

# WrStat

Braun & Co, Rast, Körner, Gwerder, Badertscher, Niedermann, Leuenberger, Walker

13. Dezember 2019

## Inhaltsverzeichnis

1	Kombinatorik	2
2	Ereignisse und ihre Wahrscheinlichkeit	3
3	Erwartungswert und Varianz	5
4	Wahrscheinlichkeitsverteilung	7
5	Schätzen <b>Skript S.??</b>	11
6	Hypothesentest <b>Skript S.??</b>	13
7	Prozessverbesserungen <b>Skript S.??</b>	15
8	Wichtige Formeln	16
9	Auswahl der Verteilung	16
10	Statistik / Wahrscheinlichkeit	17
11	Tabellen	18

# Kombinatorik

## 1.1 Produktregel

**Für-jedes-gibt-es-Regel**  $k$  Positionen müssen unabhängig von einander markiert werden, wobei  $n_i$  verschiedene Markierungen zur Verfügung stehen.

$$n_1 \cdot n_2 \cdots n_k = \prod_{i=1}^k n_i$$

**Beispiel:** Wie viele mögliche Augenzahlbilder können entstehen, wenn zwei verschiedenfarbige Würfel geworfen werden?

**Antwort:** Der erste Würfel kann  $n_1 = 6$  verschiedene Augenzahlen anzeigen, der zweite  $n_2 = 6$ . Da die beiden unabhängig sind gibt es  $n_1 \cdot n_2 = 36$  verschiedene Augenzahlbilder.

---

## 1.2 Permutation

**Grundfrage:** Auf wie viele Arten lassen sich  $n$  verschiedene Objekte anordnen? resp. Wie viele Permutationen von  $n$  Objekten gibt es?

$$P_n = n(n-1)(n-2) \dots 2 \cdot 1 = n! \quad \text{oder Rekursiv: } P_n = n \cdot P_{n-1}$$

**Beispiel:** In wie vielen verschiedenen Reihenfolgen können acht Läufer eines Rennens im Ziel eintreffen?

**Antwort:** Jede Reihenfolge ist möglich, also  $8! = 40'320$  mögliche Reihenfolgen.

---

## 1.3 Kombination

**Grundfrage:** Auf wie viele Arten kann  $k$  aus  $n$  verschiedenen Objekte auswählen.

$$\frac{n!}{k!(n-k)!} = C_k^n = \binom{n}{k}$$

**Beispiel:** Für ein Projekt stellt eine Firma mit 30 Mitarbeitern ein Team aus 5 Leuten zusammen. Auf wie viele Arten ist dies möglich?

**Antwort:** Es geht darum 5 von 30 Mitarbeitern auszuwählen, was auf  $\binom{30}{5} = 142'506$  möglich ist.

---

## 1.4 Variation

**Grundfrage:** Auf wie viele Arten kann man  $k$  mal unter  $n$  verschiedenen Objekten auswählen?

$$V_{n,k} = n^k$$



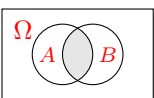
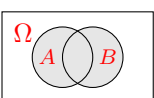
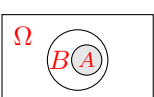
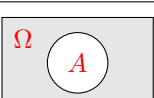
**Beispiel:** Auf wie viele Arten kann man eine Perlenkette der Länge  $k = 10$  aus  $n = 5$  Farben von Perlen herstellen?

**Antwort:** Die Variation Formel lässt sich auch über die Produktregel herleiten. Für jede Perle stehen wieder  $n$  verschiedene folge Perlen zur Auswahl.  $V_{n,k} = n^k = 5^{10} = 9'765'625$

---

## 2 Ereignisse und ihre Wahrscheinlichkeit

Begriff	Beschreibung	Modell
Elementarereignis	Der Ausgang eines Experiments	$\omega$
alle Elementarereignisse	Alle mögliche Ausgänge eines Experiments	$\Omega$
Ereignis	Teilmenge von $\Omega$ $A$ eingetreten $\Leftrightarrow$ Versuchsausgang $\omega \in A$	$A \subset \Omega$

Begriff	Beschreibung	Bild	Modell
Sicheres Ereignis	tritt immer ein		$\Omega$
Unmögliches Ereignis	kann nicht eintreten		$\emptyset = \{\}$
$A$ und $B$	Schnittmenge		$A \cap B$
$A$ oder $B$	Vereinigung		$A \cup B$
$A$ hat $B$ zur folge	$A$ ist in $B$ enthalten		$A \subset B$
nicht $A$	Komplementär Ereignis		$\Omega \setminus A$

### 2.1 Wahrscheinlichkeit & Rechenregeln

Wertebereich:	$0 \leq P(A) \leq 1$
Sicheres Ereignis:	$P(\Omega) = 1$
unmögliches Ereignis:	$P(\emptyset) = 0$
komplementär Ereignis:	$P(\bar{A}) = P(\Omega \setminus A) = 1 - P(A)$
Differenz der Ereignisse $A$ und $B$ :	$P(A \setminus B) = P(A) - P(A \cap B)$
Vereinigung zweier Ereignisse:	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$P(A) = \lim_{\text{Anzahl Versuche} \rightarrow \infty} \frac{\text{Anzahl Versuche bei der } A \text{ eingetreten ist}}{\text{Anzahl Versuche}}$$

### 2.2 Laplace-Experiment

In einem endlichen Wahrscheinlichkeitsraum  $\Omega$  haben alle Elementarereignisse die gleiche Wahrscheinlichkeit.

$$P(A) = \frac{|A|}{|\Omega|}$$

**Beispiele:** Münzen werfen wenn der Rand vernachlässigt wird, Würfeln...

## 2.3 Bedingte Wahrscheinlichkeit

Die Wahrscheinlichkeit für das Eintreten des Ereignisses  $A$  unter der Bedingung, dass das Ereignis  $B$  bereits eingetreten ist.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \underbrace{\frac{P(A) \cdot P(B)}{P(B)}}_{\text{nur wenn unabhängig}} = P(A)$$

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

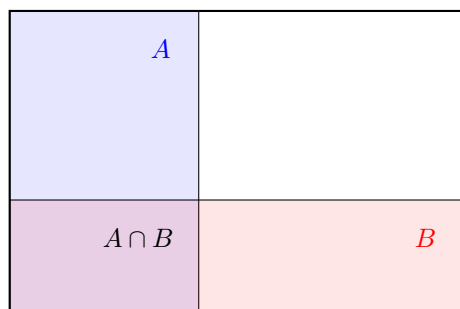
## 2.4 Unabhängige Ereignisse

Für sie gilt  $P(A \cap B) = P(A)P(B)$

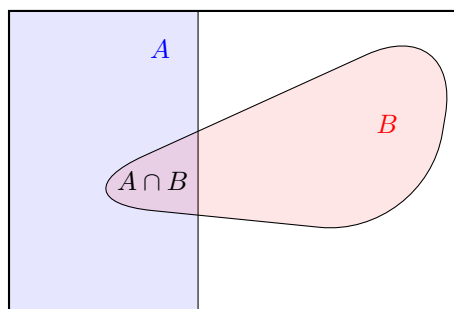
Die Tatsache, dass  $A$  eingetreten ist, hat keinen Einfluss auf die Wahrscheinlichkeit von  $B$ .

Unabhängige Ereignisse  $A$  und  $B$  liegen vor, wenn:  $P(A|B) = P(A|\bar{B})$

Wenn Ereignisse nicht gleichzeitig eintreten können, so sind sie abhängig.



Unabhängige Ereignisse



Abhängige Ereignisse

Beim Beispiel mit abhängigen Ereignissen wird  $A$  unwahrscheinlicher, wenn bereits  $B$  eingetroffen ist.

## 2.5 Totale Wahrscheinlichkeit

$$P(A) = \sum_{i=1}^n P(A|B_i) \cdot P(B_i)$$

in Matrixform:

$$\begin{pmatrix} P(A_1) \\ P(A_2) \\ \vdots \\ P(A_n) \end{pmatrix} = \underbrace{\begin{pmatrix} P(A_1|B_1) & P(A_1|B_2) & \dots & P(A_1|B_n) \\ P(A_2|B_1) & P(A_2|B_2) & \dots & P(A_2|B_n) \\ \vdots & \vdots & \ddots & \vdots \\ P(A_m|B_1) & P(A_m|B_2) & \dots & P(A_m|B_n) \end{pmatrix}}_{\text{W'keitsmatrix}} \cdot \begin{pmatrix} P(B_1) \\ P(B_2) \\ \vdots \\ P(B_n) \end{pmatrix}$$

## 2.6 Satz von Bayes

$$P(B|A) = P(A|B) \cdot \frac{P(B)}{P(A)}$$

Tauscht die Ereignisse der Bedingten Wahrscheinlichkeit.

