Stochastik

Markus Klemm.net

Wintersemester 2014/2015

Inhaltsverzeichnis

1 Zufällige Ereignisse, Wahrscheinlichkeit **2**
- 1.1 Gegenstand der Wahrscheinlichkeitstheorie . . . 2
- 1.2 Grundlegende Begriffe, Ereignisalgebra . . . 2
- 1.3 Die Wahrscheinlichkeit von Ereignissen . . . 5
  - 1.3.1 Klassische Definition . . . 5
  - 1.3.2 Axiomatische Definition der Wahrscheinlichkeit . . . 6
  - 1.3.3 Bedingte Wahrscheinlichkeit . . . 7
- 1.4 Spezielle wahrscheinlichkeitstheoretische Modelle . . . 7
  - 1.4.1 Multiplikationssatz . . . 7
  - 1.4.2 Unabhängigkeit von Ereignissen . . . 8
  - 1.4.3 Formel der totalen Wahrscheinlichkeit, BAYESsche Formel . . . 10

2 Zufällige Variable **11**
- 2.1 Grundbegriffe . . . 11
- 2.2 Diskrete Verteilungen . . . 12
  - 2.2.1 Verteilungstabelle, Erwartungswert, Streuung . . . 12
  - 2.2.2 Spezielle diskrete Verteilungen . . . 13
- 2.3 Stetige Verteilungen . . . 16
  - 2.3.1 Dichtefunktion, Erwartungswert, Streuung . . . 16
  - 2.3.2 Spezielle stetige Verteilungen . . . 16
- 2.4 Mehrdimensionale Verteilungen . . . 20
  - 2.4.1 Zufällige Vektoren . . . 20
  - 2.4.2 statistische Kennzahlen für Vektoren: Kovarianz und Korrelationskoeffizienten, sowie stochstische Unabhängigkeiten von ZG . . . 21

3 deskriptive Statistik: Grundbegriffe **23**
- 3.1 Merkmale . . . 23
- 3.2 Grundgesamtheit und Stichprobe . . . 24

4 Testtheorie **25**
4.1 Parametertests 25
4.1.1 Grundbegriffe, allgemeine Vorgehensweise 25
4.1.2 Test für Erwartungswert und Streuung bei normalverteilter GG X 27

1 Zufällige Ereignisse, Wahrscheinlichkeit

1.1 Gegenstand der Wahrscheinlichkeitstheorie

Gegenstand Untersuchung der Gesetzmäßigkeiten zufälliger Erscheinungen.

Zufällige Erscheinungen Vorgänge, die bestimmten unkontrollierten Einflüssen unterworfen sind und deren Ergebnis im Gegensatz zu deterministischen Erscheinungen im Einzelfall nicht exakt vorhergesagt werden kann.

Beispiel 1 Geg. Raum R mit Luft gefüllt, $V = 100\, m^3, p_1 = 1000\, hPa, T_1 = 250\, K(-23°C)$, Teilraum $R_0, V_0 = 1\, dm^3$

1. R hermetisch abgeschlossen, Temperatur erhöhen auf $T_2 = 300K \curvearrowright p_2 = p_1 \cdot \frac{T_2}{T_1} = 1200\, hPa$ (deterministischer Vorgang)
2. N ... Anzahl der Moleküle in R, $N \approx 3 \cdot 10^{27}$ (davon $O_2 : M = 0,6 \cdot 10^{27}$)
   n ... Anzhal der Moleküle in $R_0, n \approx 3 \cdot 10^{22}$
   - Aufenthaltsort eines bestimmten O_2-Moleküls: zufällige Erscheinung
   - Wahrscheinlichkeit, dass sich alle O_2-Moleküle in $R \backslash R_0$ befinden: $\approx 10^{-2,907 \cdot 10^{21}}$

   Ereignis möglich, aber so unwahrscheinlich, dass es praktisch nicht auftritt.

Diskussion

1. Begriffe: $H_n(A)$ ... absolute Häufigkeit von A bei n Wiederholungen
   $W_n(A) := \frac{H_n(A)}{n}$ ... relative Häufigkeit von A
2. Erfahrung: In langen Versuchsreihen schwankt die relative Häufigkeit um eine konstante Zahl.
3. Beobachtung der relativen Häufigkeit = Messverfahren zur Messung der Wahrscheinlichkeit (wie jedes Messverfahren, fehlerbehaftet, Messfehler kann beliebig verkleinert werden, wenn n hinreichen groß ist)

1.2 Grundlegende Begriffe, Ereignisalgebra

Definition 1 Ein zufälliger Versuch ist ein Vorgang, der sich (zumindest gedanklich) beliebig oft wiederholen lässt, und dessen Ergebnis im Rahmen verschiedener Möglichkeiten ungewiss ist.

Definition 2 Die Ergebnisse eines zufälligen Versuchs heißen zufällige Ereignisse. Speziell:
Ω ... sicheres Ereignis (tritt bei jeder Wiederholung auf)
Φ ... unmögliches Ereignis (tritt bei keiner Wiederholung auf)

(vorläufige) Erklärung: Grad der Gewissheit des Eintretens eines Ereignisses = Wahrscheinlichkeit (Wkt.)
Bezeichnung: P(Ereignis) = Zahl $\in [0;1], P(\Omega) = 1, P(\Phi) = 0$

Bemerkung : Jeder zufällige Versuch ist durch eine Menge Ω von Elementarerignissen ω charakterisiert. Jedem zufälligen zufälligen Ereignis A entspricht umkehrbar eindeutig eine Teilmenge A von Ω.
Oft: Idealisierte Darstellung als Menge in einer Ebene.

Ereignis	Menge
$A = \{$ ungerade Augenzahl $\}$	$A = \{1, 3, 5\}$
$\Omega = \{$ Augenzahl $< 7\}$	$\Omega = \{1, 2, 3, 4, 5, 6\}$
$\Phi = \{$ Augenzahl $= 7$ $\}$	Φ (leere Menge)

(idealer Würfel: $P(A) = \frac{3}{6}$)

Definition 3 Für zufällige Ereignisse A, B werden folgende Relationen und Operationen erklärt:

1. $A \subseteq B$, A Teilereignis von B (A zieht B nach sich (wenn A dann auch B))
2. $A = B :\Leftrightarrow A \subseteq B \wedge B \subseteq A$
3. $\bar{A}$ ... komplementäres Ereignis zu A, Negation
4. $A \cup B$ Vereinigung von A und B (A oder B)
5. $A \cap B$ Durchschnitt von A und B (A und B)
6. $A \backslash B := A \cap \bar{B}$ Differenz „A minus B“

Diskussion (Rechenregeln)

1. $A \cup \Omega = \Omega, A \cup \Phi = A, A \cap \Omega = A, A \cap \Phi = \Phi, \Phi \subseteq A \subseteq \Omega$
2. $A \subseteq B \Leftrightarrow \bar{B} \subseteq \bar{A} \Leftrightarrow A \cap B = A \Leftrightarrow A \cup B = B$
3. a) $A \cup B = B \cup A, A \cap B = B \cap A$ (Kommutativgesetz)
   b) $(A \cup B) \cup C = A \cup (B \cup C), (A \cap B) \cap C = A \cap (B \cap C)$ (Assoziativgesetz)

$$\text{Allg. } \bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n$$

tritt genau dann ein, wenn wenigstens eines der Ereignisse A_i eintritt.

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n$$

c) Formel von de MORGAN
$\overline{A \cup B} = \bar{A} \cap \bar{B} \quad \overline{A \cap B} = \bar{A} \cup \bar{B}$

Beispiel 4

1. In einem Betrieb gibt es 3 Produktionslinien (PL) gleichen Typs. Wir beobachten eine Arbeitsperiode von 16h und regestrieren, ob Störungen auftreten oder nicht.
   - $A_i :=$ { Störungen, in i-ter PL }$(i = 1, 2, 3)$ (sogenannte "‘einfache“ Ereignisse, auf einzelne PL bezogen)
   - Elementarereignisse: geordnete Zahlentupel (k_1, k_2, k_3) mit
     $k_i = \begin{cases} 1 & \ldots \quad \text{Störung(en)} \\ 0 & \ldots \quad \text{keine Störung} \end{cases}$ in i-ter PL, z.B. $A_1 = \{(1,0,0),(1,01),(1,1,0),(1,1,1)\}$
     (Falls genauere Beobachtung, etwas Anzahl/Zeitpunkte der Störungen, komplizierteres Modell)
2. Die folgenden Ereignisse $A, B, \ldots, F$ sind durch die einfachen Ereignisse A_i auszudrücken:
   $A =$ { in allen PL treten Störungen auf } $= A_1 \cap A_2 \cap A_3$
   $B =$ { in genau 2 PL treten Störungen auf} $= (A_1 \cap A_2 \cap \bar{A}_3) \cup (A_1 \cap \bar{A}_2 \cap A_3) \cup (\bar{A}_1 \cap A_2 \cap A_3)$
   $C =$ { in genau einer PL Störungen } $= (A_1 \cap \bar{A}_2 \cap \bar{A}_3) \cup (\bar{A}_1 \cap A_2 \cap \bar{A}_3) \cup (\bar{A}_1 \cap \bar{A}_2 \cap A_3)$
3. $D =$ { in keiner PL treten Störungen auf } $= \bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3$
4. $E =$ { in wenigstens einer PL treten Störungen auf } $= A_1 \cup A_2 \cup A_3$
5. $F =$ { Störungen in wenigstens 2 PL } $= B \cup A =$

Definition 4 Zwei Ereignisse A und B heißen unvereinbar, wenn $A \cap B = \Phi$ gilt
Bemerkung: Nur in diesem Fall ist das „Oder“ ($\bigcup$) gleichzeitig ein „Entweder-Oder“.

