Skript zu den Kapiteln 4 und 5 der Vorlesung

Angewandte Statistik in den Humanwissenschaften

WiSe 25/26

Prof. Dr. Michael Kohler
Fachbereich Mathematik
Technische Universität Darmstadt
kohler@mathematik.tu-darmstadt.de

“Those who ignore Statistics are condemned to reinvent it.”

BRAD EFRON

“Was war das für eine Stimme?” schrie Arthur.

“Ich weiß es nicht”, brüllte Ford zurück, “ich weiß es nicht. Es klang wie eine Wahrscheinlichkeitsrechnung.”

“Wahrscheinlichkeit? Was willst du damit sagen?”

“Eben Wahrscheinlichkeit. Verstehst du, so was wie zwei zu eins, drei zu eins, fünf zu vier. Sie sagte, zwei hoch einhunderttausend zu eins. Das ist ziemlich unwahrscheinlich, verstehst du?”

Ein Fünf-Millionen-Liter-Bottich Vanillesoße ergoß sich ohne Warnung über sie.

“Aber was soll das denn?” rief Arthur.

“Was, die Vanillesoße?”

“Nein, die Wahrscheinlichkeitsrechnung!”

DOUGLAS ADAMS

Inhaltsverzeichnis

Kapitel 4

Wahrscheinlichkeitstheorie

Ziel: Rückschlüsse aus Beobachtungen ziehen, die unter dem Einfluss von “Zufall” entstanden sind.

Dabei kann “Zufall” bei unvollständigen Beobachtungen nur gedacht sein, oder (wie z.B. bei Umfragen) künstlich bei der Erhebung der Daten eingefügt worden sein.

Beispiel: Es werden drei verschiedene Würfel geworfen. Bei fünfmaligem unbeeinflussten Werfen von jedem der drei Würfel erhält man als Ergebnisse:

1. Würfel: 6, 5, 3, 6, 1
2. Würfel: 2, 2, 7, 0, 8
3. Würfel: 4, 8, 2, 6, 1

Offensichtlich handelt es sich bei den Würfeln 2 und 3 nicht um normale Würfel, da nicht nur Zahlen aus $\{1, 2, \ldots, 6\}$ auftreten. Man erkennt aber auch, dass man keineswegs sicher sein kann, dass die drei Würfel verschieden sind.

4.1 Mathematische Beschreibung des Zufalls

Def. Ein **Zufallsexperiment** (kurz: ZE) is ein Experiment mit vorher unbestimmten Ergebnis, das im Prinzip unbeeinflusst voneinander beliebig oft wiederholt werden kann.

Bsp.: - Werfen eines Würfels
- Bestimmung der Reaktionszeit bei psychologischem Experiment einer (zufällig ausgewählten) Person.

Def. Die **Grundmenge** Ω eines ZE ist die Menge aller möglichen Ergebnisse des ZE.

Bsp.: - bei Werfen eines (echten) Würfels: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- Bei Messung der Reaktionszeit

$$\Omega = [0, \infty) \cup \{\infty\} \quad \text{(mit } \infty \text{ gleich Symbol für Unendlich (keine Reaktion))}$$

oder

$$\Omega = [0, T] \quad \text{(mit Abbruch nach Zeit } T\text{, } T > 0 \text{ fest).}$$

Def. Teilmengen der Grundmenge heißen **Ereignisse**. Die einelementigen Teilmengen heißen **Elementarereignisse**.

Bsp.: bei Würfel:

$$\begin{aligned} A_1 &= \{1\} \quad \text{(1 wurde gewürfelt)} \\ A_2 &= \{2, 4, 6\} \quad \text{(gerade Zahl wurde gewürfelt)} \\ A_3 &= \{2, 3, 4, 5, 6\} = A_1^c \quad \text{(keine 1 wurde gewürfelt)} \end{aligned}$$

Hierbei ist $A_1^c = \Omega \setminus A_1$ das sogenannte Komplement von A_1.

Bsp.: bei Reaktionszeit (in Sekunden) und $\Omega = [0, 100]$

$$\begin{aligned} A_1 &= [0, 5] \quad \text{(Reaktion innerhalb von 5 Sekunden)} \\ A_2 &= [10, 100] = [0, 10)^c \quad \text{(Mindestens 10 Sekunden bis zur Reaktion)} \end{aligned}$$

Def. Ein Ereignis **tritt ein**, falls das Ergebnis des ZE im Ereignis liegt.

ZE wird n-mal durchgeführt. Dann gibt die **absolute Häufigkeit** des Eintretens eines Ereignisses an, wie oft dabei das Ereignis eingetreten ist, die **relative Häufigkeit** des Eintretens eines Ereignisses ist die absolute Häufigkeit geteilt durch n.

Bsp.: $n = 5$-maliges Werfen eines echten Würfels ergibt die Zahlen: 5, 1, 5, 2, 4
Absolute Häufigkeit von $A = \{1, 3, 5\}$ ist 3.
Relative Häufigkeit von $A = \{1, 3, 5\}$ ist $3/5 = 0,6$.

Beobachtung aus der Praxis:

Empririsches Gesetz der großen Zahlen

Führt man ein ZE unbeeinflusst voneinander immer wieder durch, so nähert sich für große Anzahlen von Wiederholungen die relative Häufigkeit eines beliebigen Ereignisses A einer (von A abhängenden Zahl)

$$\mathbf{P}(A) \in [0, 1]$$

an. $\mathbf{P}(A)$ heißt **Wahrscheinlichkeit** (kurz: Wk.) von A.

Bsp.: Werfen eines echten Würfels, mit Laptop simulieren und auswerten.

Bez.:

$$\begin{aligned} \emptyset \quad &\ldots \quad \text{leere Menge} \\ A^c = \Omega \setminus A \quad &\ldots \quad \text{Komplement von } A \\ A \cup B \quad &\ldots \quad \text{Vereinigung von } A \text{ und } B \\ A \cap B \quad &\ldots \quad \text{Schnitt von } A \text{ und } B \end{aligned}$$

Eigenschaften von Wk.: ((i)-(v) folgen aus Eigenschaften von relativen Häufigkeiten)

(i) $0 \leq \mathbf{P}(A) \leq 1$ für alle $A \subseteq \Omega$.

(ii) $\mathbf{P}(\emptyset) = 0$, $\mathbf{P}(\Omega) = 1$.

(iii) $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$ für alle $A \subseteq \Omega$.

(iv) Für alle $A, B \subseteq \Omega$ mit $A \cap B = \emptyset$ gilt:

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

(v) Für alle $n \in \mathbb{N}$, $A_1, \ldots, A_n \subseteq \Omega$ mit $A_i \cap A_j = \emptyset$ für alle $i \neq j$ gilt:

$$\mathbf{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \cdots + \mathbf{P}(A_n).$$

(vi) Für alle $A_1, A_2, \cdots \subseteq \Omega$ mit $A_i \cap A_j = \emptyset$ für alle $i \neq j$ gilt:

$$\mathbf{P}\left(\cup_{n \in \mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)$$

(sog. σ-Additivität).

Folgerungen:

1. Für $A, B \subseteq \Omega$ mit $A \subseteq B$ gilt:

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

Denn: Aus $B = (B \setminus A) \cup A$ und $(B \setminus A) \cap A = \emptyset$ folgt mit (iv):

$$\mathbf{P}(B) = \mathbf{P}\left((B \setminus A) \cup A\right) = \mathbf{P}(B \setminus A) + \mathbf{P}(A).$$

2. Für $A, B \subseteq \Omega$ mit $A \subseteq B$ gilt:

$$\mathbf{P}(A) \leq \mathbf{P}(B).$$

Denn: Nach (1) und (i) gilt

$$\mathbf{P}(B) - \mathbf{P}(A) = \mathbf{P}(B \setminus A) \geq 0.$$

3. Für $A, B \subseteq \Omega$ gilt:

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B).$$

Denn: Es gilt

$$A \cup B = (A \setminus (A \cap B)) \cup (B \setminus (A \cap B)) \cup (A \cap B),$$

wobei die drei Mengen auf der rechten Seite paarweise leeren Schnitt haben. Mit (v) und (1) folgt:

$$\begin{aligned} \mathbf{P}(A \cup B) &= \mathbf{P}\left(A \setminus (A \cap B)\right) + \mathbf{P}\left(B \setminus (A \cap B)\right) + \mathbf{P}(A \cap B) \\ &= \mathbf{P}(A) - \mathbf{P}(A \cap B) + \mathbf{P}(B) - \mathbf{P}(A \cap B) + \mathbf{P}(A \cap B) \\ &= \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B). \end{aligned}$$

Das intuitive Verständnis von Wahrscheinlichkeiten ist oft nicht einfach. Ein berühmtes Bsp. dafür ist

Bsp.: Linda ist 31 Jahre alt, sie lebt alleine, ist selbstbewusst und sehr hübsch. Sie hat einen Magister in Philosophie. Während ihres Studiums hat sie sich intensiv mit Fragen sozialer Gerechtigkeit und Dikriminierung befasst und auch an Anti-Atom-Demonstrationen teilgenommen.

Was ist wahrscheinlicher:

1. Linda ist Bankangestellte.
2. Linda ist Bankangestellte und in der Frauenbewegung aktiv.

Nach (2) ist 1. das wahrscheinlichere Ereignis, intuitiv denken aber nach schnellem Lesen des Textes viele, dass das nicht so ist.

Bem.: Ist Ω eine Menge, so ist

$$\mathcal{P}(\Omega) = \{A \, : \, A \subseteq \Omega\} \quad \text{(sog. } \textbf{Potenzmenge} \text{ von } \Omega\text{)}$$

die Menge aller Teilmengen von Ω.

Bsp. $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$.

Def. Ein Tupel $(\Omega, \mathbf{P})$ mit $\Omega \neq \emptyset$ Menge und $\mathbf{P} : \mathcal{P}(\Omega) \to \mathbb{R}$ (d.h., $\mathbf{P}$ weist jeder Teilmenge von Ω eine reelle Zahl zu) Abbildung mit den Eigenschaften (i) bis (vi) von oben heißt **Wahrscheinlichkeitsraum** (kurz: W-Raum). In diesem Fall heißt **$\mathbf{P}$ Wahrscheinlichkeitsmaß** (kurz: W-Maß) und $\mathbf{P}(A)$ heißt **Wahrscheinlichkeit** (kurz: Wk.) von $A \subseteq \Omega$.

