



Angewandte Stochastik II

Universität Ulm
Institut für Stochastik

Vorlesungsskript
Dr. Ole Stenzel
Stand: Wintersemester 2014/15

ULM, IM OKTOBER 2014

Contents

1	Grundideen der statistischen Datenanalyse	3
1.1	Zufallsstichprobe	3
1.2	Stichprobenmittel	5
1.3	Stichprobenvarianz	7
1.4	Empirische Verteilungsfunktion	9
1.5	Dichteschätzer	12
2	Schätzung von Parametern	14
2.1	Eigenschaften von Parameterschätzern	16
2.2	Momentenmethode	17
2.3	Maximum-Likelihood-Methode	20
3	Konfidenzintervalle	24
3.1	Modellbeschreibung	24
3.2	Konfidenzintervalle bei Normalverteilung; Quantilfunktion	26
3.3	Statistische Prüfverteilungen	28
3.4	Weitere Konfidenzintervalle bei Normalverteilung	30
3.5	Zwei-Stichproben-Probleme	32
3.6	Asymptotische Konfidenzintervalle	35
4	Tests statistischer Hypothesen	38
4.1	Problemstellung und Modellbeschreibung	38
4.2	Parametrische Tests	41
4.2.1	Parameter-tests bei Normalverteilung	45
4.2.2	Zwei-Stichproben-Tests	49
4.2.3	Asymptotische Parameter-tests	51
4.3	Kolmogorow-Smirnow-Test; χ^2 -Anpassungstest	52
5	Tabellen für Verteilungsfunktionen und Quantile	58

1 Grundideen der statistischen Datenanalyse

Zu den Zielen der mathematischen Statistik gehört die Beschreibung und Untersuchung der Eigenschaften bzw. Gesetzmäßigkeiten von (großen) Datensätzen. Dabei kann es sich einerseits um sogenannte

- *reale Daten*

handeln, die sich z.B. bei der Beobachtung (Messung) von Vorgängen bzw. Strukturen in Natur, Technik oder Wirtschaft ergeben oder es können

- *synthetische Daten*

sein, die bei der Simulation solcher Vorgänge bzw. Strukturen durch Computeralgorithmen erzeugt werden.

Die grundlegende Idee der Statistik, um diese Zielstellung zu erreichen, ist die *stochastische Modellierung* der vorliegenden Daten.

Die Vorlesung nutzt die Begriffe und Ergebnisse der Wahrscheinlichkeitsrechnung, die in der vorhergehenden Vorlesung ‘Angewandte Stochastik I’ diskutiert wurden, wie:

- Ereignis und Wahrscheinlichkeit
- Zufallsvariable und Verteilung
- Erwartungswert und Varianz
- stochastische Unabhängigkeit
- Gesetz der großen Zahlen
- zentraler Grenzwertsatz

1.1 Zufallsstichprobe

Der Vektor der vorliegenden Daten (x_1, \dots, x_n) kann natürlich im Allgemeinen eine komplizierte Struktur aufweisen.

- Dabei muss der ‘Wert’ x_i nicht unbedingt eine Zahl sein, sondern x_i kann für jedes $i = 1, \dots, n$ selbst ein Vektor sein, der beispielsweise die Lage, Größe, Form und Orientierung eines geometrischen Objektes beschreiben kann.
- Im Rahmen dieser einführenden Vorlesung setzen wir jedoch stets voraus, dass - falls nicht anders angegeben - $x_i \in \mathbb{R}$ für jedes $i = 1, \dots, n$.

Wir nehmen an, dass die Daten x_1, \dots, x_n die *Realisierung* eines stochastischen Modells sind.

- Und zwar sei x_1, \dots, x_n die Realisierung einer Folge von unabhängigen und identisch verteilten Zufallsvariablen $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$, die über einem (im Allgemeinen nicht näher spezifizierten) Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) definiert sind. Die Eigenschaft, dass die Zufallsvariablen X_1, \dots, X_n unabhängig und identisch verteilt sind, wird oft mit ‘iid’ (independent and identically distributed) abgekürzt.

- Da X_1, \dots, X_n unabhängig und identisch verteilt sind, folgt insbesondere, dass für ein $\omega \in \Omega$

$$X_i(\omega) = x_i \quad \forall i \in \{1, \dots, n\}.$$

- Außerdem setzen wir stets voraus, dass $\mathbb{E}(X_i^2) < \infty$ für $i = 1, \dots, n$.

Definition 1.1

1. Der Vektor (x_1, \dots, x_n) heißt (konkrete) *Stichprobe*.
2. Die Menge \mathbb{R}^n aller (potentiell möglichen) Stichproben (x_1, \dots, x_n) heißt *Stichprobenraum*.
3. Der Zufallsvektor (X_1, \dots, X_n) heißt *Zufallsstichprobe*.
4. Für jedes $i = 1, \dots, n$ heißt x_i *Stichprobenwert* von (x_1, \dots, x_n) . Analog hierzu nennt man X_i *Stichprobenvariable* von (X_1, \dots, X_n) .
5. Die Dimension n von (x_1, \dots, x_n) bzw. (X_1, \dots, X_n) heißt *Stichprobenumfang*.

Beachte

- Die allgemeine Zielstellung der statistischen Datenanalyse, die in der Einleitung diskutiert wurde, kann nun wie folgt präzisiert werden: Aus den vorliegenden Daten x_1, \dots, x_n sollen Schlussfolgerungen über Eigenschaften der (unbekannten) Verteilung der Zufallsstichprobe (X_1, \dots, X_n) gezogen werden.
- Weil wir voraussetzen, dass die Stichprobenvariablen X_1, \dots, X_n unabhängig und identisch verteilt sind, wird die Verteilung von (X_1, \dots, X_n) eindeutig durch die (Rand-) Verteilungsfunktion $F = F_{X_i}$ einer (einzelnen) Stichprobenvariablen bestimmt. Dieses ist leicht zu sehen, da für alle $t_1, \dots, t_n \in \mathbb{R}$ gilt:

$$F_{(X_1, \dots, X_n)}(t_1, \dots, t_n) \stackrel{\text{unabh.}}{=} F_{X_1}(t_1) \cdot \dots \cdot F_{X_n}(t_n) \stackrel{\text{ident. verteilt}}{=} F(t_1) \cdot \dots \cdot F(t_n).$$

- Aus den Daten x_1, \dots, x_n sollen also Schlussfolgerungen über Eigenschaften der unbekanntenen Verteilungsfunktion F gezogen werden.

Beispiel

- Punktschätzer. Es soll anhand einer Umfrage von n Personen eine Wahlprognose für eine Stichwahl zwischen zwei Kandidaten A und B erstellt werden. Sei X die Anzahl derjenigen Personen, die Kandidat A wählen würden. Verteilungsannahme: $X \sim \text{Bin}(n, p)$. Gesucht ist eine Approximation für die unbekanntene ‘Erfolgswahrscheinlichkeit’ p anhand einer konkreten Stichprobe (x_1, \dots, x_n) .

- Konfidenzintervall. Das Füllvolumen von 0,7l Wasserflaschen sei als normalverteilt angenommen. Sei $X = (X_1, \dots, X_n)$ eine entsprechende Zufallsstichprobe, d.h. $X_i \sim N(\mu, \sigma^2)$. Es soll ein anhand einer konkreten Stichprobe (x_1, \dots, x_n) ein Intervall bestimmt werden, das die zufällige Füllmenge mit 95% Wahrscheinlichkeit enthält.
- Statistischer Test. Ein Hersteller von Lampen gibt eine durchschnittliche Lebensdauer von mindestens 1000 Stunden an. Diese Angabe soll anhand einer Realisierung (x_1, \dots, x_n) einer Zufallsstichprobe $X = (X_1, \dots, X_n)$ untersucht werden. Es wird angenommen, dass die Lebensdauern exponentiell verteilt sind mit Mittelwert μ , d.h. $X_i \sim \text{Exp}(1/\mu)$. Anhand der Stichprobe (x_1, \dots, x_n) soll überprüft werden, ob die erwartete Lebensdauer $E(X_i)$ größer gleich 1000 ist ($E(X_i) \geq 1000$).

Um Eigenschaften der Verteilungsfunktion F zu bestimmen, werden Funktionen $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ betrachtet, die der Stichprobe (x_1, \dots, x_n) die ‘Bewertung’ $\varphi(x_1, \dots, x_n) \in \mathbb{R}$ zuordnen. Dies führt zu der folgenden Begriffsbildung.

Definition 1.2 Eine Borel-messbare Abbildung $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt *Stichprobenfunktion*.

Beachte Es ist üblich, auch die zusammengesetzte Abbildung $\varphi(X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}$ mit $\varphi(X_1, \dots, X_n)(\omega) = \varphi(X_1(\omega), \dots, X_n(\omega))$ *Stichprobenfunktion* zu nennen, d.h., φ ist dann eine Funktion der Zufallsstichprobe (X_1, \dots, X_n) .

1.2 Stichprobenmittel

Wir diskutieren zunächst die Frage, wie der Erwartungswert $\mu = \mathbb{E} X_i$ der Stichprobenvariablen X_i aus den beobachteten Daten x_1, \dots, x_n bestimmt werden kann.

- Hierfür betrachten wir die Stichprobenfunktion $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ mit $\varphi(x_1, \dots, x_n) = n^{-1} \sum_{i=1}^n x_i$.
- D.h., wir betrachten das arithmetische Mittel

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

der Stichprobenwerte x_1, \dots, x_n .

- Die Zahl \bar{x}_n wird *Stichprobenmittel* der (konkreten) Stichprobe (x_1, \dots, x_n) genannt.
- Außerdem betrachten wir das arithmetische Mittel der Stichprobenvariablen X_1, \dots, X_n , dessen Eigenschaften bereits im Zusammenhang mit dem Gesetz der großen Zahlen untersucht worden sind.

Definiton 1.3 Die Zufallsvariable

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

heißt *Stichprobenmittel* der Zufallsstichprobe (X_1, \dots, X_n) .

Theorem 1.4 Es gilt

$$\mathbb{E} \bar{X}_n = \mu \quad (3)$$

und

$$\text{Var} \bar{X}_n = \frac{\sigma^2}{n}, \quad (4)$$

wobei $\mu = \mathbb{E} X_i$ und $\sigma^2 = \text{Var} X_i$ den Erwartungswert bzw. die Varianz der Stichprobenvariablen X_i bezeichnen.

Beachte

- Wegen (3) sagt man, dass bei der Schätzung von μ durch \bar{X}_n kein ‘systematischer Fehler’ begangen wird.
- Der Schätzer \bar{X}_n kann dennoch sehr ungenau sein, wobei man den in (4) gegebenen Wert σ^2/n als Kennzahl für die Schätzgenauigkeit von \bar{X}_n auffassen kann.
- Insbesondere bedeutet (4), dass die Schätzgenauigkeit mit wachsendem Stichprobenumfang n verbessert wird.
- Dabei ist σ^2 im allgemeinen ebenfalls unbekannt und muss aus den Daten geschätzt werden.