Definition 5 Ein System $\mathfrak{A}$ von Ereignisalgebra, Durchschnitt und Negation (und damit auch Differenz) abgeschlossen. (Insbesondere gehören Ω stets zu einer Algebra)

Diskussion Die Ergebnis eines zufälligen Versuchs bilden eine Ereignisalgebra.

1.3 Die Wahrscheinlichkeit von Ereignissen

1.3.1 Klassische Definition

Definition 6 Bei einem zufälligen Versuch gebe es genau N gleichmögliche Elementarereignisse $\omega_1, \ldots, \omega_N$, d.h. $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ ... Dann heißt für jedes zufällige Ereignis A die Zahl

$$P(A) := \frac{M}{N} := \frac{\text{Anzahl der für A „ günstigen Elementarereignisse}}{\text{Anzahl aller möglichen Elementarereignisse}}$$

die Wahrscheinlichkeit des zufälligen Ereignisses A, (Elementarereignis ω_i günstig für A bedeutet $\omega_i \in A$, wobei A die dem Ereignis A entsprechende Teilmenge von Ω ist)

Diskussion

1. Die Ereignisse, die dem einelementigen Teilmengen $A_i = \{\omega_i\}, i = 1, 2, \ldots, N$, entsprechen, sind atomar
2. Entscheidende Vorraussetzung für die Verwendbarkeit der klassischen Definition ist die Gleichmöglichkeit der Elementarereignisse (= Gleichwahrscheinlichkeit der atomaren Ereignissen $A_I = \{\omega_i\}$)
3. Die Ermittlung von M und N aus Def. 6 erfolgt häufig mit Hilfe der Kombinatorik

Grundaufgaben der Kombinatorik

1. Permutationen
   - P_n : Anzahl der möglichen Anordnungen von n verschiedenen Elementen $P_n = n!$
   - $P_{(n_1, n_2, \ldots, n_k)}$ ... Anzahl der möglichen Anordnungen von n Elementen, von denen jeweils $n_1, n_2, \ldots, n_k$ gleich sind $(n_1 + n_2 + \cdots + n_k = n)$
     $P_{(n_1, \ldots, n_k)} = \frac{n!}{n_1! \cdot n_2! \cdot \cdots \cdot n_k!}$
2. Kombinationen (n Elemente in Klassen zu k Elementen anordnen, ohne Berücksichtigung der Reihenfolge)
   a) ohne Wiederholung $(k \leq n) : C(n, k) = \binom{n}{k}$
   b) mit Wiederholung $C^*(n, k) = \binom{n+k-1}{k}$
3. Variationen (n Elemente in Klassen zu k Elementen anordnen, mit Berücksichtigung der Reihenfolge)
   a) ohne Wiederholung $(k \leq n) : V(n, k) = n(n-1) \cdot \cdots \cdot (n-k+1) = \frac{n!}{(n-k)!}$
   b) mit Wiederholung $V^*(n, k) = n^k$

1.3.2 Axiomatische Definition der Wahrscheinlichkeit

Vorbetrachtung

- Mangel der klassischen Definition: Nicht immer liegen gleichmögliche Elementarereignisse vor
- !Abschnitt 1.1.: $W_n(A) = \frac{H_n(A)}{n}$ ... relative Häufigkeit von A bei n Versuchswiederholungen. Im langen Versuchsreihen schwankt $W_n(A)$ um eine konstante Zahl: „$\lim\limits_{n\to\infty} W_n(A)$“ $=: P(A)$
  Zur Definition der Wiederholung ungeeignet (von Versuchsreihe abhängig)
- Aber Eigenschaften der relativen Häufigkeit
  1. $0 \leq W_n(A) \leq 1$
  2. $W_n(\Omega) = 1$
  3. $W_n(A \cup B) = W_n(A) + W_n(B)($ falls $A \cap B = \varnothing)$

Definition 7 (Axiomatische Definition der Wiederholung, Kolmogorov 1933)
Gegeben sei eine Ereignisalgebra $\mathfrak{A}$. Auf $\mathfrak{A}$ sei eine Funktion P erklärt, für die folgendes gilt:

- A1: Für jedes Ereignis $A \in \mathfrak{A}$ ist $P(A)$ erklärt und es gilt $0 \leq P(A) \leq 1$
- A2: $P(\Omega) = 1$
- A3: Für paarweise unvereinbare Ereignisse $A_i \in \mathfrak{A}$ (d.h. $A_i \cap A_j = \varnothing$ für $i \neq j$) gilt: $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$

Dann heißt $P(A)$ die Wiederholung des zufälligen Ereignisses A

Diskussion

1. Die Definition 6 (klassische Definiton) liefert ein (!) Modell eines Paares $(\mathfrak{A}, P)$, welches den Axiomen 1-3 genügt.
2. Allgemeines Vorgehen (vereinfachte Darstellung) [Theoretische Untersuchungen (Kombinatorik, physikalische Gesetze); Beobachtung der relativen Häufigkeit] $\Rightarrow$ [Für gewissen Grundereignisse sind die Wahrscheinlichkeiten exakt oder näherungsweise bekannt] $\Rightarrow$ (Rechenregeln aus A1 bis A3 ableitbar) [Wahrscheinlichkeiten für alle interessierenden Ereignisse berechenbar]

Satz 1 (Eigenschaften der Wahrscheinlichkeit)
Es seien $A, B, C, \dots$ zufällige Ereignisse. Dann gilt:

1. $P(\varnothing) = 0$
2. - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$
- allg.: $P(A_1 \cup \cdots \cup A_n) = \sum\limits_i P(A_i) - \sum\limits_{i<j} P(A_i \cap A_j) + \sum\limits_{i<j<k} P(A_i \cap A_j \cap A_k) - + \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n)$

3. $P(\bar{A}) = 1 - P(A)$
4. $A \subseteq B \Rightarrow P(A) \leq P(B)$

1.3.3 Bedingte Wahrscheinlichkeit

Definition 8 A und B seien zufällige Ereignisse, $P(B) > 0$. Dann heißt

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

die bedingte Wahrscheinlichkeit von A unter der Bedingung B.
Anmerkung des Authors: Vorsicht:

- $P(A/B)$ ... bedingte Wahrscheinlichkeit
- $A \backslash B$ ... Differenz

Diskusssion

1. Die Funktion $P(./B)$ besitzt die gleichen Eigenschaften wie die Funktion $P(.)$, z.B. $P(\bar{A}/B) = 1 - P(A/B)$ usw., vgl. Satz 1
2. Anschaulich:
   - $P(A/B)$ ... Anteil von A innerhalb der Bezugsmenge B
   - $P(A) = P(A/\Omega)$ ... Anteil von A innerhalb der Bezugsmenge Ω
3. Berechnung oft (falls gleichmögliche Elementarereignisse vorliegen) klassisch möglich.

1.4 Spezielle wahrscheinlichkeitstheoretische Modelle

1.4.1 Multiplikationssatz

Satz 2 A und B seien zufällige Ereignisse, $P(A) > 0, P(B) > 0$: Dann gilt

$$P(A \cap B) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

Allgemein: $P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \cdot \cdots \cdot P(A_n/A_1 \cap A_2 \cap \cdots \cap A_{n-1})$

Beweis: Definition 8 $\curvearrowright$ $P(A \cap B) = P(B) \cdot P(A/B)$ usw.

Beispiel 8 In einer Lostrommel befinden sich 20 Lose, davon 5 Gewinnlose. Jemand zieht 3 Lose nacheinander. Gesucht:Die Wahrscheinlichkeit 3 Gewinnlose zu ziehen.
$A_k := \{$ Gewinn bei k-ten Zug $\}, k = 1, 2, 3$
Lösung: Multiplikationssatz

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \tag{1}$$

1. Gezogenes Los wird nicht in die Trommel zurückgelegt
   (1) $\curvearrowright P(A_1 \cap A_2 \cap A_3)\frac{5}{20} \cdot \frac{4}{19} \cdot \frac{3}{18} = 0,00877\ldots$

2. gezogenes Los wird wieder in die Trommel zurückgelegt
   (1) $\curvearrowright P(A_1 \cap A_2 \cap A_3) = \frac{5}{20}^3 = 0,01562$

Diskussion

1. Anwendung des Multiplikationssatzes oft bei zufälligen Versuchen die aus aufeinanderfolgenden Teilversuchen bestehen.

2. vgl Beispiel 8
   a) ohne Zurücklegen: Ergebnis des 2. Zuges von Ergebnis des 1. Zuges abhängig
   b) mit Zurücklegen: Ergebnis des 2. Zuges wird vom Ergebnis des 1. Zuges nicht beeinflusst: $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2/A_1) = P(A_1) \cdot P(A_2)$.
   Begriff: Unabhängigkeit

1.4.2 Unabhängigkeit von Ereignissen

Definition 9 Zwei Ereignisse A und B heißen (stochastisch) unabhängig, wenn

$$P(A \cap B) = P(A) \cdot P(B)$$

gilt.

Diskussion

1. Es sei $P(B) > 0$, dann gilt: A und B sind genau dann abhängig, wenn gilt:

$$P(A/\underbrace{B}_{\text{Bedingte Whk. hängt nicht v. Bed. ab}}) = P(A)$$

2. Die Ereignisse $A_1, \ldots, A_n$ heißen (in ihrer Gesamtheit) unabhängig, wenn $P(A_{k_1} \cap A_{k_2} \cap \cdots \cap A_{k_m}) = P(A_{k_1}) \cdot P(A_{k_2}) \cdot \cdots \cdot P(A_{k_m})$ für eine beliebige Auswahl von $m(2 \leq m \leq n)$ der n Ereignisse gilt.