Bsp.: Werfen einer echten Münze. Wähle

$$\Omega = \{K, Z\}$$

(womit $\mathcal{P}(\Omega) = \{\emptyset, \{K\}, \{Z\}, \{K, Z\}\}$) und lege $\mathbf{P} : \mathcal{P}(\Omega) \to \mathbb{R}$ fest durch

- $\mathbf{P}(\emptyset) = 0$ (nach (ii))
- $\mathbf{P}(\Omega) = 1$ (nach (ii))

- $\mathbf{P}(\{K\}) = \mathbf{P}(\{Z\}) = 1/2$

  Denn: Da es sich um eine echte Münze handelt, muss $\mathbf{P}(\{K\}) = \mathbf{P}(\{Z\})$ gelten, und mit (iv) gilt weiter

$$1 = \mathbf{P}(\Omega) = \mathbf{P}(\{K\} \cup \{Z\}) = \mathbf{P}(\{K\}) + \mathbf{P}(\{Z\}).$$

Bem.: Aus technische Gründen kann man oft $\mathbf{P}(A)$ nicht für alle $A \subseteq \Omega$ sinnvoll festlegen, was wir in dieser Vorlesung aber vernachlässigen.

4.2 Der Laplacesche Wahrscheinlichkeitsraum

Def. W-Raum $(\Omega, \mathbf{P})$ mit Ω endliche Menge und

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|} \quad \text{für alle } A \subseteq \Omega$$

(mit $|A|$ = Anzahl Elemente von A) heißt **Laplacescher W-Raum**.

Bem.: Wird zu Modellierung von ZE verwendet bei denen

- nur endlich viele verschiedene Werte als Ergebnis vorkommen,
- jeder dieser Werte mit der gleichen Wahrscheinlichkeit auftritt.

Bsp.: Geworfen werden 4 Münzen (zwei 50 Cent, eine 1 Euro und eine 2 Euro Münze). Spieler bekommt als Gewinn alle Münzen, die mit Zahl oben landen. Mit welcher Wk. ist dann der Gewinn größer oder gleich einem Euro ?

Modellierung:

1. Ergebnis des ZE festlegen
   (und zwar so, dass alle Ergebnisse die gleiche Wk. haben, damit wir Laplaceschen W-Raum verwenden können).

   Dazu nummerieren wir die Münzen durch (Münze 1: 50 Cent, Münze 2: 50 Cent, Münze 3: 1 Euro, Münze 4: 2 Euro) und beschreiben das Ergebnis des ZE durch

$$\omega = (\omega_1, \omega_2, \omega_3, \omega_4) \quad \text{mit } \omega_i = \begin{cases} Z & \text{falls Münze } i \text{ mit Zahl oben landet,} \\ K & \text{sonst.} \end{cases}$$

Dann ist

$$\Omega = \{(\omega_1, \omega_2, \omega_3, \omega_4) \,:\, \omega_i \in \{K, Z\}\,(i \in \{1, \ldots, 4\})\}$$

die Grundmenge.

2. W-Raum festlegen.

   Da hier jedes der endlich vielen möglichen Ergebnisse mit der gleichen Wk. auftritt, wählen wir einen Laplaceschen W-Raum $(\Omega, \mathbf{P})$ mit $\mathbf{P} : \mathcal{P}(\Omega) \to [0,1]$ festgelegt durch

   $$\mathbf{P}(A) = \frac{|A|}{|\Omega|} \quad (A \subseteq \Omega).$$

   Hierbei gilt

   $$\Omega = 2 \cdot 2 \cdot 2 \cdot 2 = 16.$$

   Gesucht ist dann

   $$\mathbf{P}(B) = \frac{|B|}{|\Omega|} = \frac{|B|}{16}$$

   mit

   $$B = \{(\omega_1, \omega_2, \omega_3, \omega_4) \,:\, \text{Wert Münzen mit Zahl oben} \geq 1 \text{ Euro}\}.$$

3. Anzahl Elemente in B bestimmen.

   Dazu kann man z.B. in einer Tabelle alle Werte von B durchgehen (und den Wert der Münzen mit Zahl oben bestimmen), oder alternativ hier wie folgt vorgehen:

   $$\begin{aligned} B^c &= \{(\omega_1, \omega_2, \omega_3, \omega_4) \,:\, \text{Wert Münzen mit Zahl oben} < 1 \text{ Euro}\} \\ &= \{(K,K,K,K), (Z,K,K,K), (K,Z,K,K)\} \end{aligned}$$

   (da die dritte und die vierte Münze mit Kopf oben landen muss, sofern der Wert der Münzen mit Zahl oben kleiner als 1 Euro ist).

   Daher gilt

   $$|B| = |\Omega| - |B^c| = 16 - 3 = 13,$$

   also

   $$\mathbf{P}(B) = \frac{13}{16}.$$

   Alternativ erhält man das auch aus

   $$\mathbf{P}(B) = 1 - \mathbf{P}(B^c) = 1 - \frac{3}{16} = \frac{13}{16}.$$

Bem.: Im Laplaceschen W-Raum gilt immer

$$\text{Wk.} = \frac{\text{Anzahl günstiger Fälle}}{\text{Anzahl möglicher Fälle}}.$$

Bsp. Vor einer Klausur werden 60 Studierende zufällig in zwei Gruppen (SG und KG) der Größe 30 unterteilt. SG wird zu Klausurvorbereitungskurs eingeladen, KG nicht.

Abschneiden bei Klausur:

	n	Anzahl nicht bestanden	Prozent nicht bestanden
SG	30	1	3,3%
KG	30	5	16,7%

Kann man daraus schließen, dass der Kurs (auch bei zukünftigen Kursteilnehmern) die Nicht-Bestehensquote absenkt ?

Grundidee in der Statistik zur Beantwortung dieser Frage:

1. Mache Annahme, dass der Kurs keinen Einfluss hat, also unterschiedliche Nicht-Bestehensquoten nur durch zufällige Unterteilung (der Gruppe der 60 Studierender in SG und KG) entstanden sind.

   Modell dazu:
   Aus Urne mit 54 weißen und 6 schwarzen Kupgeln werden zufällig 30 Kugeln (ohne Zurücklegen) herausgezogen. Dabei wird genau eine schwarze Kugel gezogen.

2. Berechne unter dieser Annhame die Wk., dass ein Resultat beobachtet wird, dass (mindestens) so stark gegen die Annahme spricht wie das tatsächlich Beobachtete.

   Im Modell oben:
   Berechne die Wk, dass höchstens eine schwarze Kugel gezogen wird.

3. Falls diese Wk. klein ist (z.B. $\leq 0,05$), so verwerfe die Annahme in 1., andernfalls nicht.

Hilfsmittel zur Berechnugn der Wk. im obigen Urnenmodell: **Kombinatorik**.

Betrachtet wird das Ziehen von k Elementen aus einer Grundmenge Ω vom Umfang $|\Omega| = n$. Die Anzahl aller möglichen Stichproben sei N.

Dabei kann man vier verschiedene Vorgehensweisen unterscheiden, und zwar je nachdem, ob man die Elemente unmittelbar nach dem Ziehen wieder zurücklegt oder nicht, und je nachdem, ob man die Reihenfolge, in der die Elemente gezogen werden, beachtet oder nicht.

Zuerst betrachten wir das Ziehen **mit Zurücklegen** und **mit Berücksichtigung der Reihenfolge**. Hierbei wird k mal ein Element aus der Grundmenge gezogen, dabei hat man jeweils n Möglichkeiten, so dass man für die Anzahl der möglichen Stichproben erhält:

$$N = n \cdot n \cdot n \cdot \ldots \cdot n = n^k.$$

Als nächstes wird das Ziehen **ohne Zurücklegen** und **mit Berücksichtigung der Reihenfolge** betrachtet. Hier hat man für das erste Elemente n Möglichkeiten, für das zweite aber nur noch $n-1$, für das dritte $n-2$, u.s.w., und für das k-te noch $(n-k+1)$ Möglichkeiten. Damit erhält man für die Anzahl der möglichen Stichproben:

$$N = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}.$$

Dabei ist $n! = n \cdot (n-1) \cdots 1$ (gesprochen: n Fakultät) die sogenannte Fakultät von n.

Nun wird das Ziehen **ohne Zurücklegen** und **ohne Berücksichtigung der Reihenfolge** betrachtet. Ordnet man jede der dabei erhaltenen Stichproben auf alle $k!$ möglichen Weisen um, so erhält man alle Stichproben bzgl. Ziehen ohne Zurücklegen und mit Berücksichtigung der Reihenfolge.

Beispiel: Für $\Omega = \{1,2,3\}$, $n=3$ und $k=2$ erhält man die Zuordnungen

$$\begin{aligned} (1,2) &\mapsto (1,2) \text{ oder } (2,1) \\ (1,3) &\mapsto (1,3) \text{ oder } (3,1) \\ (2,3) &\mapsto (2,3) \text{ oder } (3,2) \end{aligned}$$

Daher gilt für die Anzahl der möglichen Stichproben:

$$\begin{aligned} & N \cdot k! \\ = {} & \text{Wert beim Ziehen ohne Zurücklegen und mit Berücksichtigung} \\ & \text{der Reihenfolge} \\ = {} & \frac{n!}{(n-k)!}, \end{aligned}$$

also

$$N = \frac{n!}{(n-k)! \cdot k!} = \binom{n}{k}.$$

Hierbei ist $\binom{n}{k}$ (gesprochen: n über k) der sogenannte Binomialkoeffizient.

Zum Abschluss wird noch das Ziehen **mit Zurücklegen** und **ohne Berücksichtigung der Reihenfolge** betrachtet. Hierbei gilt für die Anzahl der möglichen Stichproben:

$$N = \binom{n+k-1}{k}.$$

(ohne Beweis)

Die Ergebnisse sind in der folgenden Tabelle zusammen gefasst.

Anzahl Möglichkeiten	Ziehen mit Zurücklegen	Ziehen ohne Zurücklegen
Ziehen mit Berücksichtigung der Reihenfolge	n^k	$\frac{n!}{(n-k)!}$
Ziehen ohne Berücksichtigung der Reihenfolge	$\binom{n+k-1}{k}$	$\binom{n}{k}$

Tabelle 4.1: Grundformeln der Kombinatorik.

Damit Berechnung der Wk. im Bsp. oben:
Urne mit 60 Kugeln, davon 54 weiß und 6 schwarz. Rein zufällig werden daraus 30 Kugeln (ohne Zurücklegen) gezogen.

Wk., dass höchstens 1 davon scharz ist

$= \mathbf{P}$ [Keine Kugel schwarz] $+ \mathbf{P}$ [Genau eine Kugel schwarz] .

