \cdots (k-1) \sigma^k & \text{falls } k \text{ gerade} \\ 0 & \text{falls } k \text{ ungerade} \end{cases} \quad (5.1-6)$$

D.h.: Sind von einer **Normalverteilung** Mittelwert und Varianz bekannt, kennt man sämtliche Momente dieser Verteilung.

Ist $Y = g(X)$ eine Funktion der Zufallsvariablen X (vergleiche Abschnitt 4.2), folgt für deren Erwartungswert

$$E(Y) = E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx. \quad (5.1-7)$$

Für diskrete Zufallsvariablen ergibt sich daraus mit (4.1-12) der Spezialfall

$$E(Y) = \sum_{n=1}^{\infty} P(X = x_n) \int_{-\infty}^{\infty} g(x) \delta(x - x_n) dx$$

$$= \sum_{n=1}^{\infty} g(x_n) P(X = x_n).$$

Eine Zufallsvariable mit dem Erwartungswert $m = E(X)$ und der Varianz σ^2 wird durch

$$Y = \frac{X - m}{\sigma} \quad (5.1-8)$$

in eine standardisierte Zufallsvariable Y , die den Erwartungswert 0 und die Varianz 1 besitzt, transformiert.

Ein Wert, für den die Dichtefunktion $f(x)$ ein lokales Maximum annimmt, heißt **Modalwert** der stetigen Zufallsvariablen X .

Einen Wert x_p , der den Ungleichungen

$$P(X < x_p) \leq p, P(X > x_p) \leq 1 - p \quad (0 < p < 1) \quad (5.1-9)$$

genügt, nennen wir **p-tes Quantil**.

Ist X eine stetige Zufallsvariable, ist ein p -tes Quantil ein Wert x_p , für den

$$F(x_p) = p$$

gilt. Es kann vorkommen, daß x_p nicht eindeutig ist. Dann existieren mehrere Werte x_p , mit denen die Ungleichungen (5.1-9) erfüllt sind.

Ein Quantil der Ordnung $p = \frac{1}{2}$ heißt **Median** der Zufallsvariablen X .

Die Lage von Erwartungswert, Modalwert und Median zueinander zeigt Bild 5.1-1 für ein Beispiel. Diese drei Werte fallen für normalverteilte Zufallsvariablen zusammen.

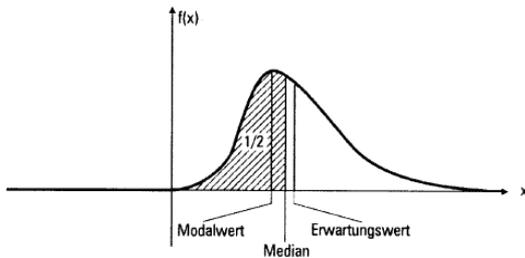


Bild 5.1-1: Lage von Erwartungswert, Modalwert und Median

Der Erwartungswert von $Y = |X - E(X)|^k$ heißt **absolutes zentrales Moment k-ter Ordnung**.

Charakteristische Funktion

Der Erwartungswert

$$\varphi(s) := E(e^{jsX}), \quad s \in \mathbb{R} \quad (5.2-1)$$

heißt **charakteristische Funktion der Zufallsvariablen X** .

$$\varphi(s) = \int_{-\infty}^{\infty} e^{jsx} f(x) dx. \quad f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(s) e^{-jsx} ds.$$

Für diskrete Zufallsvariablen folgt

$$\varphi(s) = \sum_{n=1}^{\infty} e^{jsx_n} P(X = x_n)$$

X sei $\mathcal{N}(0; 1)$ -verteilt.

$$\varphi(s) = \exp\left(-\frac{s^2}{2}\right)$$

Für die $\mathcal{N}(\mu; \sigma^2)$ -verteilte Zufallsvariable

$$Y = \sigma X + \mu$$

folgt daraus:

$$\varphi_Y(s) = e^{js\mu} \exp\left(-\frac{(\sigma s)^2}{2}\right)$$

Existiert das k -te Moment der Zufallsvariablen X , ist es durch

$$E(X^k) = \frac{\varphi^{(k)}(0)}{j^k}, \quad (k = 1, 2, \dots) \quad (5.2-6)$$

gegeben.

$$D^2(X) = \frac{\varphi''(0)}{j^2} - \left(\frac{\varphi'(0)}{j}\right)^2$$

$$= -\varphi''(0) + [\varphi'(0)]^2.$$

Ist X eine Zufallsvariable, die nur nichtnegative ganzzahlige Werte n ($n = 0, 1, 2, \dots$) annimmt, heißt

$$\psi(z) = E\{z^X\} = \sum_{n=0}^{\infty} z^n P(X = n) \quad (5.2-7)$$

$$(z \in \mathbb{C}, |z| \leq 1)$$

die **erzeugende Funktion von X** .

Wegen $\psi(1) = \sum_{n=0}^{\infty} P(X = n) = 1$ konvergiert $\psi(z)$ für $|z| \leq 1$ absolut und gleichmäßig. $\psi(z)$ ist stetig. Die Funktion $\psi(z)$ bestimmt in eindeutiger Weise die Verteilung von X , d.h. die Werte $P(X = n)$, da sie sich nur auf genau eine Weise als Potenzreihe darstellen läßt.

Existiert das k -te Moment von X , läßt es sich mit Hilfe der Ableitungen der Funktion $\psi(z)$ bis zur Ordnung k bestimmen.

Spezielle Wahrscheinlichkeitsverteilungen

Zweipunktverteilung

Die Zufallsvariable X besitzt eine Zweipunktverteilung, wenn sie genau zwei Werte x_1 und x_2 ($x_2 < x_1$) mit positiver Wahrscheinlichkeit annehmen kann. Eine andere Bezeichnung für die Zweipunktverteilung ist Binärverteilung. Insbesondere ergibt sich für $x_1 = 1, x_2 = 0$ die Null-Eins-Verteilung. Es wird ein Versuch ausgeführt, bei dem entweder das Ereignis $A = \{X = x_1\}$ oder $\bar{A} = \{X = x_2\}$ eintritt.

$$P(A) = P(X = x_1) = p$$

$$P(\bar{A}) = P(X = x_2) = 1 - p$$

$$F(x) = \begin{cases} 0 & \text{für } x < x_2 \\ 1 - p & \text{für } x_2 \leq x < x_1 \\ 1 & \text{für } x \geq x_1 \end{cases}$$

$$E(X) = x_1 p + x_2 (1 - p)$$

$$D^2(X) = (x_1 - x_2)^2 p (1 - p)$$

$$\varphi(s) = e^{jsx_2} + p [e^{jsx_1} - e^{jsx_2}]$$

Binomialverteilung

Grundlage der Binomialverteilung ist das **Bernoullische Versuchsschema**:

Es werden N voneinander unabhängige, sonst jedoch identische Versuche mit jeweils zwei Ausgängen (A, \bar{A}) durchgeführt mit

$$P(A) = p \quad (0 < p < 1).$$

Wir betrachten die diskrete Zufallsvariable X_N , die die Anzahl der Versuche (von insgesamt N), in denen A eintritt, zählt. X_N kann also die Werte $n = 0, 1, \dots, N$ annehmen.

Für $N = 1$ ist X_1 null-eins-verteilt.

$$P(X_N = K) = \binom{N}{K} p^K (1 - p)^{N-K}$$

$$F(x) = \sum_{K \leq x} \binom{N}{K} p^K (1 - p)^{N-K}$$

$$E(X_N) = Np$$

$$D^2(X_N) = Np(1 - p)$$

$$\varphi(s) = [1 + p(e^{js} - 1)]^N$$

N und p sind die Parameter der Binomialverteilung.

Die absolute Häufigkeit $X_N = h_N(A)$ eines Ereignisses A mit $p = P(A)$ in N unabhängigen Wiederholungen eines Zufallsexperiments unterliegt einer Binomialverteilung mit den Parametern N und p :

Für die relative Häufigkeit $H_N(A) = \frac{1}{N} h_N(A)$ gilt also, wenn sie als Zufallsvariable interpretiert wird,

$$E(H_N(A)) = p = P(A)$$

$$D^2(H_N(A)) = \frac{p(1 - p)}{N}$$

Polynomialverteilung

Ein Versuch werde N -mal wiederholt. Als Ergebnis jedes Versuchs kann eines der paarweise einander ausschließenden Ereignisse A_j ($j = 1, 2, \dots, r+1$) mit der Wahrscheinlichkeit $p_j = P(A_j)$, wobei $p_1 + p_2 + \dots + p_{r+1} = 1$ gilt, eintreten.

