

Grundlagen zur Wahrscheinlichkeitsrechnung

5. Dezember 2013

Inhaltsverzeichnis

1	Grundlegendes zur Wahrscheinlichkeitsrechnung	2
1.1	Allgemeine Wahrscheinlichkeitsräume und deren Eigenschaften	2
1.2	Zufallsvariablen und Verteilungen	8
1.3	Mehrdimensionale Verteilungen und Unabhängigkeit	15
1.4	Grenzwertsätze	21
1.5	Grundlegendes zum Schätzen und Testen	23

1 Grundlegendes zur Wahrscheinlichkeitsrechnung

1.1 Allgemeine Wahrscheinlichkeitsräume und deren Eigenschaften

Die Basis aller Betrachtungen bilden Wahrscheinlichkeitsräume.

1.1.1 Definition

Ein Wahrscheinlichkeitsraum wird durch ein Tripel $(\Omega, \mathcal{A}, \mathbb{P})$ dargestellt, wobei

Ω Zustandsraum

\mathcal{A} System von Teilmengen von Ω mit den Eigenschaften:

- (i) $\Omega \in \mathcal{A}$
- (ii) $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$
- (iii) $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$

\mathbb{P} Eine Funktion

$$\mathbb{P} : \mathcal{A} \ni A \rightarrow \mathbb{P}(A) \in [0, 1]$$

mit den Eigenschaften

- a) $\mathbb{P}(\emptyset) = 0$
- b) $\mathbb{P}(\Omega) = 1$
- c) Für alle paarweise disjunkten (p, d) Mengenfolgen aus $\mathcal{A} : (A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ mit $A_n \cap A_m = \emptyset$ für $n \neq m$, gilt:

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \geq 1} \mathbb{P}(A_n) \text{ , d.h.}$$

\mathbb{P} ist σ -additiv.

Das System \mathcal{A} mit den Eigenschaften (i) - (iii) nennt man ein σ -Algebra über Ω und die Funktion \mathbb{P} mit den Eigenschaften (a) - (c) ein Wahrscheinlichkeitsmaß über dem messbaren Raum (Ω, \mathcal{A}) .

Aus dieser Definition ergeben sich einige Konsequenzen und Rechenregeln, die in den folgenden beiden Bemerkungen festgehalten werden.

1.1.2 Bemerkung zum Rechnen im messbaren Raum (Ω, \mathcal{A})

Sei (Ω, \mathcal{A}) ein messbarer Raum, dann gilt:

(i) \mathcal{A} ist abgeschlossen bzgl. endlicher Vereinigung, d.h.

$$\forall A_1, A_2 \in \mathcal{A} : A_1 \cup A_2 \in \mathcal{A}$$

(ii) \mathcal{A} ist abgeschlossen bzgl. endlicher Durchschnittsbildung, d.h.

$$\forall A_1, A_2 \in \mathcal{A} : A_1 \cap A_2 \in \mathcal{A}$$

(iii) \mathcal{A} ist abgeschlossen bzgl. abzählbarer Durchschnittsbildung, d.h.

$$\forall (A_n)_{n \in \mathbb{N}} \subset \mathcal{A} : \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$$

(iv)

$$\forall A, B \in \mathcal{A} : B \setminus A \in \mathcal{A}$$

Beweis: Wir müssen (i) - (iv) zurückführen auf die Definition 1.1.1 und auf bereits bekannte Gesetzmäßigkeiten.

zu i) Def. $A_i = \emptyset$ für $i \geq 3$. Da $\emptyset = \Omega^C \in \mathcal{A}$ nach (i) und (ii) aus Def. 1.1.1, ist mit $A_1, A_2 \in \mathcal{A}$ auch die gesamte Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$. Teil (iii) liefert dann:

$$A_1, A_2 = A_1 \cup A_2 \cup \emptyset \cup \dots = \bigcup_{n \geq 1} A_n \in \mathcal{A}$$

zu ii) seien $A_1, A_2 \in \mathcal{A}$, dann folgt:

$$A_1 \cap A_2 = \left[(A_1 \cap A_2)^C \right]^C = \left[A_1^C \cup A_2^C \right]^C$$

$A_i^C \in \mathcal{A}$ gemäß Def. 1.1.1 (ii), für $i = 1, 2$. Nach dem bereits gezeigten Teil (i) dieser Bemerkung, muss dann auch $A_1^C \cup A_2^C \in \mathcal{A}$. Da \mathcal{A} abgeschlossen ist bzgl. der Komplementbildung nach Def. 1.1.1 (ii), folgt abschließend $\left[A_1^C \cup A_2^C \right]^C \in \mathcal{A}$ und damit die Beh. (ii).

zu iii) Ergibt sich analog zu (ii) unter Anwendung von Def. 1.1.1 (iii).

zu iv) Folgt sofort aus

$$B \setminus A = B \cap A^C$$

1.1.3 Bemerkung zum “Rechnen” mit \mathbb{P}

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, dann gilt:

(i) $\forall A, B \in \mathcal{A}$ mit $A \cap B = \emptyset$ gilt:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

(ii) \mathbb{P} ist monoton wachsend, d.h. $\forall A, B \in \mathcal{A}$ mit $A \subset B$ gilt:

$$\mathbb{P}(A) \leq \mathbb{P}(B)$$

(iii) Für alle $A, B \in \mathcal{A}$ mit $A \subset B$ gilt:

$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$$

Im Spezialfall $B = \Omega$ folgt damit dann $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$.