## 2.7 Google Matrix

Es gibt folgende Ereignisse:  $P(S_i) = \{\text{Ein User ist auf der Seite } i\}$  und

$P(S'_j) = \{\text{Ein User ist nach einem Klick auf der Seite } j\}$

nach einiger Zeit ergibt sich ein Gleichgewicht  $P(S_i) = P(S'_j)$

$$P(S'_j) = P(S'_j|S_1)P(S_1) + P(S'_j|S_2)P(S_2) + \dots$$

$$\begin{pmatrix} P(S'_1) \\ P(S'_2) \\ \vdots \end{pmatrix} = \underbrace{\begin{pmatrix} P(S'_1|S_1)P(S_1) & P(S'_1|S_2)P(S_2) & \dots \\ P(S'_2|S_1)P(S_1) & P(S'_2|S_2)P(S_2) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}}_H \underbrace{\begin{pmatrix} P(S_1) \\ P(S_2) \\ \vdots \end{pmatrix}}_{p \text{ (Pagerank)}} \rightarrow Hp = p$$

Wird der freie Wille noch einberechnet, dann gilt:  $H' = \alpha H + \frac{1-\alpha}{\text{Anzahl Seiten}} \mathbf{1}$  wobei  $\mathbf{1}$  nur aus 1en besteht.

### 3 Erwartungswert und Varianz

#### 3.1 Erwartungswert

Sei  $X$  eine Funktion auf  $\Omega$ , und lasse sich  $\Omega$  in endlich viele Ereignisse  $A_i$  zerlegen, auf denen  $X(\omega)$  konstant ist, dann ist der Erwartungswert von  $X$

$$\text{Erwartungswert} = \sum \text{Wert} \cdot \text{Wahrscheinlichkeit}$$

$$E(X) = \sum_{i=0}^n \underbrace{X(A_i)}_{\text{Wert}} \cdot \underbrace{P(A_i)}_{\text{W'keit}}$$

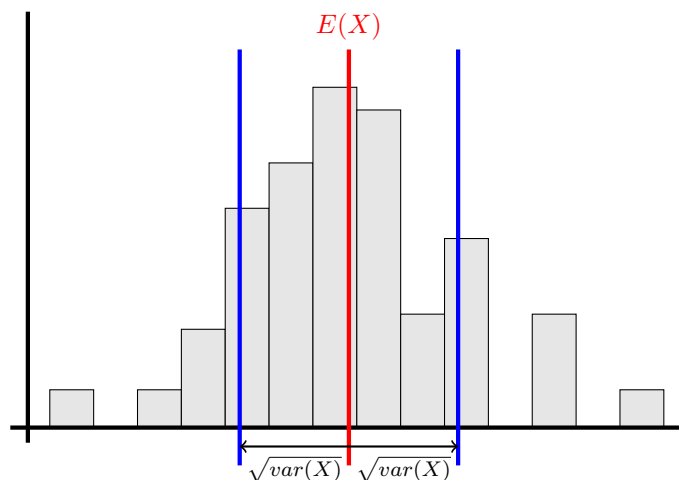
$$E(X) = \int_{-\infty}^x x \cdot \varphi(x) dx \text{ mit } \varphi(x) = \text{Dichtefunktion}$$

##### 3.1.1 Rechenregeln

$$E(X + Y) = E(X) + E(Y)$$

$$E(\lambda X + \mu) = \lambda \cdot E(X) + \mu \quad \lambda, \mu \in \mathbb{R}$$

$$E(XY) = E(X) \cdot E(Y) \quad \text{wenn } X, Y \text{ unabhängig sind}$$



#### 3.2 Varianz

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

$$\text{Standardabweichung } \sigma = \sqrt{\text{var}(X)}$$

**Achtung:** Wenn endliche Werte vorliegen muss die Formel der Stichprobenvarianz angewendet werden, Kapitel 5.2 Seite 11.

##### 3.2.1 Kovarianz

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y) = \underbrace{0}_{\text{falls } X, Y \text{ unabhängig}}$$

Ist die Kovarianz positiv so tendieren höhere  $X$ -Werte zu höheren  $Y$ -Werten.

##### 3.2.2 Rechenregeln

$$\text{var}(\lambda X) = \lambda^2 \text{var}(X) \quad \lambda, \mu \in \mathbb{R}$$

$$\text{var}(X_1 + X_2 + \dots + X_n) \neq \text{var}(nX)$$

$$\text{var}(X + Y) = \begin{cases} \text{var}(X) + \text{var}(Y) & (X, Y \text{ unabh.}) \\ \text{var}(X) + \text{var}(Y) + 2 \cdot \text{cov}(X, Y) & (X, Y \text{ abhängig}) \end{cases}$$

$$\text{var}(XY) = \text{var}(Y)\text{var}(X) + \text{var}(Y)E(X)^2 + \text{var}(X)E(Y)^2$$

#### 3.3 Erwartungswert und Varianz des arithmetischen Mittels

Es sei eine Folge von unabhängigen Zufallsvariablen  $X_1, X_2, \dots, X_n$  mit gleichem Erwartungswert  $\mu$  und gleicher Varianz  $\sigma^2$  gegeben.

$$\text{Mittelwert: } M_n = \frac{X_1 + \dots + X_n}{n}$$

$$\text{Erwartungswert: } E(X) = E(M_n) = \mu$$

$$\text{Varianz: } \text{var}(M_n) = \frac{1}{n} \text{var}(X) = \frac{\sigma^2}{n}$$

#### 3.4 Satz von Tschebyscheff

$$P(|X - E(X)| > \varepsilon) \leq \frac{\text{var}(X)}{\varepsilon^2} \quad \text{Wahrscheinlichkeit, dass } X \text{ um mehr als } \varepsilon \text{ vom Erwartungswert } E(X) \text{ abweicht.}$$

$$P(|M_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2 n} \quad \text{W'keit, dass } M_n \text{ von } n \text{ unab. ZV mit Mittelwert } \mu \text{ und Varianz } \sigma^2 \text{ mehr als } \varepsilon \text{ von } \mu \text{ abweicht.}$$

### 3.5 Regression

Allgemein:  $X, Y$  Zufallsvariable

Gesucht: Regressionsgerade  $y = ax + b$  mit min. Fehler

Fehler:  $E(Y - (aX + b)) = 0$

#### Regressionskoeffizient $r$

$r$  ist ein Mass für die Qualität der Regression (standardisiert)

$$r^2 = \frac{\text{cov}(X, Y)^2}{\text{var}(X)\text{var}(Y)} = a^2 \cdot \frac{\text{var}(X)}{\text{var}(Y)}$$

Liegt  $r$  nahe bei 1  $\hat{=}$  gute Approximation

#### Mittlerer quadratischer Fehler

$$\Delta^2 = \text{var}(Y)(1 - r^2) = \text{var}(Y) \left( 1 - \frac{\text{cov}(X, Y)^2}{\text{var}(X)\text{var}(Y)} \right)$$

#### Vorgehen: mit Fehlerberechnung

1. Tabelle mit bekannten Werten aufstellen:

Vorlage-Tabelle ist auf GitHub

$k$	$x$	$x^2$	$y$	$y^2$	$xy$
1	$x_1$	$x_1^2$	$y_1$	$y_1^2$	$x_1 y_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n$	$x_n$	$y_n^2$	$y_n$	$y_n^2$	$x_n y_n$
$\Sigma$	$\Sigma x_k$	$\Sigma x_k^2$	$\Sigma y_k$	$\Sigma y_k^2$	$\Sigma x_k y_k$
$E$	$\frac{\Sigma x_k}{n}$	$\frac{\Sigma x_k^2}{n}$	$\frac{\Sigma y_k}{n}$	$\frac{\Sigma y_k^2}{n}$	$\frac{\Sigma x_k y_k}{n}$

2. Varianzen, Kovarianz berechnen:

$$\text{var}(X) = E(X^2) - E^2(X)$$

$$\text{var}(Y) = E(Y^2) - E^2(Y)$$

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

3. Koeffizienten und Fehler der Gerade berechnen:

$$a = \frac{\text{cov}(X, Y)}{\text{var}(X)} \quad \Delta^2 = \text{var}(Y) \left( 1 - \frac{\text{cov}(X, Y)^2}{\text{var}(X)\text{var}(Y)} \right)$$

$$b = E(Y) - aE(X)$$

4. Gerade:

$$y = ax + b$$

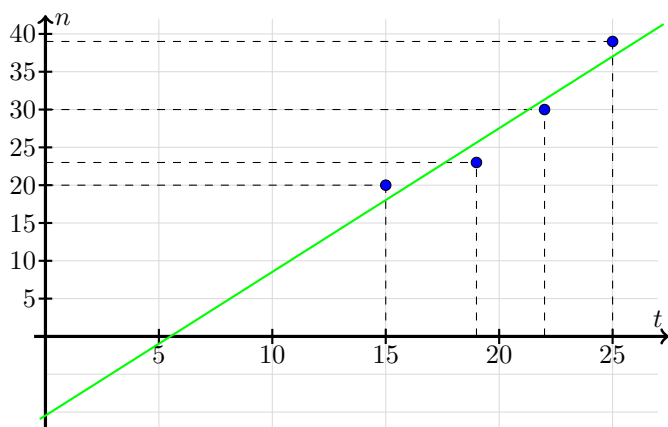
#### 3.5.1 Beispiel Regression

Je wärmer es ist, desto schneller zirpen die Grillen.