Wir betrachten die vektorwertige (vergleiche Kapitel 7), diskrete Zufallsvariable $(X_1, X_2, \dots, X_{r+1})^T$, wobei $X_j = K_j$ bedeutet, daß A_j genau K_j -mal eintritt.

$$P(X_1 = K_1, X_2 = K_2, \dots, X_{r+1} = K_{r+1}) = \frac{N!}{K_1! \cdot K_2! \cdot \dots \cdot K_{r+1}!} \cdot p_1^{K_1} p_2^{K_2} \dots p_{r+1}^{K_{r+1}},$$

wobei $K_1 + K_2 + \dots + K_{r+1} = N$ ist.

$$X_1 + X_2 + \dots + X_{r+1} = N,$$

$$X_{r+1} = N - X_r - X_{r-1} - \dots - X_1$$

$$P(X_1 = K_1, X_2 = K_2, \dots, X_r = K_r) = \frac{N!}{K_1! \cdot \dots \cdot K_r! (N - K)!} \cdot p_1^{K_1} p_2^{K_2} \dots p_r^{K_r} q^{N-K}$$

mit $K = K_1 + K_2 + \dots + K_r, q = 1 - p_1 - p_2 - \dots - p_r$

Poissonverteilung

Die diskrete Zufallsvariable X_N habe die durch (6.2-1)

$$P(X_N = K) = \binom{N}{K} p^K (1 - p)^{N-K}$$

gegebene Binomialverteilung. Gilt mit der Konstanten $\lambda > 0$ für $N = 1, 2, 3, \dots$ die Beziehung

$$p = \frac{\lambda}{N}, \quad (6.4-1)$$

dann ist

$$\lim_{N \rightarrow \infty} P(X_N = K) = \frac{\lambda^K}{K!} e^{-\lambda}. \quad (6.4-2)$$

Definition 6.4-1

Eine Zufallsvariable X , die die Werte $K = 0, 1, 2, \dots$ mit den Wahrscheinlichkeiten (6.4-2)

$$P(X = K) = \frac{\lambda^K}{K!} e^{-\lambda} \quad (\lambda > 0)$$

annimmt, ist mit dem Parameter λ **poissonverteilt**.

$$E(X) = \lambda, \quad D^2(X) = \lambda, \quad \varphi(s) = \exp(\lambda(e^{js} - 1))$$

Mit wachsendem N (bei konstantem λ) werden die graphischen Darstellungen von Binomial- und Poissonverteilung einander immer ähnlicher.

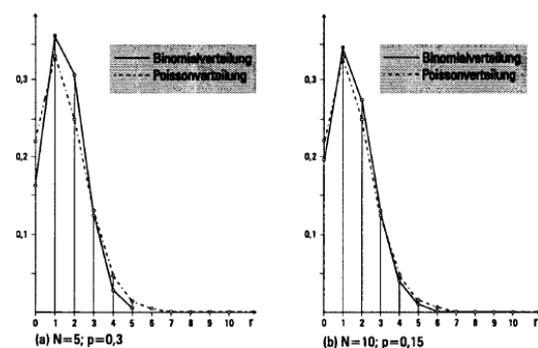


Bild 6.4-1: Approximation der Binomialverteilung durch die Poissonverteilung (nach [Fis70])

Die Einzelwahrscheinlichkeiten $p_K = P(X = K)$ einer poissonverteilten Zufallsvariablen X genügen der Rekursion

$$p_0 = e^{-\lambda}$$

$$p_K = \frac{\lambda}{K} p_{K-1} \quad K = 1, 2, \dots$$

Hypergeometrische Verteilung

In einer Urne liegen M schwarze und $N - M$ weiße Kugeln. Der Urne werden zufällig n Kugeln entnommen. Die diskrete Zufallsvariable X beschreibe die Anzahl der gezogenen schwarzen Kugeln. Es folgt (vergleiche Abschnitt 2.2)

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}; \quad k = 0, 1, 2, \dots, n. \quad (6.5-1)$$

Eine Zufallsvariable X mit Einzelwahrscheinlichkeiten nach (6.5-1) unterliegt einer hypergeometrischen Verteilung.

$$E(X) = n \frac{M}{N}, \quad D^2(X) = \frac{nM(N-M)(N-n)}{N^2(N-1)}$$

Für eine mit den Parametern $n, M, N - M$ hypergeometrisch verteilte Zufallsvariable X gilt für $p = \frac{M}{N}$ und festes n :

$$\lim_{N \rightarrow \infty} P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

für $k = 0, 1, \dots, n$

Das heißt, daß für große N (Faustregel: $\frac{n}{N} < 0,05$) die hypergeometrische Verteilung näherungsweise der Binomialverteilung mit den Parametern $p = \frac{M}{N}$ und n entspricht.

$$\lim_{N \rightarrow \infty} P(X = k) = \binom{n}{k} p^k q^{n-k} \quad \text{mit } p = \frac{M}{N}, \quad q = 1 - \frac{M}{N}.$$

Die (stetige) Gleichverteilung

Eine stetige Zufallsvariable X mit der Dichte

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x < b \\ 0 & \text{sonst} \end{cases}$$

heißt **gleichverteilt** über dem Intervall $[a, b)$.

$$F(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } a \leq x < b \\ 1 & \text{für } x \geq b \end{cases}$$

$$E(X) = \frac{1}{2}(a+b)$$

$$E(X^2) = \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3}$$

$$D^2(X) = E(X^2) - E^2(X) = \frac{(b-a)^2}{12}$$

$$\varphi(s) = \frac{e^{jsb} - e^{jsa}}{js(b-a)}$$

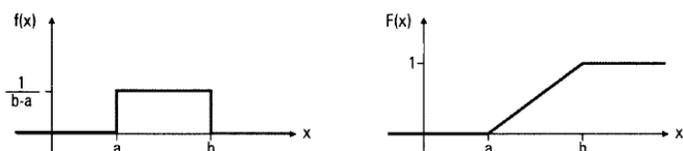


Bild 6.6-2: Dichte und Verteilungsfunktion einer über $[a, b)$ gleichverteilten Zufallsvariablen

Exponentialverteilung

Eine Zufallsvariable X mit der Dichte

$$f(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ \lambda e^{-\lambda x} & \text{für } x > 0 \end{cases}, \quad \lambda > 0$$

heißt **exponentialverteilt** mit dem Parameter λ .

$$F(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ 1 - e^{-\lambda x} & \text{für } x > 0 \end{cases}$$

$$D^2(X) = \frac{1}{\lambda^2}$$

$$E(X) = \frac{1}{\lambda}$$

$$\varphi(s) = \frac{1}{1 - \frac{js}{\lambda}}$$

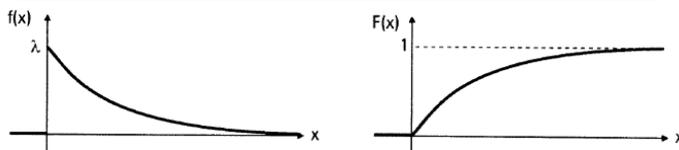


Bild 6.7-1: Dichte und Verteilungsfunktion einer exponentialverteilten Zufallsvariablen (schematisch)

Die Normalverteilung

Eine Zufallsvariable X ist **normalverteilt** (gaußverteilt), wenn ihre Dichte durch

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad \sigma > 0, \quad (6.8-1)$$

gegeben ist.

Die Normalverteilung wird mit Recht als die wichtigste Verteilung der Wahrscheinlichkeitstheorie angesehen. Ihre Bedeutung beruht vor allen Dingen darauf, daß Zufallsvariablen als normalverteilt angesehen werden können, die durch additive Überlagerung einer großen Zahl von unabhängigen zufälligen Einflüssen (d.h. Zufallsvariablen) entstehen, wobei jede der einzelnen Zufallsvariablen einen im Verhältnis zur Gesamtsumme nur unbedeutenden Beitrag liefert (vergleiche Abschnitt 7.8).

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt, \quad \sigma > 0$$

Werte von $F(x)$ bzw. $\Phi(X)$ kann man der Tabelle der Standardnormalverteilung entnehmen (Achtung: ZV muss normiert sein!)

$$\varphi(s) = e^{js\mu} \exp\left(-\frac{(\sigma s)^2}{2}\right) \quad E(X) = \mu \quad D^2(X) = \sigma^2$$

$$E[(X - \mu)^k] = \begin{cases} 1 \cdot 3 \cdot \dots \cdot (k-1) \sigma^k & \text{falls } k \text{ gerade} \\ 0 & \text{falls } k \text{ ungerade} \end{cases}$$

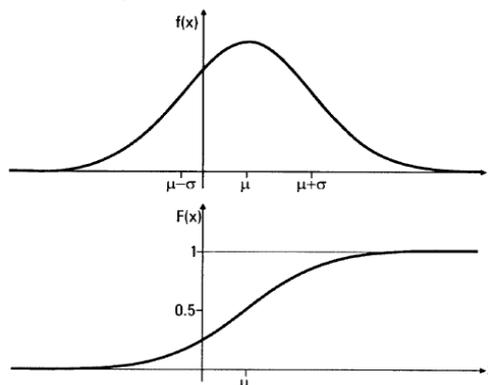


Bild 6.8-1: Dichte und Verteilungsfunktion einer $\mathcal{N}(\mu; \sigma^2)$ verteilten Zufallsvariablen

Fortsetzung Normalverteilung

Die Verteilungsfunktion $\Phi(x)$ der $\mathcal{N}(0; 1)$ -Verteilung ist tabelliert (siehe Anhang D). Mit

$$Y = \frac{X - \mu}{\sigma}$$

Normierung : meist 1. Schritt bei Aufgaben mit normalverteilten ZV

wird die $\mathcal{N}(\mu; \sigma^2)$ -verteilte Zufallsvariable X in die $\mathcal{N}(0; 1)$ -verteilte Zufallsvariable Y transformiert.

Rechenregeln für Normalverteilung

$$\Phi(-y) = 1 - \Phi(y)$$

$$P(y_1 \leq Y \leq y_2) = \Phi(y_2) - \Phi(y_1)$$

Standardnormalverteilung:

Erwartungswert $E(X) = \mu = 0$, Varianz $D^2(X) = \sigma^2 = 1$, Standardabweichung $D(X) = \sigma = 1$

Bemerkungen zur Normalverteilung

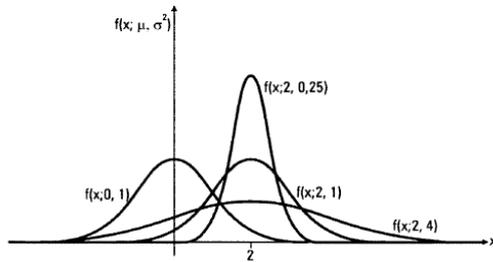


Bild 6.8-2: Einfluß der Parameter μ und σ^2 auf die Normalverteilungsdichte

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) du = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)$$

mit der (Gaußschen) Fehlerfunktion

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

Die komplementäre Fehlerfunktion ist

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt$$

$$F(x) = 1 - \frac{1}{2} \operatorname{erfc}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)$$

$$\operatorname{erf}(-x) = -\operatorname{erf}(x),$$

$$\operatorname{erfc}(-x) = 2 - \operatorname{erfc}(x),$$

$$\operatorname{erf}(0) = \operatorname{erfc}(\infty) = 0,$$

$$\operatorname{erf}(\infty) = \operatorname{erfc}(0) = 1.$$

Q-Funktion

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{t^2}{2}\right) dt$$

$$\operatorname{erfc}(x) = 2Q(\sqrt{2}x),$$

$$Q(x) = 1 - \Phi(x),$$

wobei $\Phi(x)$ die Verteilungsfunktion einer $\mathcal{N}(0; 1)$ verteilten Zufallsvariablen ist.