Neben den Formeln (3) und (4) für Erwartungswert und Varianz des Stichprobenmittels \bar{X}_n sind noch weitere Aussagen über die Verteilung von \bar{X}_n von Interesse bzw. über deren asymptotisches Verhalten für große n .

Aus dem starken Gesetz der großen Zahlen bzw. aus dem zentralen Grenzwertsatz ergibt sich unmittelbar

Theorem 1.5 Es gilt

$$P\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mu\right) = 1 \quad (5)$$

und

$$\lim_{n \rightarrow \infty} P\left(\sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \leq x\right) = \Phi(x) \quad (6)$$

für jedes $x \in \mathbb{R}$, wobei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung ist.

Beispiel

- Wir möchten die Wahrscheinlichkeit $P(|\bar{X}_n - \mu| > \varepsilon)$, dass das Stichprobenmittel \bar{X}_n um mehr als einen vorgegebenen Schwellenwert $\varepsilon > 0$ von dem zu schätzenden Wert μ abweicht, approximieren. Es gilt
-

$$\begin{aligned} P(|\bar{X}_n - \mu| > \varepsilon) &= 1 - P(|\bar{X}_n - \mu| \leq \varepsilon) \\ &= 1 - P(-\varepsilon \leq \bar{X}_n - \mu \leq \varepsilon) \\ &= 1 - P\left(-\frac{\varepsilon\sqrt{n}}{\sigma} \leq \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \leq \frac{\varepsilon\sqrt{n}}{\sigma}\right) \\ &\stackrel{(6)}{\approx} 1 - \left(\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) - \Phi\left(-\frac{\varepsilon\sqrt{n}}{\sigma}\right)\right) \\ &= 2\left(1 - \Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)\right). \end{aligned}$$

Es gilt also

$$P(|\bar{X}_n - \mu| > \varepsilon) \approx 2\left(1 - \Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right)\right) \quad (7)$$

für große n .

- Falls σ^2 unbekannt ist, dann muß σ^2 (wie bereits oben erwähnt wurde) ebenfalls aus den vorliegenden Daten geschätzt werden, um mit Hilfe von (7) eine praktikable Näherungsformel für die Wahrscheinlichkeit $P(|\bar{X}_n - \mu| > \varepsilon)$ zu erhalten.

1.3 Stichprobenvarianz

Wir untersuchen nun die Frage, wie die Varianz $\sigma^2 = \text{Var } X_i$ der Stichprobenvariablen X_i aus den beobachteten Daten x_1, \dots, x_n bestimmt werden kann. Dabei gehen wir ähnlich wie in Abschnitt 1.2 vor.

- Wir betrachten die Stichprobenfunktion $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$\varphi(x_1, \dots, x_n) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2, \quad (8)$$

wobei \bar{x}_n in (1) gegeben ist.

- Die in (8) eingeführte Größe wird *Stichprobenvarianz* der (konkreten) Stichprobe (x_1, \dots, x_n) genannt und mit s_n^2 bezeichnet.

Definition 1.6 Die Zufallsvariable

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2, \quad (9)$$

heißt *Stichprobenvarianz* der Zufallsstichprobe (X_1, \dots, X_n) .

Theorem 1.7 Es gilt

$$\mathbb{E} S_n^2 = \sigma^2. \quad (10)$$

Falls $\mathbb{E}(X_i^4) < \infty$ für $i = 1, \dots, n$, dann gilt außerdem

$$\text{Var } S_n^2 = \frac{1}{n} \left(\mu_4 - \frac{n-3}{n-1} \sigma^4 \right), \quad (11)$$

wobei $\mu_4 = \mathbb{E}(X_i^4)$ und $\sigma^4 = (\text{Var } X_i)^2$ das 4-te Moment bzw. die quadrierte Varianz der Stichprobenvariablen X_i bezeichnen.

Beweis

- Wir zeigen nur die Gültigkeit von (10). Die Gültigkeit von (11) ergibt sich durch ähnliche Überlegungen.

- Aus der Definitionsgleichung (9) von S_n^2 ergibt sich, dass

$$\begin{aligned}
 S_n^2 &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \\
 &= \frac{1}{n-1} \sum_{i=1}^n (X_i^2 - 2X_i\bar{X}_n + \bar{X}_n^2) \\
 &= \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - 2n\bar{X}_n^2 + n\bar{X}_n^2 \right) \\
 &= \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}_n^2 \right).
 \end{aligned}$$

- Wegen der Linearität des Erwartungswertes folgt hieraus, dass

$$\begin{aligned}
 \mathbb{E} S_n^2 &= \frac{1}{n-1} \left(\sum_{i=1}^n \mathbb{E}(X_i^2) - n\mathbb{E}(\bar{X}_n^2) \right) \\
 &= \frac{1}{n-1} \left(\sum_{i=1}^n (\text{Var } X_i + (\mathbb{E} X_i)^2) - n(\text{Var } \bar{X}_n + (\mathbb{E} \bar{X}_n)^2) \right) \\
 &= \frac{1}{n-1} \left(n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right) \right) \\
 &= \sigma^2,
 \end{aligned}$$

wobei sich das vorletzte Gleichheitszeichen aus (3) und (4) ergibt.

- Damit ist (10) bewiesen.

Beachte

- Wegen (10) wird bei der Schätzung von σ^2 durch S_n^2 kein systematischer Fehler begangen.

Neben den Formeln (10) und (11) für Erwartungswert und Varianz der Stichprobenvarianz S_n^2 sind erneut weitere Aussagen über die Verteilung von S_n^2 bzw. über deren asymptotisches Verhalten für große n von Interesse.

Aus dem starken Gesetz der großen Zahlen bzw. aus dem zentralen Grenzwertsatz ergibt sich

Theorem 1.8 Es gilt

$$P\left(\lim_{n \rightarrow \infty} S_n^2 = \sigma^2\right) = 1. \quad (12)$$

Falls $\mathbb{E}(X_i^4) < \infty$ für $i = 1, \dots, n$, dann gilt außerdem

$$\lim_{n \rightarrow \infty} P\left(\sqrt{n} \frac{S_n^2 - \sigma^2}{\sqrt{\mu_4 - \sigma^4}} \leq x\right) = \Phi(x) \quad (13)$$

für jedes $x \in \mathbb{R}$, wobei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung ist.

Beweis

- Wir zeigen nur die Gültigkeit von (12). Die Herleitung von (13) erfordert Hilfsmittel, die über den Rahmen dieser einführenden Vorlesung hinausgehen.
- Im Beweis von Theorem 1.7 hatten wir gezeigt, dass

$$S_n^2 = \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}_n^2 \right). \quad (14)$$

- Außerdem kann man zeigen, dass die Unabhängigkeit und identische Verteiltheit der Stichprobenvariablen X_1, \dots, X_n impliziert, dass auch die Zufallsvariablen X_1^2, \dots, X_n^2 unabhängig und identisch verteilt sind.
- Deshalb ergibt sich aus dem starken Gesetz der großen Zahlen, dass mit Wahrscheinlichkeit 1

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{\infty} X_i^2 = \mathbb{E}(X_i^2)$$

und

$$\begin{aligned} \lim_{n \rightarrow \infty} \bar{X}_n^2 &= \lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=1}^{\infty} X_i \right)^2 \\ &= \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{\infty} X_i \right)^2 \\ &= (\mathbb{E} X_i)^2. \end{aligned}$$

- Hieraus und aus (14) ergibt sich (12).

1.4 Empirische Verteilungsfunktion

Außer der Schätzung von Erwartungswert μ und Varianz σ^2 der Stichprobenvariablen X_i kann auch deren Verteilungsfunktion F aus den vorliegenden Daten x_1, \dots, x_n geschätzt werden.

- Hierfür betrachten wir für jedes $x \in \mathbb{R}$ die Stichprobenfunktion $\varphi_x : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$\varphi_x(x_1, \dots, x_n) = \frac{1}{n} |\{i : 1 \leq i \leq n, x_i \leq x\}|,$$

wobei $|A|$ so wie bisher die Anzahl der Elemente der Menge A bezeichnet.

- D.h. $\varphi_x(x_1, \dots, x_n)$ ist die relative Häufigkeit derjenigen Stichprobenwerte, die den Schwellenwert x nicht überschreiten.

Beachte

- Man kann sich leicht überlegen, dass für jeden Vektor $(x_1, \dots, x_n) \in \mathbb{R}^n$ die Abbildung

$$x \rightarrow \varphi_x(x_1, \dots, x_n) \quad (15)$$

die Eigenschaften einer Verteilungsfunktion hat.

- Die in (15) gegebene Abbildung wird deshalb *empirische Verteilungsfunktion* der (konkreten) Stichprobe (x_1, \dots, x_n) genannt.

Dies führt zu der folgenden Begriffsbildung.

Definition 1.9 Die Abbildung $\hat{F}_n : \mathbb{R} \times \Omega \rightarrow [0, 1]$ mit

$$\hat{F}_n(x, \omega) = \frac{1}{n} |\{i : 1 \leq i \leq n, X_i(\omega) \leq x\}| \quad (16)$$

heißt *empirische Verteilungsfunktion* der Zufallsstichprobe (X_1, \dots, X_n) .

Beachte

- Die in (16) gegebene Abbildung kann man als eine Familie $\{\hat{F}_n(x), x \in \mathbb{R}\}$ von Zufallsvariablen $\hat{F}_n(x) : \Omega \rightarrow [0, 1]$ auffassen.
- Eine solche Familie von Zufallsvariablen wird *empirischer Prozeß* genannt.
- Empirische Prozesse sind eine spezielle Klasse *stochastischer Prozesse*.

Theorem 1.10 Für jedes $x \in \mathbb{R}$ gilt:

1. Die Zufallsvariable $n\hat{F}_n(x)$ ist binomialverteilt mit den Parametern n und $p = F(x)$.
D.h., für $k = 0, 1, \dots, n$ gilt

$$P(n\hat{F}_n(x) = k) = \binom{n}{k} (F(x))^k (1 - F(x))^{n-k}. \quad (17)$$

2. Insbesondere gilt also

$$\mathbb{E} \hat{F}_n(x) = F(x) \quad \text{und} \quad \text{Var} \hat{F}_n(x) = \frac{1}{n} F(x)(1 - F(x)). \quad (18)$$

- 3.

$$P\left(\lim_{n \rightarrow \infty} \hat{F}_n(x) = F(x)\right) = 1. \quad (19)$$

4. Falls $0 < F(x) < 1$, dann gilt außerdem für jedes $y \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} P\left(\sqrt{n} \frac{\hat{F}_n(x) - F(x)}{\sqrt{F(x)(1 - F(x))}} \leq y\right) = \Phi(y), \quad (20)$$

wobei $\Phi(y)$ die Verteilungsfunktion der Standardnormalverteilung ist.

Beweis

- Da $n\hat{F}_n(x) = |\{i : 1 \leq i \leq n, x_i \leq x\}| = \sum_{i=1}^n \mathbb{1}_{\{X_i \leq x\}}$, können wir die Zufallsvariable $n\hat{F}_n(x)$ als die Anzahl der Erfolge beim n -maligen Münzwurf mit den identischen Erfolgswahrscheinlichkeiten $a_1 = a_2 = \dots = a_n = F(x)$ auffassen. Hieraus ergibt sich (17).