(iv) Für $A_1, A_2 \in \mathcal{A}$ gilt:

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$$

(v) Für $A_1, A_2, \dots, A_n \in \mathcal{A}$ gilt die allgemeine Ein-Ausschluss-Formel:

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \\ \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} \mathbb{P}(A_i \cap A_j \cap A_k) - \dots + (-1)^{n-1} \mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) \end{aligned}$$

(vi) Für alle $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ mit $A_1 \subset A_2 \subset A_3 \subset \dots$ gilt:

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n \geq 1} A_n\right)$$

(vii) Für alle $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ mit $A_1 \supset A_2 \supset A_3 \supset \dots$ gilt:

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n \geq 1} A_n\right)$$

(viii) Für alle $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ gilt:

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \geq 1} \mathbb{P}(A_n)$$

Beweis: Zu i) seien $A_1, A_2 \in \mathcal{A}$ mit $A_1 \cap A_2 = \emptyset$. Mit $A_k = \emptyset$ für $k \geq 3$ erhalten wir dann eine Folge $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$, die aus paarweise disjunkten Mengen besteht. Die Eigenschaften (c) und (a) aus Def. 1.1.1 garantieren dann:

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mathbb{P}(A_n) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$$

zu ii) seien $A, B \in \mathcal{A}$ mit $A \subset B$. Dann sind $B \setminus A, A$ p.d. Mengen aus \mathcal{A} und es folgt:

$$(+)\ \mathbb{P}(B) = \mathbb{P}(A \cup (B \setminus A)) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A),$$

denn $\mathbb{P}(B \setminus A) \geq 0$.

zu iii) Folgt direkt aus (+) durch Umstellen der Terme.

zu iv) seien $A_1, A_2 \in \mathcal{A}$. Man mache sich klar, dass

$$A_1 \cup A_2 = [A_1 \setminus (A_1 \cap A_2)] \cup [A_2 \setminus (A_1 \cap A_2)] \cup [A_1 \cap A_2]$$

gilt und dass die Mengen [...] auf der rechten Seite p.d. sind.

Mit dieser Darstellung folgt dann sofort:

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1 \setminus (A_1 \cap A_2)) + \mathbb{P}(A_2 \setminus (A_1 \cap A_2)) + \mathbb{P}(A_1 \cap A_2)$$

$$\stackrel{(iii)}{=} \mathbb{P}(A_1) - \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_2)$$

$$= \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$$

zu v) Mittels vollständiger Induktion und Teil (iv).

zu vi) Seien $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ mit $A_1 \subset A_2 \subset \dots$

Definiere dann:

$$A'_1 := A_1, A'_2 := A_2 \setminus A_1, \dots, A'_k := A_k \setminus \left(\bigcup_{i=1}^{k-1} A_i \right), \dots$$

Damit folgt sofort:

$$(A'_n)_{n \geq 1} \subset \mathcal{A}, \text{ die } A'_n \text{ sind p.d., } A_n = \bigcup_{k=1}^n A'_k \text{ und } \bigcup_{n \geq 1} A_n = \bigcup_{n \geq 1} A'_n.$$

Weiter gilt dann:

$$\begin{aligned} \mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) &= \mathbb{P}\left(\bigcup_{n \geq 1} A'_n\right) = \sum_{n \geq 1} \mathbb{P}(A'_n) = \lim_{N \rightarrow \infty} \sum_{n=1}^N \mathbb{P}(A'_n) \\ &= \lim_{N \rightarrow \infty} \mathbb{P}\left(\bigcup_{n=1}^N A'_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) \end{aligned}$$

zu vii) Folgt aus (vi) durch Komplementbildung.

zu viii) Sei $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$. Definiere dann

$$A'_1 := A_1, A'_2 := A_2 \setminus A_1, \dots, A'_k := A_k \setminus (A_1 \cup \dots \cup A_{k-1}), \dots$$

Dann folgt:

$$\mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) = \mathbb{P}\left(\bigcup_{n \geq 1} A'_n\right) = \sum_{n \geq 1} \mathbb{P}(A'_n) \leq \sum_{n \geq 1} \mathbb{P}(A_n)$$

□

1.1.4 Bemerkung:

Auf \mathbb{R} gibt es ein besonderes Mengensystem, das sich aus Überlegungen zur Längenmessung ergeben hat. Betrachten wir das Mengensystem \mathfrak{I} aller Intervalle, dann ist \mathfrak{I} keine σ -Algebra über \mathbb{R} ! Allerdings lässt sich aus \mathfrak{I} eine σ -Algebra konstruieren, die im Folgenden das System der Borel-Mengen genannt wird und mit \mathcal{B}^* bezeichnet ist. Genauer ist

\mathcal{B}^* die kleinste σ -Algebra, die \mathfrak{I} enthält! Dieses Mengensystem ist sehr groß und unübersichtlich. Für unsere Zwecke reicht aber die Vorstellung aus, dass \mathcal{B}^* alle “vernünftigen” Teilmengen von \mathbb{R} enthält.

1.2 Zufallsvariablen und Verteilungen

In der Regel ist $(\Omega, \mathcal{A}, \mathbb{P})$ nicht bekannt oder nur von geringem Interesse. Wir machen unsere Erfahrungen über diesen W.-Raum durch Experimente X , deren Ergebnis x wir beobachten.

1.2.1 Definition

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein W.-Raum und

$$X : \Omega \ni \omega \rightarrow X(\omega) = x \in \mathbb{R}$$

eine Funktion mit der Eigenschaft, dass

$$\forall_{t \in \mathbb{R}} : \{\omega \in \Omega : X(\omega) \leq t\} \in \mathcal{A}.$$

Dann nennt man X eine Zufallsvariable (ZV) oder auch ein Experiment.

Die Funktion

$$F : \mathbb{R} \ni t \rightarrow F(t) \equiv F_X(t) := \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq t\}) \equiv \mathbb{P}(X \leq t)$$

nennt man die Verteilungsfunktion von X .

1.2.2 Bemerkung:

Aus der Definition der ZV X kann geschlossen werden, dass für alle $B \in \mathcal{B}^*$, dem System der Borel-Mengen, $\mathbb{P}(X \in B)$ definiert ist. Des Weiteren wird dann durch $(\mathbb{R}, \mathcal{B}^*, \mathbb{P}_X)$ mit

$$\mathbb{P}_X(B) := \mathbb{P}(X \in B), \quad B \in \mathcal{B}^*$$

ein neuer W.-Raum definiert, der die Verteilung von X genannt wird. Verteilungen können sehr kompliziert sein, da \mathcal{B}^* sehr komplex und kompliziert ist. Glücklicherweise kann man aber nachweisen, dass eine Verteilung \mathbb{P}_X schon eindeutig durch die zugehörige Verteilungsfunktion F_X definiert ist.