Für die $\mathcal{N}(\mu; \sigma^2)$ -verteilte Zufallsvariable X gilt:

$$P(\mu - \sigma < X \leq \mu + \sigma) \approx 0,68$$

$$P(\mu - 2\sigma < X \leq \mu + 2\sigma) \approx 0,955$$

$$P(\mu - 3\sigma < X \leq \mu + 3\sigma) \approx 0,997$$

Anders ausgedrückt: Praktisch alle Werte von X liegen innerhalb der **3 σ -Grenzen** $\mu - 3\sigma$ und $\mu + 3\sigma$.

Weibullverteilung

Eine nichtnegative Zufallsvariable X genügt einer Weibullverteilung mit den Parametern $\beta > 0$ und $\theta > 0$, wenn sie folgende Dichte hat:

$$f(x) = \frac{\beta}{\theta} \left(\frac{x}{\theta}\right)^{\beta-1} \exp\left\{-\left(\frac{x}{\theta}\right)^\beta\right\}, \quad x \geq 0$$

$$F(x) = 1 - \exp\left\{-\left(\frac{x}{\theta}\right)^\beta\right\}, \quad x \geq 0.$$

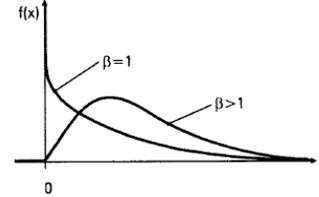


Bild 6.9-1: Einfluß des Parameters β auf die Dichte der Weibullverteilung

Erwartungswert und Varianz werden berechnet mit Hilfe der **Gammafunktion**:

$$\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du, \quad x > 0$$

$$E(X) = \theta \Gamma\left(\frac{1}{\beta} + 1\right) \quad E(X^2) = \theta^2 \Gamma\left(\frac{2}{\beta} + 1\right)$$

$$\operatorname{var}(X) = E(X^2) - E(X)^2 = \theta^2 \Gamma\left(\frac{2}{\beta} + 1\right) - \theta^2 \left[\Gamma\left(\frac{1}{\beta} + 1\right)\right]^2$$

Für $\beta = 1$ ergibt sich die Exponentialverteilung mit $\lambda = \frac{1}{\theta}$ und für $\beta = 2$ ergibt sich die Rayleighverteilung mit $\theta^2 = 2\sigma^2$.

Mehrdimensionale Zufallsvariablen

Eine vektorwertige Funktion

$$\vec{X} = \vec{X}(\xi) : \Omega \rightarrow \mathbb{R}^N, \quad (7.1-1)$$

die jedem Ergebnis $\xi \in \Omega$ einen Vektor $\vec{x} = (x_1, x_2, \dots, x_N)^T$ zuordnet, heißt **mehrdimensionale Zufallsvariable**, wenn das Urbild eines jeden Intervalls $I_{\vec{a}} = (-\infty, a_1] \times (-\infty, a_2] \times \dots \times (-\infty, a_N] \subset \mathbb{R}^N$ ein Ereignis ist:

$$\vec{X}^{-1}(I_{\vec{a}}) \in \mathfrak{B}, \quad \forall \vec{a} \in \mathbb{R}^N \quad (7.1-2)$$

Die Funktion

$$F(\vec{x}) = F(x_1, x_2, \dots, x_N) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_N \leq x_N)$$

der mehrdimensionalen Zufallsvariablen \vec{X} heißt **Verteilungsfunktion** von \vec{X} .

Durch partielle Ableitungen der Verteilungsfunktion $F(\vec{x})$ nach allen Variablen ergibt sich die **Dichte** der mehrdimensionalen Zufallsvariablen \vec{X}

$$f(\vec{x}) = \frac{\partial^N}{\partial x_1 \partial x_2 \dots \partial x_N} F(\vec{x}). \quad (7.1-4)$$

Im folgenden spezialisieren wir uns, wenn nicht ausdrücklich anders gesagt, auf den Fall zweidimensionaler Zufallsvariablen.

$$f(\vec{x}) = f(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(\vec{x})$$

Für zweidimensionale Zufallsvariablen $\vec{X} = (X_1, X_2)^T$ können beide Komponenten stetig oder diskret verteilt sein. Es kann aber auch eine der Komponenten stetig und die andere diskret verteilt sein. Sind beide Komponenten diskret verteilt, schreibt man ähnlich wie in (4.1-12) für deren „Dichte“

$$f(\vec{x}) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} P(X_1 = x_{1,n}; X_2 = x_{2,k}) \cdot \delta(x_1 - x_{1,n}; x_2 - x_{2,k})$$

zweidimensionale Normalverteilung

$$f(\vec{x}) = C \exp\left(-\frac{1}{2}Q(x_1; x_2)\right)$$

Die zweidimensionale Dichte $f(\vec{x})$ beschreibt eine Fläche im dreidimensionalen Raum, die im wesentlichen durch die quadratische Form $Q(x_1; x_2)$ charakterisiert ist. Der Einfluß der Parameter $\mu_1, \mu_2, \sigma_1, \sigma_2$ und ρ läßt sich wie folgt darstellen:

1. $Q(\mu_1; \mu_2) = 0$, d.h. $f(\vec{x})$ hat bei $\vec{x} = (\mu_1, \mu_2)^T$ ihr Maximum mit $f(\vec{x}) = C$.
2. $f(\vec{x})$ ist durch Höhenlinien (Kurven konstanter Dichte) gekennzeichnet. Diese sind gleichzeitig Höhenlinien von Q , die durch $Q(x_1; x_2) = K \geq 0$ bestimmt sind. Die Höhenlinien sind Ellipsen mit dem Mittelpunkt $\vec{\mu} = (\mu_1, \mu_2)^T$, ihre Lage in der (x_1, x_2) -Ebene wird durch die restlichen drei Parameter σ_1, σ_2, ρ bestimmt:

1. Fall: $\rho = 0$

Die Hauptachsen $A = 2a$ und $B = 2b$ der Ellipsen sind parallel zur x_1 - bzw. x_2 -Achse mit $a = \sigma_1 \sqrt{K}$, $b = \sigma_2 \sqrt{K}$

2. Fall: $\rho \neq 0$, $\sigma_1 = \sigma_2 = \sigma$

Die Hauptachsen sind gegenüber den Koordinaten um $\gamma = \frac{\pi}{4}$ gedreht. Es gilt $a = \sigma \sqrt{K(1-\rho)}$, $b = \sigma \sqrt{K(1+\rho)}$

3. Fall: $\rho \neq 0$, $\sigma_1 \neq \sigma_2$

Die Hauptachsen sind gegenüber den Koordinaten um $\gamma = \frac{1}{2} \arctan\left(\frac{2\rho\sigma_1\sigma_2}{\sigma_2^2 - \sigma_1^2}\right)$ gedreht. Es gilt

$$a = \sigma_1 \sigma_2 \sqrt{\frac{K(1-\rho^2)}{\sigma_1^2 \sin^2 \gamma + \sigma_2^2 \cos^2 \gamma + 2\rho\sigma_1\sigma_2 \sin \gamma \cos \gamma}}$$

$$b = \sigma_1 \sigma_2 \sqrt{\frac{K(1-\rho^2)}{\sigma_1^2 \cos^2 \gamma + \sigma_2^2 \sin^2 \gamma - 2\rho\sigma_1\sigma_2 \sin \gamma \cos \gamma}}$$

3. Als zweidimensionale Standardnormalverteilung ist der Fall $\mu_1 = \mu_2 = 0$; $\sigma_1^2 = \sigma_2^2 = 1$ und $-1 < \rho < 1$ anzusehen.

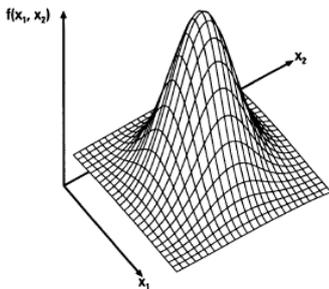


Bild 7.1-1: Dichte einer zweidimensionalen Normalverteilung

Randdichten

Die Randdichten $f_{X_1}(x_1)$ und $f_{X_2}(x_2)$ sind die Dichten der Komponenten des Zufallsvektors $\vec{X} = (X_1, X_2)^T$.