Wir ziehen dabei aus $n = 60$ Kugeln $k = 30$ Kugeln ohne Zurücklegen und ohne Beachtung der Reihenfolge. Da jedes einzelne der endlich vielen möglichen Ergebnisse mit der gleichen Wk. auftritt, können wir die gesuchte Wk. mit der Formel

$$\text{Wk.} = \frac{\text{Anzahl günstiger Fälle}}{\text{Anzahl möglicher Fälle}}$$

berechnen. Wir erhalten

$$\mathbf{P}[\text{ Keine Kugel schwarz }] = \frac{\binom{54}{30}}{\binom{60}{30}}$$

(da bei den günstigen Fällen aus 54 weißen Kugeln 30 gezogen werden) sowie

$$\mathbf{P}\,[\text{ Genau 1 Kugel schwarz }] = \frac{\binom{54}{29}\cdot\binom{6}{1}}{\binom{60}{30}}$$

(da jetzt bei den günstigen Fällen aus 54 weißen Kugeln 29 und aus 6 schwarzen Kugeln genau 1 Kugel gezogen werden). Damit

$$\text{Wk., dass höchstens 1 davon scharz ist} = \frac{\binom{54}{30}}{\binom{60}{30}} + \frac{\binom{54}{29}\cdot\binom{6}{1}}{\binom{60}{30}} \approx 0,0973 > 0,05.$$

Folgerung: Im Bsp. kann nicht ausgeschlossen werden, dass der Kurs keinerlei Einfluss hat.

4.3 Modelle für Wahrscheinlichkeiten

Betrachtet wird ZE, dessen Ergebnis reelle Zahl ist.

4.3.1 Zufallsvariablen

Im Folgenden führen wir eine Methode ein, mit der sich ein W-Raum ausgehend von einem einfachen W-Raum konstruieren lässt. Dies wird auf eine neue Schreibweise führen. die sich später (z.B. bei der Beschreibung von unbeeinflussten Wiederholungen eines ZE) als sehr nützlich erweisen wird.

Wir betrachten einen W-Raum $(\Omega, \mathbf{P})$, der ein ZE mit Ergebnis ω beschreibt.

Bsp. Unbeeinflusstes Werfen dreier echter Würfel beschreibt der W-Raum $(\Omega, \mathbf{P})$ mit

$$\Omega = \{(\omega_1, \omega_2, \omega_3) \,:\, \omega_1, \omega_2, \omega_3 \in \{1, 2, \ldots, 6\}\} = \{1, 2, \ldots, 6\}^3$$

und

$$\mathbf{P} : \mathcal{P}(\Omega) \to \mathbb{R}, \quad \mathbf{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{6^3}.$$

Die Abbildung $X : \Omega \to \mathbb{R}$ beschreibe einen Teilaspekt des ZE.

Im Bsp: Der Teilaspekt sei die Summe der Augenzahlen der 3 Würfel. Dann ist die Abbildung $X : \Omega \to \mathbb{R}$ gegeben durch

$$X((\omega_1, \omega_2, \omega_3)) = \omega_1 + \omega_2 + \omega_3.$$

Ein W-Raum, der das ZE mit Ergebnis $X(\omega)$ beschreibt, ist dann

$$(\mathbb{R}, \mathbf{P}_X)$$

mit

$$\mathbf{P}_X : \mathcal{P}(\mathbb{R}) \to \mathbb{R}, \quad \mathbf{P}_X(A) := \mathbf{P}[X \in A] := \mathbf{P}\left(\{\omega \in \Omega \,:\, X(\omega) \in A\}\right).$$

Im Bsp beschreiben wir das ZE mit Ergebnis gleich der Summe der Augenzahlen, die oben landen, mittels

$$(\mathbb{R}, \mathbf{P}_X),$$

wobei

$$\begin{aligned} \mathbf{P}_X(A) &= \mathbf{P}\left(\{(\omega_1, \omega_2, \omega_3) \in \Omega \,:\, \omega_1 + \omega_2 + \omega_3 \in A\}\right) \\ &= \frac{|\{(\omega_1, \omega_2, \omega_3) \in \Omega \,:\, \omega_1 + \omega_2 + \omega_3 \in A\}|}{6^3}. \end{aligned}$$

Hierbei haben wir zur Vereinfachung der Schreibweise die Grundmenge des W-Raumes gößer als eigentlich nötig gewählt, was aber in der Definition von $\mathbf{P}_X$ korrigiert wird (da die obige Abbildung $\mathbf{P}_X(\mathbb{R} \setminus \{3, 4, \ldots, 18\}) = 0$ erfüllt).

Interessieren wir uns z.B. für die Wk., dass die Summe der Augenzahlen gleich 17 oder 18 ist, so erhalten wir

$$\begin{aligned} \mathbf{P}_X(\{17, 18\}) &= \frac{|\{(\omega_1, \omega_2, \omega_3) \in \Omega \,:\, \omega_1 + \omega_2 + \omega_3 \in \{17, 18\}\}|}{6^3} \\ &= \frac{|\{(5,6,6), (6,5,6), (6,6,5), (6,6,6)\}|}{6^3} \\ &= \frac{4}{6^3} = \frac{1}{54}. \end{aligned}$$

Hierbei heißt:

- X **Zufallsvariable** (kurz: ZV)
- $\mathbf{P}_X : \mathcal{P}(\mathbb{R}) \to [0, 1]$ **Verteilung** von X

Damit neue Schreibweise:

bisher	neu
Ergebnis des ZE	ZV
Wk. / W-Maß	Verteilung

Im Weiteren: Konstruktion von Modellen für $\mathbf{P}_X$.

Dafür gibt es zwei Möglichkeiten:

1. $\mathbf{P}_X(B) = \mathbf{P}[X \in B]$ als **Summe** von Einzelwk. festlegen.
2. $\mathbf{P}_X(B) = \mathbf{P}[X \in B]$ als **Integral** über sogenannte Dichte festlegen.

4.3.2 Verteilungen mit Zähldichte

Zur Vereinfachung der Schreibweise nehmen wir an: Ergebnis des ZE liege in $\mathbb{N}_0$.

Wir legen Wk. $\mathbf{P}[X = k]$ von jedem $k \in \mathbb{N}_0$ fest und setzen dann für $B \subseteq \mathbb{R}$:

$$\mathbf{P}_X(B) = \mathbf{P}[X \in B] = \sum_{k \in B \cap \mathbb{N}_0} \mathbf{P}[X = k].$$

(Hier ist $\mathbf{P}_X$ Abbildung von $\mathcal{P}(\mathbb{R})$ nach $[0,1]$, wir legen daher (im Hinblick auf eine einheitliche Schreibweise) die Wk. für jede Teilmenge von $\mathbb{R}$ fest, obwohl nach Annahme oben klar ist, dass beim ZE nur Werte aus $\mathbb{N}_0$ auftreten).

Bsp. Ein echter Würfel wird solange geworfen, bis er zum ersten Mal mit 6 oben landet. X sei die zufällige Anzahl der Würfe. Wie gross ist dann

$$\mathbf{P}[X \geq 10] = \mathbf{P}[X \in \{10, 11, 12, \dots\}]\,?$$

Hier ist

$$\mathbf{P}[X \geq 10] = \sum_{k=10}^{\infty} \mathbf{P}[X = k]$$

und für $k \in \mathbb{N}$ gilt

$$\mathbf{P}[X = k] = \frac{\text{Anzahl günstiger Fälle}}{\text{Anzahl möglicher Fälle}} = \frac{5^{k-1}}{6^k} = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{k-1}.$$

Hierbei wird das k-malige Werfen eines echten Würfels betrachtet, was dem Ziehen von k Elementen aus einer Grundmenge vom Umfang $n = 6$ mit Zurücklegen und mit Betrachtung der Reihenfolge entspricht. Günstig sind die Fälle, bei denne zuerst $(k-1)$-mal eine Zahl aus $\{1, 2, \dots, 5\}$ gezogen wird und als letztes eine 6 gezogen wird.

Damit folgt:

$$\begin{aligned}\mathbf{P}[X \geq 10] &= \sum_{k=10}^{\infty} \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{k-1} = \left(\frac{1}{6} \cdot \left(\frac{5}{6}\right)^{9} + \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{10} + \ldots\right) \\ &= \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{9} \cdot \sum_{k=0}^{\infty} \left(\frac{5}{6}\right)^{k} = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{9} \cdot \frac{1}{1-\frac{5}{6}} = \left(\frac{5}{6}\right)^{9} \approx 0,19.\end{aligned}$$

Def. Eine Folge $(p_k)_{k \in \mathbb{N}_0}$ reller Zahlen mit

$$0 \leq p_k \leq 1 \quad (k \in \mathbb{N}_0) \quad \text{und} \quad \sum_{k=0}^{\infty} p_k = 1$$

heißt **Zähldichte**. In diesem Fall ist

$$\mathbf{P}_X : \mathcal{P}(\mathbb{R}) \to [0,1], \quad \mathbf{P}_X(A) = \sum_{k \in A \cap \mathbb{N}_0} p_k$$

die zu $(p_k)_{k \in \mathbb{N}_0}$ gehörende Verteilung (sog. **Verteilung mit Zähldichte** $(p_k)_{k \in \mathbb{N}_0}$).

Hierbei: p_k = Wk., dass X den Werte k annimmt $= \mathbf{P}[X = k]$.

Gebräuchliche Modelle für Verteilungen mit Zähldichte sind:

1. **Diskrete Gleichverteilung** (mit Parameter $K \in \mathbb{N}$):

$$p_k = \frac{1}{K} \quad (0 \leq k \leq K-1) \quad \text{und} \quad p_k = 0 \quad (k \geq K).$$

(Anwendung: Analog zum Laplaceschen W-Raum)

2. **Geometrische Verteilung** (mit Parameter $p \in (0,1)$):

$$p_k = (1-p)^k \cdot p \quad (k \in \mathbb{N}_0).$$

(Anwendung: Analog zum Würfelbeispiel oben, jetzt aber Anzahl Misserfolge bis zum ersten Erfolg statt wie oben Anzahl Versuche bis zum ersten Erfolg (wobei Erfolg das Würfeln der 6 ist))

3. **Binomialverteilung (kurz: $b(n,p)$)** (mit Parametern $n \in \mathbb{N}, p \in [0,1]$):

$$p_k = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad (k \in \{0,1,\ldots,n\}) \quad \text{und} \quad p_k = 0 \quad (k > n).$$

(Anwendung: vgl. Übungen, beschreibt zufälliges Verhalten der Anzahl der Erfolge bei n-maligem unbeeinflussten Durchführen eines ZE mit Erfolgswk. p beim Einzelversuch)

4. **Poissonverteilung (kurz: $\pi(\lambda)$)** (mit Parameter $\lambda \in (0,\infty)$):

$$p_k = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad (k \in \mathbb{N}_0).$$

(Anwendung: Wird zur Approximation von $b(n,p)$ mit $\lambda = n \cdot p$ verwendet, sofern n gross und p klein ist.)