1.2.3 Beispiel:

Die ZV X beschreibe das Experiment “fares Würfeln”. Dann gilt mit unseren Bezeichnungen:

$$\mathbb{P}_X(\{1\}) = \mathbb{P}(X = 1) = 1/6$$

$$\mathbb{P}_X(\{2, 3\}) = \mathbb{P}(X \in \{2, 3\}) = \mathbb{P}(\{X = 1\} \cup \{X = 2\}) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = 2/6$$

$$F_X(4) = \mathbb{P}(X \leq 4) = 4/6$$

$$F_X(-1) = \mathbb{P}(X \leq -1) = \mathbb{P}(\emptyset) = 0$$

$$F_X(17) = \mathbb{P}(X \leq 17) = 1$$

Aus den Rechenregeln für Wahrscheinlichkeiten, vgl. Bem. 1.1.3, lassen sich die folgenden Eigenschaften einer Verteilungsfunktion herleiten.

1.2.4 Bemerkung:

Sei X eine ZV mit Verteilungsfunktion $F \equiv F_X$. Dann gilt:

(i) $\forall t \in \mathbb{R} : 0 \leq F(t) \leq 1$

(ii) F ist monoton wachsend

(iii) F ist rechtsseitig stetig und besitzt linksseitige Limiten

(iv) $\lim_{t \rightarrow \infty} F(t) = 1, \lim_{t \rightarrow -\infty} F(t) = 0$

Es gibt unendlich viele Verteilungen auf $(\mathbb{R}, \mathcal{B}^*)$. Für die praktische Anwendung sind im Wesentlichen die folgenden beiden “reinen” Klassen von Bedeutung.

1.2.5 Definition

Sei X eine ZV über $(\Omega, \mathcal{A}, \mathbb{R})$ mit Verteilung $(\mathbb{R}, \mathcal{B}^*, \mathbb{P}_X)$. Falls es dann höchstens abzählbar viele Punkte $x_1, x_2, \dots \in \mathbb{R}$ gibt, mit $x_i \neq x_j$ für alle $i \neq j$, sodass

$$\sum_{i \geq 1} \mathbb{P}(X = x_i) = 1,$$

dann nennt man X eine diskrete ZV und die Funktion p mit

$$p(x_i) := \mathbb{P}(X = x_i),$$

wird Wahrscheinlichkeitsfunktion (WF) von X genannt.

1.2.6 Bemerkung:

Ist X eine diskret verteilte ZV mit WF p auf den “Massepunkten” $x_1, x_2, \dots \in \mathbb{R}$. Dann ist die Verteilung von X sehr übersichtlich und häufig auch leicht berechenbar. Sei z.B. $B \in \mathcal{B}^*$, dann gilt:

$$\mathbb{P}(X \in B) = \sum_{x_i \in B} p(x_i).$$

1.2.7 Definition:

Eine ZV X besitzt eine stetige Verteilung, falls es eine stetige, nicht negative Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$$

gibt, sodass für alle $B \in \mathcal{B}^*$:

$$\mathbb{P}(X \in B) = \int_B f(x) dx.$$

Natürgemäß muss dann auch $\int f(x) dx = 1$. Die Funktion f wird auch Dichtefunktion (DF) von X genannt.

1.2.8 Bemerkung:

Besitzt die ZV X eine stetige Verteilung, dann gilt der folgende Zusammenhang zwischen der zugehörigen Verteilungsfunktion $F \equiv F_X$ und der DF f .

(i) $F(x) = \int_{-\infty}^x f(t) dt$: für alle $x \in \mathbb{R}$

(ii) $f(x) = F'(x)$: für alle $x \in \mathbb{R}$

Beweis: zu i) folgt sofort aus

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \in]-\infty, x]) = \int_{-\infty}^x f(t) dt$$

zu ii) Bekanntlich ist F differenzierbar im Punkt $x \in \mathbb{R}$, falls

$$\lim_{k \rightarrow 0} \frac{F(x+k) - F(x)}{k}$$

existiert. $F'(x)$ ist dann dieser Grenzwert

$$\frac{F(x+k) - F(x)}{k} = \frac{1}{k} \left[\int_{-\infty}^{x+k} f(t) dt - \int_{-\infty}^x f(t) dt \right] = \frac{1}{k} \int_x^{x+k} f(t) dt \xrightarrow[k \rightarrow 0]{} f(x),$$

wobei die letzte Konvergenz aus der Hauptsatz der Differential- und Integralrechnung folgt. \square

In einem späteren Kapitel werden wir zum Modellieren wichtige diskrete und stetige ZV konkret besprechen. An dieser Stelle verzichten wir daher auf entsprechende Beispiele.

1.2.9 Definition:

Sei X eine stetige oder diskrete ZV und

$$g : \mathbb{R} \rightarrow \mathbb{R}$$

eine Abbildung. Dann ist durch

$$\Omega \ni \omega \rightarrow g(X(\omega)) \in \mathbb{R}$$

eine neue ZV definiert und wir nennen die Größe

$$\mathbb{E}(g(X)) := \begin{cases} \sum_i g(x_i) \cdot \mathbb{P}(X = x_i) & \text{falls } X \text{ diskret} \\ \int g(x) f(x) dx & \text{falls } X \text{ stetig,} \end{cases}$$

falls sie existiert, den Erwartungswert von $g(X)$. In dieser Definition bezeichnet f die Dichtefunktion von X , falls X stetig ist, und $\mathbb{P}(X = x_i)$ die WF von X im diskreten Fall.

Ist $g(x) = x^k$, $k \in \mathbb{N}$, dann nennt man

$$\mathbb{E}(g(x)) = \mathbb{E}(X^k) =: \mu_k$$

das k -te Moment von X .

Das 1-te Moment von X , also $\mathbb{E}(X)$, nennen wir den Erwartungswert von X und bez. ihn mit μ .