- Damit ist auch (18) bewiesen, da

$$\mathbb{E}(\hat{F}_n(x)) = \frac{1}{n}\mathbb{E}(n\hat{F}_n(x)) = \frac{1}{n}nF(x) = F(x)$$

und

$$\text{Var}(\hat{F}_n(x)) = \frac{1}{n^2}\mathbb{E}(n\hat{F}_n(x)) = \frac{1}{n^2}nF(x)(1-F(x)) = \frac{1}{n}F(x)(1-F(x)).$$

- Da $n\hat{F}_n(x) = |\{i : 1 \leq i \leq n, x_i \leq x\}| = \sum_{i=1}^n \mathbb{1}_{\{X_i \leq x\}}$, Wegen der Darstellungsmöglichkeit von $n\hat{F}_n(x)$ als die Anzahl der Erfolge beim n -maligen Münzwurf mit identischen Erfolgswahrscheinlichkeiten ist $n\hat{F}_n(x)$ die Summe von n unabhängigen und identisch (Bernoulli-) verteilten Zufallsvariablen. Deshalb ergibt sich (19) aus dem starken Gesetz der großen Zahlen (vgl. Theorem 4.22).
- Mit der gleichen Begründung ergibt sich (20) unmittelbar aus dem zentralen Grenzwertsatz.

Beachte

- Neben der fast sicheren Konvergenz (19), die für jedes (einzelne) $x \in \mathbb{R}$ gilt, kann man zeigen, dass sich auch die Verteilungsfunktion F insgesamt durch den empirischen Prozeß $\{\hat{F}_n(x), x \in \mathbb{R}\}$ im Sinne der fast sicheren Konvergenz approximieren läßt.
- Dabei ist die folgende Verschärfung des starken Gesetzes der großen Zahlen gemeint, die in der Literatur *Satz von Glivenko/Cantelli* genannt wird.

Theorem 1.11 Sei

$$D_n = \sup_{x \in \mathbb{R}} |\hat{F}_n(x) - F(x)|. \quad (21)$$

Dann gilt

$$P\left(\lim_{n \rightarrow \infty} D_n = 0\right) = 1. \quad (22)$$

Beweis Der Beweis von Theorem 1.11 ist tiefgehend und geht über den Rahmen dieser einführenden Vorlesung hinaus.

Ähnlich wie beim zentralen Grenzwertsatz für Summen von unabhängigen und identisch verteilten Zufallsvariablen kann man zeigen, dass auch D_n bei entsprechend gewählter Normierung gegen einen nichtdeterministischen, d.h. zufälligen Grenzwert (im Sinne der Verteilungskonvergenz) strebt.

Dies ist die Aussage des folgenden Theorems, das *Satz von Kolmogorow/Smirnow* genannt wird.

Theorem 1.12 Falls die Verteilungsfunktion F der Stichprobenvariablen X_i eine stetige Funktion ist, dann gilt für $n \rightarrow \infty$

$$\sqrt{n}D_n \xrightarrow{d} D, \quad (23)$$

wobei D eine Zufallsvariable ist, deren Verteilungsfunktion gegeben ist durch

$$P(D \leq y) = \begin{cases} \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2 y^2) & \text{für } y \geq 0, \\ 0 & \text{für } y < 0. \end{cases} \quad (24)$$

Beweis Der Beweis von Theorem 1.12 ist tiefgehend und geht über den Rahmen dieser einführenden Vorlesung hinaus.

1.5 Dichteschätzer

Wir nehmen an, dass die (konkrete) Stichprobe (x_1, \dots, x_n) die Realisierung einer Zufallsstichprobe (X_1, \dots, X_n) sei, wobei X_1, \dots, X_n unabhängig und identisch verteilte (iid) Zufallsvariablen seien mit Verteilungsfunktion F und Dichte f . Neben der Schätzung der Verteilungsfunktion (Abschnitt 1.4) ist es von Interesse, die Dichte f anhand der Stichprobe (x_1, \dots, x_n) zu schätzen. Oft bieten Dichteschätzer (Histogramme oder Kerndichteschätzer) einen intuitiven Zugang, um einen Eindruck von der Form der Dichte zu gewinnen. Dieses ist hilfreich, um herauszufinden, wie die Stichprobenvariable X_i verteilt ist, z.B. ob es sich um eine Gammaverteilung oder eine Normalverteilung handelt.

Histogramm

Ein erster sehr einfacher Schätzer für die Dichte f ist das *Histogramm*, welches wir in diesem Abschnitt einführen. Sei (x_1, \dots, x_n) eine (konkrete) Stichprobe und $c_0 < c_1 < \dots < c_k \in \mathbb{R}$ eine Folge von Intervallgrenzen, so dass die Stichprobe in (c_0, c_k) enthalten ist, d. h. $x_i \in (c_0, c_k)$ für jedes $i = 1, \dots, n$. Falls die Abstände $c_i - c_{i-1}$ konstant sind, so heißt $h = c_i - c_{i-1}$ *Bandbreite*.

Die Idee ist es nun, die relative Häufigkeit von Stichprobenwerten x_i zu betrachten, die in die jeweiligen Intervalle $(c_{i-1}, c_i]$, $i = 1, \dots, k$ fallen. Sei also

$$n_i = \sum_{j=1}^n \mathbb{1}_{(c_{i-1}, c_i]}(x_j), \quad i = 1, \dots, k$$

die Anzahl an Stichprobenwerten im Intervall $(c_{i-1}, c_i]$. Normierung durch den Stichprobenumfang n ergibt die relativen Häufigkeiten

$$f_i = \frac{n_i}{n}, \quad i = 1, \dots, k.$$

Als *Histogramm* bezeichnet man die graphische Darstellung der relativen Häufigkeiten f_i , $i = 1, \dots, k$, wobei jedem Intervall $(c_{i-1}, c_i]$ ein Rechteck mit Flächeninhalt f_i zugeordnet wird, siehe Abbildung 1. Ein Histogramm kann als Dichteschätzer interpretiert werden, da die Dichte f durch \hat{f}_n geschätzt wird, wobei

$$\hat{f}_n(x) = \sum_{i=1}^k \frac{1}{c_{i-1} - c_i} f_i \mathbb{1}_{(c_{i-1}, c_i]}(x), \quad x \in \mathbb{R}.$$

Man sieht, dass \hat{f}_n die Eigenschaften einer Dichte besitzt, d.h. der Flächeninhalt der Dichte ergibt 1 und $\hat{f}_n(x) \geq 0$ für alle $x \in \mathbb{R}$.

Als Histogramm wird ebenso die graphische Darstellung von absoluten (und nicht relativen) Häufigkeitsverteilungen bezeichnet. Dann entspricht der Flächeninhalt dem Stichprobenumfang n .

Das Histogramm als Dichteschätzer hat zwei Nachteile. Zum einen ist die geschätzte Dichte \hat{f}_n stückweise konstant - auch bei stetigen Dichten f . Zweitens hat die Wahl der Bandbreite h (bzw. die Anzahl an Intervallen) einen starken Einfluss auf das Histogramm, siehe Abbildung 2. *Kerndichteschätzer* umgehen den ersten Nachteil (bei geeigneter Wahl des Kerns) und sollen nun betrachtet werden.

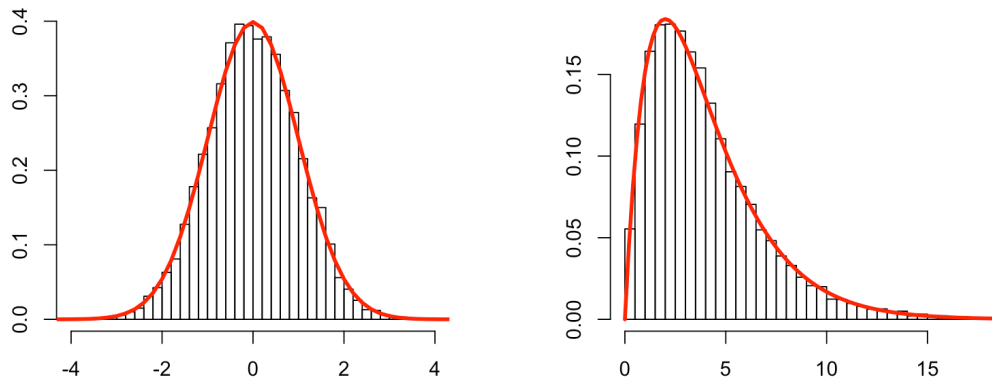


Figure 1: Histogramme von normalverteilten (links) und Gamma-verteilten (rechts) Stichproben mit einem Stichprobenumfang von $n = 10000$. Die Dichte ist als rote Kurve hinzugefügt.

Kerndichteschätzer

Definition 1.13 Ein *Kern* ist eine messbare Abbildung $K : \mathbb{R} \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

1. $K(x) \geq 0$ für alle $x \in \mathbb{R}$
2. $\int_{\mathbb{R}} K(x) dx = 1$

Beispiel Folgende Kerne werden häufig verwendet:

- Gausskern:

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right), \quad x \in \mathbb{R}$$

- Epanechnikov-Kern

$$K(x) = \begin{cases} \frac{3}{4}(1-x^2), & \text{falls } -1 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

- Bisquare-Kern

$$K(x) = \begin{cases} \frac{15}{16}(1-x^2)^2, & \text{falls } -1 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

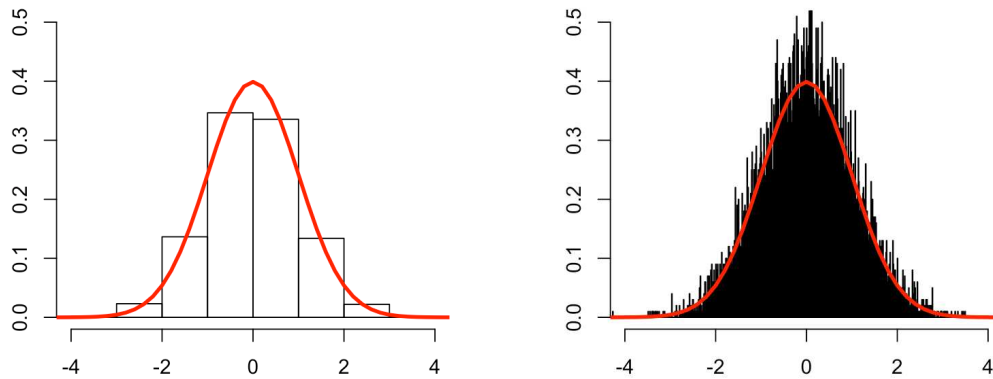


Figure 2: Histogramme von normalverteilten Stichproben (Stichprobenumfang $n = 10000$) mit einer zu gross gewählten Bandbreite (links) und einer zu klein gewählten Bandbreite (rechts). Die Dichte ist als rote Kurve hinzugefügt.