4.3.3 Verteilungen mit Dichte

Hier legen wir die Wk. von $A \subseteq \mathbb{R}$ fest, indem wir eine feste Funktion $f : \mathbb{R} \to \mathbb{R}_+$ über A integrieren, d.h. wir setzen

$$\mathbf{P}_X(A) = \mathbf{P}[X \in A] = \int_A f(x)\,dx.$$

Wegen

$$\mathbf{P}_X(\mathbb{R}) = 1$$

muss hierbei gelten:

$$\int_{\mathbb{R}} f(x)\,dx = 1.$$

Def. Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) \geq 0 \quad (x \in \mathbb{R}) \quad \text{und} \quad \int_{\mathbb{R}} f(x)\,dx = 1$$

heißt **Dichte**. In diesem Falle ist

$$\mathbf{P}_X : \mathcal{P}(\mathbb{R}) \to [0,1], \quad \mathbf{P}_X(A) = \int_A f(x)\,dx$$

die zu f gehörende Verteilung (sog. **Verteilung mit Dichte** f).

Modelle für Verteilungen mit Dichte:

1. **Gleichverteilung** (kurz: $U(a,b)$) (mit Parametern $a, b \in \mathbb{R}, a < b$):

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{für} \quad a \leq x \leq b, \\ 0 & \text{für} \quad x < a \text{ oder } x > b. \end{cases}$$

Anwendung: rein zufälliges Ziehen einer Zahl aus dem Intervall $[a,b]$.

2. **Exponentialverteilung** (kurz: $\exp(\lambda)$) (mit Parameter $\lambda \in (0, \infty)$):

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & \text{für} \quad x \geq 0, \\ 0 & \text{für} \quad x < 0. \end{cases}$$

   Anwendung: Wird zur Modellierung von Wartezeiten verwendet.

3. **Normalverteilung** (kurz: $N(a, \sigma^2)$) (mit Parametern $a \in \mathbb{R}, \sigma \in (0, \infty)$):

$$f(x) = \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \quad (x \in \mathbb{R}).$$

   Anwendung: Wird zur Approximation von Verteilung aller Art verwendet (vgl. zentraler Grenzwertsatz in Kapitel 5).

Bsp.: Eine Zahl wird rein zufällig aus dem Intervall $[0, 5]$ gezogen. Wie groß ist die Wk., dass die Zahl ≥ 3 ist ?

Sei X die (zufällig) gezogene Zahl.

Gesucht: $\mathbf{P}[X \geq 3] = \mathbf{P}[X \in [3, \infty)] = \mathbf{P}_X([3, \infty))$.

Wir bestimmen Dichte $f : \mathbb{R} \to \mathbb{R}_+$ mit

$$\mathbf{P}[X \in A] = \int_A f(x)\, dx \quad \text{für } A \subseteq \mathbb{R}.$$

Da der Wert von X immer aus $[0, 5]$ stammt, setzen wir $f(x) = 0$ für $x < 0$ oder $x > 5$.

Da die Zahl rein zufällig gezogen wird (und daher kein Teilbereich des Intervalls $[0, 5]$ beim Ziehen bevorzugt wird), wählen wir die Dichte als konstant auf $[0, 5]$. Mit

$$\int_{\mathbb{R}} f(x)\, dx = 1$$

folgt:

$$f(x) = \begin{cases} \frac{1}{5} & \text{für} \quad 0 \leq x \leq 5, \\ 0 & \text{für} \quad x < 0 \text{ oder } x > 5. \end{cases}$$

Damit

$$\mathbf{P}_X([3, \infty)) \quad = \quad \int_{[3,\infty)} f(x)\, dx = \int_3^\infty f(x)\, dx.$$

Um das Integral über die obige abschnittsweise definierte Funktion berechnen zu können, spalten wir den Integrationsbereich gemäß der abschnittsweisen Definition des Definitionsbereiches von f auf und erhalten unter Anwendung des Hauptsatzes der Differential- und Integralrechnung

$$\begin{aligned}
\mathbf{P}_X([3,\infty)) &= \int_3^5 f(x)\,dx + \int_5^\infty f(x)\,dx \\
&= \int_3^5 \frac{1}{5}\,dx + \int_5^\infty 0\,dx = \frac{1}{5}\cdot x\Big|_{x=3}^{5} + 0 \\
&= \frac{1}{5}\cdot 5 - \frac{1}{5}\cdot 3 = \frac{2}{5} = 0,4.
\end{aligned}$$

Bem.: Im Beispiel oben ist X gleichverteilt auf $[0,5]$ (kurz: $U[0,5]$-verteilt).

4.4 Der Erwartungswert

ges. "Mittlerer Wert" $\mathbf{E}X$ eines ZE mit Ergebnis X.

Motivation im Falle, dass nur $0,1,\ldots,K$ als Ergebnis vorkommen:

$x_1, \ldots, x_n \in \{0,1,\ldots,K\}$ seien n beobachtete Werte des ZE. Dann soll der Erwartungswert so definiert werden, dass für n groß gilt:

$$\mathbf{E}X \approx \frac{1}{n}\sum_{i=1}^n x_i.$$

Umsummieren der endlich vielen Summanden, also eine Umschreibung des Ausdruckes rechts der Form

$$\frac{1}{6}\cdot(1+1+3+2+3+1) = \frac{1}{6}\cdot(3\cdot 1 + 1\cdot 2 + 2\cdot 3)$$

liefert

$$\begin{aligned}
\mathbf{E}X &\approx \frac{1}{n}\sum_{k=0}^K |\{1\le i\le n : x_i = k\}|\cdot k \\
&= \sum_{k=0}^K \frac{|\{1\le i\le n : x_i = k\}|}{n}\cdot k \\
&\approx \sum_{k=0}^K k\cdot \mathbf{P}[X=k],
\end{aligned}$$

wobei wir im letzten Schritt das empirische Gesetz der großen Zahlen verwendet haben.

Dies motiviert:

Def. Ist X ZV, die mit Wk. 1 nur Werte aus $\mathbb{N}_0$ annimmt, so heißt

$$\mathbf{E}X = \sum_{k=0}^{\infty} k \cdot \mathbf{P}[X = k]$$

(sofern existent) **Erwartungswert** von X (kurz: EW von X).

ALSO: Auftretende Werte mit den zugehörigen Wk. multiplizieren und aufaddieren.

Bsp.: Ist X $\pi(\lambda)$-verteilt, d.h.

$$\mathbf{P}[X = k] = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \quad (k \in \mathbb{N}_0),$$

dann gilt:

$$\begin{aligned}
\mathbf{E}X &= \sum_{k=0}^{\infty} k \cdot \mathbf{P}[X = k] = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \\
&= \lambda \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \cdot e^{-\lambda} = \lambda \cdot 1 = \lambda,
\end{aligned}$$

wobei wir im vorletzten Schritt ausgenutzt haben, dass die Zähldichte von $\pi(\lambda)$ zu Eins aufaddiert.

Bsp.: Sei X $b(n,p)$-verteilt, d.h.

$$\mathbf{P}[X = k] = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad (k \in \{0, 1, \dots, n\}) \quad \text{und} \quad \mathbf{P}[X = k] = 0 \quad (k > n).$$

Wegen

$$\begin{aligned}
\binom{n}{k} &= \frac{n \cdot (n-1) \cdots (n-k+1)}{k \cdot (k-1) \cdots 1} \\
&= \frac{n}{k} \cdot \frac{(n-1) \cdot (n-2) \cdots (n-1-(k-1)+1)}{(k-1) \cdot (k-2) \cdots 1} \\
&= \frac{n}{k} \cdot \binom{n-1}{k-1}
\end{aligned}$$

für $1 \leq k \leq n$ gilt

$$\begin{aligned}
\mathbf{E}X &= \sum_{k=0}^{\infty} k \cdot \mathbf{P}[X = k] = \sum_{k=0}^{n} k \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} = \sum_{k=1}^{n} k \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\
&= n \cdot p \cdot \sum_{k=1}^{n} \binom{n-1}{k-1} \cdot p^{k-1} \cdot (1-p)^{n-1-(k-1)} = n \cdot p \cdot 1 = n \cdot p.
\end{aligned}$$

Hierbei haben wir im vorletzten Schritt ausgenützt, dass die Zähldichte von $b(n-1, p)$ zu Eins aufaddiert.

Im Falle einer ZV mit Dichte ersetzen wir in der Defintion oben die Summe durch ein Integral.

Def. Ist X ZV mit Dichte f, so heißt

$$\mathbf{E}X = \int_{\mathbb{R}} x \cdot f(x)\, dx$$

(sofern existent) der **Erwartungswert** von X.

ALSO: Dichte mit x multiplizieren und Resultat über $\mathbb{R}$ integrieren.

Bsp.: X gleichverteilt auf $[0, 10]$, d.h. X hat Dichte

$$f(x) = \begin{cases} \frac{1}{10} & \text{für} \quad 0 \leq x \leq 10, \\ 0 & \text{für} \quad x < 0 \text{ oder } x > 10. \end{cases}$$

Dann gilt:

$$\begin{aligned}
\mathbf{E}X &= \int_{\mathbb{R}} x \cdot f(x)\, dx = \int_{-\infty}^{\infty} x \cdot f(x)\, dx \\
&= \int_{-\infty}^{0} x \cdot 0\, dx + \int_{0}^{10} x \cdot \frac{1}{10}\, dx + \int_{10}^{\infty} x \cdot 0\, dx \\
&= 0 + \frac{1}{20} \cdot x^2 \Big|_{x=0}^{10} + 0 + 0 \\
&= \frac{1}{20} \cdot (10)^2 - \frac{1}{20} \cdot 0^2 = 5.
\end{aligned}$$

Bsp.: X $N(a, \sigma^2)$-verteilt, d.h. X hat Dichte

$$f(x) = \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \quad (x \in \mathbb{R}).$$

Dann gilt

$$
\begin{aligned}
\mathbf{E}X &= \int_{-\infty}^{\infty} x \cdot f(x)\,dx \\
&= \int_{-\infty}^{\infty} (x-a) \cdot f(x)\,dx + a \cdot \int_{-\infty}^{\infty} f(x)\,dx \\
&= 0 + a \cdot 1 = a,
\end{aligned}
$$

wobei wir beim vorletzten Gleichheitszeichen ausgenutzt haben, dass

$$(x-a) \cdot f(x) = (x-a) \cdot \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}}$$

punktsymmetrisch bzgl. $x = a$ ist, und dass f eine Dichte ist und daher zu Eins integriert.