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 \quad f_{X_2}(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1$$

Randdichten (wenn ZV diskret)

$$P(X_1 = x_{1,n}) = \sum_{k=1}^{\infty} P(X_1 = x_{1,n}; X_2 = x_{2,k})$$

$$P(X_2 = x_{2,k}) = \sum_{n=1}^{\infty} P(X_1 = x_{1,n}; X_2 = x_{2,k})$$

Bedingte Dichte

\vec{X} sei eine zweidimensionale Zufallsvariable mit der Dichte $f(\vec{x}) = f(x_1; x_2)$ und es gelte $f_{X_1}(x_1) > 0$ sowie $f_{X_2}(x_2) > 0$. Dann heißt

$$f_{X_1}(x_1|X_2 = x_2) = \frac{f(x_1; x_2)}{f_{X_2}(x_2)}$$

die bedingte Dichte von X_1 unter der Bedingung $X_2 = x_2$

$$f_{X_2}(x_2|X_1 = x_1) = \frac{f(x_1; x_2)}{f_{X_1}(x_1)}$$

ist die bedingte Dichte von X_2 unter der Bedingung $X_1 = x_1$

Eine bedingte Dichte besitzt alle Eigenschaften einer Dichte

$$F_{X_1}(x_1|X_2 = x_2) = \int_{-\infty}^{x_1} f_{X_1}(u|X_2 = x_2) du = P(X_1 \leq x_1|X_2 = x_2)$$

Formel von der totalen Wahrscheinlichkeit für Dichten

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1}(x_1|X_2 = x_2) f_{X_2}(x_2) dx_2$$

und der Satz von Bayes für Dichten

$$f_{X_2}(x_2|X_1 = x_1) = \frac{f_{X_1}(x_1|X_2 = x_2) f_{X_2}(x_2)}{\int_{-\infty}^{\infty} f_{X_1}(x_1|X_2 = x_2) f_{X_2}(x_2) dx_2}$$

Der bedingte Erwartungswert einer Zufallsvariablen X_1 unter der Bedingung $X_2 = x_2$ ist

$$E\{X_1|X_2 = x_2\} = \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1|X_2 = x_2) dx_1$$

$$E\{X_1\} = \int_{-\infty}^{\infty} E\{X_1|X_2 = x_2\} f_{X_2}(x_2) dx_2$$

Zwei Zufallsvariablen X, Y heißen **unabhängig**, wenn $f(x, y) = f_X(x) \cdot f_Y(y)$ gilt.

Diskrete Zufallsvariablen X, Y sind unabhängig, wenn $P(X = x_i; Y = y_k) = P(X = x_i)P(Y = y_k) \quad \forall i, k = 1, 2, \dots$ gilt.

Erwartungswerte für zweidimensionale Zufallsvariablen

$$E\{g(X, Y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$

$$\text{cov}(X, Y) = E\{(X - E(X))(Y - E(Y))\} = E\{(X - \mu_X)(Y - \mu_Y)\}$$

heißt Kovarianz von X und Y und

$$\rho_{X,Y} = \rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{D^2(X)D^2(Y)}} \quad \text{stets zwischen -1 und +1}$$

$$= \frac{E\{(X - \mu_X)(Y - \mu_Y)\}}{\sigma_X \sigma_Y} \quad \text{Vorzeichen ist für die Ähnlichkeit unbedeutend}$$

ist der Korrelationskoeffizient von X und Y

Der Korrelationskoeffizient $\rho_{X,Y}$ stellt ein Ähnlichkeitsmaß der Zufallsvariablen X und Y dar:

Für $|\rho_{X,Y}| = 1$ sind X und Y maximal ähnlich. Für $\rho_{X,Y} = 0$ sind sie sich komplett unähnlich, man sagt, sie seien **unkorreliert**.

Unabhängige Zufallsvariablen sind unkorreliert. Die Umkehrung dieser Aussage gilt im allgemeinen **nicht**. Haben jedoch X und Y eine Normalverteilung und hat $(X, Y)^T$ eine zweidimensionale Normalverteilung, folgt aus $\rho_{X,Y} = 0$ die Unabhängigkeit von X und Y .

$E(X + Y) = E(X) + E(Y)$, $D^2(X + Y) = D^2(X) + D^2(Y) + 2\text{cov}(X, Y)$	Sind X und Y unabhängig, folgt $\text{cov}(X, Y) = 0$, $E(X \cdot Y) = E(X) \cdot E(Y)$, $D^2(X + Y) = D^2(X) + D^2(Y)$
--	--

Kovarianzmatrix

$$\Sigma_N = \begin{bmatrix} \text{cov}(X_1, X_1) & \text{cov}(X_1, X_2) & \dots & \text{cov}(X_1, X_N) \\ \text{cov}(X_2, X_1) & \text{cov}(X_2, X_2) & \dots & \text{cov}(X_2, X_N) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_N, X_1) & \text{cov}(X_N, X_2) & \dots & \text{cov}(X_N, X_N) \end{bmatrix}$$

wichtig: $\text{cov}(X_n, X_n) = \text{var}(X_n)$

zweidimensional:

$$\Sigma_2 = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$\det \Sigma_2 = \sigma_1^2 \sigma_2^2 (1 - \rho^2)$$

$$\Sigma_2^{-1} = \frac{1}{1 - \rho^2} \begin{bmatrix} \sigma_1^{-2} & -\rho\sigma_1^{-1}\sigma_2^{-1} \\ -\rho\sigma_1^{-1}\sigma_2^{-1} & \sigma_2^{-2} \end{bmatrix} \quad Q(x_1, x_2) = (\vec{x} - \vec{\mu})^T \Sigma_2^{-1} (\vec{x} - \vec{\mu})$$

$$C = ((2\pi)^2 \det \Sigma_2)^{-\frac{1}{2}}$$

$$f(\vec{x}) = \frac{1}{2\pi \sqrt{\det \Sigma_2}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu})^T \Sigma_2^{-1} (\vec{x} - \vec{\mu}) \right\} = C \exp(-0,5 Q(x_1, x_2))$$

$$f(\vec{x}) = \frac{1}{(2\pi)^{N/2} \sqrt{\det \Sigma_N}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu})^T \Sigma_N^{-1} (\vec{x} - \vec{\mu}) \right\}$$

Funktionen 2-dim. ZV

$U_1 = g_1(X, Y), U_2 = g_2(X, Y)$ eineindeutig

$$J = \begin{vmatrix} \frac{\partial x}{\partial u_1} & \frac{\partial x}{\partial u_2} \\ \frac{\partial y}{\partial u_1} & \frac{\partial y}{\partial u_2} \end{vmatrix}$$

$h(u_1, u_2) = f[h_1(u_1, u_2); h_2(u_1, u_2)] |J|$

(a) **Summe $Z = X + Y$**

Die Dichte der Summe $Z = X + Y$ zweier unabhängiger Zufallsvariablen X und Y ist die Faltung der Einzeldichten (7.4-5).

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx \quad (7.4-5)$$

X und Y seien unabhängige Zufallsvariablen mit den charakteristischen Funktionen $\varphi_X(s)$ bzw $\varphi_Y(s)$. Die charakteristische Funktion der Zufallsvariablen $Z = X + Y$ ist dann

$$\varphi_Z(s) = \varphi_X(s) \cdot \varphi_Y(s). \quad (7.4-6)$$

(b) **Produkt $Z = X \cdot Y$**

Sind X und Y unabhängig, erhält man:

$$f_Z(z) = \int_{-\infty}^{\infty} f\left(x, \frac{z}{x}\right) \frac{1}{|x|} dx,$$

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{|x|} f_X(x) f_Y\left(\frac{z}{x}\right) dx,$$

$$F_Z(z) = \int_{-\infty}^z \int_{-\infty}^{\infty} f\left(x, \frac{u}{x}\right) \frac{1}{|x|} dx du$$

$$F_Z(z) = \int_{-\infty}^z \int_{-\infty}^{\infty} \frac{1}{|x|} f_X(x) f_Y\left(\frac{u}{x}\right) dx du$$

(c) **Quotient $Z = X/Y$**

Sind X und Y unabhängig, erhält man:

$$f_Z(z) = \int_{-\infty}^{\infty} f(yz, y) |y| dy,$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(yz) f_Y(y) |y| dy,$$

$$F_Z(z) = \int_{-\infty}^z \int_{-\infty}^{\infty} f(yu, y) |y| dy du$$

$$F_Z(z) = \int_{-\infty}^z \int_{-\infty}^{\infty} f_X(yu) f_Y(y) |y| dy du$$

Komplexwertige Zufallsvariablen

X und Y seien reellwertige Zufallsvariablen (siehe Definition 4.1-1),

dann ist mit der imaginären Einheit j $Z = X + jY$

eine komplexwertige (kurz: komplexe) Zufallsvariable.

$$\text{Varianz} \quad D^2(Z) = E \{ |Z - E(Z)|^2 \}$$

$$= E \{ (Z - E(Z))(Z - E(Z))^* \}$$

$$= D^2(X) + D^2(Y)$$

Kovarianz der komplexen Zufallsvariablen Z_1 und Z_2

$$\text{cov}(Z_1, Z_2) = E \{ (Z_1 - E(Z_1))(Z_2 - E(Z_2))^* \}$$

$$\text{cov}(Z_2, Z_1) = [\text{cov}(Z_1, Z_2)]^*$$

Transformation von Zufallszahlen

Es sei X eine Zufallsvariable mit der Verteilungsfunktion $F_X(x)$. Die Zufallsvariable $Z = g(X)$ (g eindeutig umkehrbar) ist genau dann im Intervall $[0, 1)$ gleichverteilt, wenn gilt

$$Z = g(X) = F_X(X). \quad \text{Achtung, hier ungewohnt: ZV in Verteilungsfunktion einsetzen} \quad (7.6-1)$$

Gegeben sei eine im Intervall $[0, 1)$ gleichverteilte Zufallsvariable Z und eine Zufallsvariable Y . Y hat genau dann die gewünschte Verteilungsfunktion $F_Y(y)$, wenn gilt

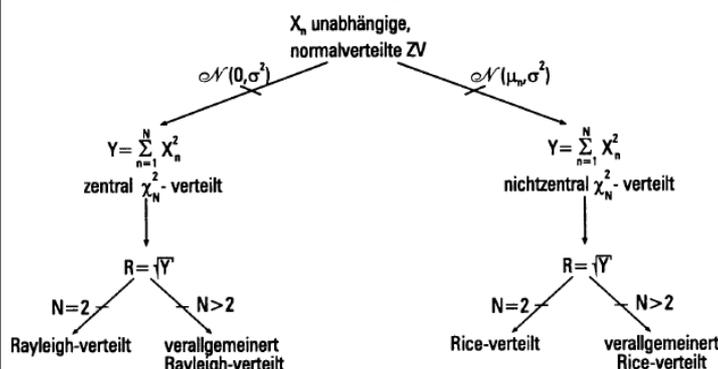
$$Y = F_Y^{-1}(Z). \quad \text{Anmerkung: } F^{-1} \text{ bei Normalverteilung nicht bekannt!}$$

Es sei eine Zufallsvariable X mit der Verteilungsfunktion $F_X(x)$ gegeben. Die Zufallsvariable Y hat die gewünschte Verteilungsfunktion $F_Y(y)$, wenn

$$Y = F_Y^{-1}(F_X(X))$$

gilt.