Definition 1.14 Sei K ein Kern, $h > 0$ ein Parameter (genannt Bandbreite) und (x_1, \dots, x_n) eine Stichprobe. Ein *Kerndichteschätzer* ist definiert durch

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad x \in \mathbb{R}.$$

Beachte

- Die Idee des Kerndichteschätzers besteht darin, jedem Stichprobenwert x_i den Anteil

$$\frac{1}{nh} K\left(\frac{x - x_i}{h}\right)$$

zuzuordnen und die geschätzte Dichte als Summe der einzelnen Beiträge zu definieren.

- Wie bei den Histogrammen spielt auch bei Kerndichteschätzern die Wahl der Bandbreite h eine entscheidende Rolle. Grundsätzlich gilt: Je größer die Bandbreite h , desto mehr werden die Daten ‘geglättet’, siehe Abbildung 3.

2 Schätzung von Parametern

In diesem Abschnitt treffen wir folgende Annahmen:

- die Verteilungsfunktion F der Stichprobenvariablen X_i gehört zu einer vorgegebenen (d.h. bekannten) *parametrischen Familie* von Verteilungsfunktionen $\{F_\theta, \theta \in \Theta\}$,

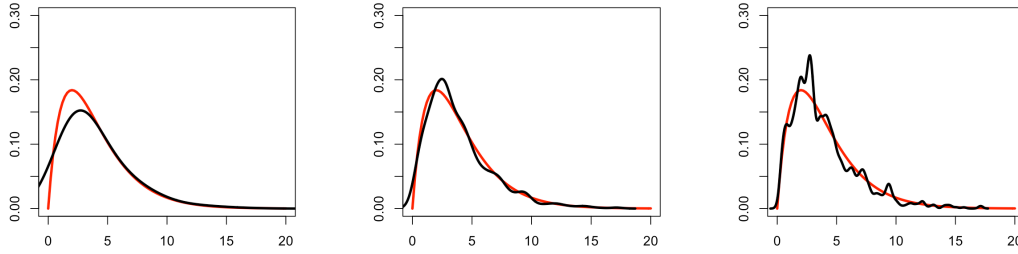


Figure 3: Kernaldichteschätzungen von Gamma-verteiltern Stichproben (Stichprobenumfang $n = 1000$) mit einer zu gross gewählten Bandbreite (links), einer ‘passenden’ Bandbreite (mittig) und einer zu klein gewählten Bandbreite (rechts). Die Dichte ist als rote Kurve hinzugefügt.

- wobei die Menge $\Theta \subset \mathbb{R}^m$ *Parameterraum* genannt wird und $m \in \mathbb{N}$ eine beliebige, jedoch vorgegebene natürliche Zahl ist.
- Mit anderen Worten: Es gelte $F = F_\theta$ für ein $\theta \in \Theta$, wobei jedoch der Parametervektor θ (bzw. ein Teil seiner Komponenten) unbekannt sei und aus den beobachteten Daten x_1, \dots, x_n geschätzt werden soll.
- Weiterhin sei der Parameterraum Θ Borel-messbar.
- Die Parametrisierung $\theta \rightarrow F_\theta$ sei identifizierbar, d.h. es gelte $F_{\theta_1} \neq F_{\theta_2}$ falls $\theta_1 \neq \theta_2$.

Wir nehmen an, dass der Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , über den die Stichprobenvariablen X_1, X_2, \dots definiert sind, der sogenannte *kanonische Wahrscheinlichkeitsraum* ist, Das heißt, wir setzen

- $\Omega = \mathbb{R} \times \mathbb{R} \times \dots = \{\omega : \omega = (\omega_1, \omega_2, \dots), \omega_i \in \mathbb{R} \ \forall i = 1, 2, \dots\}$
- und $\mathcal{F} = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \otimes \dots$,
- wobei das Wahrscheinlichkeitsmaß P gegeben ist durch:

$$P(\{\omega : \omega \in \Omega, \omega_{i_1} \leq x_{i_1}, \dots, \omega_{i_k} \leq x_{i_k}\}) = F_\theta(x_{i_1}) \cdot \dots \cdot F_\theta(x_{i_k}). \quad (25)$$

Die Stichprobenvariablen $X_i : \Omega \rightarrow \mathbb{R}$ sind gegeben durch $X_i(\omega) = \omega_i$, d.h. durch die Projektion auf die i -te Komponente ω_i von ω .

Beachte

- Das in (25) definierte Wahrscheinlichkeitsmaß P bezeichnen wir im Folgenden gelegentlich mit P_θ , um zu betonen, dass P_θ von dem (unbekannten) Parametervektor $\theta \in \Theta$ abhängt.
- Entsprechend verwenden wir die Bezeichnungen \mathbb{E}_θ und Var_θ für Erwartungswert bzw. Varianz.

- Falls keine Verwechslung möglich ist, dann bezeichnen wir auch die Verteilung einer (einzelnen) Stichprobenvariablen X_i mit P_θ , d.h. $P_\theta(X_i \leq x_i) = P_\theta((-\infty, x_i]) = F_\theta(x_i)$.
- Für jede Borel-messbare Funktion

$$\hat{\theta} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

wird der Zufallsvektor $\hat{\theta}(X_1, \dots, x_n)$ *Stichprobenfunktion* bzw. *Statistik* genannt.

- Bei Schätzung von Parametern heißt $\hat{\theta}(X_1, \dots, X_n)$ *Punktschätzer* von θ . Es wird in der Regel vorausgesetzt, dass $P_\theta(\hat{\theta}(X_1, \dots, X_n) \in \Theta) = 1$

Beispiel

- Die wichtigste parametrische Familie $\{P_\theta, \theta \in \Theta\}$ von Verteilungen, die in diesem Abschnitt betrachtet wird, ist die *Normalverteilungsfamilie*

$$\{P_\theta, \theta \in \Theta\} = \{N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\}. \quad (26)$$

In diesem Fall ist $m = 2$, $\Theta = \mathbb{R} \times (0, \infty)$ und $(\theta_1, \theta_2) = (\mu, \sigma^2)$.

- Eine andere parametrische Familie ist die *Binomialverteilungsfamilie*

$$\{P_\theta, \theta \in \Theta\} = \{Bin(n, p), n \in \mathbb{N}, p \in [0, 1]\}. \quad (27)$$

In diesem Fall ist $m = 2$, $\Theta = \mathbb{N} \times [0, 1]$ und $(\theta_1, \theta_2) = (n, p)$.

Beachte Weitere Beispiele von Familien parametrischer Verteilungen sind die

- Poisson-Verteilung,
- Exponentialverteilung,
- (diskrete bzw. stetige) Gleichverteilung.

2.1 Eigenschaften von Parameterschätzern

- Wir betrachten eine beliebige Stichprobenfunktion $\hat{\theta} : \mathbb{R}^n \rightarrow \Theta \subset \mathbb{R}^m$
- und den zugehörigen Schätzer $\hat{\theta}(X_1, \dots, X_n)$ des Parametervektors θ .

Definition 2.1

- Es gelte $\mathbb{E}_\theta |\hat{\theta}(X_1, \dots, X_n)| < \infty$ für jedes $\theta \in \Theta$, wobei $|\hat{\theta}(X_1, \dots, X_n)|$ die Norm des Zufallsvektors $\hat{\theta}(X_1, \dots, X_n)$ bezeichnet.
- Dann heißt der Schätzer $\hat{\theta}(X_1, \dots, X_n)$
 1. *erwartungstreu* oder *unbiased*, falls $\mathbb{E}_\theta \hat{\theta}(X_1, \dots, X_n) = \theta$ für jedes $\theta \in \Theta$,
 2. *asymptotisch erwartungstreu*, falls $\lim_{n \rightarrow \infty} \mathbb{E}_\theta \hat{\theta}(X_1, \dots, X_n) = \theta$ für jedes $\theta \in \Theta$.
 3. Die Differenz $\mathbb{E}_\theta \hat{\theta}(X_1, \dots, X_n) - \theta$ heißt *Bias* (Synonym: *Verzerrung*).

4. Falls $\mathbb{E}_\theta(|\hat{\theta}(X_1, \dots, X_n) - \theta|^2) < \infty$, dann heißt $\mathbb{E}_\theta(|\hat{\theta}(X_1, \dots, X_n) - \theta|^2)$ *MSE* (mean squared error) bzw. *erwartete quadratische Abweichung* des Schätzers $\hat{\theta}(X_1, \dots, X_n)$ von dem zu schätzenden Parametervektor $\theta \in \Theta \subset \mathbb{R}^m$.

Beispiel

- Die Stichprobenvariablen X_i seien normalverteilt, d.h. $X_i \sim N(\mu, \sigma^2)$.
- Aus (3) und (10) ergibt sich, dass der Schätzer (\bar{X}_n, S_n^2) des Parametervektors (μ, σ^2) erwartungstreu ist.

Definition 2.2 Der Schätzer $\hat{\theta}(X_1, \dots, X_n)$ des Parametervektors θ heißt

1. *schwach konsistent*, falls für $n \rightarrow \infty$

$$\hat{\theta}(X_1, \dots, X_n) \xrightarrow{\text{st}} \theta,$$

d.h.

$$\lim_{n \rightarrow \infty} P_\theta(|\hat{\theta}(X_1, \dots, X_n) - \theta| > \varepsilon) = 0 \quad \forall \varepsilon > 0$$

für jedes $\theta \in \Theta$.

2. *stark konsistent*, falls für $n \rightarrow \infty$

$$\hat{\theta}(X_1, \dots, X_n) \xrightarrow{\text{f.s.}} \theta,$$

d.h.

$$P_\theta\left(\lim_{n \rightarrow \infty} \hat{\theta}(X_1, \dots, X_n) = \theta\right) = 1$$

für jedes $\theta \in \Theta$.

Beispiel

- Die Stichprobenvariablen X_i seien normalverteilt, d.h. $X_i \sim N(\mu, \sigma^2)$.
- Dann ergibt sich aus (5) und (12), dass der Schätzer

$$\hat{\theta}(X_1, \dots, X_n) = (\bar{X}_n, S_n^2)$$

des Parametervektors $\theta = (\mu, \sigma^2)$ sowohl schwach als auch stark konsistent ist.

2.2 Momentenmethode

- Die Momentenmethode ist eines der ältesten Verfahren zur Gewinnung von Schätzern für die unbekanntenen Komponenten des Parametervektors $\theta = (\theta_1, \dots, \theta_m)$.
- Die Methode wurde von Karl Pearson (1857-1936) gegen Ende des 19. Jahrhunderts eingeführt und
- beruht auf dem Vergleich von Momenten der Stichprobenvariablen X_i, \dots, X_n mit den entsprechenden empirischen Momenten der konkreten Stichprobe (x_1, \dots, x_n) .

Modellannahmen

- Es gelte $\mathbb{E}_\theta(|X_i|^r) < \infty$ für jedes $\theta \in \Theta$ und für eine natürliche Zahl $r \geq m$.
- Ausserdem nehmen wir an, dass für jedes $k = 1, \dots, r$ das k -te Moment $\mu_k = \mathbb{E}_\theta(X_i^k)$ der Stichprobenvariablen X_i eine (bekannte) Funktion des Parametervektors $\theta = (\theta_1, \dots, \theta_m)$ ist.
- D.h., für jedes $k = 1, \dots, r$ gelte

$$\mu_k = g_k(\theta_1, \dots, \theta_m), \quad (28)$$

wobei $g_k : \mathbb{R}^m \rightarrow \mathbb{R}$ eine Borel-messbare Funktion sei.