Eigenschaften des Erwartungswertes:

a) Monotonie des Erwartungswertes

Sind X_1, X_2 ZVen mit $X_1 \leq X_2$ (d.h., Wert beim 1. ZE ist immer kleiner als Wert beim 2. ZE), so gilt:

$$\mathbf{E}X_1 \leq \mathbf{E}X_2$$

(d.h., 1. Mittelwert $\leq$ 2. Mittelwert)

b) Linearität des Erwartungswertes

Sind X_1, X_2 ZVen und sind $a, b \in \mathbb{R}$, so gilt:

$$\mathbf{E}\left(a \cdot X_1 + b \cdot X_2\right) = a \cdot \mathbf{E}X_1 + b \cdot \mathbf{E}X_2$$

(d.h. Mittelwert einer Linearkombination zufälliger Werte ist die Linearkombination der Mittelwerte).

Verallgemeinerung der bisherigen Formeln:

1. Ist X ZV, die mit Wk. 1 nur Werte in $\mathbb{N}_0$ annimmt, so setzen wir für $h : \mathbb{R} \to \mathbb{R}$

$$\mathbf{E}\left[h(X)\right] = \sum_{k=0}^{\infty} h(k) \cdot \mathbf{P}[X = k]$$

(sofern existent).

2. Ist X ZV mit Dichte f, so setzen wir für $h : \mathbb{R} \to \mathbb{R}$
$$\mathbf{E}\left[h(X)\right] = \int_{\mathbb{R}} h(x) \cdot f(x)\, dx$$
(sofern existent).

Bsp. Ist X_1 auf $\{0, 1, \ldots, 4\}$ gleichverteilt, d.h.
$$\mathbf{P}[X_1 = k] = \begin{cases} \frac{1}{5} & \text{für} \quad k \in \{0, 1, \ldots, 4\}, \\ 0 & \text{für} \quad k > 4, \end{cases}$$
so gilt
$$\begin{aligned} \mathbf{E}(X_1^2) &= \sum_{k=0}^{\infty} k^2 \cdot \mathbf{P}[X = k] = \sum_{k=0}^{4} k^2 \cdot \frac{1}{5} \\ &= \frac{1}{5} \cdot (0^2 + 1^2 + 2^2 + 3^2 + 4^2) = \frac{30}{5} = 6. \end{aligned}$$
Ist X_2 gleichverteilt auf $[0, 10]$, d.h. X_2 hat Dichte
$$f(x) = \begin{cases} \frac{1}{10} & \text{für} \quad 0 \leq x \leq 10, \\ 0 & \text{für} \quad x < 0 \text{ oder } x > 10, \end{cases}$$
so gilt
$$\begin{aligned} \mathbf{E}\left(e^{X_2}\right) &= \int_{-\infty}^{\infty} e^x \cdot f(x)\, dx = \int_0^{10} e^x \cdot \frac{1}{10}\, dx \\ &= \frac{1}{10} \cdot e^x \Big|_{x=0}^{10} = \frac{1}{10} \cdot \left(e^{10} - 1\right). \end{aligned}$$

4.5 Die Varianz

ges.: Beschreibung der Schwankung eines zufälligen Wertes um seinen Mittelwert.

Wir verwenden dazu die mittlere quadratische Abweichung zwischen dem zufälligem Wert und seinem Mitelwert.

Def. Ist X ZV mit existierendem Erwartungswert $\mathbf{E}X \in \mathbb{R}$, so heißt
$$V(X) = \mathbf{E}\left(|X - \mathbf{E}X|^2\right)$$
die **Varianz** von X.

Bem.: Wir betrachten oben das Quadrat der (zufälligen) Abweichung $X - \mathbf{E}X$, da:

- Erwartungswert der Abweichung wäre Null,
- durch Verwendung des Quadrates werden Formeln (s.u.) einfacher.

Bem.: Beachte: $|a|^2 = a^2$ für $a \in \mathbb{R}$.

Bsp. X sei die Augenzahl beim (zufälligem) Werfen eines echten Würfels. D. g.

$$\mathbf{E}X = \sum_{k=1}^{6} k \cdot \mathbf{P}[X = k] = \sum_{k=1}^{6} k \cdot \frac{1}{6} = \frac{7}{2}$$

und unter Verwendung der Bezeichnung $h(x) = (x - 7/2)^2$ folgt

$$\begin{aligned} V(X) &= \mathbf{E}\left(|X - 7/2|^2\right) = \mathbf{E}h(X) = \sum_{k=1}^{6} h(k) \cdot \mathbf{P}[X = k] \\ &= \sum_{k=1}^{6} \left(k - \frac{7}{2}\right)^2 \cdot \frac{1}{6} = \cdots = \frac{35}{12} \approx 2.92. \end{aligned}$$

Bsp. Ist X $N(a, \sigma^2)$-verteilt, d.h. X hat die Dichte

$$f(x) = \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \quad (x \in \mathbb{R}),$$

so gilt:

$$\begin{aligned} V(X) &= \mathbf{E}\left((X - \mathbf{E}X)^2\right) = \mathbf{E}\left((X - a)^2\right) \\ &= \int_{-\infty}^{\infty} (x-a)^2 \cdot \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \, dx = \int_{-\infty}^{\infty} (x-a) \cdot \frac{(x-a)}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \, dx. \end{aligned}$$

Mit Hilfe von partieller Integration (wobei der erste Faktor abgeleitet und der Rest integriert wird) erhalten wir

$$\begin{aligned} V(X) &= (x-a) \cdot \frac{-\sigma^2}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \Big|_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty} 1 \cdot \frac{(-1) \cdot \sigma^2}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-a)^2}{2 \cdot \sigma^2}} \, dx \\ &= 0 + \sigma^2 \cdot \int_{-\infty}^{\infty} f(x) \, dx = \sigma^2. \end{aligned}$$

Satz Ist X ZV mit existierendem Erwartungswert $\mathbf{E}X \in \mathbb{R}$, so gilt

$$V(X) = \mathbf{E}(X^2) - (\mathbf{E}X)^2.$$

Beweis. Unter Verwendung der 2. binomischen Formel und der Linearität des Erwartungswertes folgt:

$$\begin{aligned} V(X) &= \mathbf{E}\left((X - \mathbf{E}X)^2\right) \\ &= \mathbf{E}\left(X^2 - 2 \cdot \mathbf{E}X \cdot X + (\mathbf{E}X)^2\right) \\ &= \mathbf{E}\left(X^2\right) - 2 \cdot \mathbf{E}X \cdot \mathbf{E}X + \mathbf{E}\left((\mathbf{E}X)^2\right) \\ &= \mathbf{E}(X^2) - 2 \cdot (\mathbf{E}X)^2 + (\mathbf{E}X)^2 \\ &= \mathbf{E}(X^2) - (\mathbf{E}X)^2, \end{aligned}$$

wobei wir beim vorletzten Gleichheitszeichen ausgenutzt haben, dass eine Zufallsvariable, die einen festen Wert (hier: $(\mathbf{E}X)^2$) mit Wahrscheinlichkeit Eins annimmt, genau diesen Wert auch als Erwartungswert hat. □

Im Bsp. oben:

$$\mathbf{E}(X^2) = \sum_{k=1}^{6} k^2 \cdot \frac{1}{6} = \cdots = \frac{91}{6}$$

und damit

$$V(X) = \mathbf{E}(X^2) - (\mathbf{E}X)^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \cdots = \frac{35}{12}.$$

Satz Ist X ZV mit existierendem Erwartungswert $\mathbf{E}X \in \mathbb{R}$, so gilt für beliebige $a, b \in \mathbb{R}$:

$$V(a \cdot X + b) = a^2 \cdot V(X).$$

Beweis. Unter zweimaliger Verwendung der Linearität des Erwartungswertes folgt:

$$\begin{aligned} V(a \cdot X + b) &= \mathbf{E}\left(|a \cdot X + b - \mathbf{E}(a \cdot X + b)|^2\right) \\ &= \mathbf{E}\left(|a \cdot X + b - (a \cdot \mathbf{E}(X) + b)|^2\right) \\ &= \mathbf{E}\left(|a \cdot X - a \cdot \mathbf{E}(X)|^2\right) \\ &= a^2 \cdot \mathbf{E}\left(|X - \mathbf{E}(X)|^2\right) \\ &= a^2 \cdot V(X). \end{aligned}$$

□

4.6 Unabhängigkeit

W-Raum $(\Omega, \mathbf{P})$, $A, B \subseteq \Omega$.

Wann beeinflussen sich A und B gegenseitig nicht, d.h., wann ändert sich durch das Beobachten des Eintretens von A nicht die Wk. von B, und umgekehrt ?

Dazu: ZE n mal durchführen. n_A bzw. n_B bzw. $n_{A\cap B}$ sei die Anzahl des Eintretens von A bzw. B bzw. $A \cap B$ bei diesem n-maligem Durchführen des ZE.

Wenn sich A und B im obigen Sinne nicht beeinflussen, dann sollte für n groß gelten:

$$\frac{n_{A\cap B}}{n_B} \approx \frac{n_A}{n} \quad \text{und} \quad \frac{n_{A\cap B}}{n_A} \approx \frac{n_B}{n},$$

d.h. die relative Häufigkeit des Eintretens von A unter allen denjenigen Ergebnissen, bei denen B eingetreten ist, sollte ungefähr gleich der relativen Häufigkeit von A bei allen n Ergebnissen sein. Und die relative Häufigkeit des Eintretens von B unter allen denjenigen Ergebnissen, bei denen A eingetreten ist, sollte ungefähr gleich der relativen Häufigkeit von B bei allen n Ergebnissen sein.

Dies ist äquivalent zu

$$\frac{n_{A\cap B}}{n} \approx \frac{n_A}{n} \cdot \frac{n_B}{n}.$$

Letzteres führt nach dem empirischen starken Gesetz der großen Zahlen für n groß auf die Forderung:

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

Def. a) W-Raum $(\Omega, \mathbf{P})$. Ereignisse A und B heißen **unabhängig**, falls gilt:

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

b) ZVen X_1, X_2 heißen **unabhängig**, falls für alle $A, B \subseteq \mathbb{R}$ gilt:

$$\mathbf{P}\left[X_1 \in A, X_2 \in B\right] = \mathbf{P}\left[X_1 \in A\right] \cdot \mathbf{P}\left[X_2 \in B\right].$$

Bem. b) bedeutet anschaulich: Werte von X_1 und X_2 beeinflussen sich gegenseitig nicht.