Aus normalverteilten abgeleitete ZV



Satz 7.7-1

Die Zufallsvariablen X_1 und X_2 seien unabhängig und besitzen eine $N(\mu_n; \sigma_n^2)$ -Verteilung; $n = 1, 2$. Die Summe $Y = X_1 + X_2$ hat dann eine $N(\mu_1 + \mu_2; \sigma_1^2 + \sigma_2^2)$ -Verteilung.

(i) Die Erweiterung der Aussage von Satz 7.7-1 ist trivial: Sind die X_n $N(\mu_n; \sigma_n^2)$ -verteilt und unabhängig, hat

$$Y = \sum_{n=1}^N X_n$$

eine $N\left(\sum_{n=1}^N \mu_n; \sum_{n=1}^N \sigma_n^2\right)$ -Verteilung.

(ii) Die Klasse unabhängiger normalverteilter Zufallsvariablen ist gegenüber Summenbildung abgeschlossen.

Gesetze der großen Zahlen und Grenzwertsätze

Indikatorfunktion

$$I_n(\xi) = \begin{cases} 1 & \text{für } \xi \in A \\ 0 & \text{sonst} \end{cases}$$

Satz 7.8-1 (Tschebyscheffsche Ungleichung)

Es sei $c \in \mathbb{R}$ beliebig und X eine Zufallsvariable mit endlicher Varianz. Dann gilt für jedes $\epsilon > 0$: $P\{|X - c| \geq \epsilon\} \leq \frac{1}{\epsilon^2} E\{[X - c]^2\}$

Satz 7.8-2 (Bernoullisches Gesetz der großen Zahlen)

Ist X_1, X_2, \dots eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit $P(X_n = 1) = p, P(X_n = 0) = 1 - p$ ($0 < p < 1$),

$$\text{gilt für alle } \epsilon > 0 \quad \lim_{N \rightarrow \infty} P \left\{ \left| \frac{1}{N} \sum_{n=1}^N X_n - p \right| < \epsilon \right\} = 1.$$

Die Wahrscheinlichkeit $P(a)$ ist nicht der Grenzwert der relativen Häufigkeit. Man sagt, die relative Häufigkeit **konvergiere in Wahrscheinlichkeit** gegen die Wahrscheinlichkeit.

Satz 7.8-3 (Chintschinsches Gesetz der großen Zahlen)

Ist X_1, X_2, \dots eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit $E(X_n) = \mu < \infty$, folgt für alle $\epsilon > 0$:

$$\lim_{N \rightarrow \infty} P \left\{ \left| \frac{1}{N} \sum_{n=1}^N X_n - \mu \right| < \epsilon \right\} = 1$$

Grundlagen stochastischer Prozesse (SP)

Definition 8.1-1

Ein **stochastischer Prozeß** $X(t, \xi)$ ist eine mit dem Parameter t indizierte Familie von Zufallsvariablen.

- (i) t wird im allgemeinen als Zeit interpretiert, d.h. t ist ein kontinuierlicher Parameter.
- (ii) Zu jedem (festen) Zeitpunkt t_0 ist $X_{t_0}(\xi) = X(t_0, \xi)$ eine Zufallsvariable.
- (iii) Wird der Zufallsparameter ξ in $X(t, \xi)$ festgehalten, ergibt sich eine Zeitfunktion $X(t, \xi_0)$, die als **Realisierung** oder **Pfad** des Prozesses bezeichnet wird. Im allgemeinen ist die Anzahl aller möglichen Pfade (überabzählbar) unendlich groß. Bild 8.1-1 zeigt drei Pfade eines stochastischen Prozesses.
- (iv) Statt $X(t, \xi)$ schreiben wir im folgenden $X(t)$, bedenken dabei jedoch immer, daß es sich um einen stochastischen Prozeß handelt.

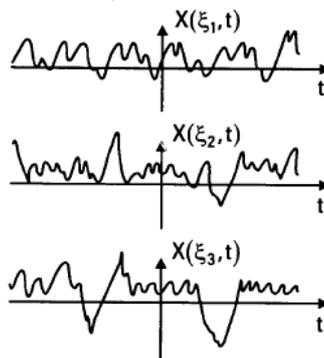


Bild 8.1-1: Realisierungen eines stochastischen Prozesses

Falls $f(x_{t_1+h}, x_{t_2+h}, \dots, x_{t_N+h}) = f(x_{t_1}, x_{t_2}, \dots, x_{t_N})$ für jedes h und jedes N gilt, ist $X(t)$ ein **stark stationärer Prozeß**.

Sämtliche Dichten eines stark stationären Prozesses sind gegenüber beliebigen Verschiebungen der Zeitachse invariant.

Scharmittelwerte

k-tes Moment $E\{X^k(t_n)\} = \int_{-\infty}^{\infty} x_{t_n}^k f(x_{t_n}) dx_{t_n}$

Im allgemeinen hängt also das k -te Moment (8.2-1) vom Zeitpunkt t_n ab. Ist der zugrundeliegende Prozeß jedoch stark stationär, gilt also $f(x_{t_n}) = f(x_{t_n+h}) \forall h$ und $\forall t_n$, sind die Momente nicht zeitabhängig.

Autokorrelationsfunktion (AKF)

$$\varphi_{XX}(t_1, t_2) = E\{X(t_1)X(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_{t_1} x_{t_2} f(x_{t_1}, x_{t_2}) dx_{t_1} dx_{t_2}$$

bei stark stationären gilt für AKF:

$$\varphi_{XX}(t_1, t_2) = \varphi_{XX}(t_2 - t_1) = \varphi_{XX}(\tau)$$

$$\varphi_{XX}(-\tau) = \varphi_{XX}(\tau)$$

Def. (schwach) stationär:

Erwartungswert (1.Moment) konstant und es gilt

$$\varphi_{XX}(t_1, t_2) = \varphi_{XX}(t_2 - t_1) = \varphi_{XX}(\tau)$$

Jeder stark stationäre Prozeß ist schwach stationär.
Die Umkehrung gilt i.A. nicht!

$$\varphi_{XX}(0) = E\{X^2(t)\} \text{ ist die } \mathbf{mittlere Leistung}$$

Autokovarianzfunktion

$$c_{XX}(t_1, t_2) = E\{(X(t_1) - \mu(t_1))(X(t_2) - \mu(t_2))\} \\ = \varphi_{XX}(t_1, t_2) - \mu(t_1)\mu(t_2)$$

$$\text{mit } \mu(t_n) = E\{X(t_n)\}; n = 1, 2$$

Bei stationären SP gilt:

$$c_{XX}(t_1, t_2) = c_{XX}(t_2 - t_1) = c_{XX}(\tau) = \varphi_{XX}(\tau) - \mu^2$$

Satz 7.8-4 (Zentraler Grenzwertsatz):

X_1, X_2, \dots sei eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit

$$E(X_n) = m < \infty \text{ und } D^2(X_n) = d^2 < \infty,$$

dann gilt mit $S_N = \sum_{n=1}^N X_n$ für jedes $x \in \mathbb{R}$

$$\lim_{N \rightarrow \infty} P\left\{ \frac{S_N - Nm}{\sqrt{Nd}} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{y^2}{2}\right) dy. \quad (7.8-10)$$

Bemerkung:

Die Folge der Verteilungsfunktionen der standardisierten Summen

$$Z_N = \frac{S_N - Nm}{\sqrt{Nd}}$$

konvergiert für $N \rightarrow \infty$ gegen die Verteilungsfunktion einer $\mathcal{N}(0;1)$ -verteilten Zufallsvariablen.

Der zentrale Grenzwertsatz gilt unter weiterreichenden als den hier erwähnten Bedingungen. So kann (bei Einführung anderer Randbedingungen) auf die Forderung verzichtet werden, daß die Zufallsvariablen X_1, X_2, \dots identisch verteilt sind [Rén71].

Satz 7.8-5 (Satz von de Moivre-Laplace):

Ist S_N eine binomialverteilte Zufallsvariable mit den Parametern N und p , gilt für jedes $x \in \mathbb{R}$

$$\lim_{N \rightarrow \infty} P\left\{ \frac{S_N - Np}{\sqrt{Np(1-p)}} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{y^2}{2}\right) dy. \quad (7.8-11)$$

$S_N = \sum_{n=1}^N X_n$ ist die Summe (identisch) Null-Eins-verteilter Zufallsvariablen (Bernoullisches Versuchsschema).

Als Faustregel wird im allgemeinen angegeben, daß die Approximation der Binomialverteilung mit der Normalverteilung für praktische Zwecke ausreicht, wenn $D^2(S_N) \geq 9$ gilt.