Lösungsansatz

- Für jedes $k = 1, \dots, r$, sei

$$\hat{m}_k = \frac{1}{n} \sum_{i=1}^n x_i^k$$

das k -te empirische Moment.

- Löse das Gleichungssystem

$$\hat{m}_k = g_k(\theta_1, \dots, \theta_m), \quad k = 1, \dots, r \quad (29)$$

nach den Unbekannten $\theta_1, \dots, \theta_m$ auf, wobei

$$\hat{\theta}(x_1, \dots, x_n) = \left(\hat{\theta}_1(x_1, \dots, x_n), \dots, \hat{\theta}_m(x_1, \dots, x_n) \right) \quad (30)$$

die Lösung sei.

- Dabei wird vorausgesetzt, dass $\hat{\theta}(x_1, \dots, x_n)$ eindeutig bestimmt ist und dass die Abbildung

$$(x_1, \dots, x_n) \rightarrow \hat{\theta}(x_1, \dots, x_n)$$

eine Stichprobenfunktion ist, die den Stichprobenraum \mathbb{R}^n in den Parameterraum $\Theta \subset \mathbb{R}^m$ abbildet.

Definition 2.3 Der Zufallsvektor $\hat{\theta}(X_1, \dots, X_n)$ heißt *Momentenschätzer* (bzw. *M-Schätzer*) des Parametervektors θ , wobei in (30) die Zufallsstichprobe (X_1, \dots, X_n) anstelle der konkreten Stichprobe (x_1, \dots, x_n) eingesetzt wird.

Beachte

- Es gibt Beispiele parametrischer Verteilungsfamilien, so dass das Gleichungssystem (29) für $r = m$ keine eindeutig bestimmte Lösung besitzt. Dann ist $r > m$, d.h., die Anzahl der betrachteten Momente kann größer als die Anzahl der (unbekannten) Parameterkomponenten sein. In vielen Fällen ist jedoch das Gleichungssystem (29) für $r = m$ eindeutig lösbar.
- Mit Hilfe des starken Gesetzes der großen Zahlen kann die folgende Begründung für die Anwendung der Momentenmethode gegeben werden.

- Für das empirische Moment $\hat{m}_k(X_1, \dots, X_n)$ der Zufallsstichprobe (X_1, \dots, X_n) gilt

$$\hat{m}_k(X_1, \dots, X_n) \xrightarrow{\text{f.s.}} \mu_k, \quad (31)$$

falls $n \rightarrow \infty$.

- Falls die (28) gegebene Abbildung $g = (g_1, \dots, g_m) : \Theta \rightarrow C$ mit $C = g(\Theta) = \{g(\theta) : \theta \in \Theta\} \subset \mathbb{R}^m$ eineindeutig ist und falls die Umkehrabbildung $g^{-1} : C \rightarrow \Theta$ stetig ist, dann gilt wegen (31) für den M-Schätzer $\hat{\theta}(X_1, \dots, X_n)$, dass für jedes $\theta \in \Theta$ mit Wahrscheinlichkeit 1

$$\lim_{n \rightarrow \infty} \hat{\theta}(X_1, \dots, X_n) = \theta. \quad (32)$$

- D.h. der Schätzer $\hat{\theta}(X_1, \dots, X_n)$ ist stark konsistent.

Beispiel Die Stichprobenvariablen X_i seien normalverteilt, d.h., $X_i \sim N(\mu, \sigma^2)$. Dann ist

- $m = 2$ mit $(\theta_1, \theta_2) = (\mu, \sigma^2)$
- Außerdem ist

$$\mu_1 = \mu = g_1(\mu, \sigma^2) \quad \text{und} \quad \mu_2 = \sigma^2 + \mu^2 = g_2(\mu, \sigma^2).$$

- Das Gleichungssystem (29) hat also die Form

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n x_i &= \mu \\ \frac{1}{n} \sum_{i=1}^n x_i^2 &= \hat{\sigma}^2 + \hat{\mu}^2. \end{aligned}$$

- Durch Umstellen ergibt sich

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

und

$$\begin{aligned} \hat{\sigma}^2 &= \left(\frac{1}{n} \sum_{i=1}^n x_i^2 \right) - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2 \\ &\vdots \\ &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2. \end{aligned}$$

- D.h., bei normalverteilten Stichprobenvariablen X_i ergeben sich mit der Momentenmethode die Schätzer

$$\hat{\mu}(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i \quad (33)$$

und

$$\hat{\sigma}^2(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \quad (34)$$

für die Modellparameter μ bzw. σ^2 .

Beachte

- Aus den Theoremen 1.4 und 1.5 ergibt sich, dass der in (33) gewonnene Schätzer $\hat{\mu}(X_1, \dots, X_n)$ des Parameters μ erwartungstreu und konsistent ist.
- Aus den Theoremen 1.7 und 1.8 ergibt sich, dass der in (34) gewonnene Schätzer $\hat{\sigma}^2(X_1, \dots, X_n)$ des Parameters σ^2 asymptotisch erwartungstreu und konsistent ist.
- Wegen des im Vergleich zu S_n^2 modifizierten Normierungsfaktors $1/n$ ist $\hat{\sigma}^2(X_1, \dots, X_n)$ jedoch *nicht* erwartungstreu.
- Somit ist auch der Schätzer $(\hat{\mu}(X_1, \dots, X_n), \hat{\sigma}^2(X_1, \dots, X_n))$ des Parametervektors (μ, σ^2) lediglich asymptotisch erwartungstreu und konsistent.

Beispiel Die Stichprobenvariablen X_i seien Gamma-verteilt, d.h., $X_i \sim \Gamma(b, p)$ mit $b > 0, p > 0$. Die Dichte der Gammaverteilung ist gegeben durch

$$f_{b,p}(x) = \begin{cases} \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} & \text{für } x > 0, \\ 0 & \text{sonst.} \end{cases}$$

- Dann ist $m = 2$ mit $(\theta_1, \theta_2) = (b, p)$.
- Man kann ausrechnen, dass

$$\mu_1 = \frac{p}{b}, \quad \mu_2 = \frac{p(p+1)}{b^2}.$$

- Hieraus ergibt sich das Gleichungssystem

$$\hat{m}_1 = \frac{p}{b}, \quad \hat{m}_2 = \frac{p(p+1)}{b^2}$$

mit der Lösung

$$\hat{b} = \frac{\hat{m}_1}{\hat{m}_2 - (\hat{m}_1)^2}, \quad \hat{p} = \frac{(\hat{m}_1)^2}{\hat{m}_2 - (\hat{m}_1)^2}$$

2.3 Maximum-Likelihood-Methode

- Eine andere Methode zur Gewinnung von Schätzern für die unbekanntenen Komponenten des Parametervektors $\theta = (\theta_1, \dots, \theta_m)$ ist die Maximum-Likelihood-Methode.
- Genauso wie bei der Momentenmethode wird auch bei der Maximum-Likelihood-Methode das Ziel verfolgt, θ so zu schätzen, dass eine möglichst gute Anpassung der Modellcharakteristiken P_θ bzw. F_θ an die beobachteten Daten erreicht wird.
- Dabei wird bei der Maximum-Likelihood-Methode die Wahrscheinlichkeit für das Eintreten bestimmter (infinitesimaler) Ereignisse maximiert.
- Als der ‘Erfinder’ der Maximum-Likelihood-Methode gilt Sir Ronald Aylmer Fisher (1890–1962).

Wir betrachten nur die beiden (grundlegenden) Fälle, dass die Stichprobenvariablen X_1, \dots, X_n entweder diskret oder absolutstetig sind. D.h., für jedes $\theta \in \Theta$ gelte entweder

- $P_\theta(X_i \in C) = 1$ für eine abzählbare Menge $C \subset \mathbb{R}$,
- wobei wir mit $\{p(x; \theta), x \in C\}$ die Wahrscheinlichkeitsfunktion von X_i bezeichnen; $p(x; \theta) = P_\theta(X_i = x)$

oder

- $F_\theta(x) = \int_{-\infty}^x f(y; \theta) dy$ für jedes $x \in \mathbb{R}$,
- wobei $f(\cdot; \theta)$ die Dichte von X_i ist.

Definition 2.4 Die Abbildung $L : \mathbb{R}^n \times \Theta \rightarrow [0, \infty)$ sei durch die folgende Vorschrift gegeben.

- Falls X_i diskret ist, dann gelte

$$L(x_1, \dots, x_n; \theta) = p(x_1; \theta) \dots p(x_n; \theta) \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n, \theta \in \Theta. \quad (35)$$

- Falls X_i absolutstetig ist, dann gelte

$$L(x_1, \dots, x_n; \theta) = f(x_1; \theta) \dots f(x_n; \theta) \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n, \theta \in \Theta. \quad (36)$$

Für jeden Vektor $(x_1, \dots, x_n) \in \mathbb{R}^n$ heißt die Abbildung $\theta \rightarrow L(x_1, \dots, x_n; \theta)$ die *Likelihood-Funktion* der Stichprobe (x_1, \dots, x_n) .

Die Idee der Maximum-Likelihood-Methode besteht nun darin, für jede (konkrete) Stichprobe (x_1, \dots, x_n) einen Parametervektor $\theta \in \Theta$ zu bestimmen, so dass der Wert $L(x_1, \dots, x_n; \theta)$ der Likelihood-Funktion maximiert wird. Etwas vereinfacht bedeutet dies, dass der Parameter θ so gewählt wird, dass die (konkrete) Stichprobe (x_1, \dots, x_n) möglichst 'plausibel' erscheint. Betrachten wir den Fall, dass die Stichprobenvariablen X_1, \dots, X_n diskret sind. Dann ist der Wert der Likelihood-Funktion $L(x_1, \dots, x_n; \theta) = p(x_1; \theta) \dots p(x_n; \theta)$ gerade die Wahrscheinlichkeit, dass bei festem θ die (konkrete) Stichprobe (x_1, \dots, x_n) gezogen wird. Im diskreten Fall wird also die Wahrscheinlichkeit, die (konkrete) Stichprobe (x_1, \dots, x_n) zu ziehen, maximiert.

Definition 2.5 Sei $\hat{\theta} : \mathbb{R}^n \rightarrow \Theta \subset \mathbb{R}^m$ eine Stichprobenfunktion mit

$$L(x_1, \dots, x_n; \theta) \leq L(x_1, \dots, x_n; \hat{\theta}(x_1, \dots, x_n)) \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n, \theta \in \Theta. \quad (37)$$

Der Zufallsvektor $\hat{\theta}(X_1, \dots, X_n)$ wird dann *Maximum-Likelihood-Schätzer (ML-Schätzer)* von θ genannt.