Folgende Daten wurden erhoben:

$T$ [°C]	$N$ [Zirpen/15 Sekunden]
15	20
19	23
22	30
25	39

Es wird vermutet, dass die Anzahl der Zirplaute pro 15 Sekunden linear von der Temperatur abhängt. Finden Sie eine solche Gesetzmässigkeit und beurteilen Sie ihre Qualität.



1. Tabelle ausfüllen:

$k$	$t$	$t^2$	$n$	$n^2$	$tn$
1	15	225	20	400	300
2	19	361	23	529	437
3	22	484	30	900	660
4	25	625	39	1521	975
$\Sigma$	81	1695	112	3350	2372
$E$	20.25	423.75	28	837.5	593

2. Varianzen und Kovarianzen berechnen:

$$\text{var}(t) = E(t^2) - E^2(t) = 423.75 - 20.25^2 = 13.6875$$

$$\text{var}(n) = E(n^2) - E^2(n) = 837.5 - 28^2 = 53.5$$

$$\text{cov}(t, n) = E(tn) - E(t)E(n) = 593 - 20.25 \cdot 28 = 26$$

3. Koeffizienten und Fehler der Gerade berechnen:

$$a = \frac{\text{cov}(t, n)}{\text{var}(t)} = \frac{26}{13.6875} = 1.8995$$

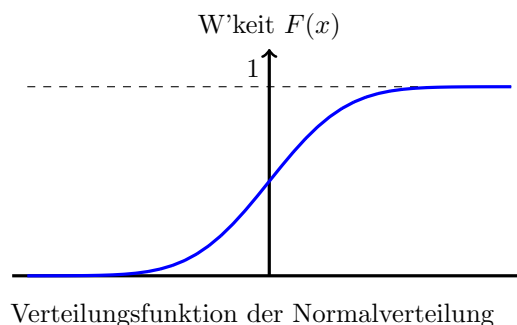
$$b = E(n) - aE(t) = 28 - 1.8995 \cdot 20.25 = -10.4649$$

$$r = \sqrt{\frac{\text{cov}(t, n)^2}{\text{var}(t)\text{var}(n)}} = \sqrt{\frac{26^2}{13.6875 \cdot 53.5}} = 0.9608$$

## 4 Wahrscheinlichkeitsverteilung

### 4.1 Verteilungsfunktion

allgemein	diskret	kontinuierlich
$P(X \leq x) = F(x)$	$= \sum_{k=-\infty}^x p_k$	$= \int_{-\infty}^x \varphi(\tilde{x}) d\tilde{x}$
$P(X > x) = 1 - P(X \leq x)$		
$P(a \leq X < b) = F(b) - F(a)$	$= \sum_{k=a}^b p_k$	$= \int_a^b \varphi(\tilde{x}) d\tilde{x}$



#### 4.1.1 Eigenschaften

Bei einem Sprung gilt: **Sprunghöhe = Wahrscheinlichkeit des Wertes x**

$$\underbrace{\mathbb{D}(F) = \mathbb{R}}_{\text{Definitionsbereich}}$$

$$\underbrace{\mathbb{W}(F) \in [0, 1]}_{\text{Wertebereich}}$$

$$F(-\infty) = 0$$

$$F(\infty) = 1$$

$$F(x) \text{ ist monoton steigend}$$

### 4.2 Wahrscheinlichkeitsdichte

**Dichtefunktion oder Wahrscheinlichkeitsdichte:**

$$\varphi(x) = F'(x)$$

**Bei Sprungstellen von F(x):**

$\varphi(x)$  = Dirac mit Gewichtung der Sprunghöhe

**Allgemein gilt:**

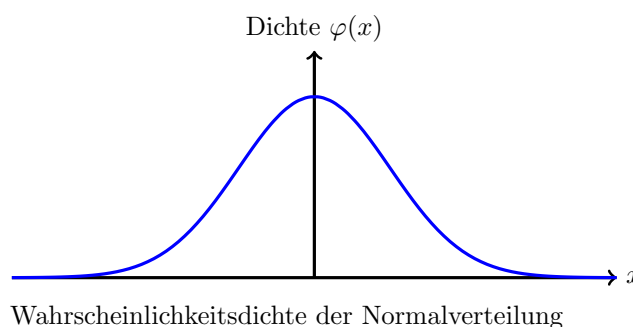
$$\int_{-\infty}^{\infty} \varphi(x) dx = 1$$

**Erwartungswert:** Wert mal W'keit

$$E(\mathbf{X}) = \int_{-\infty}^{\infty} \mathbf{x} \cdot \varphi(x) dx \quad \text{bzw.} \quad \sum_{-\infty}^{\infty} \mathbf{x} \cdot p_k$$

$$E(\mathbf{X}^2) = \int_{-\infty}^{\infty} \mathbf{x}^2 \cdot \varphi(x) dx \quad \text{bzw.} \quad \sum_{-\infty}^{\infty} \mathbf{x}^2 \cdot p_k$$

$$E(\mathbf{X}^N) = \int_{-\infty}^{\infty} \mathbf{x}^N \cdot \varphi(x) dx \quad \text{bzw.} \quad \sum_{-\infty}^{\infty} \mathbf{x}^N \cdot p_k$$



Wahrscheinlichkeitsdichte der Normalverteilung

### 4.3 Rechenregeln für $\varphi$ und $F$

**Gegeben:**  $X, Y$  Zufallsvariablen und  $\varphi_X, \varphi_Y$  bekannt

**Verteilungsfunktion:**

$$F_{X+a}(x) = F_X(x - a)$$

$$F_{\lambda X}(x) = F_X\left(\frac{x}{\lambda}\right)$$

$$F_{X+Y}(x) = F_X * \varphi_Y(y) = F_Y * \varphi_X(x)$$

$$F_{\sqrt{X}}(x) = F_X(x^2)$$

$$F_{X^2}(x) = F_X(\sqrt{x}) - F_X(-\sqrt{x})$$

**Dichte:**

$$\varphi_{X+a}(x) = \varphi_X(x - a)$$

$$\varphi_{\lambda X}(x) = \varphi_X\left(\frac{x}{\lambda}\right) \frac{1}{\lambda}$$

$$\varphi_{X+Y}(x) = \varphi_X * \varphi_Y(x)$$

$$\varphi_{\sqrt{X}}(x) = 2x\varphi_X(x^2)$$

$$\varphi_{X^2}(x) = \frac{1}{2}x^{-\frac{1}{2}}(\varphi_X(\sqrt{x}) + \varphi_X(-\sqrt{x}))$$

#### 4.3.1 Algorithmus Bsp.

1. Definition von  $F$  anwenden:  $F_{\lambda X}(x) = P(\underbrace{\lambda X \leq x}_*)$

2. Bedingung \* umformen:  $P(X \leq \frac{x}{\lambda}) = F_X(\frac{x}{\lambda})$

3. für Dichte:  $\frac{d}{dx}$

$$\varphi_{\lambda X}(x) = \frac{d}{dx} F_{\lambda X}(x) = \frac{d}{dx} F_X\left(\frac{x}{\lambda}\right) = \varphi_X\left(\frac{x}{\lambda}\right) \frac{1}{\lambda}$$

#### 4.3.2 Maximalwert eines Intervalls

$X_1, \dots, X_i$  sind auf dem Intervall  $[0, l]$  mit  $F_X(x)$  verteilt

$$M = \max\{X_1, \dots, X_i\}$$

$$F_M(x) = F_X(x)^n$$

#### 4.3.3 Median Skript S.??

Der Median  $med(X)$  von  $X$  ist eine Zufallsvariable, welche für  $F(med(X)) = \frac{1}{2}$  ist.

## 4.4 Normalverteilung

Viele kleine, unabhängige Zufallsvariable sammeln sich zu einer normalverteilten Zufallsvariable.

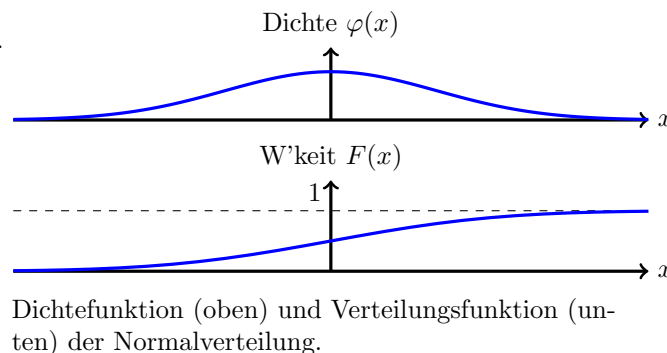
$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} = N(\mu; \sigma)$$

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \int_{-\infty}^x e^{-\frac{(\tilde{x}-\mu)^2}{2\sigma^2}} d\tilde{x}$$

Addieren von Normalverteilungen:

$$N(\mu_1; \sigma_1) + N(\mu_2; \sigma_2) = N(\mu_1 + \mu_2; \sqrt{\sigma_1^2 + \sigma_2^2})$$

Für  $F(-x)$  gilt  $F(-x) = 1 - F(x)$



### 4.4.1 Standardisierung

Erwartungswert:  $E(X) = \mu$  (=0 bei Standardnormalver.)