Oft wird bei der Approximation der Binomialverteilung eine Stetigkeitskorrektur vorgenommen [Hen97].
Statt

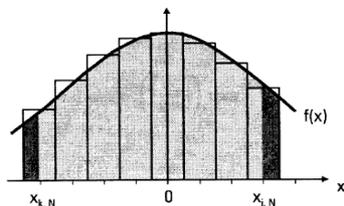
$$P(k < S_N \leq l) \approx \Phi\left(\frac{l - Np}{\sqrt{Np(1-p)}}\right) - \Phi\left(\frac{k - Np}{\sqrt{Np(1-p)}}\right) \\ = \Phi(x_{l,N}) - \Phi(x_{k,N})$$

wird oft

$$P(k < S_N \leq l) \approx \Phi\left(\frac{l - Np + \frac{1}{2}}{\sqrt{Np(1-p)}}\right) - \Phi\left(\frac{k - Np - \frac{1}{2}}{\sqrt{Np(1-p)}}\right)$$

benutzt.

Vergleicht man das Histogramm der standardisierten Binomialverteilung mit der standardisierten Gaußdichte $f(x)$, wird die Motivation dieser Korrektur offensichtlich (Bild 7.8-3).



In den Rahmen der hier zitierten Grenzwertaussagen gehört auch der in Abschnitt 6.4 bewiesene Satz 6.4-1, der den Zusammenhang zwischen Binomial- und Poissonverteilung kennzeichnet. Dieser Satz wird deshalb auch als Poisson'scher Grenzwertsatz bezeichnet.

Eine experimentelle Bestätigung des Satzes von de Moivre-Laplace stellt das **Galton'sche Brett** (Bild 7.8-4) dar [Rén71].

Satz 7.8-4 (Zentraler Grenzwertsatz):
 X_1, X_2, \dots sei eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit
 $E(X_n) = m < \infty$ und $D^2(X_n) = d^2 < \infty$,
 dann gilt mit $S_N = \sum_{n=1}^N X_n$ für jedes $x \in \mathbb{R}$

$$\lim_{N \rightarrow \infty} P\left\{ \frac{S_N - Nm}{\sqrt{Nd}} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{y^2}{2}\right) dy. \quad (7.8-10)$$

Definition 8.1-1
 Ein **stochastischer Prozeß** $X(t, \xi)$ ist eine mit dem Parameter t indizierte Familie von Zufallsvariablen.
 (i) t wird im allgemeinen als Zeit interpretiert, d.h. t ist ein kontinuierlicher Parameter.
 (ii) Zu jedem (festen) Zeitpunkt t_0 ist $X_{t_0}(\xi) = X(t_0, \xi)$ eine Zufallsvariable.
 (iii) Wird der Zufallsparameter ξ in $X(t, \xi)$ festgehalten, ergibt sich eine Zeitfunktion $X(t, \xi_0)$, die als **Realisierung** oder **Pfad** des Prozesses bezeichnet wird. Im allgemeinen ist die Anzahl aller möglichen Pfade (überabzählbar) unendlich groß. Bild 8.1-1 zeigt drei Pfade eines stochastischen Prozesses.
 (iv) Statt $X(t, \xi)$ schreiben wir im folgenden $X(t)$, bedenken dabei jedoch immer, daß es sich um einen stochastischen Prozeß handelt.

Bild 8.1-1: Realisierungen eines stochastischen Prozesses

Falls $f(x_{t_1+h}, x_{t_2+h}, \dots, x_{t_N+h}) = f(x_{t_1}, x_{t_2}, \dots, x_{t_N})$ für jedes h und jedes N gilt, ist $X(t)$ ein **stark stationärer Prozeß**.
 Sämtliche Dichten eines stark stationären Prozesses sind gegenüber beliebigen Verschiebungen der Zeitachse invariant.

Scharmittelwerte
k-tes Moment $E\{X^k(t_n)\} = \int_{-\infty}^{\infty} x_{t_n}^k f(x_{t_n}) dx_{t_n}$
 Im allgemeinen hängt also das k -te Moment (8.2-1) vom Zeitpunkt t_n ab. Ist der zugrundeliegende Prozeß jedoch stark stationär, gilt also $f(x_{t_n}) = f(x_{t_n+h}) \forall h$ und $\forall t_n$, sind die Momente nicht zeitabhängig.

Autokorrelationsfunktion (AKF)

$$\varphi_{XX}(t_1, t_2) = E\{X(t_1)X(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_{t_1} x_{t_2} f(x_{t_1}, x_{t_2}) dx_{t_1} dx_{t_2}$$
 bei stark stationären gilt für AKF:

$$\varphi_{XX}(t_1, t_2) = \varphi_{XX}(t_2 - t_1) = \varphi_{XX}(\tau)$$

$$\varphi_{XX}(-\tau) = \varphi_{XX}(\tau)$$

Def. (schwach) stationär:
 Erwartungswert (1.Moment) konstant und es gilt

$$\varphi_{XX}(t_1, t_2) = \varphi_{XX}(t_2 - t_1) = \varphi_{XX}(\tau)$$
 Jeder stark stationäre Prozeß ist schwach stationär.
 Die Umkehrung gilt i.A. nicht!

$$\varphi_{XX}(0) = E\{X^2(t)\} \text{ ist die } \mathbf{mittlere Leistung}$$

Autokovarianzfunktion

$$c_{XX}(t_1, t_2) = E\{(X(t_1) - \mu(t_1))(X(t_2) - \mu(t_2))\} \\ = \varphi_{XX}(t_1, t_2) - \mu(t_1)\mu(t_2)$$
 mit $\mu(t_n) = E\{X(t_n)\}; n = 1, 2$
 Bei stationären SP gilt:

$$c_{XX}(t_1, t_2) = c_{XX}(t_2 - t_1) = c_{XX}(\tau) = \varphi_{XX}(\tau) - \mu^2$$

Der stochastische Prozeß $X(t)$ heißt **normal** (oder **Gaußprozeß**), wenn für jedes N und beliebige Zeitpunkte t_1, t_2, \dots, t_N der Zufallsvektor $[X(t_1), X(t_2), \dots, X(t_N)]^T$ eine N -dimensionale Normalverteilung besitzt.

Für normalverteilte SP sind schwache und starke Stationarität dasselbe!

Kreuzkorrelationsfunktion (KKF)

$$\varphi_{XY}(t_1, t_2) = E\{X(t_1)Y(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_{t_1} y_{t_2} f_{XY}(x_{t_1}, y_{t_2}) dx_{t_1} dy_{t_2}$$

Kreuzkovarianzfunktion

$$c_{XY}(t_1, t_2) = \varphi_{XY}(t_1, t_2) - \mu_X(t_1)\mu_Y(t_2)$$

SP $X(t)$, $Y(t)$ heißen **gemeinsam stationär** wenn sowohl $X(t)$ als auch $Y(t)$ stationär ist und ihre KKF nur von Zeitdifferenzen abhängt. Dann gilt:

$$\varphi_{XY}(t_1, t_2) = \varphi_{XY}(\tau), \quad c_{XY}(t_1, t_2) = c_{XY}(\tau)$$

$$\varphi_{XY}(-\tau) = \varphi_{YX}(\tau)$$

Def: SP $X(t)$, $Y(t)$ heißen **stochastisch unabhängig** wenn gilt:

$$f_{XY}(x_{t_1}, x_{t_2}, \dots, x_{t_N}; y_{t_1}, y_{t_2}, \dots, y_{t_M}) = f_X(x_{t_1}, x_{t_2}, \dots, x_{t_N}) \cdot f_Y(y_{t_1}, y_{t_2}, \dots, y_{t_M})$$

unkorreliert

$$\varphi_{XY}(t_1, t_2) = E\{X(t_1)\} \cdot E\{Y(t_2)\}$$

orthogonal.

$$\varphi_{XY}(t_1, t_2) = E\{X(t_1)Y(t_2)\} = 0$$

Für unkorrelierte Prozesse verschwindet die Kreuzkovarianzfunktion identisch, das heißt es gilt $c_{XY}(t_1, t_2) = 0 \forall t_1, t_2$.

Komplexwertige SPs

$$Z(t) = X(t) + jY(t) \quad (8.3-1)$$

ist ein **komplexwertiger** (kurz: **komplexer**) **stochastischer Prozeß**, wenn sowohl $X(t)$ als auch $Y(t)$ reellwertige Zufallsprozesse sind.

$$\begin{aligned} \text{AKF } \varphi_{ZZ}(t_1, t_2) &= E\{Z(t_1) \cdot Z^*(t_2)\} \\ &= \{\varphi_{XX}(t_1, t_2) + \varphi_{YY}(t_1, t_2) + j[\varphi_{YX}(t_1, t_2) - \varphi_{XY}(t_1, t_2)]\} \end{aligned}$$

Sind die Prozesse $X(t)$ und $Y(t)$ gemeinsam stationär, ist auch $Z(t)$ stationär und es gilt mit $\tau := t_2 - t_1$

$$\varphi_{ZZ}(t_1, t_2) = \varphi_{ZZ}(t_2 - t_1) = \varphi_{ZZ}(\tau).$$

$$\varphi_{ZZ}^*(\tau) = \varphi_{ZZ}(-\tau)$$

KKF:

Sind $Z(t) = X(t) + jY(t)$ und $W(t) = U(t) + jV(t)$ zwei komplexe SPs,

So gilt

$$\begin{aligned} \varphi_{ZW}(t_1, t_2) &= E\{Z(t_1)W^*(t_2)\} \\ &= \varphi_{XU}(t_1, t_2) + \varphi_{YV}(t_1, t_2) + j[\varphi_{YU}(t_1, t_2) - \varphi_{XV}(t_1, t_2)] \end{aligned}$$

Sind $Z(t)$ und $W(t)$ gemeinsam stationär, folgt:

$$\varphi_{ZW}^*(\tau) = \varphi_{WZ}(-\tau)$$

Zeitmittelwerte

Der **stark stationäre stochastische Prozeß** $X(t)$ ist **ergodisch**, wenn alle seine statistischen Eigenschaften aus einer einzigen Realisierung $x(t)$ abgeleitet werden können.