3. A und B seien unabhängig, dann sind auch A und $\bar{B}$, $\bar{A}$ und B sowie $\bar{A}$ und $\bar{B}$ unabhängig, analog für mehr als 2 Ereignisse von A und B sind unabhängig, d.h. $P(A \cap B) = P(A) \cdot P(B)$

4. Vorsicht: Man unterscheide A und B sind unvereinbar, d.h. $A \cap B = \varnothing$

5. Veranschaulichung des Begriffes Unabhängigkeit
   Produkt mit 2 möglichen Fehlern, z.B. Videokassetten, Fehler 1: Schlechte Bildqualität, Fehler 2: schlechte Tonqualität.
   $A := \{$ Produkt besitzt Fehler 1 $\}, B := \{$ Produkt besitzt Fehler 2 $\}$

   a) Hersteller 1
      $P(A) = 20\%\ P(B) = 10\%, P(A \cap B) = 5\%, P(A/B) = 50\%$

      - 20% aller Erzeugnisse besitzen Fehler 1
      - Unter den Produkten mit Fehler 2 besitzt die Hälfte (50%) auch den Fehler 1, d.h. unter diesen Produkten tritt Fehler 1 häufiger auf $\rightarrow$ Stochastische Abhängigkeit zwischen 1 und 2
        $P(A \cap B) = 0,05 \neq P(A)P(B) = 0,02$

   b) Hersteller 2
      $P(A) = 20\%, P(B) = 10\%, P(A \cap B) = 2\%, P(A/B) = 20\%$

      Anteil von A unter allen Produkten = 20%, aber auch Anteil innerhalb von B ist 20% $\curvearrowright$ Unabhängigkeit der beiden Fehler
      $P(A \cap B) = 0,02 = P(A) \cdot P(B)$

Satz 3 $A_1, A_2, \ldots, A_n$ seien (in ihrer Gesamtheit) unabhängig. Dann gilt für $A = A_1 \cup A_2 \cup \cdots \cup A_n$

$$P(A) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot \cdots \cdot P(\bar{A}_n)$$

Beweis:

1. $\bar{A} = \bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n$ (de Morgan)
2. $P(\bar{A}) = P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot \cdots \cdot P(\bar{A}_n)$ (Unabhängigkeit)
3. $P(A) = 1 - P(\bar{A})$

Beispiel 9 Drei Jäger schießen gleichzeitig, unabhängig voneinander auf einen Fuchs.

- Jäger 1 trifft mit Wahrscheinlichkeit $0,8$
- J $2 : 0,75$
- J 3: $0,2$

Wie groß ist die Wahrscheinlichkeit, dass der Fuchs getroffen wird?
$A := \{$ Fuchs wird getroffen $\}$
$A_i = \{$ Jäger i trifft den Fuchs $\}(i = 1, 2, 3)$
$A = A_1 \cup A_2 \cup A_3 \quad (A_i$unabhängig$)$

$$\underbrace{\curvearrowright}_{\text{Satz 3}} P(A) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot P(\bar{A}_3) = 1 - 0,2 \cdot 0,25 \cdot 0,8 = \underline{\underline{0,96}}$$

1.4.3 Formel der totalen Wahrscheinlichkeit, Bayessche Formel

Satz 4 (Formel der totalen Wahrscheinlichkeit)
Es sei $A_1, \dots, A_n$ ein vollständiges System paarweise unvereinbarer Ereignisse (d.h. $A_i \cap A_j = \varnothing$ falls $i \neq j \wedge A_1 \cup A_2 \cup \dots \cup A_n = \Omega$)
Dann gilt für ein beliebiges Ereignis B:

$$P(B) = \sum_{i=1}^{n} P(A_i) \cdot P(B/A_i)$$

$= P(A_1) \cdot P(B/A_1) + \dots + P(A_n) \cdot P(B/A_n)$
Beweis: $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_n \cap B)$
$P(B) = P(A_1 \cap B) + \dots + P(A_n \cap B) = P(A_1) \cdot P(B(A_1) + \dots + P(A_n) \cdot P(B/A_n)$

Beispiel 11 Die 3 Jäger aus Beispiel 9 gehen erneut auf die Jagdt. Die Trefferwahrscheinlichkeiten sind

- J1: $0,85$
- J2: $0,75$
- J3: $0,2$

Diesmal schießt nur 1 Jäger, der durch das Los ermittelt wird

1. Wie groß ist die Wahrscheinlichkeit, dass der Fuchs getroffen wird?
2. Der Fuchs wurde getroffen. Mit welcher Wahrscheinlichkeit war Jäger 3 der Schütze?

$B = \{$ Fuchs wurde getroffen $\}$
$A_i = \{$ Jäger i wird ausgelost $\} (i = 1, 2, 3)$

1. $P(B) = P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + P(A_3) \cdot P(B/A_3)$
   $= \frac{1}{3} \cdot 0,85 + \frac{1}{3} \cdot 0,75 + \frac{1}{3} \cdot 0,2 = 0,6$
2. $P(A_3/B) = \frac{P(A_3 \cap B)}{P(B)} = \frac{P(A_3) \cdot P(B/A_3)}{P(B)} = \frac{\frac{1}{3} \cdot 0,2}{0,6} = 0,111\dots$

Satz 5 Es gilt unter den Vorraussetzungen des Satzes 4:

$$P(A_j/B) = \frac{P(A_j) \cdot P(B/A_j)}{P(B)} \quad (j = 1, \dots, n)$$

(Bayessche Formel für die Rückschlusswerte $P(A_j/B)$

Diskussion Anwendung der Sätze 4 und 5 oft bei zufälligen Versuchen, die aus 2 aufeinanderfolgenden Teilversuchen bestehen. Im Beispiel 11: 1. Teilversuch: Auslosen, 2. Teilversuch: Schießen

2 Zufällige Variable

2.1 Grundbegriffe

- Zufälliger Versuch $\rightarrow$ zufällige Ereignisse $\rightarrow$ Wahrscheinlichkeit
- Ω ... Menge aller Elementarereignisse

Definition 1 Ist jedem Elementarereignis ω eine reelle Zahl $X(\omega)$ zugeordnet, so heißt die dadurch erklärte Funktion X (reelle) Zufallsgröße.

Bemerkungen

1. Der funktionelle Zusammenhang $\omega \rightarrow X(\omega)$ ist im allgemeinem uninteressant.
2. Von Interesse ist dagegen die Wahrscheinlichkeit, dass die Zufallsgröße (ZG) einen bestimmten Wert annimmt bzw. in ein vorgegebenes Intervall fällt.
3. Dazu ist die sogenannte Verteilungsfunktion (VF) nützlich.

Definition 2 Die Funktion $F_X(x) = P(X \leq x),\ x \in \mathbb{R}$ heißt Verteilungsfunktion der ZG X.

Diskussion

1. $F_X(x)$ ist die Wahrscheinlichkeit, dass die ZG X einen Wert annimmt, der $\leq x$ ist.
2. Eigenschaften Eine Funktion $F(x), x \in \mathbb{R}$, ist genau dann VF einer ZG X, wenn folgendes gilt:
   a) $0 \leq F(x) \leq 1$
   b) $\forall x_1, x_2 \in \mathbb{R} : x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$ (Monotonie)
   c) $\lim\limits_{x \to -\infty} F(x) = 0,\ \lim\limits_{x \to \infty} F(x) = 1$
   d) $\lim\limits_{x \to x_0 + 0} F(x) = F(x_0)$ (rechtsseitige Stetigkeit)
3. Beispiele (diskrete und stetige Verteilungen)
4. Ist $F_X(x)$ bekannt, so lassen sich alle interessanten Wahrscheinlichkeiten berechnen z.B. gilt
   a) $P(X = a) = F_X(a) - \lim\limits_{x \to a-0} F(x)$
   b) $P(X > a) = 1 - F_X(a)$ usw.

2.2 Diskrete Verteilungen

2.2.1 Verteilungstabelle, Erwartungswert, Streuung

Definition 3 Eine ZG X heißt diskret verteilt, wenn sie nur endlich viele oder abzählbar unendlich viele Werte $x_0, x_1, x_2, \dots$ mit den Wahrscheinlichkeiten $p_0, p_1, p_2, \dots$ annimmt. Sprechweise auch X ist diskrete ZG.

Diskussion

1. Verteilungstabelle

Werte	x_0	x_1	x_2	$\dots$
Wahrscheinlichkeiten	p_0	p_1	p_2	$\dots$

Dabei

$$p_i = P(X = x_1), \ \sum_i p_i = 1 \quad (p_1 > 0)$$

2. Graphische Darstellung (Stabdiagramm)
3. $P(a \leq X \leq b) = \sum\limits_{i:a \leq x_i \leq b} p_i$

Definition 4 Es sei X eine diskrete ZG mit der Verteilungstabelle

x_0	x_1	x_2	$\dots$
p_0	p_1	p_2	$\dots$

Dann werden definiert:

1. $$\underbrace{EX}_{\mu_x} := \sum_i x_i p_i \dots \text{ Erwartungswert (Mittelwert von } X\text{)}$$
2. $$\underbrace{D^2X}_{\text{var}(X)} := \sum_i (x_i - EX)^2 p_i \dots \text{Streuung (Varianz, Dispersion) von } X$$
3. $$\sigma_x := \sqrt{D^2X} \dots \text{Standardabweichung von } X$$

Diskussion

1. D^2X ist die mittlere quadratische Abweichung einer ZG von ihrem Erwartungswert. Es gilt
$$D^2X = E(X - EX)^2 = E(X^2) - (EX)^2$$
(Formel gilt auch im stetigem Fall!)
Die Streuung ist eine „rechnerische Größe", keine anschauliche Bedeutung.
2. Für eine beliebige Funktion $g(x)$ gilt:
$$Eg(X) = \sum_i g(x_i) p_i, \text{ z.B. } E(X^2) = \sum_i x_i^2 p_i$$

2.2.2 Spezielle diskrete Verteilungen

1. Hypergeometrische Verteilung

Definition 5 Die ZG X heißt hypergeometrisch verteilt mit den ganzzahligen Parametern N, M und $n (0 < M \leq N, 0 < n \leq N)$, wenn sie die Werte $x_m = m$ mit den Wahrscheinlichkeiten

$$p_m := P(X = m) = \frac{\binom{M}{m} \cdot \binom{N-M}{n-m}}{\binom{N}{n}}, m = 0, 1, \ldots, n$$

auch mit Kurzschreibweise $X \in H(N, M, n)$

Diskussion

a) Anwendung Stichprobe ohne Zurücklegung (z.B. Qualitätskontrolle, Lotto)

Allgemein: N Objekte, davon M mit bestimmten Merkmal (z.B. Ausschuss, Gewinnzahl), n Objekte entnehmen.
X ... Anzahl der Objekte unter den n entnommenen, die das Merkmal besitzen $\curvearrowright X \in H(N, M, n)$, vgl. auch Ü.A. 2.7.