Varianz :  $var(X) = \sigma^2$  (=1 bei Standardnormalver.)

$$Z = \frac{X - \mu}{\sigma} \quad \text{mit } E(Z) = 0 \text{ und } var(Z) = 1$$

$$F_Z(z) = P(Z \leq z) = F_X(\sigma z + \mu)$$

$$F_X(x) = F_Z\left(\frac{x-\mu}{\sigma}\right)$$

$$\varphi_Z(z) = \sigma \cdot \varphi_X(\sigma z + \mu)$$

$$\varphi_X(x) = \frac{1}{\sigma} \cdot \varphi_Z\left(\frac{x-\mu}{\sigma}\right)$$

68% der Werte liegen im Intervall  $[\mu - \sigma, \mu + \sigma]$

95% in  $[\mu - 2\sigma, \mu + 2\sigma]$

99.7% in  $[\mu - 3\sigma, \mu + 3\sigma]$

0.5pt

**Rezept:** (Berechnen der W'keit das  $X \in [x_{min}, x_{max}]$  )

1. W'keit Formel hinschreiben

$$P(x_{min} \leq X \leq x_{max})$$

2. Standardisieren:

$$P_Z\left(\frac{x_{min} - \mu}{\sigma} \leq Z \leq \frac{x_{max} - \mu}{\sigma}\right)$$

3. In Verteilungsfunktion einfüllen:

$$P_Z = F\left(\frac{x_{max} - \mu}{\sigma}\right) - F\left(\frac{x_{min} - \mu}{\sigma}\right)$$

4. Werte aus der Tabelle auf Seite 18 auslesen:

**Bei negativen Werten:**

$$F(-0.7) = 1 - F(0.7) = 1 - 0.7580 = 0.2420$$

Ist der gesuchte Wert **zwischen zwei Werten** der Tabelle, dann der Wert aus den benachbarten Werten abschätzen. Bsp:

$$F(0.72) = 0.7642, \quad F(0.73) = 0.7673 \rightarrow F(0.725) \approx 0.765$$

**Beispiel:**

Ein Signal mit normalverteilten Werten  $X$  mit Erwartungswert 1 und Standardabweichung 2 wird über einen Eingangverstärker geleitet, der Werte zwischen  $\pm 3.3$  verarbeiten kann, bevor er übersteuert wird. Wie häufig wird der Verstärker übersteuert?

$$\mu = 1, \sigma = 2$$

$$1. P(-3.3 \leq X \leq 3.3)$$

$$2. P\left(\frac{-3.3 - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{3.3 - \mu}{\sigma}\right) \\ P\left(\frac{-3.3 - 1}{2} \leq Z \leq \frac{3.3 - 1}{2}\right) \\ P(-2.15 \leq Z \leq 1.15)$$

$$3. F(1.15) - F(-2.15) = F(1.15) - 1 + F(2.15)$$

$$4. 0.8749 + 0.9842 - 1 = 0.8591$$

### 4.4.2 Zentraler Grenzwertsatz

$X_1, X_2, \dots, X_n$  sind lauter identisch verteilte (nicht notwendig normalverteilt!) unabhängige Zufallsvariablen mit demselben Erwartungswert  $\mu$  und derselben Varianz  $\sigma^2$  und mit  $Z = \frac{X - \mu}{\sigma}$ . Dann hat die Summe

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i$$

den Erwartungswert  $n\mu$  und die Varianz  $n\sigma^2$ .

Die damit verbundene standardisierte ( $E(S_n) = 0, var(S_n) = 1$ ) Variable  $S_n$  ist somit wie folgt definiert:

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu}{\sigma} = \frac{1}{\sqrt{n} \cdot \sigma} \left[ \left( \sum_{i=1}^n X_i \right) - n\mu \right] = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

Für  $n \rightarrow \infty$  strebt die Verteilung von  $S_n$  gegen die Standardnormalverteilung.



## 4.5 Exponentialverteilung

Zur Ermittlung der Dauer bis zum Ausfall/Zerfall von Bauteilen/Stoffen ohne Gedächtnis

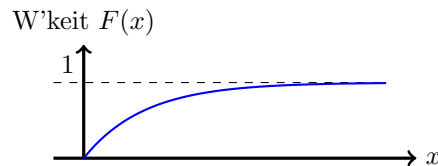
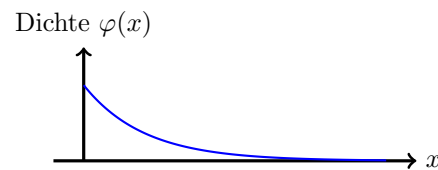
(W'keit, dass  $X$  in der nächsten Minute defekt geht = const.).

Beispiele :

- Lebensdauer von Atomen beim radioaktiven Zerfall
- Lebensdauer von Bauteilen, Maschinen & Geräten (MTBF - Mean Time Between Failure =  $\frac{1}{a}$ )

gilt  $P(X \leq t) = P(X \leq t_0 + t | X > t_0)$

$P(X > t) = P(X > t_0 + t | X > t_0)$



Dichtefunktion (oben) und Verteilungsfunktion (unten) der Exponentialverteilung.

Dichtefunktion und Verteilungsfunktion

$$\varphi(x) = \begin{cases} ae^{-ax} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad \begin{matrix} X < x \text{ Ausfall/Zerfall} \\ X > x \text{ kein Ausfall/Zerfall} \end{matrix}$$

$$F(x) = \begin{cases} 1 - e^{-ax} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Erwartungswert und Varianz

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

$$\text{var}(X) = \frac{1}{a^2}$$

Lebensdauer von mehreren **unabhängigen** Bauteilen

$$\begin{aligned} P(X > x) &= P(X_1 > x_1 \cap X_2 > x_2 \cap X_3 > x_3 \dots) \\ &= P(X_1 > x_1) \cdot P(X_2 > x_2) \cdot P(X_3 > x_3) \dots \\ &= (1 - P(X_1 \leq x_1)) \cdot (1 - P(X_2 \leq x_2)) \cdot (1 - P(X_3 \leq x_3)) \dots \\ &= (1 - F(x_1)) \cdot (1 - F(x_2)) \cdot (1 - F(x_3)) \dots \end{aligned}$$

## 4.6 Hypergeometrische Verteilung

Ist die Wahrscheinlichkeit dass in einer  $m$  Elemente umfassenden Stichprobe aus einer Grundgesamtheit von  $n$  Elementen, von denen  $r$  eine spezielle Eigenschaft besitzen,  $k$  Elemente mit der Eigenschaft zu finden sind.

$$p(k) = P(X = k) = \frac{\binom{r}{k} \binom{n-r}{m-k}}{\binom{n}{m}} \quad \text{für } 0 \leq k \leq r \text{ und } k \leq m$$

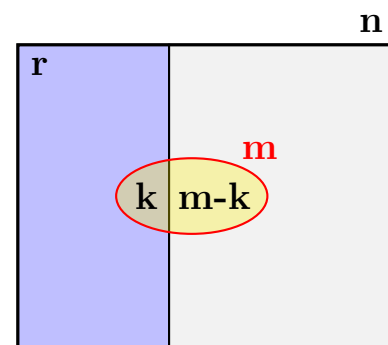
Erwartungswert:  $E(X) = m \frac{r}{n}$

Varianz:  $\text{var}(X) = m \frac{r(n-r)(n-m)}{n^2(n-1)}$

**Beispiel:**

Lotto,  $n = 45$  Zahlen,  $r = 6$  (die gezogenen Zahlen),  $m = 6$  (meine Zahlen)

$$P(X = 4) = P(\text{Ein Vierer}) = \frac{\binom{6}{4} \binom{39}{2}}{\binom{45}{6}} = 0.001364$$



## 4.7 Poissonverteilung **Skript S.??**

$$P_\lambda(k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{W'keit dass } k \text{ Ereignisse im Intervall } [0, x] \text{ auftreten}$$

Erwartungswert:  $E(X) = \lambda$

Varianz:  $\text{var}(X) = \lambda$

$$P(X < k) \leq \sum_0^k P_\lambda(k) = \sum_0^k \frac{\lambda^k}{k!} e^{-\lambda}$$

$$P(X > k) = 1 - P(X < k)$$

$x$  = Anzahl Versuche

$\lambda$  = Ereignisse pro Intervall im Mittel

### Anwendungsbeispiele:

Für die Häufigkeiten seltener Ereignisse. Anzahl Anrufe bei einer Telefonzentrale in einer gewissen Periode. Anzahl grosse Versicherungsschäden in einer gewissen Periode. Anzahl Jobs, die bei einem Server ankommen. Anzahl Ereignisse in einem Zeitintervall. Anzahl Lokomotiven der SBB, die in der nächsten Woche einen Defekt haben. Anzahl der Gewinner mit 4 Richtigen im Lotto.