1.2.10 Bemerkung:

(i) In der letzten Definition wird einschränkend bemerkt, dass $\mathbb{E}(g(X))$ existieren muss, wenn man $\mathbb{E}(g(X))$ betrachtet. Man kann beweisen, dass die Existenz von $\mathbb{E}(g(X))$ genau dann gesichert ist, wenn

$$\mathbb{E}(|g(X)|) < \infty,$$

wobei $\mathbb{E}(|g(X)|)$ stets gemäß der definierenden Formeln zu berechnen ist, d.h. man muss $|g(x_i)|$ im diskreten Fall bzw. $|g(x)|$ im stetigen Fall in die Formeln einsetzen.

Der Erwartungswert einer ZV X stellt ein Zentrum der Verteilung \mathbb{P}_X dar, um den herum die Daten “streuen”.

Diese Streuung wird durch die Varianz beschrieben.

1.2.11 Definition:

Sei X eine ZV mit $\mathbb{E}(X^2) < \infty$ bez. mit μ den Erwartungswert von X , dann wird durch

$$VAR(X) := \mathbb{E}((X - \mu)^2) \equiv \sigma^2$$

die Varianz von X definiert und mit σ^2 bezeichnet.

1.2.12 Lemma:

Für die Varianz σ^2 einer ZV X gilt:

$$\sigma^2 = \mathbb{E}(X^2) - \mu^2$$

Beweis: Wir beweisen die Behauptung nur für den stetigen Fall, d.h. wir nehmen an, dass X eine stetige Verteilung mit Dichte f hat. Dann gilt:

$$\sigma^2 = \mathbb{E}((X - \mu)^2) = \int (x - \mu)^2 f(x) dx = \int (x^2 - 2x\mu + \mu^2) f(x) dx$$

$$\begin{aligned}
&= \underbrace{\int x^2 f(x) dx}_{=\mathbb{E}(X^2)} - 2\mu \underbrace{\int x \cdot f(x) dx}_{=\mu} + \mu^2 \underbrace{\int f(x) dx}_{=1} \\
&= \mathbb{E}(X^2) - \mu^2
\end{aligned}$$

Neben μ gibt es weitere “zentrale Lageparameter” einer Verteilung, die in der folgenden Definition zusammengefasst werden.

1.2.13 Definition:

Sei X eine stetige (diskrete) ZV mit Dichtefunktion f (WF p). Bezeichne ferner mit F die zugehörige Verteilungsfunktion von X .

- i) Die Maximalstelle von f , falls sie existiert, wird Modalwert genannt. Beachte, dass die Maximalstelle nicht eindeutig sein muss!
- ii) Die “Umkehrfunktion” von F ,

$$F^{-1}(u) := \inf \{x : F(x) \geq u\}, \quad 0 < u < 1,$$

wird Quantilfunktion genannt und die Größe

$$F^{-1}(1/2) := \inf \{x : F(x) \geq 1/2\}$$

definiert den Median der Verteilung.

1.2.14 Bemerkung:

- (i) Die Quantilfunktion F^{-1} von F ist eine Verallgemeinerung des Begriffs der Umkehrfunktion. Dies ist notwendig, denn F muss nicht notwendigerweise umkehrbar sein! Ist F aber umkehrbar, dann stimmt die Umkehrfunktion mit der Quantilfunktion überein!
- (ii) Die Quantilfunktion hat eine zentrale Bedeutung beim Erzeugen von ZV auf dem Computer. Es gilt nämlich

$$(+)\ F^{-1}(U) \sim F,$$

wobei U eine auf dem Einheitsintervall gleichverteilte ZV ist, d.h. U hat die Dichte

$$f(u) := \begin{cases} 1 & : 0 \leq u \leq 1 \\ 0 & : \text{sonst} \end{cases}$$

(+) besagt nun: Man erhält eine nach F verteilte ZV X , indem man sich eine gleichverteilte (auf dem Einheitsintervall) ZV U erzeugt und dann

$$X := F^{-1}(U)$$

berechnet. Wie man solch ein U erzeugt, wird in einem anderen Kapitel behandelt.

1.3 Mehrdimensionale Verteilungen und Unabhängigkeit

In der Grundvorlesung hatten wir den Begriff der bedingten Wahrscheinlichkeit kennengelernt:

$$(+)\ \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

$\mathbb{P}(A|B)$ bezeichnet die W. dass A eintritt, wenn wir wissen, dass B eingetreten ist.

Hat A aber “nichts” mit B zu tun, d.h. A und B sind “unabhängig” voneinander, dann nutzt die Information, dass B eingetreten ist, nichts. In diesem Fall sollte daher

$$(++)\ \mathbb{P}(A|B) = \mathbb{P}(A)$$

sein. Falls also A und B unabhängig voneinander sind, dann ergeben $(+)$ und $(++)$

$$\mathbb{P}(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

oder aber

$$(+++)\ \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B).$$

$(+++)$ ist dann die definierende Gleichung für die Unabhängigkeit der Ereignisse A und B .

1.3.1 Definition

Es seien X, Y ZV über $(\Omega, \mathcal{A}, \mathbb{P})$. Dann wird durch

$$(X, Y) : \Omega \ni \omega \rightarrow (X(\omega), Y(\omega)) \in \mathbb{R}^2$$

ein zweidimensionaler Zufallsvektor (ZVe) definiert. Auf \mathbb{R}^2 kann wieder ein entsprechendes System von Borel-Mengen, mit $\mathcal{B}(\mathbb{R}^2)$ bez., definiert werden und in Analogie zur Definition 1.2.1 bezeichnet man wieder mit

$$\mathbb{P}_{X,Y}(B) := \mathbb{P}(X, Y) \in \mathcal{B}, B \in \mathcal{B}(\mathbb{R}^2)$$

die Verteilung von (X, Y) ; auch gemeinsame Verteilung genannt.