2. Mit $p := \frac{M}{N}$ (Anteilswert) gilt

$$EX = np,\ D^2X = \frac{N-n}{N-1} np(1-p)$$

Beispiel 4 In einer Lostrommel befinden sich 20 Lose, davon 5 Gewinnlose (vgl. Beispiel 8, Kap. 1), Jemand zieht 3 Lose (ohne Zurücklegung). Wie groß ist die Wahrscheinlichkeit, dass sich darunter genau 2 Gewinnnlose befinden?

Lösung $X \ldots$ Anzahl der Gewinnlose unter den 3 gezogenen, $X \in H(N, M, n)$ mit $N = 20, M = 5, n = 3$

$P(X = 2) = \cdots = 0,1316$

3. Binomialverteilung

Definition 6 Die ZG X heißt binomialverteilt mit den Parametern n und $p (n \in \mathbb{N}^*, 0 < p < 1)$, wenn sie die Werte $x_m = m$ mit den Wahrscheinlichkeiten

$$p_m = P(X = m) = \binom{n}{m} p^m (1-p)^{n-m}, m = 0, 1, 2, \ldots, n$$

annimmt. Kurz $X \in B(n, p)$

Diskussion

1. Es git
$$EX = np, D^2X = np(1-p)$$

2. Anwendung
   - Stichprobe mit Zurücklegung:dabei $p = \frac{M}{N}$
   - Angenäherte Berechnung der hypergeometrischen Verteilung $H(N, M, n)$ für große N durch die leichter handhabbare Binomialverteilung:
   $$H(N, M, n) \approx B(n, p) \text{ mit } p = \frac{M}{N} \text{ falls } \frac{n}{N} \leq 0,05$$
   (bei großem N ist es praktisch bedeutungslos, ob mit oder ohne Zurücklegung gearbeitet wird)
   - Bernoulli-Schema: Es werden n unabhängige Wiederholungen eines Versuches durchgeführt. Bei jeder Wiederholung wird festgestellt, ob ein bestimmtes Ereignis A eintritt oder nicht. Es sei $p := P(A)$ bei jeder einzelnen Wiederholung.
   X ... Anstelle der Versuche, bei denen A eintritt $\curvearrowright X \in B(n, p)$
   Denn: $A_i = \{$ bei i-ter Wiederholung tritt A ein $\} \curvearrowright \{X = m\} =$
   $\underbrace{(A_1 \cap \cdots \cap A_m \cap \bar{A}_{m+1} \cap \cdots \cap \bar{A}_n)}_{\text{unabhängig}} \cup \cdots \cup \underbrace{(\bar{A}_1 \cap \cdots \cap \bar{A}_{n-m} \cap A_{n-m+1} \cap \cdots \cap A_n)}_{\text{paarweise unvereinbar}}$
   $\curvearrowright P(X = m) = P(\ldots) + \cdots + P(\ldots)$
   $= p^m(1-p)^{n-m} \cdots + \cdots + p^m(1-p)^{n-m}$ (Anzahl der Summanden $= \binom{n}{m}$)
   $= \binom{n}{m} p^m (1-p)^{n-m}$

Beispiel 5 Ein Massenprodukt mit einem Ausschussanteil von 3% wird in Packungen zu 20 Stück verkauft. Wie groß ist die Wahrscheinlichkeit, dass eine Packung höchstens 2 Ausschusstücke enthält?

Lösung: $X \ldots$ Anzahl der Ausschusstücke in einer Packung von 20 Stück. $X \in B(n, p), n = 20, p = 0,03\%$

$P(X \leq 2) = p_0 + p_1 + p_2 = 0,979 = 98\%$

3. Poisson-Verteilung

Definition 7 Die ZG X heißt Poisson-verteilt mit dem Parameter $\lambda > 0$, wenn sie die Werte $x_m = m$ mit den Wahrscheinlichkeiten
$$p_m = P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda}, m = 0, 1, 2, \ldots$$
annimmt. Kurz $X \in P(\lambda)$

Diskussion

a) Es gilt: $EX = \lambda,\ D^2X = \lambda$

b) • Es gilt

$$B(n,p) \approx P(\lambda) \text{ mit } \lambda = np, \text{ falls etwa } n \geq 60 \wedge p \leq 0,1$$

• Anwendung

– Bedienungstheorie, Zuverlässigkeitstheorie

– Anzahl der eintreffenden Kunden, Forderungen pro Zeiteinheit

– Anzahl der Störungen im Produktionsprozess pro Zeiteinheit

Beispiel 6 In einer Produktionsanlage trifft im Durchschnitt alle 5 Stunden eine Störung auf (d.h. im Mittel 0,2 Störungen pro Stunde). Die Zahl der Störungen in einem bestimmten Zeitraum kann als Poisson-verteilt angesehen werden. Wie groß ist die Wahrscheinlichkeit, dass in einer 8-stündigen Schicht mehr als 2 Störungen auftreten?

Lösung: $X \ldots$ Anzahl der Störungen in 8 h
$X \in P(\lambda)$ mit $\lambda = EX = 8 \cdot 0,2 = 1,6$

$P(X > 2) = 1 - P(X \leq 2) = 1 - (p_0 + p_1 + p_2)$
$= 1 - (\frac{1,6^0}{0!} + \frac{1,6^1}{1!} + \frac{1,6^2}{2!})e^{-1,6} = 0,2166 \approx 22\%$

4. Weitere diskrete Verteilungen

• diskrete gleichmäßige Verteilung

$$P(X = x_m) = \frac{1}{n}, m = 1, 2, \ldots, n$$

, z.B. Augenzahl beim Werfen mit einem idealen Würfel $n = 6$

• negative Binomialverteilung (Parameter $r \in \mathbb{N}^*, p \in (0;1)$)

$$P(X = m) = \binom{m+r-1}{m}(1-p)^m \cdot p^r, m = 0, 1, 2, \ldots$$

$= \binom{-r}{m} \cdot (p-1)^m p^r$
$X \ldots$ Anzahl der Misserfolge ($\bar{A}$) vor dem r-ten Erfolg (A) beim Bernoulli-Schema bei unbeschränkter Zahl der Wiederholungen, dabei $P(A) = p$

$$EX = \frac{(1-p)r}{p}$$

• speziell $r = 1 \curvearrowright$ geometrische Verteilung
$X \ldots$ Anzahl der Misserfolge vor dem ersten Erfolg beim Bernoulli-Schema

$$p_m = P(X = m) = (1-p)^m \cdot p, m = 0, 1, 2, \ldots$$

Diskussion $\sum\limits_{m=0}^{\infty} p_m = \underbrace{\sum_{m=0}^{\infty}(1-p)^m}_{\text{geom. Reihe mit q=1-p, An.Gl}a=p, s=\frac{a}{1-p}=1} p = \frac{p}{1-(1-p)} = 1$

2.3 Stetige Verteilungen

2.3.1 Dichtefunktion, Erwartungswert, Streuung

Definition 8 Eine ZG X heißt stetig verteilt, wenn sie alle Werte aus einem Intervall annehmen kann und eine sogenannte Dichtefunktion $f_X(x) \geq 0$ mit

$$F_X(x) = \int\limits_{-\infty}^{x} f_X(t)\,dt$$

existiert. (X ist stetige ZG)

Diskussion

1. Dichte $f_X(x) = F'_X(x) \quad f_X(x) \geq 0$
   $\int\limits_{-\infty}^{\infty} f_X(x)\,dx = 1$

2. Es gilt $P(a \leq X \leq b) = F_X(b) - F_X(a) = \int\limits_{a}^{b} f_X(x)\,dx$ insbesondere gilt $P(X = a) = 0 \quad \forall a \in \mathbb{R}$

Definition 9 Es sei X eine stetige ZG mit der Dichte $f(x)$. Dann werden analog Def. 4 erklärt: $EX = \int\limits_{-\infty}^{\infty} xf(x)\,dx,\ D^2X = \int\limits_{-\infty}^{\infty} (x-EX)^2 f(x)\,dx$

Bemerkung: Für eine beliebige Funktion $g(x)$ gilt:
$Eg(x) = \int\limits_{-\infty}^{\infty} g(x)f(x)\,dx$

2.3.2 Spezielle stetige Verteilungen

1. Normalverteilung (Gauss-Verteilung)

   Definition 10 Die ZG X heißt normalverteilt mit den Parametern μ und $\sigma^2 (\mu \in \mathbb{R}, \sigma > 0)$, wenn sie die Dichte

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

   besitzt.
   Kurzschreibweise: $X \in N(\mu, \sigma^2)$

Diskussion

a) Es gilt
$$EX = \mu, D^2X = \sigma^2$$

b) Verteilungsfunktion ist nicht in geschlossener Form darstellbar (Integraldarstellung bzw. unendliche Reihe)

c) Es gilt:
$$X \in N(\mu, \sigma^2) \Rightarrow Z := \frac{X - \mu}{\sigma} \in N(0;1) \tag{2}$$