## 4.8 Binomialverteilung **Skript S.??**

Wird angewendet bei einem Experiment mit nur zwei Ausgängen (Ereignis mit W'keit  $p$  tritt ein, Ereignis tritt nicht ein). Eine Zufallsvariable mit diskreten Werten  $k \in \{0, \dots, n\}$  heisst binomialverteilt zum Parameter  $p$ , wenn die Wahrscheinlichkeit des Wertes  $k$  wie folgt ist:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \mu = E(X) = p \cdot n \quad \sigma^2 = \text{var}(X) = n \cdot p(1-p)$$

$n$ : Versuche       $k$ : k-mal erfolgreich       $p$ : Wahrscheinlichkeit

Approximation mit Normalverteilung:  $P(a \leq x \leq b) \simeq \underbrace{P(a - 0.5 \leq x \leq b + 0.5)}_{\text{Normalverteilung}}$

**Beispiel:** Wie hoch ist die Wahrscheinlichkeit, dass bei 350 Leuten genau  $k$  ( $k \leq 350$ ) heute Geburtstag haben?

$$P(k) = \binom{350}{k} \left(\frac{1}{365}\right)^k \left(\frac{364}{365}\right)^{350-k}$$

## 4.9 Gleichverteilung

**Stetig **Skript S.??****

$$\varphi(x) = \begin{cases} 0 & x < a \\ \frac{1}{b-a} & x \in [a, b] \\ 0 & x > b \end{cases}$$

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b] \\ 1 & x > b \end{cases}$$

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

Varianz:  $\text{var}(X) = \frac{(b-a)^2}{12}$

**Diskret **Skript S.??****

$$F(x) = \begin{cases} 0 & x \leq 1 \\ \frac{|x|}{n} & 1 \leq x \leq n \\ 1 & x \geq n \end{cases}$$

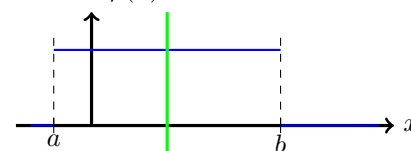
Wahrscheinlichkeit:  $p(k) = \frac{1}{n}$

Erwartungswert:  $E(X) = \frac{n+1}{2}$

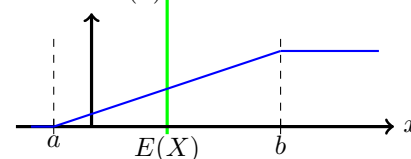
Varianz:  $\text{var}(X) = \frac{n^2-1}{12}$

$E(X^2) = \frac{2n^2+3n+1}{6}$

Dichte  $\varphi(x)$



W'keit  $F(x)$



Verteilungsfunktion(oben) und Wahrscheinlichkeitsdichte (unten) der Gleichverteilung

## 4.10 Potenzgesetze (Power-Laws) **Skript S.??**

Die Normalverteilung beschreibt (physikalische) Grössen, die vor allem in einer bestimmten Grössenordnung vorkommen. Die Potenzgesetze dienen dazu, Grössen welche einen grossen Wertebereich annehmen können, zu beschreiben.

$$\varphi(x) = \begin{cases} \frac{\alpha-1}{x_{min}^{1-\alpha}} \cdot x^{-\alpha} & x > x_{min} \\ 0 & \text{sonst} \end{cases} \quad \text{mit } \alpha > 1$$

$$E(X) = \frac{\alpha-1}{\alpha-2} \cdot x_{min}$$

$$E(X^2) = \frac{\alpha-1}{\alpha-3} \cdot x_{min}^2$$

$$\text{var}(X) = \left( \frac{\alpha-1}{\alpha-3} - \left( \frac{\alpha-1}{\alpha-2} \right)^2 \right) \cdot x_{min}^2$$

$$x_{\frac{1}{2}} = 2^{\frac{1}{\alpha-1}} \cdot x_{min}$$

Eine nach dem Potenzgesetz verteilte Zufallsvariable kann man daran erkennen, dass die Dichtefunktion in doppelt logarithmischer Darstellung eine Gerade ist. Wegen  $\log p(x) = -\alpha \log x + \log C$  ist die Steigung der Geraden  $-\alpha$ .

Der Parameter  $\alpha$  kann mithilfe eines Maximum-Likelihood Schätzers bestimmt werden. Es gilt:  $\alpha = 1 + \frac{n}{\sum_{i=1}^n \log \frac{x_i}{x_{min}}}$

## 5 Schätzen **Skript S.??**

### 5.1 Konsistente Schätzer **Skript S.??**

Ein Schätzer ist konsistent, wenn  $\lim_{n \rightarrow \infty} = E(X)$  ergibt

Der Mittelwert der Stichprobe ist ein konsistenter Schätzer.  $\lim_{n \rightarrow \infty} = \frac{X_1 + \dots + X_n}{n} = E(X)$

Der Schätzer  $\bar{X} = \frac{X_1 + \dots + X_n}{n}$  heisst der Stichprobenmittelwert der Stichprobe  $X_1, \dots, X_n$ .

### 5.2 Erwartungstreue Schätzer **Skript S.??**

Ein Schätzer ist erwartungstreu, wenn  $E(\text{Schätzer}) = E(\text{realer Wert})$

Ist der Stichprobenmittelwert ein konsistenter Schätzer, aber er ist sogar erwartungstreu:  $E(\mu(X_1, \dots, X_n)) = \frac{E(X_1) + \dots + E(X_n)}{n} = E(X)$

Erwartungstreue Schätzer für  $var(x)$  ist:

$$S^2 = \frac{n}{n-1} \left( \underbrace{\frac{1}{n} \sum X_i^2}_{E(X^2)} - \underbrace{\left( \frac{1}{n} \sum X_i \right)^2}_{E(X)^2} \right)$$

Stichprobenvarianz, empirische Varianz

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$\bar{X} = M_n$  heisst Stichprobenmittelwert

#### 5.2.1 Kleinstmöglicher Fehler

$$E\left(\left(E(X) - \frac{x_1 + \dots + x_n}{n}\right)^2\right) = \text{minimal}$$

### 5.3 Maximum Likelihood Schätzer **Skript S.??**

Sinn des Likelihoodschäzers ist einen unbekannten Parameter  $\vartheta$  einer Dichtefunktion  $\phi(x, \vartheta)$  zu schätzen.

$$L(x_1, \dots, x_n; \vartheta) = \phi(x_1, \vartheta) \cdot \dots \cdot \phi(x_n, \vartheta) \implies \frac{d}{d\vartheta} L(x_1, \dots, x_n; \vartheta) = 0 \implies \vartheta = ? \text{ (Maximum-Likelihood-Schätzer)}$$

Für eine normalverteilte Grösse lautet die Likelihood Funktion:  $L(x_1, \dots, x_n; \vartheta) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \vartheta)^2}$

Der unbekannte Parameter  $\vartheta$  kann nun durch suchen des Maximums der Funktion ermittelt werden ( $\vartheta$  wird variiert). Die Funktion wird maximal, wenn die Summe im Exponent minimal wird. Das  $\vartheta$ , das die Summe minimiert, kann durch **ableiten nach  $\vartheta$  und null setzen ermittelt** werden. Es können auch Stichprobenvarianz  $S^2$  oder Ähnliches ermittelt werden.

### 5.4 Verteilung der Schätzwerte **Skript S.??**

$X_1, \dots, X_n$  sind unabhängige, normalverteilte Zufallsvariablen mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$ . Dann gilt

1.  $\bar{X}$  und  $S^2$  sind unabhängig
2.  $\bar{X} = \frac{X_1 + \dots + X_n}{n}$  ist normalverteilt mit  $E(\bar{X}) = \mu$  und  $var(\bar{X}) = \frac{\sigma^2}{n}$
3.  $\frac{n-1}{\sigma^2} \cdot S^2$  ist  $\chi_{n-1}^2$ -verteilt mit  $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

## 5.5 Konfidenzintervall von Messwerten

### 5.5.1 Konfidenzintervall Skript S.??

Ein Intervall  $[L(X_1, \dots, X_n), R(X_1, \dots, X_n)]$  heisst ein  $1 - \alpha$ -Konfidenzintervall für den Parameter  $\vartheta$ , wenn der wahre Wert des Parameters  $\vartheta$  höchstens mit Wahrscheinlichkeit  $\alpha$  ausserhalb des Intervalls liegt.

Es gilt:  $P(L \leq \vartheta \leq R) = 1 - \alpha$

### 5.5.2 Bei bekannter Varianz $\sigma^2$

Falls Varianz  $\sigma^2$  von Messwerten bekannt ist, handelt es sich bei  $\bar{X}$  um **normalverteilte** Zufallsvariable mit Varianz  $\sigma^2/n$ .

Also kann sehr einfach ein  $x$  für das Konfidenzintervall gefunden werden:  $P\left(\left|\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right| \leq x\right) = 1 - \alpha \Rightarrow F(x) = 1 - \frac{\alpha}{2}$

Daraus ergibt sich folgendes Konfidenzintervall  $\mu \in \left[\bar{X} - x \frac{\sigma}{\sqrt{n}}, \bar{X} + x \frac{\sigma}{\sqrt{n}}\right]$  mit W'keit  $1 - \alpha$

### 5.5.3 Bei geschätzter Varianz $S^2$

#### t-Verteilung

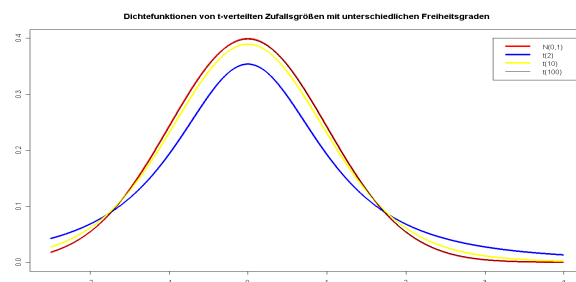
Der Mittelwert  $\left(\frac{x_1 + \dots + x_n}{n}\right)$  normalverteilter Daten ist t-Verteilt, wenn Varianz mit Stichprobenvarianz geschätzt wurde. Ab einer gewissen Anzahl Messungen ( $n \geq 30$ ) kann näherungsweise auch mit der Normalverteilung gerechnet werden.