Es lässt sich zeigen, dass für $A, B \in \mathcal{B}^*$ stets

$$A \times B \in \mathcal{B}(\mathbb{R}^2)$$

gilt und man schreibt dann auch für $\mathbb{P}_{X,Y}(A \times B)$:

$$\mathbb{P}_{X,Y}(A \times B) = \mathbb{P}((X, Y) \in A \times B) = \mathbb{P}(\{X \in A\} \cap \{Y \in B\}) \equiv \mathbb{P}(X \in A, Y \in B).$$

1.3.2 Definition

Zwei ZV X, Y über $(\Omega, \mathcal{A}, \mathbb{P})$ sind genau dann unabhängig von einander, wenn

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \cdot \mathbb{P}(Y \in B), \text{ für alle } A, B \in \mathcal{B}^*.$$

Analog zum eindimensionalen Fall lassen sich auch hier diskrete und stetige Verteilungen betrachten.

1.3.3 Definition

Seien X, Y ZV über $(\Omega, \mathcal{A}, \mathbb{P})$ mit Verteilung $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mathbb{P}_{X,Y})$. Falls es dann höchstens abzählbar viele Punkte $(x_1, y_1), (x_2, y_2), \dots \in \mathbb{R}^2$ gibt mit $(x_i, y_i) \neq (x_j, y_j)$ für $i \neq j$, sodass

$$\sum_i \mathbb{P}(X = x_i, Y = y_i) = 1,$$

dann nennt man (X, Y) einen diskreten ZVe mit zugehöriger Wahrscheinlichkeitsfunktion (gemeinsame WF)

$$p(x_i, y_i) := \mathbb{P}(X = x_i, Y = y_i).$$

1.3.4 Definition

Ein zweidimensionaler ZVe (X, Y) besitzt eine stetige Verteilung, falls es eine stetige, nicht negative Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0}$$

gibt, sodass für alle $A, B \in \mathcal{B}^*$:

$$\mathbb{P}(X \in A, Y \in B) = \int_A \int_B f(x, y) dy dx$$

f wird wieder Dichtefunktion (oder gemeinsame Dichte) genannt.

1.3.5 Lemma

Gegeben sind ZV X, Y mit gemeinsamer Dichte $f(x, y)$ (WF $p(x_i, y_i)$). Dann gilt:

(i) X hat die Dichte (WF)

$$f_X(x) := \int f(x, y) dy$$
$$\left(p_X(x_i) = \sum_j p(x_i, y_j) \right)$$

Analog hat Y die Dichte (WF)

$$f_Y := \int f(x, y) dx$$
$$\left(p_Y(y_i) = \sum_j p(x_j, y_i) \right)$$

(ii) X ist genau dann unabhängig von Y , wenn

$$f(x, y) = f_X(x) \cdot f_Y(y)$$
$$(p(x_i, y_i) = p_X(x_i) \cdot p_Y(y_i))$$

Beweis: Zu i) Zeige hier nur den stetigen Fall für ZV X . Da $f : \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0}$, ist f_X eine nicht-negative Funktion auf \mathbb{R} . Nach Def. 1.2.7 reicht es daher zu zeigen, dass

$$\mathbb{P}(X \in B) = \int_B f_X(x) dx, \forall B \in \mathcal{B}^*$$

gilt. Sei dazu $B \in \mathcal{B}^*$ beliebig. Dann gilt:

$$\mathbb{P}(X \in B) = \mathbb{P}(X \in B, Y \in \mathbb{R}) = \int_B \int_{\mathbb{R}} f(x, y) dy dx = \int_B f_X(x) dx.$$

Zu ii) Zeige hier nur, dass aus $f(x, y) = f_X(x)f_Y(y)$ die Unabhängigkeit von X und Y folgt. Seien dazu $A, B \in \mathcal{B}^*$ beliebig. Dann gilt:

$$\mathbb{P}(X \in A, Y \in B) = \int_A \int_B f(x, y) dy dx = \int_A \int_B f_X(x) \cdot f_Y(y) dy dx$$

$$\begin{aligned}
&= \int_A f_X(x) \cdot \left[\int_B f_Y(y) dy \right] dx = \int_A f_X(x) \cdot \mathbb{P}(Y \in B) dx \\
&= \mathbb{P}(Y \in B) \cdot \int_A f_X(x) dx = \mathbb{P}(Y \in B) \cdot \mathbb{P}(X \in A).
\end{aligned}$$

□

1.3.6 Bemerkung

Die in den Definitionen 1.3.1-1.3.4 eingeführten Begriffe lassen sich auf bel. $n \in \mathbb{N}$ sinngemäß übertragen. Gleiches gilt für Lemma 1.3.5.

1.3.7 Definition

Seien X, Y ZV über $(\Omega, \mathcal{A}, \mathbb{P})$ mit $\mathbb{E}(X^2) < \infty$ und $\mathbb{E}(Y^2) < \infty$. Ferner bezeichne $\mu_X := \mathbb{E}(X)$ und $\mu_Y := \mathbb{E}(Y)$. Dann wird durch

$$COV(X, Y) := \mathbb{E}((X - \mu_X) \cdot (Y - \mu_Y))$$

die Kovarianz von X und Y bzw. durch

$$COV(X, Y) := \frac{COV(X, Y)}{\sqrt{\sigma_X^2 \cdot \sigma_Y^2}}$$

die Korrelation von X und Y bezeichnet. Dabei ist $\sigma_X^2 = VAR(X)$ bzw. $\sigma_Y^2 := VAR(Y)$.

Das folgende Lemma enthält einige nützliche Regeln zum Rechnen mit \mathbb{E} und VAR .

1.3.8 Lemma

Gegeben seien X, Y ZV über $(\Omega, \mathcal{A}, \mathbb{P})$ und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

- (i) $\mathbb{E}(\alpha X + \beta Y) = \alpha \mathbb{E}(X) + \beta \mathbb{E}(Y)$
- (ii) $VAR(\alpha X) = \alpha^2 VAR(X)$
- (iii) $COV(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y)$
- (iv) Falls X unabhängig ist von Y :

$$(a) \quad VAR(X + Y) = VAR(X) + VAR(Y)$$

$$(b) \text{ } COV(X, Y) = 0$$

Beweis: Wir betrachten hier nur den stetigen Fall und nehmen an, dass (X, Y) die gemeinsame Dichte $f(x, y)$ habe. Gemäß Lemma 1.3.5 hat dann X die Dichte $f_X(x) = \int f(x, y) dy$ bzw. Y die Dichte $f_Y(y) = \int f(x, y) dx$.