Man kann zeigen:

Sind X_1 und X_2 **unabhängige** ZVen mit existierenden Varianzen, so gilt:

$$V(X_1 + X_2) = V(X_1) + V(X_2).$$

ALSO: Bei Unabhängigkeit ist die Varianz einer Summe gleich der Summe der Varianzen.

Gilt aber ohne Unabhängigkeitsvoraussetzung nicht, denn

$$V(X + X) = 4 \cdot V(X) \neq V(X) + V(X)$$

für $V(X) \neq 0$.

Kapitel 5

Statistische Testverfahren

5.1 Einführung

Ziel: Hypothesen über die Realität empirisch “belegen” oder “widerlegen”.

Bsp. für Fragestellungen:

1. Schätzen Examenskandidaten ihre eigene Leistung richtig ein ?
2. Reden Frauen mehr als Männer ?
3. Welche Auswirkungen haben Computerspiele auf die Spielenden ?

Vorgehen:

i) Hypothese formulieren

ii) Methode der Datenerhebung festlegen

iii) Mathematisches Modell der Daten festlegen

iv) Hypothese im mathematischen Modell formulieren

v) Statistisches Testverfahren auswählen

vi) Daten erheben

vii) Statistisches Testverfahren anwenden

viii) Resultat auf Realität übertragen

In der Praxis: Hauptaufwand bei vi)

Manchmal sind Hypothese und/oder Daten schon vorgegeben.

Beachte: a) Daten erst ansehen, nachdem das statistische Testverfahren festgelegt wurde (!)

b) Statt Hypothese "belegen" / "widerlegen" liefert Statistik nur: Hypothese statistisch signifikant oder nicht statistisch signifikant (!) (vgl.: *Statistics means never having to say you are certain!*).

Zu i) Wir formulieren zwei entgegengesetze Hypothesen:
Nullhypothese H_0 und *Gegen- / Alternativhypothese* H_1

In Bsp. 1:
H_0: Examenskandidaten schätzen ihre Leistung eher zu gut ein.
H_1: Examenskandidaten schätzen ihre Leistung zu schlecht ein.

In Bsp. 2:
H_0: Frauen reden nicht mehr als Männer.
H_1: Frauen reden mehr als Männer.

Hierbei kommt das statistisch zu Sichernde in die Gegenhypothese (!)

Zu ii)

In Bsp. 1:

n Studierende in Klausur fragen, wieviele Punkte sie ihrer Meinung nach erzielt haben, und

$$x_i = \text{Echte Punktzahl} \ - \ \text{geschätzte Punktzahl bei Student } i$$

ermitteln für $i = 1, \ldots, n$.

In Bsp. 2:

Bei n männlichen Studierenden und m weiblichen Studierenden die Anzahlen gesprochener Worte $x_1, \ldots, x_n$ und $y_1, \ldots, y_m$ in einem festen Zeitraum ermitteln.

Zu iii)

Daten als Realisierungen von unabhängigen und identisch verteilten ZVn auffassen:

Def. $(\Omega, \mathbf{P})$ W-Raum, X_1, X_2, ...: $\Omega \to \mathbb{R}$ reelle ZVn.
a) X_1, X_2, ... **unabhängig** :$\Leftrightarrow$ Für alle $n \in \mathbb{N}$ und alle $B_1, \ldots, B_n \subseteq \mathbb{R}$ gilt:

$$\mathbf{P}[X_1 \in B_1, \ldots, X_n \in B_n] = \mathbf{P}[X_1 \in B_1] \cdots \mathbf{P}[X_n \in B_n].$$

b) X_1, X_2, ... **identisch verteilt** :$\Leftrightarrow$ $\mathbf{P}_{X_1} = \mathbf{P}_{X_2} = \ldots$

Entsprechend für X_1, ..., X_n.

Anschaulich:
ZVn unabhängig: Werte beeinflussen sich gegenseitig nicht
ZVn identisch verteilt: Werte werden nach dem gleichen Prinzip erzeugt.

Damit Annahme in Bsp. 1:

x_1, ..., x_n sind Realisierungen (beobachtete Werte) von ZVn X_1, ..., X_n, die unabhängig und identisch verteilt sind (*Einstichprobenproblem*).

Und Annahme in Bsp. 2:

$x_1, \ldots, x_n, y_1, \ldots, y_m$ sind Realisierungen von ZVn $X_1, \ldots, X_n, Y_1, \ldots, Y_m$, die unabhängig sind und bei denen X_1, ..., X_n identisch verteilt sind und Y_1, ..., Y_m identisch verteilt sind (*Zweistichprobenproblem*).

Weiter schränken wir die Klasse der betrachteten Verteilungen ein, z.B. in Bsp. 1 durch die Annahme:

$\mathbf{P}_{X_1}$ ist eine Normalverteilung mit unbekanntem EW $a \in \mathbb{R}$ und unbekannter Varianz σ^2 ($\sigma \in \mathbb{R}_+ \setminus \{0\}$).

Häufig wird hierbei die Normalverteilung gewählt, denn es gilt:

Satz (Zentraler Grenzwertsatz).

Sind X_1, X_2, ... unabhängige identisch verteilte reelle ZVn mit $\sigma^2 = V(X_1) > 0$ existent, so verhält sich

$$Z_n = \frac{\sum_{i=1}^n X_i - \mathbf{E}(\sum_{i=1}^n X_i)}{\sqrt{\mathbf{Var}(\sum_{i=1}^n X_i)}} = \frac{\sum_{i=1}^n X_i - n \cdot \mathbf{E}X_1}{\sqrt{n} \cdot \sigma}$$

annähernd wie eine $N(0,1)$-verteilte ZV.

Genauer gilt für jedes $x \in \mathbb{R}$:

$$\lim_{n\to\infty} \mathbf{P}[Z_n \leq x] = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot e^{-t^2/2} dt.$$

Anschaulich: Summen unabhängiger identisch verteilter ZVn sind annähernd normalverteilt.

In Bsp. 1:
Schätzung der Studierenden ist Summe einzelner Schätzungen, folglich normalverteilt.

Bsp. zu ZGWS am Rechner: Echten Würfel n-mal werfen, $x_1, \ldots, x_n$ seien die gewürfelten Zahlen. Dann verhält sich

$$\frac{\sum_{i=1}^{n} x_i - n \cdot 3,5}{\sqrt{n} \cdot \sqrt{35/12}}$$

wie Realisierung einer $N(0,1)$-verteilten Zufallsvariable. Davon N Werte am Rechner erzeugen, und Histogramm dieser Werte mit Dichte von $N(0,1)$ vergleichen.

Damit mathematisches Modell

- in Bsp. 1: $x_1, \ldots, x_n$ sind Realisierungen von unabh. id. verteilten ZVn $X_1, \ldots, X_n$ mit X_1 $N(a, \sigma^2)$-verteilt, wobei $a \in \mathbb{R}$, $\sigma > 0$ unbekannt sind.
- in Bsp. 2: $x_1, \ldots, x_n, y_1, \ldots, y_m$ sind Realisierungen von unabhängigen ZVn $X_1, \ldots, X_n, Y_1, \ldots, Y_m$, mit $X_1, \ldots, X_n$ identisch verteilt, $Y_1, \ldots, Y_m$ identisch verteilt, X_1 $N(a_X, \sigma_X^2)$-verteilt und Y_1 $N(a_Y, \sigma_Y^2)$-verteilt, wobei $a_X \in \mathbb{R}$, $\sigma_X > 0$ $a_Y \in \mathbb{R}$, $\sigma_Y > 0$ unbekannt sind.

Zu iv)

In Bsp. 1 (mit x_i = echte Punktzahl - geschätzte Punktzahl, $X_1 \sim N(a, \sigma^2)$):

$H_0 : a \leq 0$

$H_1 : a > 0$

In Bsp. 2 (mit $X_1 \sim N(a_X, \sigma_X^2)$, $Y_1 \sim N(a_Y, \sigma_Y^2)$):

$H_0 : a_X \geq a_Y$

$H_1 : a_X < a_Y$

Zu v) Statistisches Testverfahren festlegen

In Bsp. 1:

Geg.: Realisierungen $x_1, \dots, x_n$ von unabhängigen $N(a, \sigma^2)$-verteilten ZVn.

Zu entscheiden ist zwischen den beiden Hypothesen $H_0 : a \leq 0$ und $H_1 : a > 0$.

Ein **statistischer Test** ist eine Abbildung

$$\varphi : \mathbb{R}^n \to \{0, 1\}.$$

Gilt $\varphi(x_1, \dots, x_n) = 1$ so entscheiden wir uns für H_1.

Gilt $\varphi(x_1, \dots, x_n) = 0$ so entscheiden wir uns für H_0.

Mögliche Fälle:

	H_0 richtig	H_1 richtig
$\varphi(x_1, \dots, x_n) = 0$	richtig	falsch, sog. β-**Fehler** (bzw. Fehler 2. Art)
$\varphi(x_1, \dots, x_n) = 1$	falsch, sog. α-**Fehler** (bzw. Fehler 1. Art)	richtig

Im Falle $a \leq 0$ ist

$$\mathbf{P}_{(a,\sigma^2)}[\varphi(X_1, \dots, X_n) = 1]$$

die Wk. des Auftretens des α-Fehlers, und im Falle $a > 0$ ist

$$\mathbf{P}_{(a,\sigma^2)}[\varphi(X_1, \dots, X_n) = 0]$$

die Wk. des Auftretens des β-Fehlers. (Hierbei ist mit $\mathbf{P}_{(a,\sigma^2)}$ gemeint, dass in der Wk. $X_1, \dots, X_n$ unabhängig $N(a, \sigma^2)$-verteilt sind, und für jeden Wert von (a, σ^2) existiert die Wk. eines $\alpha-$ oder eines $\beta-$Fehlers.)

Problem: Es existiert kein Test, bei dem sowohl die α-Fehlerwk. als auch die β-Fehlerwk. minimal sind im Vergleich zu allen anderen Tests.

Ausweg: Wir geben Schranke für die α-Fehlerwk. vor ($\alpha \in (0, 1)$, oft $\alpha = 0.05$), betrachten nur Tests, deren α-Fehlerwk. alle kleiner oder gleich dieser Schranke sind (sog. Tests zum Niveau α), und nehmen unter diesen Test denjenigen, für den die maximale β-Fehlerwk. am kleinsten ist.