#### Checkliste

- 1)  $\bar{X}, S$  als Schätzungen aus  $x_i$  bestimmen
- 2)  $t$  aus  $t$ -Tabelle ( $k = n - 1$ ) für  $1 - \frac{\alpha}{2}$  = W'keit für eine Seite
- 3) Intervall  $\left[\bar{X} - t \frac{S}{\sqrt{n}}, \bar{X} + t \frac{S}{\sqrt{n}}\right]$ ,  $(1 - \alpha)$  Konfidenzintervall

#### Anwendung

$\frac{\bar{X} - \mu}{S/\sqrt{n}}$  t-Verteilt



$\lim_{x \rightarrow \infty} = 0$  aber langsamer wie bei Gaussverteilung

Beispiel: 10 Messungen ergeben Durchschnittswert 4,7 und eine Standardabweichung 0,1. Finde ein 99% Konfidenzintervall für  $\mu$ .

Finde t:

$k = n - 1$	...	0,995
$\vdots$	$\vdots$	$\vdots$
9	...	3,2498

$$\left[\bar{X} - 3,2498 \frac{S}{\sqrt{n}}, \bar{X} + 3,2498 \frac{S}{\sqrt{n}}\right] \Rightarrow \left[4,7 - 3,2498 \frac{0,1}{\sqrt{10}}, 4,7 + 3,2498 \frac{0,1}{\sqrt{10}}\right]$$

$\mu \in [4,5072, 4,8028]$  mit Wahrscheinlichkeit 99%

## 6 Hypothesentest **Skript S.??**

### 6.1 Grundsätze

- “Man braucht Aussagen, die man widerlegen könnte.”
- Irrtum ist möglich (Irrtumsw'keit  $\alpha$ )
- Beweis durch Widerlegen des Gegenbeweises

### 6.2 Vorgehen (Nullhypothese)

1. Hypothese, die der Test **widerlegen** soll  
→ Nullhypothese  $\Rightarrow$  Keine Wirkung/Effekt
2. Festlegung der Irrtumswahrscheinlichkeit  $\alpha = 0.05, 0.01, \dots$  (Niveau= $1-\alpha$ )
3. Testgrösse  $X$ , W'keitsverteilung  $\rightarrow$  Nur Werte bis zum getesteten Ereignis betrachten! (Nicht Zukunft mit einbeziehen)
4. Bestimmung der Schranken  $x_{krit}$  für:
  - Einseitiger Test  $P(X > x_{krit}) = \alpha$
  - Zweiseitiger Test  $P(|X| > x_{krit}) = \alpha/2$
5. Wert für  $F\left(\frac{x_{krit}-\mu}{\sigma}\right)$  aus Tabelle 11.1 (S. 18)
6. Falls Messungen ergeben  $X > x_{krit} \Rightarrow$  Hypothese **falsch** mit W'keit  $1 - \alpha$

### 6.3 Testen einer diskreten Verteilung **Skript S.??**

#### 6.3.1 $\chi^2$ -Test mit k-möglichen Ausgängen

Mögliche Ausgänge:  $I_i, i = 1, \dots, k$

Wahrscheinlichkeit von Ausgang  $i$ :  $P(X \in I_i) = p_i$

$n$  Beobachtungen, davon jeweils  $n_i$  mit Ausgang  $i$

$$D = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i} \quad \text{ist } \chi_{k-1}^2, \text{ mit } k-1 \text{ Freiheitsgrade}$$

#### Durchführung des $\chi^2$ -Tests

1. Daten erfassen: Damit der Test optimal funktioniert muss folgende Bedingung erfüllt sein:  $n_i \geq 5 \forall i$

2. Diskrepanz  $D$  berechnen

$i$	Ausgang	$p_i$	$n_i$	$(n_i - np_i)$	$(n_i - np_i)^2 / np_i$
1					
2					
3					
4					
5					
			$n$		$D = \sum$

3. Schwellenwert für  $D$

$x_{1-\alpha}$  für  $F_{\chi_{k-1}^2}(x) = 1 - \alpha$  aus  $\chi_{k-1}^2$ -Tabelle lesen.

Wenn  $\alpha$  nicht vorgegeben, dann  $\alpha$  z.B. 0.1 oder 0.05 wählen.

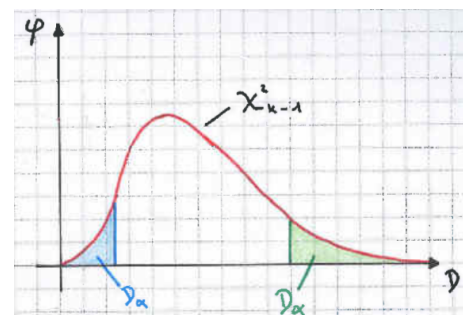
Anzahl Freiheitsgrade = Anzahl Ausgänge  $- 1 = k - 1$

$p = 1 - \alpha$        $\alpha$  = Irrtumswahrscheinlichkeit

$D \geq x_{1-\alpha}$  Hypothese ist unwahrscheinlich!

$D < x_{1-\alpha}$  Hypothese nicht widerlegbar!

$D \leq x_\alpha$  Daten möglicherweise "fabriziert"!



- Nachteile:
- Grob verpixeltes Bild der Verteilung
  - wenn wenig Messwerte  $\Rightarrow$  geringe Aussagekraft

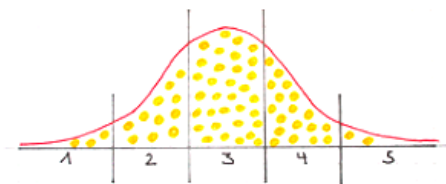
## 6.4 Testen einer stetigen Verteilung **Skript S.??**

### 6.4.1 Mit $\chi^2$ -Test

Der  $\chi^2$ -Test kann im Prinzip nur diskrete Wahrscheinlichkeitsverteilungen testen.

→ Will man eine stetige Verteilung testen, muss man zunächst Klassen von Werten bilden. → Anschliessend deren Wahrscheinlichkeiten berechnen und dann prüfen ob die künstlich diskrete Verteilung im  $\chi^2$ -Test Bestand hat.

⇒ Achtung: Durch die Klassenbildung wird eine künstliche Diskretisierung eingeführt.

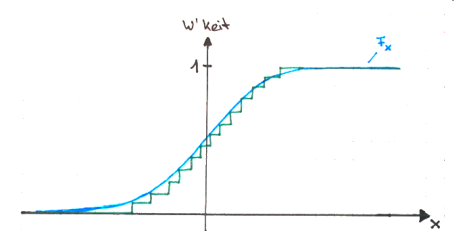


### 6.4.2 Kolmogorov-Smirnov Test

**Nullhypothese** Messwerte  $x_1, \dots, x_n$  haben Verteilungsfunktion  $F_X$

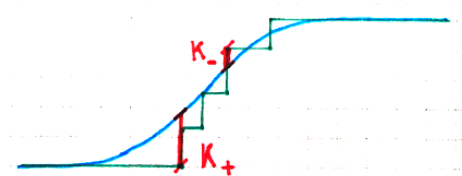
**Idee:** Vergleiche Verteilungsfunktionen (statt der Dichtefunktionen)

**Daten:**  $X$  Zufallsvariable, Verteilungsfunktion  $F_X$ ,  $n$  Messungen ergeben Stichprobe  $x_i$



$F_X$  theoretische Verteilungsfunktion von  $X$

empirische Verteilungsfunktion  $\frac{\text{Anzahl}\{x_i \leq x\}}{n}$



Testgrösse

$$\max (F_{emp}(x) - F_X(x))$$

$$\min (F_{emp}(x) - F_X(x))$$

#### Durchführen des Kolmogorov-Smirnov Test

1. Werte  $x_1, \dots, x_n$  in aufsteigender Reihenfolge sortieren
2.  $K_n^\pm$  berechnen

Mit Tabelle:

$$K_n^+ = \sqrt{n} \max_{1 \leq i \leq n} \left( \frac{i}{n} - F_X(x_i) \right)$$

$$K_n^- = \sqrt{n} \max_{1 \leq i \leq n} \left( F_X(x_i) - \frac{i-1}{n} \right)$$

$i$	$x_i$	$i/n$	$F_X(x_i)$	$(i-1)/n$
1	$\min(x_i)$			
	$\vdots$			
n	$\max(x_i)$			

3. Finde  $t_{n,1-\alpha}, t_{n,\alpha}$  in der Tabelle 11.3 (S. 19)
4. Falls  $K_n^+ > t_{n,1-\alpha}$  oder  $K_n^- < t_{n,\alpha}$ , verwerfe die Hypothese, dass  $X$  die Verteilungsfunktion  $F_X$  hat.