- Es sei $\Phi(x)$ die VF der sogenannten standardisierten Normalverteilung (NV) $N(0,1)$.
- $\Phi(x)$ ist tabelliert.
- Jede beliebige NV lässt sich wegen (2) mit Hilfe der Φ-Funktion ausdrücken
- Eigenschaften der Φ-Funktion:

$$\Phi(x) = \frac{1}{2}, \quad \Phi(-x) = 1 - \Phi(x)$$

d) Es sei $X \in N(\mu, \sigma^2)$. Dann gilt:

i. $F_X(x) = \Phi(\frac{x-\mu}{\sigma})$

ii. $P(a \leq X \leq b) = \Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})$
speziell: $P(X \geq a) = 1 - \Phi(\frac{a-\mu}{\sigma})$
$P(X \leq b) = \Phi(\frac{b-\mu}{\sigma})$

iii. $P(|X - \mu| \leq a) = 2\Phi(\frac{a}{\sigma}) - 1 \quad (a > 0)$
speziell

$a = \sigma :$	$P(\|X - \mu\| \leq \sigma) = 2\Phi(1) - 1$	$= 0,6827$
$a = 2\sigma :$	$P(\|X - \mu\| \leq 2\sigma)$	$= 0,9545$
$a = 3\sigma :$	$P(\|X - \mu\| \leq 3\sigma)$	$= 0,9973$

e) Anwendung

- Messfehler
- geometrische oder physikalische Kenngrößen von Produkten (Länge, Masse, Widerstand, ...)
- biologische Merkmale
- allgemein: Summe einer großen Anzahl kleinerer unabhängiger Größen $\curvearrowright$ NV

f) Vorausschau auf statistische Methoden

ZufallsgrößeX	$\underbrace{\rightarrow}_{n\text{-mal beobachten}}$	Stichprobe $x_1, \dots, x_n$
theor. Erwartungsw μ	Schätzw. für μ (empirischer Erw.wert)	$\bar{x} = \frac{x_1+x_2+\cdots+x_n}{n}$
theor. Streuung σ^2	Schätzw. für σ^2 (empirische Streuung)	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$
(Modell)	(Beobachtung)	

Histogramm

- Es existieren Testverfahren zur Überprüfung, ob eine bestimmte Verteilung vorliegt oder nicht.

Beispiel 8 Ein Drehteil besitzt einen Soll-Durchmesser von 500 mm, die Toleranzgrenzen sind 499,6 und 500,3 (alle Maße in mm). Die von der Maschine produzierten Teile besitzen in Wirklichkeit einen Durchmesser der normalverteilt mit $\mu = 500$ und $\sigma = 0,2$ ist (siehe Diskussion 7).

Wie groß ist die Wahrscheinlichkeit, dass ein solches Teil

i. innerhalb der Toleranzgrenzen liegt

ii. Ausschuss ist, d.h. dass der Durchmesser kleiner als die untere Toleranzgrenze ist?

iii. Wie genau muss die Maschine arbeiten, d.h. wie groß darf die Standardabweichung σ höchstens sein, damit höchstens 1% der produzierten Teile Ausschuss sind? (Standardabweichung ist ein Qualitätsmerkmal der Maschine, spezifisch für jede einzelne Maschine)

Lösung: X ... Durchmesser (in mm)
$X \in N(\mu, \sigma^2), \mu = 500, \underbrace{\sigma = 0,2}_{\text{Aufg. i und ii}}$

i. $P(499,6 \leq X \leq 500,3) = \Phi(\frac{500{,}3-500}{0{,}2}) - \Phi(\frac{499{,}6-500}{0{,}2})$
$= \Phi(1,5) - \underbrace{\Phi(-2)}_{(1-\Phi(2)} = \underline{\underline{0,91044}}$

ii. $P(X \leq 499,6) = F_X(499,6) = \Phi(\frac{499{,}6-500}{0{,}2}) = \Phi(-2) = 1 - \Phi(2) =$ $\underline{\underline{0,02275}}$

iii. $f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2\right)$ mit $\mu = 500$
$N(500, \sigma^2)$ mit $P(X < 499,6) = 0,01$ Gesucht: σ
X mit μ zentrieren: $X - \mu$ dann mit σ normieren $\frac{X-\mu}{\sigma}$
(X ist standardverteilt, $N(0,1)$-Verteilung, Dichte $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

$P(x < 499,6) = P(\frac{X-500}{\sigma} \leq \frac{499,6-500}{\sigma}) = \Phi(\frac{-0,4}{\sigma}) = \int\limits_{-\infty}^{} \frac{-0,4}{\sigma}\varphi(x)\,dx =$
$0,01$
Umkehrfunktion: $\frac{-0,4}{\sigma} = \Phi^{-1}(0,01)$
$\Phi^{-1}(0,01) =$ Quantil der Ordnung $0,01 = \underbrace{z_\alpha}_{<0}, \alpha = 0,01$

Quantil der Ordnung γ Kurz $x_\gamma,\ 0 < \gamma < 1$

Definition: $P(X < x_\gamma) \leq \gamma \leq P(X \leq x_\gamma)$
stetige Verteilung: $P(X \leq x_\sigma) = F(x_\gamma) = \gamma$

2. Exponentialverteilung

Definition 11 Die ZG X heißt exp.-verteilt mit dem Parameter $\lambda(\lambda > 0)$, wenn sie die Dichte

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

besitzt. Kurz $X \in E(\lambda)$
Verteilungsfunktion $F(X) = \int\limits_{-\infty}^{x} f(t)\,dt = \begin{cases} 1 - e^{-\lambda \cdot x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

Diskussion

a) Es gilt $E(X) = \frac{1}{x}, D^2(X) = Var(x) = \frac{1}{x^2}$ d.h. Erwartungswert und Standardabweichung stimmen stets überein: $\frac{1}{x}$
b) Intervallwahrscheinlichkeit $P(a \leq X \leq b) = F(b) - F(a) = e^{-x \cdot a} - e^{-x \cdot b}$
c) Anwendungen: Bedienungstheorie, Zuverlässigkeitstheorie, zufällig. Lebensdauer, Bedienzeiten, Zeit zwischen zwei ankommenden Forderungen in einem Bedienungssystem (auch ankommende Aufträge, Zugriffe, ...)
Die X_i sind stochast. unabhängige ZG, $X_i \in E(\lambda)$ wenn Y die zufällige Anzahl der Zugriffe im Zeitintervall ΔT ist, dann hat Y eine Poissionverteilung $Y \in P(\lambda \cdot \Delta t)$. Mittlere Anzahl der Zugriffe im Zeitintervall Δt ist $E(Y) = \lambda \cdot \Delta t$, Auskunftsrate $= \lambda$ (Anzahl der Zugriffe pro Zeiteinheit)
Mittlere Zeitdauer zwischen zwei Forderungen $E(X_i) = \frac{1}{x}$ (Zugriffe, Auskünfte)

Beispiel 9 Ein System bestehe aus drei parallel geschalteten Elementen die unabhängig voneinander arbeiten.
Aus statisk. Untersuchungen sei bekann, dass die Lebensdauer der einzelnen Elemente exp.-verteilt sind mit dem Erwartungswert 1000 h.

i. Wie groß ist die Wahrscheinlichkeit, dass ein einzelnes Element höchstens 500 h funktioniert?

ii. Wie groß ist die Wahrscheinlichkeit, dass das System mindestens 500 h funktioniert?

iii. Für welchen Zeitraum beträgt die Zuverlässigkeit des Systems 99% ?

Lösung: X_i ... Lebensdauer des i-ten Elements ($i = 1, 2, 3$)
$X_i \in E(\lambda), \lambda = \frac{1}{E(X_i)} = \frac{1}{1000} = 0,001\, h^{-1}$
X ... Lebensdauer des Systems
$F(x) = P(X \leq x) = P(\{X_1 \leq x\} \cap \{X_2 \leq x\} \cap \{X_3 \leq x\})$
$= P(X_1 \leq x) \cdot P(X_2 \leq x) \cdot P(X_3 \leq x)$
$= F_{X_1}(x) \cdot F_{X_2}(x) \cdot F_{X_3}(x) = (1 - e^{-\lambda \cdot x})^3, x \geq 0$

i. $P(X_i \leq 500) = 1 - e^{-\lambda \cdot 500} = 1 - e^{-0,5} \approx 0,3935$

ii. $P(X_i \geq 500) = 1 - P(X_i < 500) = 1 - P(X_i \leq 500) = e^{-0,5}$
für das System: $P(X \geq 500) = 1 - P(X \leq 500) = 1 - (1 - e^{-0,5})^3 \approx 0,9391$

iii. Ansatz: $P(X \geq t) = 0,99$, gesucht t (System soll mindestens bis Zeitpunkt t funktionieren), $\lambda = 0,001$
analog ii) statt 500 jetzt t $\Rightarrow P(X \geq t) = 1 - (1 - e^{-\lambda \cdot t})^3 = 0,99$
$\curvearrowright 0,01 = (1 - e^{-\lambda t})^3$
$\sqrt[3]{0,01} = 1 - e^{-\lambda t} \quad t = -1000 \cdot \ln(1 - \sqrt[3]{0,01}) \approx 242,6\, h$

3. Chi-Quadraht-Verteilung

2.4 Mehrdimensionale Verteilungen

2.4.1 Zufällige Vektoren

Definition 12

$$\vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = (X_1, X_2, \ldots, X_n)^T$$

heißt

1. diskreter Zufalls-Vektor, wenn alle Komponenten $X_1, \ldots, X_n$ Zufallsgrößen sind.