Die Ergodizität läßt sich im allgemeinen nicht nachprüfen. Man hilft sich hier mit der **Ergodenhypothese**, mit der die Ergodizität einfach postuliert wird, oder man begnügt sich mit eingeschränkten Aussagen.

Es seien $g: \mathbb{R} \rightarrow \mathbb{R}$ eine reellwertige Funktion und $x(t)$ ein Pfad des stark stationären stochastischen Prozesses $X(t)$.

$$\overline{g[x(t)]} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T g[x(t)] dt \quad (8.4-1)$$

heißt **zeitlicher Mittelwert der Realisierung** $x(t)$ bezüglich der Funktion g .

Der **stark stationäre stochastische Prozeß** $X(t)$ heißt **ergodisch bezüglich** g , wenn $E\{g(X(t))\}$ existiert und

$$\overline{g[x(t)]} = E\{g(X(t))\} \quad (8.4-3)$$

gilt.

k-tes Moment:

$$m^{(k)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x^k(t) dt \quad (8.4-4)$$

k-tes zentrales Moment:

$$\mu^{(k)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - m^{(1)}]^k dt \quad (8.4-5)$$

Autokorrelationsfunktion:

$$\varphi_{XX}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t+\tau) dt \quad (8.4-6)$$

Autokovarianzfunktion:

$$c_{XX}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - m^{(1)}][x(t+\tau) - m^{(1)}] dt \quad (8.4-7)$$

Zwei stark stochastische Prozesse sind **gemeinsam ergodisch**, wenn beide Prozesse gemeinsam stark stationär und ergodisch sind und wenn auch für ihre gemeinsamen Momente die Vertauschbarkeit von Schar- und Zeitmittelwerten gegeben ist. Dann folgt:

Kreuzkorrelationsfunktion:

$$\varphi_{XY}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)y(t+\tau) dt \quad (8.4-8)$$

Kreuzkovarianzfunktion:

$$c_{XY}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - m_x^{(1)}][y(t+\tau) - m_y^{(1)}] dt \quad (8.4-9)$$

Das Leistungsdichtespektrum

$X(t)$ sei ein stationärer stochastischer Prozeß mit der Autokorrelationsfunktion $\varphi_{XX}(\tau)$. Dann heißt

$$\Phi_{XX}(f) = \int_{-\infty}^{\infty} \varphi_{XX}(\tau) e^{-j2\pi f\tau} d\tau \quad (8.5-1)$$

$$\Phi_{XX}(f) \geq 0 \quad \forall f$$

Leistungsdichtespektrum des Prozesses $X(t)$.

mit inverser FT:

$$\varphi_{XX}(\tau) = \int_{-\infty}^{\infty} \Phi_{XX}(f) e^{j2\pi f\tau} df$$

Mittlere Leistung

$$\varphi_{XX}(0) = \int_{-\infty}^{\infty} \Phi_{XX}(f) df = E\{X^2(t)\} \geq 0$$

Für reelle stationäre SP: $\Phi_{ZZ}(f) = \Phi_{ZZ}^*(f)$

Das Leistungsdichtespektrum ist in jedem Fall eine nichtnegative reellwertige Funktion.

Kreuz-Leistungsdichtespektrum:

$$\Phi_{XY}(f) = \int_{-\infty}^{\infty} \varphi_{XY}(\tau) e^{-j2\pi f\tau} d\tau$$

konj. komplex: $\Phi_{XY}^*(f) = \Phi_{YX}(f)$

für gemeinsam stationäre reelle: $\Phi_{YX}(f) = \Phi_{XY}(-f)$

Zeitdiskrete Zufallsprozesse

Ein zeitdiskreter Zufallsprozess $X(n)$ ist eine Familie zufälliger Folgen $\{x(n)\}$. $X(n)$ kann z.B. durch äquidistante Abtastung eines zeitkontinuierlichen Prozesses $X(t)$ entstanden sein ($n = n\Delta t$). Die Eigenschaften von $X(t)$ und $X(n)$ sind daher einander sehr ähnlich.

$$k\text{-tes Moment } E\{X^k(n)\} = \int_{-\infty}^{\infty} x_n^k f(x_n) dx_n$$

$$\text{Autokorrelationsfolge } \varphi_{XX}(n, k) = E\{X(n)X(k)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_n x_k f(x_n, x_k) dx_n dx_k$$

Autokovarianzfolge

$$c_{XX}(n, k) = \varphi_{XX}(n, k) - E\{X(n)\}E\{X(k)\}$$

Für stationäre SPs:

$$\varphi_{XX}(n, k) = \varphi_{XX}(k - n)$$

$$c_{XX}(n, k) = c_{XX}(k - n) = \varphi_{XX}(k - n) - [E\{X(n)\}]^2$$

mittlere Leistung $\varphi_{XX}(0) = E\{X^2(n)\}$

Leistungsdichtespektrum

$$\Phi_{XX}(f) = \sum_{m=-\infty}^{\infty} \varphi_{XX}(m) e^{-j2\pi f m}$$

$$\varphi_{XX}(m) = \int_{-f_p/2}^{f_p/2} \Phi_{XX}(f) e^{j2\pi f m} df$$

Da $\Phi_{XX}(f)$ die Fouriertransformierte eines zeitdiskreten Signals ist, ist $\Phi_{XX}(f)$ periodisch mit der Periode $f_p = \frac{1}{\Delta t}$. D.h. es gilt

$$\Phi_{XX}(f + k \cdot f_p) = \Phi_{XX}(f) \quad \text{für } k = \pm 1, \pm 2, \dots,$$

wobei für $\Delta t = 1$ natürlich auch $f_p = 1$ ist.

Spezielle SPs

Gaußsches Weißes Rauschen

Unter einem (reellen) **weißen Gaußschen Rauschen** $\tilde{X}(t)$ versteht man einen mittelwertfreien, stationären Gaußprozess, dessen Leistungsdichtespektrum auf der gesamten Frequenzachse konstant ist (Bild 9.1-1(a)):

$$\text{Leistungsdichtespektrum } \Phi_{\tilde{X}\tilde{X}}(f) = \frac{N_o}{2} \quad \forall f$$

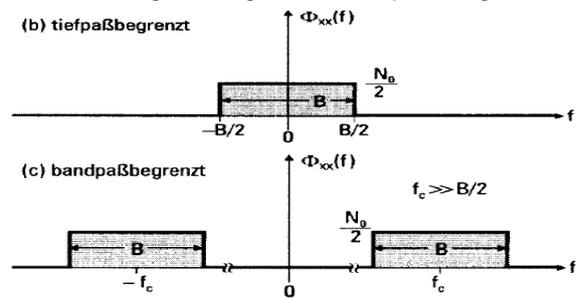
$$\text{AKF (9.1-2) } \varphi_{\tilde{X}\tilde{X}}(\tau) = \frac{N_o}{2} \delta(\tau)$$

Bemerkungen zum Gaußschen Weißen Rauschen

- Mittlere Leistung ist nicht endlich
- Realisierungen sind Distributionen

Aufgrund von (9.1-2) sind für einen weißen Gaußschen Rauschprozess $\tilde{X}(t)$ die Zufallsvariablen $\tilde{X}(t_1)$ und $\tilde{X}(t_2)$ unkorreliert und damit, da es sich um einen Gaußprozess handelt, unabhängig, wenn nur $\tau = t_2 - t_1 \neq 0$ gilt.

In der Anwendung benötigt: tief-/bandpassbegrenzt:



Für tiefpassbegr. Rauschen gilt: (9.1-4)

$$\Phi_{XX}(f) = \begin{cases} \frac{N_o}{2} & \text{für } |f| \leq \frac{B}{2} \\ 0 & \text{sonst} \end{cases} \quad \varphi_{XX}(\tau) = \frac{N_o}{2} \frac{\sin(\pi B\tau)}{\pi\tau}$$

$$\sigma_X^2 = \varphi_{XX}(0) = \frac{N_o}{2} B$$

Die Realisierungen sind beliebig oft diff'bare Fkt.

Aufgrund von (9.1-4) sind für ein tiefpaßbegrenzt reelles weißes Rauschen $X(t)$ die Zufallsvariablen $X(t_1)$ und $X(t_2)$ unkorreliert und damit, da es sich um einen Gaußprozess handelt, unabhängig, wenn $\tau = t_2 - t_1 = \frac{n}{B}$ ($n \in \mathbb{Z}, n \neq 0$) gilt.

$$\varphi_{XX}(\tau) \xrightarrow{B \rightarrow \infty} \frac{N_o}{2} \delta(\tau)$$

$$F(x) = \frac{1}{\sqrt{\pi N_o B}} \int_{-\infty}^x \exp\left(-\frac{u^2}{N_o B}\right) du$$

Ist $X(t)$ ergodisch, kann die Amplitudenverteilung des tiefpaßbegrenzten reellen weißen Gaußschen Rauschens aus einer einzigen Realisierung ermittelt werden (vergleiche Bild 9.1-2).