## 6.5 Vergleichen von Mittelwerten (t-Test) **Skript S.??**

**Nullhypothese**  $\Rightarrow E(X) = E(Y)$

**Daten**  $\mathbf{n}$  Messungen von  $x_i \Rightarrow \bar{X}; S_X^2$

$\mathbf{m}$  Messungen von  $y_i \Rightarrow \bar{Y}; S_Y^2$

**Testgrösse**  $T = \frac{\bar{X} - \bar{Y}}{\sqrt{(n-1) \cdot S_X^2 + (m-1) \cdot S_Y^2}} \cdot \sqrt{\frac{nm(n+m-2)}{n+m}}$

$T$  ist t-verteilt

**Schwellenwert**  $t_{krit}$  aus Tabelle 11.4 mit Freiheitsgrad  $n+m-2$  und  $p = 1-\alpha$

**Test** Falls  $T > t_{krit}$  wird die Hypothese  $E(X) = E(Y)$  verworfen



### 7.2.2 Vorgehen

1. Bestimmung des Zustandsvektors  $x_k$  und des Messvektors  $z_k$
2. Aufstellen der Messmatrix  $H_k$  aus:  $z_k = H_k x_k$
3. Berechnen der Systementwicklung  $\varphi_k$  damit  $x_{k+1} = \varphi_k x_k$
4. Aufstellen der Fehler-Kovarianzmatrixen  $Q_k$  und  $R_k$

5. **Vorhersage-Schritt:**  $\hat{x}_{k+1|k} = \varphi_k \hat{x}_k \quad P_{k+1|k} = \varphi_k P_k \varphi_k^t + Q_k$

6. **Korrektur-Schritt:**  $\hat{x}_k = (I - K_k H_k) \hat{x}_{k|k-1} + K_k z_k \quad P_k = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^t + K_k R_k K_k^t$

7. **Bestimmung des optimalen  $K_k$ :**  $K_k = P_{k|k-1} H_k^t (H_k P_{k|k-1} H_k^t + R_k)^{-1}$

8. Damit wird  $P_k$  vereinfacht zu:  $P_k = (I - K_k H_k) P_{k|k-1} \rightarrow P_k$  bleibt variabel

### 7.2.3 Vereinfachung

In der Praxis kann meist davon ausgegangen werden, dass  $\varphi_k$ ,  $R_k$ ,  $Q_k$  und  $H_k$  konstant sind. Damit werden auch  $P_k$  und  $K_k$  konstant und die Vorhersage wird vereinfacht zu:  $\hat{x}_k = \hat{x}_{k|k-1} + K(z_k - H \hat{x}_{k|k-1})$

## 8 Wichtige Formeln

### 8.1 Reihenentwicklungen

<b>Geometrische Reihe</b>	$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad  x  < 1$
	$\sum_{k=0}^{\infty} k x^k = x \sum_{k=1}^{\infty} k x^{k-1} = \frac{x}{(1-x)^2} \quad x \neq 1$
<b>Binominalreihe</b>	$\sum_{n=0}^{\infty} \binom{\alpha}{n} x^n = (1+x)^\alpha \quad x \in (-1, 1)$
<b>E-Funktion</b>	$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$

## 9 Auswahl der Verteilung

Gleichverteilung	Alle möglichen Werte haben die gleiche Wahrscheinlichkeit
Exponentialverteilung	Dauer von zufälligen Zeitintervallen ohne Gedächtnis
Normalverteilung	Viele kleine, unabhängige Zufallsprozesse sammeln sich zu einer normalverteilten Zufallsvariable
Binomialverteilung	Experiment mit zwei Ausgängen
Hypergeometrische Verteilung	Wahrscheinlichkeit, dass in einer Stichprobe, Elemente einer Teilmenge zu finden sind
Poissonverteilung	Häufigkeiten seltener Ereignisse
t-Verteilung	Varianz aus Stichproben geschätzt (für grosse $n$ kann Normalverteilung verwendet werden)
$\chi^2$ -Verteilung	Test ob eine Zufallsvariable Normalverteilt ist. Anwendung bei Hypothesentests



## 10 Statistik / Wahrscheinlichkeit

### 10.1 Funktionen

<i>mean</i> ({...})	Berechnet das arithmetische Mittel der Elemente der Liste.
<i>mean</i> ({...}, {...})	Mit einer zweiten Liste lassen sich die Elemente einzeln gewichten.
<i>mean</i> ( <i>A</i> )	Gibt einen Zeilenvektor mit den arith. Mitteln der Spalten zurück.
<i>mean</i> ( <i>A</i> , <i>B</i> )	Mit einer Matrix <i>B</i> lassen sich die Elemente von <i>A</i> gewichten.
<i>median</i> ({...})	Berechnet den Median der Elemente der Liste.
<i>median</i> ( <i>A</i> )	Gibt einen Zeilenvektor mit den Medianwerten der Spalten zurück.
<i>stdDev</i> ({...})	Berechnet die Standardabweichung $\sigma$ der Liste
<i>variance</i> ({...})	Berechnet die Varianz $\sigma^2$ der Liste
<i>nCr</i> ( <i>n</i> , <i>k</i> )	Binominalkoeffizient $\binom{n}{k}$ - funktioniert auch für Listen und Matrizen
<i>nPr</i> ( <i>n</i> , <i>k</i> )	Anzahl Möglichkeiten unter Berücksichtigung der Reihenfolge <i>k</i> Elemente aus <i>n</i> auszuwählen.
<i>OneVar</i> <i>L1</i> , [ <i>L2</i> ], [ <i>L3</i> ], [ <i>L4</i> ] <i>ShowStat</i>	Berechnet die Statistiken der Liste <i>L1</i> . Die Statistik wird mit <i>ShowStat</i> eingeblendet. Folgende Werte werden berechnet: $\bar{x}$ , $\sum x$ , $\sum x^2$ , $\sigma x$ , ... Optionale Listen: <i>L2</i> : Häufigkeit, <i>L3</i> : Klassencodes, <i>L4</i> : Klassenliste
<i>TwoVar</i> <i>L1</i> , <i>L2</i> , [ <i>L3</i> ], [ <i>L4</i> ], [ <i>L5</i> ] <i>ShowStat</i>	Gleich wie <i>OneVar</i> , einfach für 2 Variablen. <i>L1</i> , <i>L2</i> : Variablen <i>X</i> und <i>Y</i> , <i>L3</i> – 5: wie bei <i>OneVar</i>

### 10.2 Regression

Zur Berechnung einer Regression muss eine Liste ({...}) die x-Werte enthalten und eine zweite Liste die y-Werte. Der Befehl *LinReg* *L1*, *L2* berechnet die lineare Regression. Mit *ShowStat* werden die berechneten Werte angezeigt. Es ist auch möglich, die Datenpunkte und die Regressionskurve zu plotten: *Regeq*(*x*)  $\rightarrow$  *y1*(*x*) und *NewPlot* 1, 1, *L1*, *L2* Optional können weitere Listen angegeben werden: *L3*: Häufigkeit, *L4*: Klassencodes, *L5*: Klassenliste, wobei alle Listen ausser *L5* die gleiche Dimension besitzen müssen. 'Iterationen' gibt die maximale Anzahl Lösungsversuche an. (standardmässig: 64)

Lineare Regression	<i>LinReg</i> <i>L1</i> , <i>L2</i> , [ <i>L3</i> ], [ <i>L4</i> , <i>L5</i> ]
Logarithmische Regression	<i>LnReg</i> <i>L1</i> , <i>L2</i> , [ <i>L3</i> ], [ <i>L4</i> , <i>L5</i> ]
Logistische Regression	<i>Logistic</i> <i>L1</i> , <i>L2</i> , [Iterationen], [ <i>L3</i> ], [ <i>L4</i> , <i>L5</i> ]
Potenz-Regression	<i>PowerReg</i> <i>L1</i> , <i>L2</i> , [ <i>L3</i> ], [ <i>L4</i> , <i>L5</i> ]
Quadratische Polynomische Regression	<i>QuadReg</i> <i>L1</i> , <i>L2</i> , [ <i>L3</i> ], [ <i>L4</i> , <i>L5</i> ]
Kubische Regression	<i>CubReg</i> <i>L1</i> , <i>L2</i> , [ <i>L3</i> ], [ <i>L4</i> , <i>L5</i> ]
Polynomische Regression 4-ter Ordnung	<i>QuartReg</i> <i>L1</i> , <i>L2</i> , [ <i>L3</i> ], [ <i>L4</i> , <i>L5</i> ]

### 10.3 Zufallszahlen

<i>RandSeed</i> 1147	Setzt die Ausgangsbasis (Seed) für den Zufallszahl-Generator
<i>rand</i> ()	Gibt eine Zufallszahl zwischen 0 und 1 zurück.
<i>rand</i> ( <i>n</i> )	Gibt eine Zufallszahl zwischen 0 und <i>n</i> (für <i>n</i> pos.) bzw. zwischen <i>n</i> und 0 (für <i>n</i> neg.) zurück.
<i>randMat</i> ( <i>n</i> , <i>m</i> )	Erzeugt eine ganzzahlige Matrix mit <i>n</i> Zeilen und <i>m</i> Spalten mit Werten $-9 < x < +9$ .
<i>randNorm</i> ( <i>a</i> , <i>sd</i> )	Gibt eine reelle Zufallszahl um den Mittelwert <i>a</i> mit der Standardabweichung <i>sd</i> aus.
<i>randPoly</i> ( <i>x</i> , <i>n</i> )	Erzeugt ein Polynom der Variable <i>x</i> der Ordnung <i>n</i> mit Koeffizienten $-9 < x < +9$