Beachte

- Es gibt Beispiele parametrischer Verteilungsfamilien, so dass der mittels (37) definierte Maximum-Likelihood-Schätzer $\hat{\theta}(X_1, \dots, X_n)$ nicht eindeutig bestimmt ist.
- In vielen Fällen lässt sich das Optimierungsproblem (37) jedoch durch Differenzieren eindeutig lösen.
- Anstelle (37) wird dabei oft die (äquivalente) Bedingung

$$\log L(x_1, \dots, x_n; \theta) \leq \log L(x_1, \dots, x_n; \hat{\theta}(x_1, \dots, x_n)), \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n, \theta \in \Theta \quad (38)$$

betrachtet.

- Für jeden Vektor $(x_1, \dots, x_n) \in \mathbb{R}^n$ wird die Abbildung $\theta \rightarrow \log L(x_1, \dots, x_n; \theta)$ die *Loglikelihood-Funktion* der Stichprobe (x_1, \dots, x_n) genannt.
- Sie bietet den Vorteil, dass beim Übergang zur Loglikelihood-Funktion die Produkte in (35) bzw. (36) in (einfacher handhabbare) Summen übergehen: Damit wird

$$L(x_1, \dots, x_n; \theta) = p(x_1; \theta) \dots p(x_n; \theta) \quad \text{bzw.} \quad L(x_1, \dots, x_n; \theta) = f(x_1; \theta) \dots f(x_n; \theta)$$

zu

$$\log L(x_1, \dots, x_n; \theta) = \sum_{i=1}^n \log p(x_i; \theta) \quad \text{bzw.} \quad \log L(x_1, \dots, x_n; \theta) = \sum_{i=1}^n \log f(x_i; \theta).$$

Die notwendige Bedingung zur Bestimmung eines Maximum lautet dann:

$$\frac{\partial \log L(x_1, \dots, x_n; \theta)}{\partial \theta_i} = 0 \quad \forall i = 1, \dots, m.$$

Beispiele

1. Wer war der Absender?

- Eine Warenlieferung eines unbekanntes Herstellers bestehe aus 12 Exemplaren eines Artikels.
- Dabei sei festgestellt worden, dass eines der 12 Exemplare Ausschuss ist.
- Es sei bekannt, dass nur drei potentiell mögliche Hersteller in Frage kommen und dass deren Lieferungen erfahrungsgemäss jeweils einen Ausschussanteil von $\theta_1 = 0,05$, $\theta_2 = 0,10$ bzw. $\theta_3 = 0,15$ aufweisen.
- *Frage:* Welcher der drei Hersteller war vermutlich der Absender der Warenlieferung?
- *Modell:* Betrachten die Stichprobenvariablen X_1, \dots, X_{12} mit

$$X_i(\omega) = \begin{cases} 1, & \text{falls das } i\text{-te Exemplar der Lieferung Ausschuss ist} \\ 0 & \text{sonst} \end{cases}$$

und die Familie der drei Bernoulli-Verteilungen $\{\text{Bin}(1, \theta_1), \text{Bin}(1, \theta_2), \text{Bin}(1, \theta_3)\}$, d.h. $m = 1$ und $\Theta = \{\theta_1, \theta_2, \theta_3\}$.

- *Lösung:* Die Stichprobenfunktion

$$\hat{\theta} : \{0, 1\}^{12} \rightarrow \{\theta_1, \theta_2, \theta_3\}$$

wird so gewählt, dass für jeden Vektor $(x_1, \dots, x_{12}) \in \{0, 1\}^{12}$ mit $|\{i : x_i = 1\}| = 1$ die Wahrscheinlichkeit

$$P_\theta((X_1, \dots, X_{12}) = (x_1, \dots, x_{12})) = \theta(1 - \theta)^{11}$$

maximal ist.

- Es gilt

θ	$P_\theta((X_1, \dots, X_{12}) = (x_1, \dots, x_{12}))$
0,05	0,028
0,10	0,031
0,15	0,025

- Das Maximum 0,031 steht in der zweiten Zeile dieser Tabelle.
- Also ist $\hat{\theta}(x_1, \dots, x_{12}) = \theta_2$, d.h., der Hersteller mit dem Ausschussanteil $\theta_2 = 0.10$ war vermutlich der Absender der Lieferung.

2. Bernoulli-Verteilung

- Betrachten die Familie $\{P_\theta, \theta \in \Theta\} = \{\text{Bin}(1, p), p \in [0, 1]\}$ der Bernoulli-Verteilungen.
- Dann gilt

$$p(x; p) = \begin{cases} p^x(1-p)^{1-x}, & \text{falls } x \in \{0, 1\} \\ 0 & \text{sonst} \end{cases}$$

- Die Likelihood-Funktion L ist also gegeben durch

$$L(x_1, \dots, x_n; p) = \begin{cases} \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i}, & \text{falls } (x_1, \dots, x_n) \in \{0, 1\}^n \\ 0 & \text{sonst} \end{cases}$$

- Falls $x_1 = \dots = x_n = 0$ bzw. $x_1 = \dots = x_n = n$, dann sieht man leicht, dass die Abbildung $p \rightarrow L(x_1, \dots, x_n; p)$ an der Stelle $p = 0$ bzw. $p = 1$ ein (eindeutig bestimmtes) Maximum hat.
- Sei nun $(x_1, \dots, x_n) \in \{0, 1\}^n$ mit $0 < \sum_{i=1}^n x_i < n$. Dann ist

$$p \rightarrow \log L(x_1, \dots, x_n; p) = \left(\sum_{i=1}^n x_i \right) \log p + \left(n - \sum_{i=1}^n x_i \right) \log(1-p)$$

eine stetige Funktion im Intervall $(0, 1)$, und es gilt

$$\lim_{p \rightarrow 0} \log L(x_1, \dots, x_n; p) = -\infty \quad \text{bzw.} \quad \lim_{p \rightarrow 1} \log L(x_1, \dots, x_n; p) = -\infty.$$

- Die Abbildung $p \rightarrow \log L(x_1, \dots, x_n; p)$ hat also ein Maximum im Intervall $(0, 1)$.
- Durch Differenzieren nach p ergibt sich

$$\frac{\partial \log L(x_1, \dots, x_n; p)}{\partial p} = \left(\sum_{i=1}^n x_i \right) \frac{1}{p} - \left(n - \sum_{i=1}^n x_i \right) \frac{1}{1-p}.$$

- Weil die Gleichung

$$\left(\sum_{i=1}^n x_i \right) \frac{1}{p} - \left(n - \sum_{i=1}^n x_i \right) \frac{1}{1-p} = 0$$

die (eindeutig bestimmte) Lösung

$$\hat{p}(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i \quad \left(= \bar{x}_n \right)$$

hat, nimmt die Abbildung $p \rightarrow \log L(x_1, \dots, x_n; p)$ an der Stelle $p = \bar{x}_n$ ihr Maximum an.

- Also ist der Maximum-Likelihood-Schätzer für den Parameter p gegeben durch

$$\hat{p}(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i \quad \left(= \bar{X}_n \right).$$

3. Gleichverteilung

- Betrachten die Familie $\{P_\theta, \theta \in \Theta\} = \{[0, b], b > 0\}$ von Gleichverteilungen.
- Dann gilt

$$f(x; b) = \begin{cases} \frac{1}{b}, & \text{falls } 0 \leq x \leq b \\ 0 & \text{sonst} \end{cases}$$

- Die Likelihood-Funktion L ist somit gegeben durch

$$L(x_1, \dots, x_n; b) = \begin{cases} \frac{1}{b^n}, & \text{falls } 0 \leq x_1, \dots, x_n \leq b \\ 0 & \text{sonst} \end{cases}$$

- Weil die Abbildung $b \rightarrow L(x_1, \dots, x_n; b)$ monoton fallend ist für $b > \max\{x_1, \dots, x_n\} \geq 0$, ergibt sich der Maximum-Likelihood-Schätzer

$$\hat{b}(X_1, \dots, X_n) = \max\{X_1, \dots, X_n\}$$

für den Parameter b .

3 Konfidenzintervalle

3.1 Modellbeschreibung

In diesem Abschnitt setzen wir (so wie in Abschnitt 2) voraus, dass

- die Verteilungsfunktion F der Stichprobenvariablen X_i zu einer vorgegebenen *parametrischen Familie* von Verteilungsfunktionen $\{F_\theta, \theta \in \Theta\}$ gehört; $\Theta \subset \mathbb{R}^m$.

Dabei nehmen wir (zur Vereinfachung der Darlegungen) an, dass

- jeweils nur eine einzelne Komponente θ_j des Parametervektors $\theta = (\theta_1, \dots, \theta_m)$ aus den beobachteten Daten x_1, \dots, x_n geschätzt werden soll; $j \in \{1, \dots, m\}$.

Genauso wie in Abschnitt 2 nehmen wir an, dass

- der Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , über dem die Stichprobenvariablen X_1, X_2, \dots definiert sind, der kanonische Wahrscheinlichkeitsraum dieser Zufallsvariablen ist.

Anstelle eine (einzelne) Stichprobenfunktion zu betrachten, wie wir es in den Abschnitten 1 und 2 tun, betrachten wir nun

- zwei Stichprobenfunktionen $\underline{\theta} : \mathbb{R}^n \rightarrow \mathbb{R}$ und $\bar{\theta} : \mathbb{R}^n \rightarrow \mathbb{R}$, so dass

$$\underline{\theta}(x_1, \dots, x_n) \leq \bar{\theta}(x_1, \dots, x_n) \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n. \quad (39)$$

Definition 3.1 Sei $\gamma \in (0, 1)$ eine beliebige, jedoch fest vorgegebene Zahl. Dann heißt das zufällige Intervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ *Konfidenzintervall* für θ_j zum Niveau γ , falls

$$P_\theta(\underline{\theta}(X_1, \dots, X_n) \leq \theta_j \leq \bar{\theta}(X_1, \dots, X_n)) \geq \gamma \quad (40)$$

für jedes $\theta \in \Theta$.

Beachte

- Die Stichprobenfunktionen $\underline{\theta}$ und $\bar{\theta}$ sind nicht eindeutig bestimmt.
- Sie sollten so gewählt werden, dass das Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ bei gegebenem Niveau γ möglichst kurz ist.
- Weil das Konfidenzintervall möglichst kurz sein soll, wird anstelle von (40) manchmal auch die folgende Bedingung betrachtet:

$$\inf_{\theta \in \Theta} P_\theta(\underline{\theta}(X_1, \dots, X_n) \leq \theta_j \leq \bar{\theta}(X_1, \dots, X_n)) = \gamma. \quad (41)$$

- Wenn (bei vorgegebenem θ) die Verteilung der Zufallsvariablen $\underline{\theta}(X_1, \dots, X_n)$ und $\bar{\theta}(X_1, \dots, X_n)$ schwierig handhabbar bzw. unbekannt ist, d.h., wenn es nicht möglich ist, die Gültigkeit von (40) bzw. (41) nachzuweisen, dann kann beispielsweise die folgende Bedingung betrachtet werden:

$$\lim_{n \rightarrow \infty} P_\theta(\underline{\theta}(X_1, \dots, X_n) \leq \theta_j \leq \bar{\theta}(X_1, \dots, X_n)) \geq \gamma \quad (42)$$

für jedes $\theta \in \Theta$.