Tiefpassbegr. komplexes Rauschen:

Das tiefpaßbegrenzte komplexe weiße Gaußsche Rauschen $Z(t) = X(t) + jY(t)$ besitzt das Leistungsdichtespektrum

$$\Phi_{ZZ}(f) = \begin{cases} N_o & \text{für } |f| \leq \frac{B}{2} \\ 0 & \text{sonst} \end{cases} \quad (9.1-6)$$

Realteilprozess $X(t)$ und Imaginärteilprozess $Y(t)$ sind stochastisch unabhängige tiefpaßbegrenzte reelle weiße Gaußsche Rauschprozesse mit der selben mittleren Leistung $\sigma_X^2 = \varphi_{XX}(0) = \sigma_Y^2 = \varphi_{YY}(0) = \frac{N_o}{2} B$.

$$\text{AKF } \varphi_{ZZ}(\tau) = N_o \frac{\sin \pi B\tau}{\pi\tau}$$

Mittl. Leistung $\sigma_Z^2 = \varphi_{ZZ}(0) = \varphi_{XX}(0) + \varphi_{YY}(0) = 2\sigma_X^2 = N_o B$

Poissonprozess

Forderungen:

Wahrscheinlichkeiten für eintreffende Pakete:

$$P\{X(t + \Delta t) - X(t) = 1\} = \lambda \Delta t + o(\Delta t)$$

$$P\{X(t + \Delta t) - X(t) = 0\} = 1 - \lambda \Delta t + o(\Delta t)$$

Wahrscheinlichkeit ist unabhängig von der Lage von Δt , Prozess ist gedächtnislos

Mit $X(0) = 0$ gibt $X(t)$ die Anzahl der im Intervall der Dauer t eingetroffenen Pakete an.

$$\text{Es gilt für } t \geq 0 \quad P\{X(t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

d.h. $X(t)$ folgt einer Poissonverteilung

$$E\{X(t)\} = \lambda t \quad D^2\{X(t)\} = \lambda t$$

Der Poissonprozess ist nichtstationär (folgt aus $E\{X(t)\}$)

$$\text{Mittlere Ankunftsrate: } \lambda = \frac{E\{X(t)\}}{t}$$

Die ZV $X(t)$ ist eng um den Erwartungswert verteilt.

$$\frac{D\{X(t)\}}{E\{X(t)\}} = \frac{1}{\sqrt{\lambda t}}$$

Schätzwert $\lambda = \frac{T_1}{t}$ Mit n: Zahl der Pakete von Zeit 0 bis t

Für einen Poissonschen Ankunftsprozess $X(t)$ besitzt die Zufallsvariable T eine Exponentialverteilung, d. h. die Dichte von T ist (vergleiche (6.7-1))

$$f_T(\tau) = \begin{cases} 0 & \text{für } \tau \leq 0 \\ \lambda e^{-\lambda\tau} & \text{für } \tau > 0 \end{cases}, \lambda > 0. \quad (9.2-4)$$

$$P\{T > \tau\} = P\{X(\tau) = 0\} = e^{-\lambda\tau}$$

$$P\{T \leq \tau\} = F_T(\tau) = 1 - e^{-\lambda\tau} \quad (\tau > 0)$$

$$E\{T\} = \frac{1}{\lambda} \quad D^2\{T\} = \frac{1}{\lambda^2}$$

Bei einer mittleren Ankunftsrate λ ist die mittlere Differenz Zwischen zwei Ankünften $1/\lambda$.

$$P\{N(t, t + \Delta t) = 0\} = 1 - \lambda\Delta t + o(\Delta t) \quad \text{mit } \lambda = \sum_{m=1}^M \lambda_m$$

$$P\{N(t, t + \Delta t) = 1\} = \lambda\Delta t + o(\Delta t)$$

Summen unabhängiger Poissonprozesse bilden also wiederum Poissonprozesse.

Markoffprozesse, Markoffketten

$X(t)$ heißt Markoffscher Prozeß, wenn

$$P\{X(t_{m+1}) \leq x_{m+1} | X(t_m) \leq x_m, \dots, X(t_{m-k}) \leq x_{m-k}\}$$

$$= P\{X(t_{m+1}) \leq x_{m+1} | X(t_m) \leq x_m\} \quad (9.3-1)$$

für jedes m und jedes k sowie beliebige Zeitpunkte $t_{m-k} < t_{m-k+1} < \dots < t_{m+1}$ gilt.

Die Zukunft eines Markoffprozesses hängt nur von der Gegenwart und nicht von der Vergangenheit ab.

Definiere:

Zustandsmenge Z (ganzzahlig)

Zeitparametermenge $T = \{t_0, t_1, \dots, t_m, \dots\}$

(abzählbar, nicht unbedingt äquidistant)

Der zustands- und zeitdiskrete stochastische Prozeß $X(t)$ heißt **Markoffkette**, wenn

$$P\{X(t_{m+1}) = i_{m+1} | X(t_m) = i_m, \dots, X(t_0) = i_0\}$$

$$= P\{X(t_{m+1}) = i_{m+1} | X(t_m) = i_m\} \quad (9.3-2)$$

$\forall m > 2$ und $\forall i_0, i_1, \dots, i_{m+1} \in Z$ gilt.

Übergangswahrscheinlichkeiten k -ter Stufe

$$P\{X(t_{m+k}) = j | X(t_m) = i\} = p_{ij}(t_m, t_{m+k})$$

Die Markoffkette $X(t)$ heißt **homogen**, wenn für beliebige Zustände i, j und beliebige Zeitpunkte t_m, t_{m+1} die Übergangswahrscheinlichkeiten $p_{ij}(t_m, t_{m+1}) = p_{ij}$ nicht von der Zeit abhängen.

Außerdem muss die Zustandsmenge endlich sein!

Übergangsmatrix:	$\bar{P} =$	p_{11}	p_{12}	\dots	p_{1N}
		p_{21}	p_{22}	\dots	p_{2N}
		\vdots	\vdots		\vdots
		p_{N1}	p_{N2}	\dots	p_{NN}
		Zeilensumme muss stets 1 sein!			

Eine $(N \times N)$ -Matrix $P' = [p_{ij}]$, für deren Elemente

$$p_{ij} \geq 0 \quad \forall i, j \quad \text{und} \quad \sum_{j=1}^N p_{ij} = 1 \quad \forall i$$

gelten, heißt **stochastische Matrix**. Ihre Zeilenvektoren sind **stochastische Vektoren**.

Sind \bar{A} und \bar{B} stochastische Matrizen, ist auch $\bar{C} = \bar{A} \cdot \bar{B}$ eine stochastische Matrix.

Ein **gerichteter Graph** ist ein Mengenpaar (B_G, F_G) , wobei $B_G \neq \emptyset$ eine Zustandsmenge und $F_G \subseteq B_G \times B_G$ eine Menge von Übergängen ist. Wird jedem Übergang eine Übergangswahrscheinlichkeit p_{ij} mit

$$0 \leq p_{ij} \leq 1, \quad \sum_j p_{ij} = 1$$

zugeordnet, entsteht ein **bewerteter Graph**, der **Übergangsgraph** einer **homogenen Markoffkette**.

Jede homogene Markoffkette kann als Irrfahrt auf einem bewerteten Graphen interpretiert werden.

Ein **Zustand i** heißt **absorbierend**, wenn $p_{ii} = 1$ gilt. Die Menge R der absorbierenden Zustände heißt **Rand**. $Z - R$ ist die Menge der **inneren Zustände**. Eine **Markoffkette** heißt **absorbierend**, wenn $R \neq \emptyset$ und R von jedem inneren Zustand aus erreichbar ist.

Für eine absorbierende Markoffkette endet die Irrfahrt in einem Zustand des Randes R .

P_i : die Wahrscheinlichkeit vom Zustand i aus in U absorbiert zu werden,

$$P_i = \sum_{k=1}^N p_{ik} P_k$$

$$\vec{p}(k) = \vec{p}^T(0) \cdot P^k = \vec{p}^T(1) \cdot P^{k-1} = \vec{p}^T(2) \cdot P^{k-2} \dots$$

Mit $p^T(0)$: Zustandsvektor zu Beginn
 P : Übergangsmatrix (siehe links unten)

$$\vec{p}(t+1) = \vec{p}^T(t) \cdot P$$

mittlere Dauer m_i der Irrfahrt bis zur Absorption im Rand

$$m_i = 1 + \sum_{k=1}^N p_{ik} m_k \quad m_i = 0 \text{ auf dem Rand}$$

Zyklostationäre Prozesse

$$Z(t) = \sum_{n=-\infty}^{\infty} A(nT) g(t - nT) \quad (\text{SP}) \quad \begin{matrix} A(nT) \text{ stationär} \\ g \text{ reell} \end{matrix}$$

$$\varphi_{AA}(kT) = E\{A(nT)A^*((n+k)T)\}$$

$$E\{Z(t)\} = m_A \sum_{n=-\infty}^{\infty} g(t - nT)$$

$$\varphi_{ZZ}(t, t + \tau) = E\{Z(t)Z^*(t + \tau)\}$$

$$= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \varphi_{AA}((n-m)T) \cdot g(t - nT)g(t + \tau - mT)$$

(i) $E\{Z(t + kT)\} = E\{Z(t)\} \quad \forall k \in \mathbb{Z} \quad \text{und}$

(ii) $\varphi_{ZZ}(t + kT, t + \tau + kT) = \varphi_{ZZ}(t, t + \tau) \quad \forall k \in \mathbb{Z}$.

Folgerung: Sowohl $E\{Z(t)\}$ als auch $\varphi_{ZZ}(t, t + \tau)$ sind mit T periodisch.

Ein **stochastischer Prozeß $Z(t)$** , dessen Erwartungswert und dessen Autokorrelationsfunktion periodisch mit derselben Periode T sind, heißt (**schwach**) **zyklostationär**.

Die Autokorrelationsfunktion zyklostationärer Prozesse $\varphi_{ZZ}(t, t + \tau)$ hängt von $t + \tau$ und t ab!

$$\bar{\varphi}_{ZZ}(\tau) = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \varphi_{ZZ}(t, t + \tau) dt \quad \begin{matrix} \text{Über eine Periode} \\ \text{Gemittelte AKF} \end{matrix}$$

↓ Fourier-Transf.

$$\bar{\Phi}_{ZZ}(f) = \int_{-\infty}^{\infty} \bar{\varphi}_{ZZ}(\tau) e^{-j2\pi f\tau} d\tau \quad \begin{matrix} \text{Mittleres Leistungs-} \\ \text{dichtespektrum} \end{matrix}$$