2. stetiger Zufalls-Vektor, wenn die Komponenten eine gemeinsame Dichte $f_X(x_1, x_2, \ldots, x_n) \geq 0$ besitzen, d.h. $\mathcal{P}(\vec{X} \in B \in \mathbb{R}^n) =$,
$$B = \{ \underbrace{\int \ldots \int}_{n\text{-faches Integral}} f_X(x_1, \ldots, x_n)\, dx_1 \ldots dx_n \}$$
$$B = \{(x_1, x_2)^T | a \leq x_1 \leq b, c \leq x_2 \leq d\} : \mathcal{P}(\vec{X} \in B) = \int_{x_1=a}^{b} \int_{x_2=c}^{d} f_{\vec{X}}(x_1, x_2)\, dx_1\, dx_2$$

Diskussion $n = 1$: Verteilungstabelle:

$\begin{array}{c|cccc} & x_0 & x_1 & x_2 & \dots \\ \hline & p_0 & p_1 & p_2 & \dots \end{array}$ mit $p_n = \mathcal{P}(X = x_k)$

$\mathcal{P}(a \leq X \leq b) = \sum\limits_{k:a \leq x_k \leq b} p_k$

Verteilungstabelle als stochastische Matrix $\underline{P} : \sum\limits_i \sum\limits_j p_{ij} = 1 \wedge p_{ij} \geq 0$

Randverteilungen: $\mathcal{P}(X = x_i) = \mathcal{P}(\{X = x_i\} \cap \{Y \text{ bel.}\}) = \mathcal{P}(((X \quad Y) = (x_i \quad y_j), i \text{ fest}, j = 0, 1, 2, \dots) = \sum\limits_j p_{ij} = p_{i.}$, bzw. für $Y : \mathcal{P}(Y = y_j) = p_{.j}$

Randverteilungen für stetige Komponenten: Randdichte für X $X f_X(x) = \int\limits_{y=-\infty}^{\infty} f_X(x, y)\, dy$ bzw. $Y : f_Y(y) = \int\limits_{x=-\infty}^{\infty} f_X(x, y)\, dx$

aus bekannten Randverteilungen von X und Y (Fall $n = 2$) läßt sich i. Allg. nicht die Verteilung des Vektors $\underline{X} = \begin{pmatrix} X \\ Y \end{pmatrix}$ rekonstruieren.

z.B. diskrete Größen, aus $p_{0.}, p_{1.}, \dots,$ und $p_{.0}, p_{.1}, p_{.2}, \dots$ bekommt man i. Allg. nicht die Matrix $\underline{P} = (p_{ij})$ (Der Autor entschuldigt sich für die konfusen Mitschriften während der Vertretung des regulären Dozenten)

Beispiel 10 X ... zufäll. Anzahl der techn. Durchsichten eines PKW eines best. Typs zwischen 0 und 15000 km. Y ... zufäll. Anzahl der Motorproblemen dieser PKW zwischen 0 und 15000 km.

2.4.2 statistische Kennzahlen für Vektoren: Kovarianz und Korrelationskoeffizienten, sowie stochstische Unabhängigkeiten von ZG

Definition 13 $\vec{X} = \begin{pmatrix} X \\ Y \end{pmatrix}$ sei ein zufällg. Vektor. Dann heißen die Kennzahlen

$$\operatorname{cov}(X, Y) := E[(X - E(X)) \cdot (Y - E(Y))] = E(X \cdot Y) - (EX) \cdot (EY)$$

$$\varrho_{(x,y)} := \frac{\operatorname{cov}(X, Y)}{D(X) \cdot D(Y)}$$

X und Y heißen unkorreliert, wenn $\varrho(X, Y) = 0$ gilt. Beweis: Es gilt stets $-1 \leq \varrho(X, Y) \leq 1$.

Definition 14 $\underline{X} = \begin{pmatrix} X \\ Y \end{pmatrix}$ zufäll. Vektor. X und Y heißen stochast. unabhängig, wenn für bel. Intervalle I_1 für X und I_2 für Y gilt: $\mathcal{P}(\underbrace{\{X \in I_1\} \cap \{Y \in I_2\}}_{\text{gemeins. Verteil.}}) = \underbrace{\mathcal{P}(X \in I_1) \cdot \mathcal{P}(Y \in I_2)}_{\text{Faktorisierung (Randvertl. gen.)}}$

d.h. diskrete ZG: für alle Gitterpunkte $\begin{pmatrix} X_i \\ Y_j \end{pmatrix} : p_{ij} = p_{i.} \cdot p_{.j} \quad \forall i, j$ bzw. stetige ZG, gemein. Dichte $f_X(x, y) = f_X(x) \cdot f_Y(y) \quad \forall \begin{pmatrix} X \\ Y \end{pmatrix} \in \mathbb{R}^2$

Bemerkung: Stochast. Unabhängigkeit $\Rightarrow$ Unkorreliertheit. Unkorreliert $\Rightarrow$ Stochast. Unabhängigkeit nur bei Normalverteilung.

Diskussion: $\operatorname{cov}(X, Y) = E[(X - E(X)) \cdot (Y - E(Y))] \ldots$ Kovarianz
Varianz $= \{D^2(X) = E[(X - E(X))^2] = E[(X - E(X)) \cdot (X - E(X))], D^2(Y) = E[(Y - E(Y))^2] = \ldots\}$

Eigenschaften mit dem „E-Operator“:

1. $E(const.) = const.$ kurz $E(c) = c$
2. $E(\alpha \cdot X + \beta \cdot Y) = \alpha \cdot E(X) + \beta \cdot E(Y) \quad \alpha, \beta \in \mathbb{R}$, feste Zahlen
3. $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + \cdots + E(X_n) = \sum\limits_{k=1}^{n} E(X_k)$

D-Operator:

1. $D^2(X) = 0 \Leftrightarrow \exists c \in \mathbb{R} : \mathcal{P}(X = c) = 1$ keine Streuung
2. $D^2(\alpha \cdot X) = a^2 \cdot D^2(X), \alpha \in \mathbb{R}$ fest vorgegeben
3. $D^2(X \pm Y) = D^2(X) + D^2(Y) \pm 2 \cdot \operatorname{cov}(X, Y)$
4. $X_1, X_2, \ldots, X_n$ paarweise unkorreliert: $D^2(X_1 + X_2 + \ldots) = \sum\limits_{k=1}^{n} D^2(X_k)$

Diskussion $\varrho(x, y)$ ist ein Maß für die lineare Abhängigkeit von X zu Y. Der Grad des linearen Zusammenhangs wird über das Bestimmtheitsmaß

$$B = \varrho^2(x, y)$$

beschrieben. $B = 1 = 100\%$ bedeutet strenger linearer Zusammenhang, hingegen $B = 0$ bedeutet kein linearer Zusammenhang. (Weder linearer noch nicht linearer Zusammenhang), X und Y stochastisch unabhängig.

Zur Geraden $y = a_1 \cdot x + a_0$. Es gilt für die optimale Gerade: $a_1 = \frac{\partial Y}{\partial X} \cdot \varrho(x, y) = \frac{\operatorname{cov}(X,Y)}{D(Y)D(X)} \cdot \frac{\partial Y}{\partial Y} = \frac{\operatorname{cov}(X,Y)}{D^2(X)}$ da $\partial X = D(X), \partial Y = D(Y)$ und $a_0 = E(Y) - a_1 \cdot E(X)$. Die optimale Gerade heißt Regressionsgerade von Y bzgl X.
Es sei $g(X, Y)$ eine zufällige Funktion X und Y, z.B. $g(s, t) = s^2 \cdot t^3 \Rightarrow g(X, Y) = X^2 \cdot Y^3$ usw.

Was ist dann E(g(X,Y))?
Diskrete ZG: Wahrscheinlichkeitsmassen p_{ij} auf den Gitterpunkten $\begin{pmatrix} x_i \\ y_i \end{pmatrix} : E(g(X, Y)) = \sum\limits_i \sum\limits_j g(x_i, y_i) \cdot p_{ij}$
Stetige ZG: gemeinsame Dichtefunktion: $f_{\vec{X}}(x, y) : E(g(X, Y)) = \int\limits_{X=-\infty}^{\infty} \int\limits_{Y=-\infty}^{\infty} f_{\vec{X}}(x, y) \cdot g(x, y)\, dy dx$

Definition 15 Es sei $\vec{X} = (X_1, X_2, \dots, X_n)^T$

1. in Verallgemeinerung der Streuung einer ZG X jetzt die Kovarianzmatrix $\vec{K}_{\vec{X}} := E((\vec{X} - \vec{\mu}) \cdot (\vec{X} - \vec{\mu})^T)$

$\vec{\mu} = E(\vec{X})$
subparagraphBeispiel 14 Die ZEit, die zur Überprüfung bestimmter elektronischer Bauelemente benötigt wird, beträgt im Mittel 1,248 min, Standardabw. sei 0,54635 min. (Grundlage Stichprobenerhebung) vgl. Aufgabe (3) 3 der Aufg. sammlung.

1. Wie großt ist die Wkt., dass in 130 min genau 100 Bauelemente geprüft werden können?

2. Welche Anzahl von Bauelementen kann in 2 Stunden mit mindests. 95% Sicherheit geprüft werden?

Lösung Sei X_i die zufällige individuelle Prüfzeit des ersten Bauelements (Wkt.-verteilung von X_i muss nicht bekannt sein. Ausreichend ist die Kenntnis von $E(X_i)$ und $D^2(X_i)$) Sei $\mu = E(X_i) = 1,248, D^2(X_i) = \sigma_x^2 = 0,54635^2\,[\mathrm{min}^2]$ Vorraussetzung: individuelle Prüfzeiten stochast. unabhängig.
Gesamtprüfzeit $S_{100} = \sum\limits_{i=1}^{100} X_i \underbrace{\approx}_{\text{ZGWS}} \in N(n\mu, n \cdot \sigma_x^2)$

- $P(S_{100} \leq 130) \approx \Phi(\frac{130-n\cdot\mu}{\sqrt{n\cdot\sigma_x^2}}) = \Phi(\frac{130-100\cdot 1,248}{\sqrt{100}\cdot 0,54635}) = 0,83 = 83\%$. Die geforderte Überprüfung von 100 Bauelementen in der vorgegebenenen Zeit wird mit einer Wkt. von 83 % erfüllt. In 17 % der angeordneten Überprüfungen von 100 Bauelementen wird die Vorgabezeit 130 min überschritten werden.