- Falls (42) gilt, dann nennt man $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ *Konfidenzintervall* für θ_j zum asymptotischen Niveau γ (bzw. kurz *asymptotisches Konfidenzintervall*).
- Übliche Werte für γ sind $\gamma = 0.99, \gamma = 0.95$ oder $\gamma = 0.9$.

Die *praktische Berechnung* eines (konkreten) Konfidenzintervalls $(\underline{\theta}(x_1, \dots, x_n), \bar{\theta}(x_1, \dots, x_n))$ für θ_j auf der Basis einer (konkreten) Stichprobe (x_1, \dots, x_n) besteht aus den folgenden Schritten:

1. Bestimmen zwei Stichprobenfunktionen $\underline{\theta} : \mathbb{R}^n \rightarrow \mathbb{R}$ und $\bar{\theta} : \mathbb{R}^n \rightarrow \mathbb{R}$, so dass
 - (39) gilt und
 - eine der Bedingungen (40)–(42) erfüllt ist.
2. Berechnen die Funktionswerte $\underline{\theta}(x_1, \dots, x_n)$ und $\bar{\theta}(x_1, \dots, x_n)$.

3.2 Konfidenzintervalle bei Normalverteilung; Quantilfunktion

In diesem Abschnitt nehmen wir an, dass die Stichprobenvariablen X_1, \dots, X_n normalverteilt sind, d.h., es gelte

$$\{P_\theta, \theta \in \Theta\} = \{N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\}.$$

Dabei benötigen wir die folgende

Definition 3.2

- Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine beliebige Verteilungsfunktion.
- Die Funktion $F^{-1} : [0, 1] \rightarrow \mathbb{R}$ mit

$$F^{-1}(y) = \inf\{x : F(x) \geq y\} \quad (43)$$

heißt *Quantilfunktion* von F .

- Sei $\gamma \in (0, 1)$. Die Zahl $F^{-1}(\gamma)$ wird dann γ -*Quantil* von F genannt.

Beachte

- Wenn $F : \mathbb{R} \rightarrow \mathbb{R}$ eine beliebige monoton wachsende rechtsstetige Funktion ist (die nicht unbedingt eine Verteilungsfunktion sein muß), dann heißt die in (43) definierte Funktion F^{-1} *verallgemeinerte inverse Funktion* von F .
- Quantilfunktionen sind also spezielle verallgemeinerte inverse Funktionen.

Bei der Bestimmung von Konfidenzintervallen für normalverteilte Stichprobenvariablen wird insbesondere die Quantilfunktion der Standardnormalverteilung benötigt.

Beachte

- Das γ -Quantil der Verteilungsfunktion Φ der Standardnormalverteilung wird mit z_γ bezeichnet.
- Mit anderen Worten: Für jedes $\gamma \in (0, 1)$ ist z_γ die Lösung der *Quantilgleichung* $\Phi(z_\gamma) = \gamma$.
- Für $\gamma \geq 0,50$ kann man das γ -Quantil z_γ aus Tabelle 1 entnehmen.
- Aus der Symmetrieeigenschaft $\Phi(-x) = 1 - \Phi(x)$ für jedes $x \in \mathbb{R}$ ergibt sich außerdem, dass

$$z_\gamma = -z_{1-\gamma} \quad (44)$$

für jedes $\gamma \in (0, 1)$.

Beispiel *Konfidenzintervall für den Erwartungswert μ (bei bekannter Varianz σ^2)*

- Wir nehmen an, dass $X_i \sim N(\mu, \sigma^2)$ für ein (unbekanntes) $\mu \in \mathbb{R}$ und ein (bekanntes) $\sigma^2 > 0$.
- Man kann leicht zeigen, dass $\sqrt{n}(\bar{X}_n - \mu)/\sigma \sim N(0, 1)$.

- Hieraus ergibt sich, dass für $\alpha \in (0, 1)$

$$P\left(z_{\alpha/2} \leq \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \leq z_{1-\alpha/2}\right) = 1 - \alpha.$$

da

$$\begin{aligned} P\left(z_{\alpha/2} \leq \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \leq z_{1-\alpha/2}\right) &= \Phi(z_{1-\alpha/2}) - \Phi(z_{\alpha/2}) \\ &= 1 - \frac{\alpha}{2} - \frac{\alpha}{2} \\ &= 1 - \alpha \end{aligned}$$

- Unter Berücksichtigung von (44) ergibt sich somit, dass

$$P\left(-z_{1-\alpha/2} \leq \sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \leq z_{1-\alpha/2}\right) = 1 - \alpha$$

bzw.

$$P\left(\bar{X}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

- Also ist mit

$$\underline{\theta}(X_1, \dots, X_n) = \bar{X}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \text{bzw.} \quad \bar{\theta}(X_1, \dots, X_n) = \bar{X}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \quad (45)$$

ein Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für μ zum Niveau $\gamma = 1 - \alpha$ gegeben.

- Für die Länge $\ell(X_1, \dots, X_n)$ dieses Konfidenzintervalls gilt

$$\begin{aligned} \ell(X_1, \dots, X_n) &= \bar{\theta}(X_1, \dots, X_n) - \underline{\theta}(X_1, \dots, X_n) \\ &= z_{1-\alpha/2} \frac{2\sigma}{\sqrt{n}}. \end{aligned}$$

- Hieraus ergibt sich insbesondere, dass $\ell(X_1, \dots, X_n)$ nicht vom Zufall abhängt.
- Außerdem erkennen wir, dass die Länge $\ell(X_1, \dots, X_n)$ des in (45) gegebenen Konfidenzintervalls klein ist, falls das Niveau $\gamma = 1 - \alpha$ klein, die Varianz σ^2 klein bzw. der Stichprobenumfang n groß ist.
- Man kann sich leicht überlegen, dass $\ell(X_1, \dots, X_n) \leq \varepsilon$ gilt, falls

$$n \geq \left(\frac{2\sigma z_{1-\alpha/2}}{\varepsilon}\right)^2, \quad (46)$$

wobei $\varepsilon > 0$ ein vorgegebener Schwellenwert ist.

Beachte

- Für eine (konkrete) Stichprobe (x_1, \dots, x_n) , die wir als Realisierung der Zufallsstichprobe (X_1, \dots, X_n) auffassen, ergibt sich nun das ‘konkrete’ Konfidenzintervall $(\underline{\theta}(x_1, \dots, x_n), \bar{\theta}(x_1, \dots, x_n))$ durch Einsetzen in (45).

- Falls beispielsweise für $\gamma = 0.95$ und $\sigma = 0.10$ ein solches Konfidenzintervall für die folgenden Daten 41.60, 41.48, 42.34, 41.95, 41.86, 42.18, 41.72, 42.26, 41.81, 42.04 ermittelt werden soll,
- dann entnehmen wir zunächst das Quantil $z_{0,975} = 1,96$ aus Tabelle 1
- und erhalten somit für $n = 10$

$$\begin{aligned}\underline{\theta}(x_1, \dots, x_n) &= \bar{x}_n - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 41.924 - 1.96 \frac{0.1}{\sqrt{10}} = 41.862 \\ \bar{\theta}(x_1, \dots, x_n) &= \bar{x}_n + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = 41.924 + 1.96 \frac{0.1}{\sqrt{10}} = 41.986\end{aligned}$$

- Somit ergibt sich das (konkrete) Konfidenzintervall (41.862, 41.986) für den Erwartungswert μ .

3.3 Statistische Prüfverteilungen

Außer der Quantilfunktion der Standardnormalverteilung werden bei der Bestimmung von Konfidenzintervallen bei normalverteilten Stichprobenvariablen die Quantilfunktionen weiterer Verteilungen benötigt.

Dabei werden insbesondere sogenannte *Prüfverteilungen* betrachtet, die wie folgt definiert sind.

Definition 3.3 Seien $r, s \in \mathbb{N}$ beliebige Zahlen und seien $X_0, X_1, \dots, X_r, \dots, X_{r+s}$ unabhängige und $N(0, 1)$ -verteilte Zufallsvariable. Dann heißt die Verteilung von

- $U_r = \sum_{i=1}^r X_i^2$ χ^2 -Verteilung mit r Freiheitsgraden (Schreibweise: $U_r \sim \chi_r^2$)
- $V_r = X_0 / \sqrt{\frac{1}{r} \sum_{i=1}^r X_i^2}$ t -Verteilung mit r Freiheitsgraden (Schreibweise: $V_r \sim t_r$)
- $W_{r,s} = \frac{1}{r} \sum_{i=1}^r X_i^2 / \frac{1}{s} \sum_{i=r+1}^{r+s} X_i^2$ F -Verteilung mit (r, s) Freiheitsgraden (Schreibweise: $W_{r,s} \sim F_{r,s}$)

Theorem 3.4 Sei $r \in \mathbb{N}$ eine beliebige Zahl. Dann sind die Dichten der Zufallsvariablen $U_r \sim \chi_r^2$ und $V_r \sim t_r$ gegeben durch

$$f_{U_r}(x) = \begin{cases} \frac{x^{(r-2)/2} e^{-x/2}}{2^{r/2} \Gamma(r/2)}, & \text{falls } x > 0 \\ 0 & \text{sonst} \end{cases} \quad (47)$$

bzw.

$$f_{V_r}(x) = \frac{\Gamma((r+1)/2)}{\Gamma(r/2)} \frac{1}{\sqrt{r\pi} (1+x^2/r)^{(r+1)/2}} \quad (48)$$

für jedes $x \in \mathbb{R}$, wobei $\Gamma : (0, \infty) \rightarrow (0, \infty)$ die Gammafunktion mit

$$\Gamma(z) = \int_0^{\infty} e^{-y} y^{z-1} dy, \quad z > 0 \quad (49)$$

bezeichnet; $\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(z+1) = z\Gamma(z)$.

Wir bestimmen nun die (gemeinsame) Verteilung des Stichprobenmittels \bar{X}_n und der Stichprobenvarianz S_n^2 bei normalverteilten Stichprobenvariablen X_1, \dots, X_n .

Theorem 3.5 Sei (X_1, \dots, X_n) eine normalverteilte Zufallsstichprobe mit $X_i \sim N(\mu, \sigma^2)$. Dann sind

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

unabhängige Zufallsvariablen, und es gilt

$$\bar{X}_n \sim N(\mu, \sigma^2/n) \quad \text{bzw.} \quad \frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2. \quad (50)$$

Beweis

- Der Nachweis der Unabhängigkeit von \bar{X}_n und S_n^2 erfordert Hilfsmittel, die über den Rahmen dieser einführenden Vorlesung hinausgehen.
- Andererseits kann man sich leicht überlegen, dass $\bar{X}_n \sim N(\mu, \sigma^2/n)$.
- Wir skizzieren hier lediglich eine Begründung für die Gültigkeit von

$$\frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2.$$

- Offenbar gilt die Identität

$$\begin{aligned} \sum_{i=1}^n (X_i - \mu)^2 &= \sum_{i=1}^n ((X_i - \bar{X}_n) + (\bar{X}_n - \mu))^2 \\ &= \sum_{i=1}^n (X_i - \bar{X}_n)^2 + 2(\bar{X}_n - \mu) \cdot \underbrace{\sum_{i=1}^n (X_i - \bar{X}_n)}_{=0} + n(\bar{X}_n - \mu)^2 \\ &= \sum_{i=1}^n (X_i - \bar{X}_n)^2 + n(\bar{X}_n - \mu)^2 \end{aligned}$$

- Hieraus ergibt sich, dass

$$\underbrace{\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2}_{\sim \chi_n^2} = \frac{\sum_{i=1}^n (X_i - \bar{X}_n)^2}{\sigma^2} + \underbrace{\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}\right)^2}_{\sim \chi_1^2}.$$

- Weil die beiden Summanden auf der rechten Seite dieser Gleichung unabhängig sind, ergibt sich nun aus der Definition der χ^2 -Verteilung, dass der erste Summand auf der rechten Seite χ_{n-1}^2 -verteilt ist.