Folgerung: Ist nach Anwendung dieses Tests $\varphi(x_1, \ldots, x_n) = 1$ (und wir lehnen H_0 ab), so ist im Falle, dass H_0 richtig ist, ein Fehler aufgetreten. Ein solcher Fehler tritt aber nur mit Wk. $\leq \alpha$ auf.
Sprechweise: H_1 *ist statistisch signifikant.*

Ist dagegen $\varphi(x_1, \ldots, x_n) = 0$ (und wir lehnen H_0 nicht ab), so ist im Falle das H_1 richtig ist ein Fehler aufgetreten. Mit welcher Wk. ein solcher Fehler auftritt wissen wir aber leider nicht!
Sprechweise: H_1 *ist statistisch nicht signifikant.*

Daher: Das statistisch zu Sichernde kommt in Gegenhypothese (also in H_1).

5.2 Der t-Test von Student

Macht Aussagen über EW normalverteilter Daten.

Varianten davon:

- ein- oder zweiseitiger Test

- Ein- oder Zweistichprobenproblem

Einseitiger t-Test für eine Stichprobe

geg.: $x_1, \ldots, x_n \in \mathbb{R}$, $a_0 \in \mathbb{R}$, $\alpha \in (0, 1)$

Annahme an die Daten: $x_1, \ldots, x_n$ sind Realisierungen von unabhängig $N(a, \sigma^2)$-verteilten ZVn $X_1, \ldots, X_n$, wobei $a \in \mathbb{R}$, $\sigma > 0$ unbekannt sind.

Zu testen: $H_0 : a \leq a_0$ vs. $H_1 : a > a_0$ zum Niveau α.

Vorgehen: Schätze zunächst $a = \mathbf{E}X_1$ durch

$$\frac{1}{n} \sum_{i=1}^{n} X_i.$$

Plausibel, denn es gilt:

Satz (Starkes Gesetz der großen Zahlen):

Sind $X_1, X_2, \ldots$ unabhängige identisch verteilte reelle ZVn mit existierendem EW $\mathbf{E}X_1$, so gilt:

$$\mathbf{P}\left[\frac{1}{n}\sum_{i=1}^{n} X_i \to \mathbf{E}X_1 \quad (n \to \infty)\right] = 1.$$

Idee: Setze sodann

$$\varphi(x_1, \ldots, x_n) = \begin{cases} 1, & \text{falls } \frac{1}{n}\sum_{i=1}^{n} x_i - a_0 > c, \\ 0 & \text{sonst,} \end{cases}$$

mit $c \in \mathbb{R}$ geeignet.

Damit φ Test zum Niveau α ist (d.h., alle α-Fehler sind $\leq \alpha$) muss gelten:

$$\mathbf{P}_{(a,\sigma^2)}\left[\frac{1}{n}\sum_{i=1}^{n} X_i - a_0 > c\right] \leq \alpha$$

für alle $a \leq a_0$ und alle $\sigma > 0$.

Im Hinblick auf β-Fehler sollte c möglichst klein sein (damit H_1 möglichst selten abgelehnt wird).

Idee: Wähle c minimal so, dass Wk. oben für $a = a_0$ gerade α ist.

Setze

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)^2.$$

Man kann zeigen:
Sind $X_1, \ldots, X_n$ unabhängig $N(a_0, \sigma^2)$-verteilt, so ist

$$\frac{\sqrt{n}}{\sqrt{S^2}} \cdot \left(\frac{1}{n}\sum_{i=1}^{n} X_i - a_0\right) \quad t_{n-1}\text{-verteilt,}$$

d.h., t-verteilt mit $n-1$ Freiheitsgraden.

Daher gilt

$$\mathbf{P}_{(a_0,\sigma^2)}\left[\frac{\sqrt{n}}{\sqrt{S^2}} \cdot \left(\frac{1}{n}\sum_{i=1}^{n} X_i - a_0\right) > t_{n-1;\alpha}\right] = \alpha \quad \text{für alle } \sigma > 0,$$

wobei $t_{n-1;\alpha}$ das sog. α-Fraktil von t_{n-1} ist, d.h. für Vf. F von t_{n-1} gilt: $F(t_{n-1;\alpha}) = 1 - \alpha$.

Hierbei heißt die Funktion $F : \mathbb{R} \to \mathbb{R}$ mit $F(x) = \mathbf{P}[X \leq x]$ **Verteilungsfunktion** (kurz: Vf.) von X bzw. $\mathbf{P}_X$.

Dies führt auf den sog. **einseitigen t-Test von Student** (für eine Stichprobe):

$$\varphi(x_1, \dots, x_n) = \begin{cases} 1, & \text{falls } \frac{1}{n} \sum_{i=1}^n x_i > a_0 + \frac{s_n}{\sqrt{n}} \cdot t_{n-1;\alpha}, \\ 0 & \text{sonst,} \end{cases}$$

wobei

$$s_n^2 = \frac{1}{n-1} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2.$$

In Bsp. 1:
Datenerhebung mit $n = 15$ ergibt $x_1, \dots, x_{15}$ mit

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 6,4 \quad \text{und} \quad s_n^2 = \frac{1}{n-1} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2 = 61,7.$$

Wir testen

$$H_0 : a \leq 0 \quad \text{vs.} \quad H_1 : a > 0$$

zum Niveau $\alpha = 0,05$.

Es gilt: $t_{n-1;\alpha} = t_{14;0,05} \approx 2.14$ (von Rechner bzw. Tabelle).

Wir berechnen

$$a_0 + \frac{s_n}{\sqrt{n}} \cdot t_{n-1;\alpha} = 0 + \frac{\sqrt{61,7}}{\sqrt{15}} \cdot 2,14 \approx 4,34.$$

Also gilt

$$\bar{x} = 6,4 > a_0 + \frac{s_n}{\sqrt{n}} \cdot t_{n-1;\alpha}$$

und der t-Test von Student ergibt

$$\varphi(x_1, \dots, x_n) = 1.$$

Folgerung: H_1 ist statistisch signifikant.

Aussage in der Realität:
Examenskandidaten schätzen ihre Leistung eher zu schlecht ein.

Einseitiger t-Test für zwei Stichproben

Bsp. 2: Sprechen Frauen mehr als Männer ?

Wir testen

H_0: Frauen reden nicht mehr als Männer

vs.

H_1: Frauen reden mehr als Männer.

Dazu: Von n Frauen und m Männer werden die Anzahlen $x_1, \ldots, x_n$ bzw. $y_1, \ldots, y_m$ der in einem festen Zeitraum gesprochenen Wörter ermittelt.

Annahme an die Erzeugung der Daten:
$x_1, \ldots, x_n, y_1, \ldots, y_m$ sind Realisierungen von unabhängigen ZVn $X_1, \ldots, X_n, Y_1, \ldots, Y_m$, wobei $X_1, \ldots, X_n$ $N(a_X, \sigma_X^2)$-verteilt und $Y_1, \ldots, Y_m$ $N(a_Y, \sigma_Y^2)$-verteilt sind, mit $a_X, a_Y \in \mathbb{R}$ und $\sigma_X, \sigma_Y > 0$ unbekannt.

Hypothesen im mathematischen Modell: $H_0 : a_X \leq a_Y$, $H_1 : a_X > a_Y$.

Naheliegend: Wir lehnen H_0 ab, falls

$$\frac{1}{n} \sum_{i=1}^{n} x_i - \frac{1}{m} \sum_{j=1}^{m} y_j > c.$$

Problem: Wahl von c ?

Vereinfachende Annahme: $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ (gleiche Varianzen).

Wir können σ_X^2 bzw. σ_Y^2 schätzen durch

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 \quad \text{mit} \quad \bar{X} = \frac{1}{n} \sum_{j=1}^{n} X_j$$

und

$$S_Y^2 = \frac{1}{m-1} \sum_{i=1}^{m} (Y_i - \bar{Y})^2 \quad \text{mit} \quad \bar{Y} = \frac{1}{m} \sum_{j=1}^{n} Y_j.$$

Wir schätzen $\sigma^2 = \sigma_X^2 = \sigma_Y^2$ durch die sog. **gepoolte Stichprobenvarianz**

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{j=1}^{m} (Y_j - \bar{Y})^2}{n+m-2} = \frac{(n-1) \cdot S_X^2 + (m-1) \cdot S_Y^2}{n+m-2}.$$

Man kann zeigen: Gilt $a_X = a_Y$, so ist

$$\frac{1}{\sqrt{\frac{1}{n}+\frac{1}{m}}} \cdot \frac{1}{\sqrt{S^2}} \cdot (\bar{X} - \bar{Y}) \quad t_{n+m-2} - \text{verteilt.}$$

Damit Ablehnung von H_0, falls

$$\frac{1}{\sqrt{\frac{1}{n}+\frac{1}{m}}} \cdot \frac{1}{\sqrt{S^2}} \cdot (\bar{X} - \bar{Y}) > t_{n+m-2;\alpha},$$

wobei $t_{n+m-2;\alpha}$ das α-Fraktil der t-Verteilung mit $n+m-2$-Freiheitsgraden ist.

Dies führt auf den *einseitigen t-Test für zwei Stichproben*:

Lehne

$$H_0 : a_X \leq a_Y$$

ab, falls

$$\frac{1}{\sqrt{\frac{1}{n}+\frac{1}{m}}} \cdot \frac{1}{\sqrt{s^2}} \cdot \left(\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{m}\sum_{j=1}^{m} y_j \right) > t_{n+m-2;\alpha},$$

wobei

$$\begin{aligned} s^2 &= \frac{(n-1)\cdot s_x^2 + (m-1)\cdot s_y^2}{n+m-2}, \\ s_x^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad \text{mit} \quad \bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j, \\ s_y^2 &= \frac{1}{m-1}\sum_{i=1}^{m}(y_i - \bar{y})^2 \quad \text{mit} \quad \bar{y} = \frac{1}{m}\sum_{j=1}^{n} y_j. \end{aligned}$$

In Bsp. 2 sind die beobachteten Daten: $n = 210$, $\bar{x} = 16215$, $s_x^2 = 7301^2$, $m = 186$, $\bar{y} = 15669$, $s_y^2 = 8663^2$.

Wir wählen $\alpha = 0,05$.

Dann gilt $t_{n+m-2;\alpha} = t_{394;0,05} \approx 1,649$.

Mit

$$\begin{aligned}
&\frac{1}{\sqrt{\frac{1}{n}+\frac{1}{m}}} \cdot \frac{1}{\sqrt{s^2}} \cdot \left(\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{m}\sum_{j=1}^{m} y_j\right) \\
&= \frac{1}{\sqrt{\frac{1}{210}+\frac{1}{186}}} \cdot \frac{1}{\sqrt{\frac{209\cdot 7301^2+185\cdot 8663^2}{210+186-2}}} \cdot (16215-15669) \\
&\approx 0,6864 < t_{n+m-2;\alpha}
\end{aligned}$$

folgt:

H_0 kann zum Niveau α nicht abgelehnt werden, d.h., H_1 ist statistisch nicht signifikant.