- Vorgabezeit von 130, auf 120 min heruntergesetzt, Vorgabewkt. von 83 % auf 95 % erhöhen. Welche Anzahl n an Bauelementen darf höchstens noch vorgegeben werden. Ansatz: $P(S_n \leq 100) \geq 0,95 \Leftrightarrow \Phi(\frac{120-n\cdot\mu}{\sqrt{n\cdot\sigma_x^2}}) \geq 0,95 = \Phi(z_{0,95})$

  Die Verteilungsfunktion $n = \Phi(z)$ ist streng monoton wachsend $\Rightarrow \frac{120-n\cdot\mu}{\sqrt{n}\cdot\sigma_x} \geq z_{0,95} = 1,645 \Rightarrow \underbrace{120 - n\cdot\mu}_{+} \geq \underbrace{z\cdot\sqrt{n}\cdot\sigma}_{+} \Rightarrow (120 - n\cdot\mu)^2 \geq z^2 \cdot n\sigma^2$

  $n^2 \cdot \mu^2 - 240n\cdot\mu + 120^2 - z^2 \cdot n \cdot \sigma^2 \geq 0 \Rightarrow n^2 + \frac{z^2\cdot\sigma^2 - 240\mu}{\mu^2} \cdot n + (\frac{120}{\mu}^2 \geq 0 \rightarrow n_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{x} - q}$ (mit $n > 0$) Lösung $n \approx 89,3 \Rightarrow n \leq 89$

3 deskriptive Statistik: Grundbegriffe

3.1 Merkmale

Merkmal: zufällige Größe (X),, die beobachtet wird

Merkmalsausprägungen: konkrete Werte von X, die in der Datenerhebung auftreten können

Klassifikation der Merkmale

- - quantitative Merkmale
  - qualitative Merkmale (verbal deshalb in Zahlen kodieren)
- - Nominalskala (keine Ordnungsstruktur wie größer als) z.B. Geschlecht, Konfession
  - Ordinalskala (Rangordnung) z.B. Schulnoten
  - metrische Skala: Rangordnung und zusätzlich sind die Abstände zwischen den Ausprägungen sinnvoll interpretierbar, z.B. Einkommen
- - diskretes Merkmal
  - stetiges Merkmal
  - quasistetiges Merkmal (z.B. Digitalisierung)

3.2 Grundgesamtheit und Stichprobe

Grundgesamtheit (machnmal Grundgesamtheit X): beinhaltet alle für die statist. Erhebung (Datensammlung) relevanten Informationen

Definition 1 Eine ZGR X, durch die ein bestimmtes Merkmal beschrieben wird, heißt Grundgesamtheit X

Diskussion

1. Die GG X ist wahrscheinlichkeitstheoretisch vollständig diskretisierbar, z.B. wenn deren Verteilungsfkt. $n = F(x) = P(X \leq x)$ (eindim.), mehrdim. analog (Zufallsvector $\vec{x}$ betrachten)

Beispiel 5 200 CD-Rohlinge eines bestimmten Fabrikats wurden einer Qualitätsprüfung unterzogen. Dabei erwiesen sich 12 als unbrauchbar. Man gebe zum Konfidenzniveau 95 % einen konkreten Vertrauensbereich für den unbekannten Ausschussanteil p dieser CD-Rohlinge an.

Lösung: $n = 200, w_n = \frac{12}{200} = 0,06$
$\curvearrowright n_p(1-p) \approx nw_n(1-w_n) = 200 \cdot 0,06 \cdot 0,94 = 11,28 > 9$
Näherung (a) anwendbar
$1 - \alpha = 0,95 \curvearrowright \alpha = 0,05 \curvearrowright z = z_{1-\frac{\alpha}{2}} = z_{0,975} = 1,96$

Diskussion

- Länge des Konfidenzintervalls (für großes n): $b-a \approx \frac{2z\cdot\sqrt{p(1-p)}}{\sqrt{n}} \leq \frac{2}{\sqrt{n}}$ Verdopplung der Genauigkeit erfordert 4-fachen Stichprobenumfang
- Für kleine n siehe Merkblatt Konfidenzintervalle

4 Testtheorie

Problem:

1. Geg. Stichprobe $X_1, \dots, X_n$ aus GG X
2. Aufgabe: Annahmen (Hypothesen) über die unbekannte Verteilung der GG X überprüfen

- Fall: Verteilungsfunktion: bis auf Parameter Θ bekannt, Hypothese betrifft nur Parameter Θ (z.B. $\Theta = \Theta_0$, wobei Θ_0 ... Sollwert) Parametertests(vgl. 3.1.)
- Verteilungstyp unbekannt, nicht parametrische Tets (vgl. 3.2)

Test-Prinzip Entscheidung zwischen der Hypothese (H_0) und einer sogenannten Alternative (H_1)

Vorgehensweise

- Vorgabe einer Irrtumswahrscheinlichkeit $\alpha \in (0,1), \alpha$ klein (oft 0,05, auch 0,01, 0,1) Wahrscheinlichkeit H_0 abzulehnen obwohl H_0 richtig
- Nullhypothese H_0 und Alternativhypothese H_1 angeben.
- Konstruktion einer Testgröße $T = T(X_1, \dots, X_n)$ deren Verteilung bzw. Gültigkeit von H_0 bekannt ist
- Angabe muss kritischen Bereichs K derart, dass unter H_0 gilt: $P(T \in K) \leq \alpha$ Wahrscheinlichkeit von K ist nicht eindeutig. K soll die Werte enthalten die für die Alternative H_1 sprechen
- Entscheidungsregel: Gilt für die konkrete Stichprobe $x_1, \dots, x_n : t := T(x_1, \dots, x_n) \in K$, dann wird H_0 zugunsten von H_1 abgelehnt, anderenfalls ist gegen H_0 nichts einzuwenden (H_0 ist damit nicht bestätigt)

4.1 Parametertests

4.1.1 Grundbegriffe, allgemeine Vorgehensweise

Beispiel 1 (Zur Demonstration der allg. Vorgehensweise)
X ... Flüssigkeitsmenge, die von einem Abfüllautomaten pro Flasche angegeben wird [ml] , Sollwert $\mu_0 = 500, X \in N(\mu, \sigma^2), \mu$ unbekannt, σ^2 bekannt ($\sigma = 5$). Zu überprüfen ist, ob der Sollwert f ein Mittel einhalten wird. Irrtumswahrscheinlichkeit 5 %. Eine Stichprobe vom Umfang $n = 20$ ergab $\bar{x} = 498ml$.

Lösung:

1. $\alpha = 0,05$

2. • $H_0 : \mu = \underbrace{500}_{\mu_0}$
   • Für die Alternative gibt es 3 Varianten, welche sinnvoll ist, hängt vom Anwender ab

   a) $H_1 : \mu \neq 500$: (z.B. für unabhängigen Beobachter, Gutachter, Abweichung nicht oben und unten kritisch)
   b) $H_1' : \mu > 500$ (für Betreiber des Automaten wichtig, zuviel abgefüllt!)
   c) $H_1'' : \mu < 500$ (für den Verbraucher wichtig)

3. Testgröße $T = \frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} \underbrace{\in}_{\text{falls } H_0} N(0;1)$ konkreter Wert $t = \frac{498-500}{5}\sqrt{20} = -1,78$

4. Krit. Bereich

5. Entscheidung
   a) $t \notin K \curvearrowright$ gegen H_0 nichts einzuwenden
   b) $t \notin K \curvearrowright$ gegen H_0 nichts einzuwenden
   c) $t \in K \curvearrowright H_0$ wird zugunsten von $H_1'' : \mu < 500$ abgelehnt, Wktk. einer Fehlentscheidung 5 % (statistische Sicherheit $1 - \alpha = 95\%$

Diskussion

1. Bei einseitiger Fragestellung wird oft folgende Form der Nullhypothese verwendet: Variante
   a) $H_0 : \Theta \leq \Theta_0, H_1 : \Theta > \Theta_0$
   b) $H_0 : \Theta \geq \Theta_0, H_1 : \Theta < \Theta_0$

   Dann Verteilung von T nur bei Gültigkeit der Gleichheitszeichen bekannt (i.A.) vgl. Schritt 3, Im jedem Fall ist aber $P(T \in K) \leq \alpha$ (unter H_0)

2. Es besteht ein enger Zusammenhang zwischen Konfidenzschätzungen und Paramtetertests Bsp.: Test (zweiseitig), $X \in N(\mu, \sigma^2), \Theta = \mu, \sigma^2$ bekannt, $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0 \Leftrightarrow$ (zweiseitiges Konfidenzintervall für μ

   Es gilt $\mu_0 \notin I \Leftrightarrow T \in K$ D.h. Ablehnung der Nullhypothese genau dann, wenn das Konfidenzintervall den Sollwert μ_0 nicht überdeckt

3. Mögliche Fehler bei Tests:
   a) Fehler 1. Art H_0 wird abgelehnt, obwohl H_0 richtig
   b) Fehler 2. Art H_0 wird nicht abgelehnt, obwohl H_0 falsch

4. Die Wkt. für das Auftreten eines Fehlers 1. Art ist höchstens gleich der Irrtumswkt α

- α heißt auch Signifikanzniveau (es wird getestet ob wesentliche) Abweichungen vom Sollwert auftreten
- Ein Test gemäß 1-5 heißt auch Signifikanztest

5. Analyse des Fehlers 2. Art am Beispiel

$$X \in N(\mu, \sigma^2), \sigma^2 \text{ bekannt}, H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0$$