Zu i):

$$\mathbb{E}(\alpha X + \beta Y) = \int \int (\alpha x + \beta y) f(x, y) dx dy =$$

$$\int \alpha x f_X(x) dx + \int \beta y f_Y(y) dy = \alpha \mathbb{E}(X) + \beta \mathbb{E}(Y)$$

Zu ii):

$$\begin{aligned} VAR(\alpha X) &= \mathbb{E} \left([\alpha X - \mathbb{E}(\alpha X)]^2 \right) \stackrel{(i)}{=} \mathbb{E} \left([\alpha X - \alpha \mathbb{E}(X)]^2 \right) \\ &= \mathbb{E} \left(\alpha^2 [X - \mathbb{E}(X)]^2 \right) \stackrel{(i)}{=} \alpha^2 \mathbb{E} \left([X - \mathbb{E}(X)]^2 \right) = \alpha^2 VAR(X) \end{aligned}$$

Zu iii):

$$\begin{aligned} COV(X, Y) &= \mathbb{E}([X - \mathbb{E}(X)][Y - \mathbb{E}(Y)]) \\ &\stackrel{(i)}{=} \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \mathbb{E}(Y) \\ &\stackrel{(i)}{=} \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \mathbb{E}(Y) \\ &= \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \mathbb{E}(Y) \end{aligned}$$

Zu iv): Zeige zunächst

$$(+)\ X, Y \text{ unabh.} \Rightarrow \mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

$$\text{Zu } (+): \mathbb{E}(X \cdot Y) = \int \int x \cdot y f(x, y) dx dy$$

Gemäß Lemma 1.3.5 (ii) gilt aber unter Unabhängigkeit:

$$f(x, y) = f_X(x) \cdot f_Y(y)$$

und damit dann weiter

$$\mathbb{E}(X \cdot Y) = \int \int x \cdot y f_X(x) f_Y(y) dx dy = \int x f_X(x) dx \cdot \int y f_Y(y) dy = \mathbb{E}(X) \cdot \mathbb{E}(Y)$$

also (+).

Kommen wir nun zur Formel für die Varianz:

$$\begin{aligned} \text{VAR}(X + Y) &= \mathbb{E}([X + Y]^2) - (\mathbb{E}(X + Y))^2 \\ &= \mathbb{E}(X^2 + 2XY + Y^2) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\ &= \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(X \cdot Y) - \left[(\mathbb{E}(X))^2 + 2\mathbb{E}(X)\mathbb{E}(Y) + (\mathbb{E}(Y))^2 \right] \\ &\stackrel{(+)}{=} \mathbb{E}(X^2) - (\mathbb{E}(X))^2 + \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 \\ &= \text{VAR}(X) + \text{VAR}(Y) \end{aligned}$$

Für Teil (b) gilt zusammen mit (+) und (iii)

$$\text{COV}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y) \stackrel{(+)}{=} 0$$

1.3.9 Definition

Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von ZV über $(\Omega, \mathcal{A}, \mathbb{P})$ mit der Eigenschaft, dass für alle $N \in \mathbb{N}$ gilt: X_1, \dots, X_N sind unabhängige ZV, dann nennt man $(X_n)_{n \in \mathbb{N}}$ eine Folge von unabhängigen ZV. Besitzen darüber hinaus noch alle ZV der Folge die gleiche Verteilung, dann sprechen wir von einer IID Folge (Independent and Identically Distributed).

1.4 Grenzwertsätze

Im Rahmen statistischer Anwendungen sind die folgenden zwei Sätze von größter Wichtigkeit.

1.4.1 Zentraler Grenzwertsatz (Central Limit Theorem CLT)

Sei $(X_n)_{n \in \mathbb{N}}$ eine IID Folge von ZV über $(\Omega, \mathcal{A}, \mathbb{P})$ mit $\mu := \mathbb{E}(X_1)$ und $\sigma^2 := \text{VAR}(X_1)$. Dann gilt:

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}(\sqrt{n}(\bar{X}_n - \mu) / \sigma \leq t) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp(-u^2/2) du \right| \rightarrow 0, \quad (1)$$

für $n \rightarrow \infty$. Dabei bezeichnet $\bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i$ das arithmetische Mittel von X_1, \dots, X_n .

Ferner bez. hier $\sup_{t \in \mathbb{R}} A(t) := \sup_{t \in \mathbb{R}} \{A(t) : t \in \mathbb{R}\}$ mit $A(t) := \left| \mathbb{P}(\dots \leq t) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \dots \right|$.

Die Größe

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

ist die Dichtefunktion der Standardnormalverteilung.

1.4.2 Bemerkung

(i) Die Konvergenz unter (1) ist äußerst bemerkenswert, denn sie besagt, dass, egal, wie die X -Daten verteilt sind, die Verteilungsfunktion von

$$\sqrt{n} \cdot \frac{\bar{X}_n - \mu}{\sigma}$$

immer gegen die Standardnormalverteilung konvergiert.

(ii) Bezeichnen wir mit

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

die sogenannte Stichprobenvarianz, dann gilt die Aussage (1) auch, wenn wir σ durch

$\sqrt{S_n^2}$ ersetzen, d.h.

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P} \left(\sqrt{n}(\bar{X} - \mu) / \sqrt{S_n^2} \leq t \right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp \left(-\frac{u^2}{2} \right) du \right| \rightarrow o, \quad (2)$$

für $n \rightarrow \infty$.