## 11 Tabellen

© Prof. Dr. Andreas Müller

### 11.1 Quantilen der Normalverteilung

$p$	$x$
0.75	0.6745
0.8	0.8416
0.9	1.2816
0.95	1.6449
0.975	1.9600
0.99	2.3263
0.995	2.5758
0.999	3.0902
0.9995	3.2905

### 11.2 Verteilungsfunktion der Normalverteilung

$x$	+0.00	+0.01	+0.02	+0.03	+0.04	+0.05	+0.06	+0.07	+0.08	+0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

### 11.3 Quantilen für den Kolmogorov-Smirnov-Test

$n$	$p = 0.01$	$p = 0.05$	$p = 0.1$	$p = 0.25$	$p = 0.5$	$p = 0.75$	$p = 0.9$	$p = 0.95$	$p = 0.99$
1	0.01000	0.05000	0.10000	0.25000	0.50000	0.75000	0.90000	0.95000	0.99000
2	0.01400	0.06749	0.12955	0.29289	0.51764	0.70711	0.96700	1.09799	1.27279
3	0.01699	0.07919	0.14714	0.31117	0.51469	0.75394	0.97828	1.10166	1.35889
4	0.01943	0.08789	0.15899	0.32023	0.51104	0.76419	0.98531	1.13043	1.37774
5	0.02152	0.09471	0.16750	0.32490	0.52449	0.76741	0.99948	1.13916	1.40242
6	0.02336	0.10022	0.17385	0.32717	0.53193	0.77028	1.00520	1.14634	1.41435
7	0.02501	0.10479	0.17873	0.32804	0.53635	0.77552	1.00929	1.15373	1.42457
8	0.02650	0.10863	0.18256	0.32802	0.53916	0.77971	1.01346	1.15859	1.43272
9	0.02786	0.11191	0.18560	0.32745	0.54109	0.78246	1.01731	1.16239	1.43878
10	0.02912	0.11473	0.18803	0.32975	0.54258	0.78454	1.02016	1.16582	1.44397
11	0.03028	0.11718	0.19000	0.33304	0.54390	0.78633	1.02249	1.16885	1.44837
12	0.03137	0.11933	0.19160	0.33570	0.54527	0.78802	1.02458	1.17139	1.45207
13	0.03239	0.12123	0.19291	0.33789	0.54682	0.78966	1.02649	1.17357	1.45527
14	0.03334	0.12290	0.19396	0.33970	0.54856	0.79122	1.02823	1.17552	1.45810
15	0.03424	0.12439	0.19482	0.34122	0.55002	0.79259	1.02977	1.17728	1.46060
16	0.03509	0.12573	0.19552	0.34250	0.55123	0.79377	1.03113	1.17888	1.46283
17	0.03589	0.12692	0.19607	0.34360	0.55228	0.79482	1.03237	1.18032	1.46483
18	0.03665	0.12799	0.19650	0.34454	0.55319	0.79578	1.03351	1.18162	1.46664
19	0.03738	0.12895	0.19684	0.34535	0.55400	0.79667	1.03457	1.18282	1.46830
20	0.03807	0.12982	0.19709	0.34607	0.55475	0.79752	1.03555	1.18392	1.46981
30	0.04354	0.13510	0.20063	0.35087	0.56047	0.80362	1.04243	1.19164	1.48009
50	0.05005	0.13755	0.20794	0.35713	0.56644	0.80988	1.04933	1.19921	1.48969
100	0.05698	0.14472	0.21370	0.36331	0.57269	0.81634	1.05627	1.20666	1.49864
200	0.06049	0.14887	0.21816	0.36784	0.57725	0.82099	1.06117	1.21180	1.50458

## 11.4 Quantilen der t-Verteilung

$k = n - 1$	0.75	0.8	0.9	0.95	0.975	0.99	0.995
1	1.0000	1.3764	3.0777	6.3138	12.7062	31.8205	63.6567
2	0.8165	1.0607	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.7649	0.9785	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.7407	0.9410	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.7267	0.9195	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.7176	0.9057	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	0.8960	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.7064	0.8889	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	0.8834	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	0.8791	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	0.8755	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	0.8726	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	0.8702	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	0.8681	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	0.8662	1.3406	1.7531	2.1314	2.6025	2.9467
16	0.6901	0.8647	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	0.8633	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	0.8620	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	0.8610	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	0.8600	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	0.8591	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.6858	0.8583	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	0.8575	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	0.8569	1.3178	1.7109	2.0639	2.4922	2.7969
25	0.6844	0.8562	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	0.8557	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	0.8551	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	0.8546	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	0.8542	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	0.8538	1.3104	1.6973	2.0423	2.4573	2.7500
50	0.6794	0.8489	1.2987	1.6759	2.0086	2.4033	2.6778
100	0.6770	0.8452	1.2901	1.6602	1.9840	2.3642	2.6259
500	0.6750	0.8423	1.2832	1.6479	1.9647	2.3338	2.5857
$10^3$	0.6747	0.8420	1.2824	1.6464	1.9623	2.3301	2.5808
$10^4$	0.6745	0.8417	1.2816	1.6450	1.9602	2.3267	2.5763
$10^5$	0.6745	0.8416	1.2816	1.6449	1.9600	2.3264	2.5759
$10^6$	0.6745	0.8416	1.2816	1.6449	1.9600	2.3264	2.5758

## 11.5 Quantilen der $\chi^2$ -Verteilung

$k = n - 1$	$p = 0.01$	$p = 0.05$	$p = 0.1$	$p = 0.25$	$p = 0.5$	$p = 0.75$	$p = 0.9$	$p = 0.95$	$p = 0.99$
1	0.000	0.004	0.016	0.102	0.455	1.323	2.706	3.841	6.635
2	0.020	0.103	0.211	0.575	1.386	2.773	4.605	5.991	9.210
3	0.115	0.352	0.584	1.213	2.366	4.108	6.251	7.815	11.345
4	0.297	0.711	1.064	1.923	3.357	5.385	7.779	9.488	13.277
5	0.554	1.145	1.610	2.675	4.351	6.626	9.236	11.070	15.086
6	0.872	1.635	2.204	3.455	5.348	7.841	10.645	12.592	16.812
7	1.239	2.167	2.833	4.255	6.346	9.037	12.017	14.067	18.475
8	1.646	2.733	3.490	5.071	7.344	10.219	13.362	15.507	20.090
9	2.088	3.325	4.168	5.899	8.343	11.389	14.684	16.919	21.666
10	2.558	3.940	4.865	6.737	9.342	12.549	15.987	18.307	23.209
11	3.053	4.575	5.578	7.584	10.341	13.701	17.275	19.675	24.725
12	3.571	5.226	6.304	8.438	11.340	14.845	18.549	21.026	26.217
13	4.107	5.892	7.042	9.299	12.340	15.984	19.812	22.362	27.688
14	4.660	6.571	7.790	10.165	13.339	17.117	21.064	23.685	29.141
15	5.229	7.261	8.547	11.037	14.339	18.245	22.307	24.996	30.578
16	5.812	7.962	9.312	11.912	15.338	19.369	23.542	26.296	32.000
17	6.408	8.672	10.085	12.792	16.338	20.489	24.769	27.587	33.409
18	7.015	9.390	10.865	13.675	17.338	21.605	25.989	28.869	34.805
19	7.633	10.117	11.651	14.562	18.338	22.718	27.204	30.144	36.191
20	8.260	10.851	12.443	15.452	19.337	23.828	28.412	31.410	37.566
21	8.897	11.591	13.240	16.344	20.337	24.935	29.615	32.671	38.932
22	9.542	12.338	14.041	17.240	21.337	26.039	30.813	33.924	40.289
23	10.196	13.091	14.848	18.137	22.337	27.141	32.007	35.172	41.638
24	10.856	13.848	15.659	19.037	23.337	28.241	33.196	36.415	42.980
25	11.524	14.611	16.473	19.939	24.337	29.339	34.382	37.652	44.314
26	12.198	15.379	17.292	20.843	25.336	30.435	35.563	38.885	45.642
27	12.879	16.151	18.114	21.749	26.336	31.528	36.741	40.113	46.963
28	13.565	16.928	18.939	22.657	27.336	32.620	37.916	41.337	48.278
29	14.256	17.708	19.768	23.567	28.336	33.711	39.087	42.557	49.588
30	14.953	18.493	20.599	24.478	29.336	34.800	40.256	43.773	50.892
50	29.707	34.764	37.689	42.942	49.335	56.334	63.167	67.505	76.154
100	70.065	77.929	82.358	90.133	99.334	109.141	118.498	124.342	135.807
500	429.388	449.147	459.926	478.323	499.333	520.950	540.930	553.127	576.493
1000	898.912	927.594	943.133	969.484	999.333	1029.790	1057.724	1074.679	1106.969