$\curvearrowright$ krit. Bereich $K = (z_{1-\alpha}; \infty)$

- Für beliebiges $\mu \in \mathbb{R}$ werden erklärt

Operationscharakteristik	$OC(\mu) := P(H_0 \text{ wird nicht abgelehnt}) = P(T \notin K)$
Gütefunktion	$g(\mu) := P(H_0 \text{ wird abgelehnt}) = 1 - OC(\mu) = P(T \in K)$

- Es gilt: $\bar{X} \in N(\mu; \frac{\sigma^2}{n}) \curvearrowright T = \frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} \in N(\frac{\mu-\mu_0}{\sigma}\sqrt{n}, 1) \curvearrowright OC(\mu) = P(T \notin K) = P(T \leq z_{1-\alpha}) = \Phi(z_{1-\alpha} - \frac{\mu-\mu_0}{\sigma}\sqrt{n})$
- OC ist von n abhängig. Es gilt für jedes $\mu > \mu_0 (\equiv H_1)$ $\lim\limits_{n\to\infty} OC(\mu) = 0$
- Der Stichprobenumfang n lässt sich so bestimmen, dass für $\mu \geq \mu_1 > \mu_0$ gilt: $OC(\mu) \leq \beta : OC(\mu) \leq \beta \Leftrightarrow z_{1-\alpha} - \frac{\mu-\mu_0}{\sigma}\sqrt{n} = \Phi^{-1}(\beta) = z_\beta = -z_{1-\beta}$
  $\curvearrowright n \geq (\frac{z_{1-\alpha}+z_{1-\beta}}{\mu_1-\mu_0} \cdot \sigma)^2$
  μ_1 und β sind vorgebar, bei Überschreitung von μ_1 (wesentliche Überschreitung des Sollwerts) beträgt die Wkt für Fehler 2. Art höchstens β, für unwesentliche Überschreitungen (zwischen μ_0 und μ_1) trifft das nicht zu. Hier liegt die Wkt für einen Fehler 2. Art zwischen β und $1 - \alpha$.

6. Zur Entscheidungsregel

Ergebnis	Entscheidung	Fehlentscheidung falls	Wkt. Fehlentscheidung	stat. Sicherheit
$t \in K$	H_0 zugunsten von H_1 ablehnen	H_0 richtig	$\leq \alpha$ (Fehler 1. Art)	$\geq 1 - \alpha$
$t \notin K$	gegen H_0 nicht einzuwenden	H_0 falsch	$< 1 - \alpha$	-

Bei einseitiger Fragestellung ist es zweckmäßig, die vermutete bzw. zu beweisende Aussage als Alternative zu wählen

Bemerkung Indirekter Beweis (Logik), H_1 ist zu beweisen, Annahme des Gegenteils H_0 auf Widerspruch führen $\curvearrowright$ H_0 falsch, H_1 wahr (deterministisch, 100 % Sicherheit)

4.1.2 Test für Erwartungswert und Streuung bei normalverteilter GG X

Testgrößen und ihre Verteilung sowie zugehörige kritische Bereiche s. Merkblatt Parametertests

Beispiel Auf einen Drehautomaten werden Zylinder hergestellt. Der Durchmesser kann als normalerweilt angesehen werden. Die Streuung σ^2 ist ein Gütemaß für den Drehautomaten. Der Hersteller des Automaten gibt an, dass die Standardabweichung σ höchstens 0,03 mm beträgt. Der Betreiber des Automaten zweifelt dies an und möchte bei einer statischen Sicherheit von 95 % das Gegenteil beweißen (d.h. $\sigma > 0,03$). Dazu werden von 40 herstellen Zylindern die Durchmesser kontrolliert. Es ergibt sich $\bar{x} = 50,03mm, s = 0,097mm$. Lässt sich die Vermutung des Betriebs bestätigen?

Lösung: X^2- Streuungstest

1. Irrutmswkt $\alpha = 0,05$
2. $H_0 : \sigma^2 \leq \sigma_0^2, H_1 : \sigma^2 > \sigma_0^2$ (mit $\sigma_0 = 0,03$)
3. Testgröße vgl. Merkblatt $T = \frac{(n-1)S^2}{\sigma_0^2}$, konkreter Wert $t = \frac{39 \cdot 0,037^2}{0,03^2} = 59,32$
4. Krit. Bereich, vgl. Merkblatt $K = (X^2_{n-1,1-\alpha}; \infty) = (X^2_{39;0,95}; \infty) = (54,57; \infty)$
5. Entscheidung: $t \in K$, H_0 wird zugunsten von $H_1(\sigma > 0,03)$ abgelehnt, mit 95 % iger statistischer Sicherheit lässt sich die Vermutung des Betreibers bestätigen.

Diskussion

1. Die Entscheidung Ablehnung von H_0 oder nicht, hängt vom gewählten Signifikanzniveau α ab. Hätte man z.B. $\alpha = 0,01$ gewählt, so erhielte man $t \notin K = (62,43; \infty)$, d.h. gegen H_0 (Behauptung des Herstellers) ist nichts einzuwenden bzw. H_1 (die Behauptung des Betreibers) ließe sich nicht mit 99 % Sicherheit nachweisen.
2. Selbstverständlich ist das Niveau α vor der konkreten Durchführung des Tests festzulegen
3. Derjenige α-Wert, für den eine Grenze des kritischen Bereiches mit dem konkreten Wert t der Testgröße übereinstimmt, d.h. die Grenzstelle zwischen Ablehnung und Nichtablehnung heißt auch p-Wert.
   Damit $p < \alpha \curvearrowright$ Ablehnung von H_0
   $p \geq \alpha \curvearrowright$ gegen H_0 ist nichts einzuwenden.

   Im Beispiel 2 ergibt sich $p = 0,0195 < \alpha = 0,05$ Entscheidung wie oben! Die Angabe des p-Wertes erfolgt bei vielen Software-Paketen sowie TR anstelle des kritischen Bereiches.

Beispiel 4 Bei 100 Bauelementen der gleichen Art werde die Lebensdauer überprüft. Eine statistische Auswertung ergab:

1. $\bar{x} = 1203,1\ h, s = 614\ h$

2. Häufigkeitstabelle

0; 500	500; 1000	1000; 1500	1500; 2000	2000; 2500	2500; 3000	3000; 3500	3500; 4000
11	29	27	23	7	2	0	1

Man überprüfe bei einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$, ob die Grundgesamtheit X als expotentialverteilt angesehen werden kann.

Lösung: X^2-Anpassungstest

1. $\alpha = 0,05$
2. $H_0 : X \in E(\lambda), F(x) = F_0(x)$ mit $F_0(x) = \begin{cases} 1 - e^{-\lambda x} & \text{für} \quad x \geq 0 \\ 0 & \text{für} \quad x < 0 \end{cases}$, dabei $\lambda = \frac{1}{\bar{x}} = 8,312 \cdot 10^{-4}$(Max-Likeltrod-Schätzung für λ, vgl. ÜA B9) $H_1 = \bar{H_0}$
3. Testgröße

$K_j = [a_j, a_{j+1})$	$F_0(a_j) = 1 - e^{-\lambda a_j}$	$p_j = F_0(a_j + 1) - F_0(a_j)$	$n \cdot P_j$	h_j
$[0; 500)$	0	0,340	34,0	11

$$T = \sum_{j=1}^{k} \frac{(H_j - np_j)^2}{np_j} \underbrace{\tilde{\in}}_{(H_0)} X^2_{k-1-m}$$

k=7, m=1

konkreter Wert: $t = \frac{(11-34,0)^2}{34,0} + \cdots + \frac{(2-7,0)^2}{7,0} + \frac{(1-5,5)^2}{5,5} = 52,9$

4. $K = (X^2_{k-1-m;1-\alpha}; \infty) = (11,07; \infty)$
$X^2_{5;0,95} = 11,07$
5. $t \in K, H_0$ wird abgelehnt, d.h. die Lebensdauer ist mit 95 %iger Sicherheit nicht expotentialverteilt

Diskussion zum Beispiel 4

- Dichte der Expotentialverteilung
- Bessere Anpassung an das Histogramm z.B. durch die Rayleigh-Verteilung (vgl. Beispiel 2, Kap. 2.1.)
Dichte: $f_0(x) = \begin{cases} \frac{2x}{\lambda} e^{-\frac{x^2}{\lambda}} & \text{für} \quad x \geq 0 \\ 0 & \text{für} \quad x < 0 \end{cases}$
- Ein analog durchgeführter Test führt bei dieser Verteilung auf $t \notin K$, d.h. es lässt sich nicht mit 95%iger Sicherheit widerlegen, dass X Rayleigh-verteilt ist.

Beispiel 5 100 Würfe mit einer Münze ergaben $58 \times \{\text{Zahl}\}$. Man überprüfe mit einer Irrtumstswkt. von $0,05$, ob die Münze als ideal (symmetrisch) angesehen werden kann, d.h. ob die beiden möglichen Versuchsausgänge { Wappen } bzw. { Zahl } gleichwahrscheinlich sind.

Lösung: X^2-Anpassungstest Versuchsergebnis: $E_1 := \{\text{Wappen}\}, E_2 := \{\text{Zahl}\}$

1. $\alpha = 0,05$
2. $H_0 : P(E_1) = P(E_2) = 0,5, H_1 = \bar{H}_0$
3. Testgröße

Versuchserg	p_j	np_j	h_j
E_1	$0,5$	50	42
E_2	$0,5$	50	58

Diskussion: Falls im Beispiel 5 116 $\times$ {Zahl} bei 200 Würfen auftritt (gleiche relative Häufigkeit), dann $t = 5,12 \in K$ Ablehnung von H_0

4.1.3 Weitere parameterfreie Tests

- Kolmogorov-Test (Test auf Unterliegen einer stetigen VF: $F_0(x)$, benötigt wird die Urliste)
- X^2-Unabhängigkeitstest (zur Überprüfung der Unabhängigkeit zweier Merkmale X und Y auf der Basis einer zweidimensionalen Stichprobe $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$)