Der zweite wichtige Satz aus der Wahrscheinlichkeitstheorie im Rahmen von statistischen Betrachtungen ist

1.4.3 Kolmogorovs Gesetz der großen Zahlen (strong law of large numbers SLLN)

Sei $(X_n)_{n \in \mathbb{N}}$ eine IID Folge von ZV über $(\Omega, \mathcal{A}, \mathbb{P})$ mit $\mathbb{E}(|X_i|) < \infty$. Dann gilt:

$$\mathbb{P} \left(\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mathbb{E}(X) \right) = 1$$

1.5 Grundlegendes zum Schätzen und Testen

Generell versuchen wir in der Statistik etwas über die unbekannte Verteilung einer ZV X zu erfahren. Eine erste Beschreibung der Verteilung, die aber noch recht ungenau ist, wird durch $\mathbb{E}(X) = \mu$ und $\sigma^2 = \text{VAR}(X)$ gegeben. Haben wir jetzt n solcher IID ZV, also X_1, \dots, X_n , die alle die gleiche Verteilung wie X haben, dann erhalten wir mit

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i$$

bzw.

$$S_n^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \quad (3)$$

zwei neue ZV, die als Schätzer (genauer Punktschätzer) für μ bzw. σ^2 benutzt werden. Dass \bar{X}_n und S_n^2 sinnvoll sind, folgt aus dem SLLN, denn

$$\mathbb{P}(\bar{X}_n \rightarrow \mu) = 1 \text{ bzw. } \mathbb{P}(S_n^2 \rightarrow \sigma^2) = 1$$

Diese beiden Schätzer sind außerdem noch erwartungstreu, denn

$$\mathbb{E}(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} \cdot n \cdot \mu = \mu$$

und

$$\mathbb{E}(S_n^2) = \sigma^2$$

wie man durch direktes Nachrechnen überprüfen kann.

Des Weiteren gilt mit Lemma 1.3.8:

$$\text{VAR}(\bar{X}_n) = \text{VAR}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot \text{VAR}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{VAR}(X_i) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

Die unter (3) angegebenen “Statistiken” sind Punktschätzer für den Erwartungswert bzw. für die Varianz. In vielen Fällen interessieren wir uns aber nicht nur für den Schätzwert, sondern auch für die “Genauigkeit”. Diese wird u.a. durch die Varianz des Schätzers angegeben. Eine andere Form, diese Genauigkeit zu spezifizieren, ist das Konfidenzintervall.

1.5.1 Definition

Gegeben seien X_1, \dots, X_n ZV über $(\Omega, \mathcal{A}, \mathbb{P})$, die die gleiche Verteilung haben und ein Konfidenzniveau $1 - \alpha$, wobei $0 < \alpha < 1$. Unter einem Konfidenzintervall für den Erwartungswert μ ,

$$K_n := K_n(X_1, \dots, X_n)$$

verstehen wir einen von X_1, \dots, X_n abhängigen (also zufälligen) Bereich K_n mit der Eigenschaft

$$\mathbb{P}(\mu \in K_n) = 1 - \alpha.$$

Falls nur sichergestellt werden kann, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}(\mu \in K_n) = 1 - \alpha$$

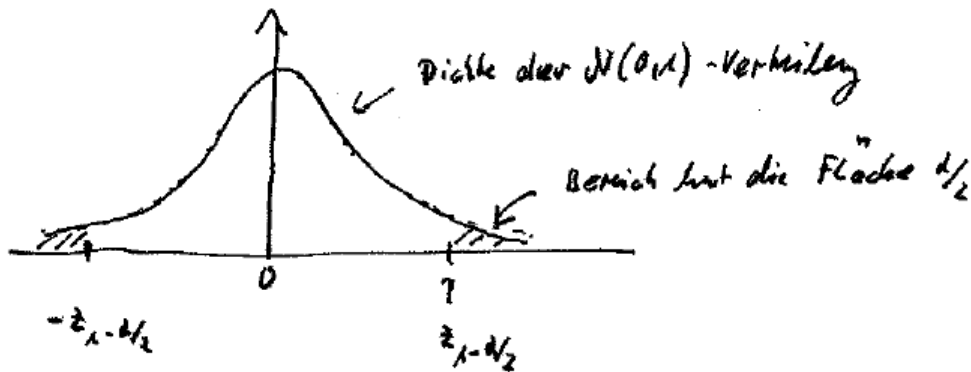
gilt, dann sprechen wir von einem asymptotischen Konfidenzintervall.

1.5.2 Beispiel

Konstruktion eines asymptotischen KI. Gegeben sei eine IID Folge von ZV über $(\Omega, \mathcal{A}, \mathbb{P})$, also $(X_n)_{n \in \mathbb{N}}$, mit $\mathbb{E}(X_1^2) < \infty$. Gemäß (2) aus dem CLT folgt dann, dass die Verteilung von

$$\sqrt{n} \frac{\bar{X}_n - \mu}{\sqrt{S_n^2}}$$

gegen die Standardnormalverteilung konvergiert.



Wähle für $z_{1-\frac{\alpha}{2}} = \Phi^{-1}(1 - \alpha/2)$ das $1 - \frac{\alpha}{2}$ Quantil (vgl. Def. 1.2.12) der Standardnormalverteilung, deren Verteilungsfunktion hier mit Φ bezeichnet ist.

Der CLT garantiert dann, dass

$$\begin{aligned} & \mathbb{P} \left(-z_{1-\frac{\alpha}{2}} \leq \sqrt{n} \cdot \frac{\bar{X}_n - \mu}{\sqrt{S_n^2}} \leq z_{1-\frac{\alpha}{2}} \right) \\ & \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-z_{1-\frac{\alpha}{2}}}^{z_{1-\frac{\alpha}{2}}} \exp \left(-\frac{u^2}{2} \right) du \\ & = \Phi \left(z_{1-\frac{\alpha}{2}} \right) - \Phi \left(-z_{1-\frac{\alpha}{2}} \right) \\ & = \left(1 - \frac{\alpha}{2} \right) - \frac{\alpha}{2} = 1 - \alpha \end{aligned}$$

Ferner folgt durch einfaches Umstellen:

$$-z_{1-\frac{\alpha}{2}} \leq \sqrt{n} \cdot \frac{\bar{X}_n - \mu}{\sqrt{S_n^2}} \leq z_{1-\frac{\alpha}{2}} \Leftrightarrow \bar{X}_n - \frac{\sqrt{S_n^2} \cdot z_{1-\frac{\alpha}{2}}}{\sqrt{n}} \leq \mu \leq \bar{X}_n + \frac{\sqrt{S_n^2} \cdot z_{1-\frac{\alpha}{2}}}{\sqrt{n}}$$