5.3 Der F-Test

Beim t-Test für zwei Stichproben müssen die Varianzen in beiden Stichproben gleich sein. Diese Annahme lässt sich mit dem F-Test überprüfen.

geg.: Realisierungen $x_1, \ldots, x_n$, $y_1, \ldots, y_m$ von unabhängigen ZVn $X_1, \ldots, X_n$, $Y_1, \ldots, Y_m$, wobei $X_1, \ldots, X_n$ $N(a_X, \sigma_X^2)$-verteilt und $Y_1, \ldots, Y_m$ $N(a_Y, \sigma_Y^2)$-verteilt sind, mit $\mu_X, \mu_Y \in \mathbb{R}$ und $\sigma_X, \sigma_Y > 0$ unbekannt.

Zu testen ist:

$$H_0 : \sigma_X^2 = \sigma_Y^2 \quad \text{vs.} \quad H_1 : \sigma_X^2 \neq \sigma_Y^2.$$

Wir schätzen σ_X^2 bzw. σ_Y^2 durch

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)^2$$

bzw.

$$S_Y^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(Y_i - \frac{1}{m}\sum_{j=1}^{m} Y_j\right)^2.$$

Dann sprechen Werte von

$$T(X_1, \ldots, Y_m) = \frac{S_X^2}{S_Y^2}$$

nahe bei Eins für H_0.

Man kann zeigen: Bei Gültigkeit von H_0 ist $T(X_1,\ldots,Y_m)$ $F_{n-1,m-1}$-verteilt.

Beim F-Test wird daher H_0 abgelehnt, falls $T(X_1,\ldots,Y_m)$ größer als das $\alpha/2$-Fraktil der $F_{n-1,m-1}$-Verteilung oder kleiner als das $(1-\alpha/2)$-Fraktil der $F_{n-1,m-1}$-Verteilung ist.

Anwendung in Bsp. 2:
Ist die Annahme gleicher Varianzen dort gerechtfertigt ?

Wir testen

$$H_0 : \sigma_X^2 = \sigma_Y^2 \quad \text{vs.} \quad H_1 : \sigma_X^2 \neq \sigma_Y^2$$

zum Niveau $\alpha = 0,05$ mit dem F-Test.

Gegebene Daten sind:

$n = 210$, $s_x^2 = 7301^2$, $m = 186$, $s_y^2 = 8663^2$.

Hier gilt

$$T(x_1,\ldots,y_m) = \frac{s_x^2}{s_y^2} = \frac{7301^2}{8663^2} \approx 0,7102,$$

sowie

$$F_{n-1,m-1;\alpha/2} = F_{209,185;0,025} \approx 1,32$$

und

$$F_{n-1,m-1;1-\alpha/2} = F_{209,185;0,975} \approx 0,75.$$

Wegen $T(x_1,\ldots,y_m) < F_{n-1,m-1;1-\alpha/2}$ kann H_0 hier abgelehnt werden, d.h., H_1 ist statistisch signifikant.

Folgerung: Die Anwendung des zweiseitigen t-Test auf diese Daten ist *nicht* zulässig!

5.4 Power und Effektstärke

Wir betrachten den einseitigen t-Test von Student:

Teste

$$H_0 : \mu \leq \mu_0 \quad \text{versus} \quad H_1 : \mu > \mu_0$$

zum Niveau α mittels

$$\varphi(x_1,\ldots,x_n) = \begin{cases} 1, & \text{falls } \frac{\sqrt{n}}{\sqrt{s^2}} \cdot \left(\frac{1}{n}\sum_{i=1}^n x_i - \mu_0\right) > t_{n-1;\alpha}, \\ 0 & \text{sonst,} \end{cases}$$

wobei

$$s^2 = \frac{1}{n-1}\sum_{i=1}^n \left(x_i - \frac{1}{n}\sum_{j=1}^n x_j\right)^2, \quad t_{n-1;\alpha} = \alpha\text{-Fraktil von } t_{n-1}.$$

Seien $X_1,\ldots,X_n$ unabhängig $N(\mu,\sigma^2)$-verteilt. Dann heißt

$$\beta_\varphi : \mathbb{R} \times \mathbb{R}_+ \setminus \{0\} \to [0,1]$$

mit

$$\begin{aligned} \beta_\varphi((\mu,\sigma^2)) &= \mathbf{P}_{(\mu,\sigma^2)}\left[\varphi(X_1,\ldots,X_n) = 1\right] \\ &= \mathbf{P}_{(\mu,\sigma^2)}\left[\frac{\sqrt{n}}{\sqrt{S^2}} \cdot \left(\frac{1}{n}\sum_{i=1}^n X_i - \mu_0\right) > t_{n-1;\alpha}\right] \end{aligned}$$

die **Gütefunktion** des Tests φ.

$\beta_\varphi((\mu,\sigma^2))$ ist die Wahrscheinlichkeit der Annahme von H_1 bei Vorliegen des Parameters (μ,σ^2).

Ist (μ,σ^2) so, dass $\mu \le \mu_0$ gilt (also H_0 stimmt), dann ist $\beta_\varphi((\mu,\sigma^2))$ ein α-Fehler, also gilt dann

$$\beta_\varphi((\mu,\sigma^2)) \le \alpha.$$

Ist dagegen (μ,σ^2) so, dass $\mu > \mu_0$ gilt (also H_1 stimmt), dann heißt $\beta_\varphi((\mu,\sigma^2))$ **Power** des Tests φ, und

$$1 - \beta_\varphi((\mu,\sigma^2))$$

ist ein β-Fehler.

Hierbei gilt:

1. $\beta_\varphi((\mu,\sigma^2))$ monoton wachsend in μ (für $\sigma > 0$ fest).
2. $\mu \mapsto \beta_\varphi((\mu,\sigma^2))$ stetig (für $\sigma > 0$ fest).
3. $\beta_\varphi((\mu_0,\sigma^2)) = \alpha$.

4. $\lim_{\mu\to\infty} \beta_\varphi((\mu, \sigma^2)) = 1$.

Probleme:

1.) Ist $\mu > \mu_0$ mit μ nahe bei μ_0 fest, so gilt für n groß:

$$\frac{\sqrt{n}}{\sqrt{S^2}} \cdot \left(\frac{1}{n} \sum_{i=1}^{n} X_i - \mu_0 \right) \approx \frac{\sqrt{n}}{\sigma} \cdot (\mu - \mu_0)$$

und die rechte Seite wird beliebig groß für n groß.

Dies impliziert:

$$\mathbf{P}_{(\mu,\sigma^2)} \left[\varphi(X_1, \ldots, X_n) = 1 \right] \to 1 \quad (n \to \infty).$$

Folglich wird für n groß H_1 statistisch signifikant, obwohl μ nur etwas größer als μ_0 ist (und eigentlich kein für die Praxis relevanter Effekt vorliegt).

2.) Da $\mu \mapsto \beta_\varphi((\mu, \sigma^2))$ stetig von μ abhängt, gilt für $\mu > \mu_0$ mit μ nahe bei μ_0:

$$\beta_\varphi((\mu, \sigma^2)) \approx \beta_\varphi((\mu_0, \sigma^2)) = \alpha,$$

also ist in diesem Fall (für n fest) der β-Fehler gleich

$$1 - \beta_\varphi((\mu, \sigma^2)) \approx 1 - \alpha$$

und damit groß!

Ausweg: Mache Annahme über die sog. **Effektstärke**

Dazu definiere

$$d = \frac{\mu - \mu_0}{\sigma} \quad \ldots \quad \text{Cohen's } d$$

d gibt an, wie stark μ von μ_0 abweicht in Vielfachen von σ.

Sprechweise:

$d = 0,2$ ... kleine Effektstärke
$d = 0,5$ ... mittlere Effektstärke
$d = 0,8$ ... große Effektstärke

Trick: Die Gütefunktion hängt nur von d ab, denn:

$$\begin{aligned}\beta_\varphi((\mu,\sigma^2)) &= \mathbf{P}_{(\mu,\sigma^2)}\left[\frac{\sqrt{n}}{\sqrt{S^2}}\cdot\left(\frac{1}{n}\sum_{i=1}^n X_i-\mu_0\right)>t_{n-1;\alpha}\right]\\ &= \mathbf{P}_{(\mu,\sigma^2)}\left[\frac{\sqrt{n}}{\sqrt{S^2/\sigma^2}}\cdot\left(\frac{1}{n}\sum_{i=1}^n \frac{X_i-\mu}{\sigma}+\frac{\mu-\mu_0}{\sigma}\right)>t_{n-1;\alpha}\right]\\ &= \mathbf{P}_{(0,1)}\left[\frac{\sqrt{n}}{\sqrt{S^2}}\cdot\left(\frac{1}{n}\sum_{i=1}^n X_i+d\right)>t_{n-1;\alpha}\right]\end{aligned}$$

da

$$S^2/\sigma^2=\frac{1}{n-1}\cdot\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}-\frac{1}{n}\sum_{j=1}^n\frac{X_j-\mu}{\sigma}\right)^2$$

und

$$\frac{X_i-\mu}{\sigma}\sim N(0,1)\quad\text{(ohne Beweis).}$$

Hierbei ist der Wert der Gütefunktion monoton wachsend in d.

Folgerung: Wenn wir wissen, wie groß d ist, berechnet sich die Power des Tests gemäß

$$\mathbf{P}\left[\frac{\sqrt{n}}{\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\frac{1}{n}\sum_{j=1}^n X_j\right)^2}\cdot\left(\frac{1}{n}\sum_{i=1}^n X_i-d\right)>t_{n-1;\alpha}\right]$$

mit $X_1, \ldots, X_n$ unabhängig $N(0,1)$, und diese Wk. kann man in Abhängigkeit von n am Rechner berechnen.

Die Idee ist nun, vor Beginn der Datenerhebung den Stichprobenumfang n so groß zu wählen, dass obige Wk. einen festen Wert übersteigt, also z.B. $\geq 0,8$ ist (für $\alpha = 0,05$).

Wenn wir jetzt unseren Test anwenden, und unsere Annahme an d stimmt, so folgt:

α-Fehler $\leq \alpha = 0,05$
β-Fehler $\leq 1-0,8 = 0,2$.

Also: Sowohl eine Entscheidung für H_1 als auch eine Entscheidung für H_0 ist für die Praxis bedeutsam, da wir beidesmal die auftretenden Fehlerwk. kontrolliert haben!

Bsp. mit R: Power in Abhängigkeit von Effektstärke und Stichprobenumfang approximativ berechnen.

Index