Also

$$K_n := \left[\bar{X}_n - \frac{\sqrt{S_n^2} \cdot z_{1-\frac{\alpha}{2}}}{\sqrt{n}}, \bar{X}_n + \frac{\sqrt{S_n^2} \cdot z_{1-\frac{\alpha}{2}}}{\sqrt{n}} \right]$$

und $\mathbb{P}(\mu \in K_n) \rightarrow 1 - \alpha$, für $n \rightarrow \infty$. \square

Ein weiteres großes und wichtiges Gebiet in der Statistik ist die Testtheorie. Die Grundidee eines Tests ist die stochastische Variante eines “indirekten Beweises” aus der Mathematik:

- (i) Wir wollen eine Behauptung B beweisen, bzw. statistisch nachweisen.
- (ii) Wir definieren als Nullhypothese H_0 : “nicht B gilt” und als Alternative H_1 : “ B gilt”.
- (iii) Wir machen n Experimente, X_1, \dots, X_n , und überprüfen, ob diese für H_1 sprechen. (Dieser Schritt entspricht der Herleitung des Widerspruchs im indirekten Beweis)

Anders als beim indirekten Beweis können beim Testen auch Fehler auftreten, die sich aus der Natur des Zufalls zwangsläufig ergeben. Man unterscheidet hier den Fehler 1-ter Art, d.h. H_0 abzulehnen obwohl H_0 richtig ist, vom Fehler 2-ter Art, d.h. H_1 abzulehnen, obwohl H_1 richtig ist. Im Konkreten ist ein maximaler Fehler 1-ter Art vorgegeben und der Test muss so konstruiert sein, dass dieses “Testniveau” (maximaler Fehler 1-ter Art) eingehalten wird. Falls wir den Fehler 1-ter Art nur asymptotisch ($n \rightarrow \infty$) garantieren können, dann sprechen wir von einem asymptotischen Test.

1.5.3 Beispiel

Gegeben sei $(X_n)_{n \in \mathbb{N}}$ IID Folge über $(\Omega, \mathcal{A}, \mathbb{P})$ mit $\mathbb{E}(X_1^2) < \infty$. Bezeichne mit $\mu := \mathbb{E}(X_1)$. Getestet werden soll $H_0 : \mu \leq 17$ gegen $H_1 : \mu > 17$ zum Testniveau $\alpha = 5\%$ (max. Fehler 1-ter Art).

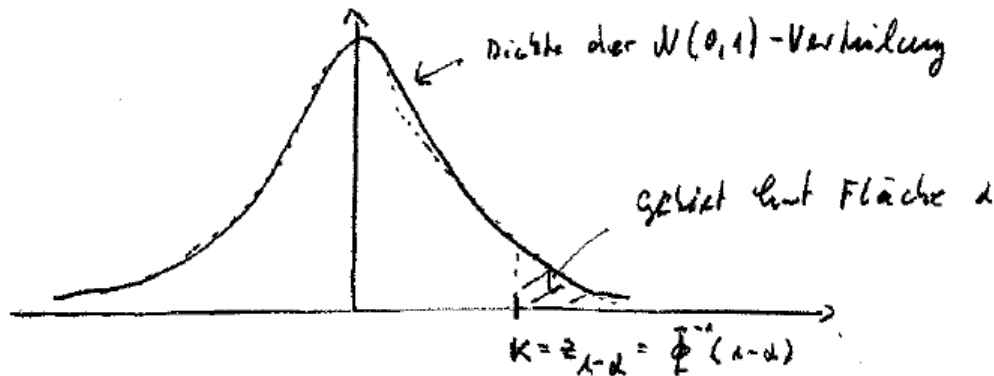
Das μ , das am nächsten zu H_1 liegt, wäre hier $\mu = 17$. Falls $\mu = 17$, so besagt das SLLN, dass

$$\sqrt{n} \cdot \frac{\bar{X}_n - 17}{\sqrt{S_n^2}} =: T_n$$

sich asymptotisch wie die Standardnormalverteilung Φ verhält. Falls $\mu > 17$, so besagt das SLLN, dass

$$\mathbb{P}(T_n > K) \rightarrow 1, n \rightarrow \infty$$

für jedes konstante K gelten muss.



Als Entscheidungsregel erhalten wir jetzt:

H_0 wird abgelehnt, falls $T_n > z_{1-\alpha}$ (das $1 - \alpha$ Quantil der $\mathcal{N}(0, 1)$ -Verteilung)

H_0 wird beibehalten, falls $T_n \leq z_{1-\alpha}$

Falls unser Test nicht dazu geführt hat, dass wir H_0 ablehnen konnten, d.h. wir haben keinen Widerspruch zu H_0 gefunden, dann bedeutet das nicht, dass H_0 richtig ist!

In der Regel ist es unbefriedigend, wenn ein Test nicht H_0 ablehnt. Es gibt aber Situationen, bei denen man genau solch ein Ergebnis erhofft. Im Folgenden skizzieren wir zwei Situationen dieser Art.

(i) Wir wollen einen Test anwenden, der voraussetzt, dass die vorliegenden Daten einer Normalverteilung entstammen. Zur Überprüfung dieser Voraussetzung testen wir:

H_0 : Daten sind normalverteilt

gegen

H_1 : Daten sind nicht normalverteilt

Hierzu wenden wir einen Anpassungstest (goodness-of-fit) an und hoffen, dass H_0 nicht abgelehnt wird.

(ii) Wir haben einen neuen Zufallszahlengenerator für die Gleichverteilung auf $[0, 1]$ gebaut und wollen testen, ob dieser auch funktioniert. Also:

H_0 : Daten sind auf $[0, 1]$ gleich verteilt

gegen

H_1 : Daten sind nicht auf $[0, 1]$ gleich verteilt

Auch hier würde man wieder einen Anpassungstest (χ^2 -Test oder Kolmogorov-Smirnov-Test) durchführen und hoffen, dass H_0 nicht abgelehnt wird.