Korollar 3.6 Sei (X_1, \dots, X_n) eine normalverteilte Zufallsstichprobe mit $X_i \sim N(\mu, \sigma^2)$. Dann gilt

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \sim t_{n-1}. \quad (51)$$

Beweis Weil die Zufallsvariablen \bar{X}_n und S_n^2 unabhängig sind (vgl. Theorem 3.5), kann man sich leicht überlegen, dass auch die Zufallsvariablen $\sqrt{n}(\frac{\bar{X}_n - \mu}{\sigma})$ und $\sqrt{\frac{1}{n-1} \frac{(n-1)S_n^2}{\sigma^2}}$ unabhängig sind. Deshalb ergibt sich unmittelbar aus der Definition 3.3 der t-Verteilung, dass

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} = \frac{\sqrt{n} \frac{\bar{X}_n - \mu}{\sigma}}{\sqrt{\frac{1}{n-1} \frac{(n-1)S_n^2}{\sigma^2}}} \sim t_{n-1}.$$

Beachte

- Das γ -Quantil der χ^2 -Verteilung mit r Freiheitsgraden wird mit $\chi_{r,\gamma}^2$ bezeichnet, vgl. Tabelle 2.
- Das γ -Quantil der t-Verteilung mit r Freiheitsgraden wird mit $t_{r,\gamma}$ bezeichnet, vgl. Tabelle 3.
- Analog zu der Symmetrieeigenschaft (44) der Quantile der Standardnormalverteilung gilt

$$t_{r,\gamma} = -t_{r,1-\gamma} \quad (52)$$

für beliebige $r \in \mathbb{N}$ und $\gamma \in (0, 1)$.

3.4 Weitere Konfidenzintervalle bei Normalverteilung

In diesem Abschnitt nehmen wir erneut an, dass die Stichprobenvariablen X_1, \dots, X_n normalverteilt sind, d.h., es gelte

$$\{P_\theta, \theta \in \Theta\} = \{N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\}.$$

Wir diskutieren die Herleitung von Konfidenzintervallen für μ bzw. σ^2 .

Beispiel *Konfidenzintervall für den Erwartungswert μ (bei unbekannter Varianz σ^2)*

- Im Unterschied zu dem in Abschnitt 3.2 diskutierten Beispiel nehmen wir nun an, dass $X_i \sim N(\mu, \sigma^2)$ für ein (unbekanntes) $\mu \in \mathbb{R}$ und ein (unbekanntes) $\sigma^2 > 0$.
- Das in (45) konstruierte Konfidenzintervall für μ ist jetzt nicht geeignet, weil in (45) die unbekannte Größe σ vorkommt.
- Wir nutzen vielmehr das Ergebnis von Korollar 3.6 und erkennen, dass für jedes $\alpha \in (0, 1)$

$$P\left(t_{n-1,\alpha/2} \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \leq t_{n-1,1-\alpha/2}\right) = 1 - \alpha \quad (53)$$

bzw. wegen (52)

$$P\left(-t_{n-1,1-\alpha/2} \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \leq t_{n-1,1-\alpha/2}\right) = 1 - \alpha$$

- Diese Gleichung wird nun so umgeformt, dass ein Konfidenzintervall für μ entsteht:

$$\begin{aligned}
 & P\left(-t_{n-1,1-\alpha/2} \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \leq t_{n-1,1-\alpha/2}\right) = 1 - \alpha \\
 \Leftrightarrow & P\left(-t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}} \leq \bar{X}_n - \mu \leq t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}}\right) = 1 - \alpha \\
 \Leftrightarrow & P\left(-t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}} - \bar{X}_n \leq -\mu \leq t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}} - \bar{X}_n\right) = 1 - \alpha \\
 \Leftrightarrow & P\left(\bar{X}_n - t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}} \leq \mu \leq \bar{X}_n + t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}}\right) = 1 - \alpha.
 \end{aligned}$$

D.h.

$$P\left(\bar{X}_n - t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}} \leq \mu \leq \bar{X}_n + t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}}\right) = 1 - \alpha.$$

- Also ist mit

$$\underline{\theta}(X_1, \dots, X_n) = \bar{X}_n - t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}} \quad \text{bzw.} \quad \bar{\theta}(X_1, \dots, X_n) = \bar{X}_n + t_{n-1,1-\alpha/2} \frac{S_n}{\sqrt{n}} \quad (54)$$

ein (symmetrisches) Konfidenzintervall für μ zum Niveau $\gamma = 1 - \alpha$ gegeben.

Beachte

- Dieses Beispiel (und auch die anderen Beispiele) kann dahingehend verallgemeinert werden, dass anstelle von (53) die folgende Gleichung für beliebige $\alpha_1, \alpha_2 \in [0, 1/2)$ mit $\alpha_1 + \alpha_2 = \alpha \in (0, 1)$ betrachtet wird:

$$P\left(t_{n-1,\alpha_2} \leq \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n} \leq t_{n-1,1-\alpha_1}\right) = 1 - \alpha. \quad (55)$$

- Hieraus ergibt sich dann das (asymmetrische) Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für μ zum Niveau $\gamma = 1 - \alpha$ mit

$$\underline{\theta}(X_1, \dots, X_n) = \bar{X}_n - t_{n-1,1-\alpha_1} \frac{S_n}{\sqrt{n}} \quad \text{bzw.} \quad \bar{\theta}(X_1, \dots, X_n) = \bar{X}_n + t_{n-1,1-\alpha_2} \frac{S_n}{\sqrt{n}}. \quad (56)$$

- Für $\alpha_1 = 0$ ergibt sich insbesondere das sogenannte *einseitige Konfidenzintervall* $(-\infty, \bar{\theta}(X_1, \dots, X_n))$ für μ zum Niveau $\gamma = 1 - \alpha$ mit

$$\bar{\theta}(X_1, \dots, X_n) = \bar{X}_n + t_{n-1,1-\alpha} \frac{S_n}{\sqrt{n}}.$$

- Analog ergibt sich für $\alpha_2 = 0$ das einseitige Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \infty)$ für μ zum Niveau $\gamma = 1 - \alpha$ mit

$$\underline{\theta}(X_1, \dots, X_n) = \bar{X}_n - t_{n-1,1-\alpha} \frac{S_n}{\sqrt{n}}.$$

Ähnliche Überlegungen führen nun zu Konfidenzintervallen für σ^2 bei bekanntem bzw. unbekanntem μ .

Beispiel Konfidenzintervall für die Varianz σ^2 (bei bekanntem Erwartungswert μ)

- Weil $(X_i - \mu)/\sigma \sim N(0, 1)$, ergibt sich aus der Definition 3.3 der χ^2 -Verteilung, dass $n\tilde{S}_n^2/\sigma^2$ eine χ^2 -verteilte Zufallsvariable ist mit n Freiheitsgraden, wobei

$$\tilde{S}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2.$$

- Für beliebige $\alpha_1, \alpha_2 \in [0, 1/2)$ mit $\alpha_1 + \alpha_2 = \alpha \in (0, 1)$ gilt also

$$P\left(\chi_{n, \alpha_2}^2 \leq \frac{n\tilde{S}_n^2}{\sigma^2} \leq \chi_{n, 1-\alpha_1}^2\right) = 1 - \alpha. \quad (57)$$

- Dies ergibt dann das Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für σ^2 zum Niveau $\gamma = 1 - \alpha$ mit

$$\underline{\theta}(X_1, \dots, X_n) = \frac{n\tilde{S}_n^2}{\chi_{n, 1-\alpha_1}^2} \quad \text{bzw.} \quad \bar{\theta}(X_1, \dots, X_n) = \frac{n\tilde{S}_n^2}{\chi_{n, \alpha_2}^2}. \quad (58)$$

Beispiel Konfidenzintervall für die Varianz σ^2 (bei unbekanntem Erwartungswert μ)

- Aus Theorem 3.5 ergibt sich, dass für beliebige $\alpha_1, \alpha_2 \in [0, 1/2)$ mit $\alpha_1 + \alpha_2 = \alpha \in (0, 1)$

$$P\left(\chi_{n-1, \alpha_2}^2 \leq \frac{(n-1)S_n^2}{\sigma^2} \leq \chi_{n-1, 1-\alpha_1}^2\right) = 1 - \alpha. \quad (59)$$

- Dies ergibt das Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für σ^2 zum Niveau $\gamma = 1 - \alpha$ mit

$$\underline{\theta}(X_1, \dots, X_n) = \frac{(n-1)S_n^2}{\chi_{n-1, 1-\alpha_1}^2} \quad \text{bzw.} \quad \bar{\theta}(X_1, \dots, X_n) = \frac{(n-1)S_n^2}{\chi_{n-1, \alpha_2}^2}. \quad (60)$$

3.5 Zwei-Stichproben-Probleme

In Verallgemeinerung der Situation, die bisher betrachtet wurde, sollen nun gleichzeitig zwei Datensätze $(x_{11}, \dots, x_{1n_1})$ und $(x_{21}, \dots, x_{2n_2})$ stochastisch modelliert werden, wobei die Stichprobenumfänge $n_1, n_2 \in \mathbb{N}$ beliebige Zahlen sind. Dabei nehmen wir an, dass die (konkreten) Stichproben $(x_{11}, \dots, x_{1n_1})$ und $(x_{21}, \dots, x_{2n_2})$ Realisierungen von zwei Zufallsstichproben $(X_{11}, \dots, X_{1n_1})$ bzw. $(X_{21}, \dots, X_{2n_2})$ sind.

Wir betrachten nun also zwei (unendliche) Folgen X_{11}, X_{12}, \dots und X_{21}, X_{22}, \dots von Zufallsvariablen und nehmen an, dass

- die (zweidimensionalen) Zufallsvektoren X_1, X_2, \dots mit $X_i = (X_{1i}, X_{2i})$ unabhängig und identisch verteilt sind,

- die (gemeinsame) Verteilungsfunktion F von X_i zu einer vorgegebenen parametrischen Familie von Verteilungsfunktionen $\{F_\theta, \theta \in \Theta\}$ gehört, $\Theta \subset \mathbb{R}^m$,
- der Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , über dem die Zufallsvektoren X_1, X_2, \dots definiert sind, der kanonische Wahrscheinlichkeitsraum dieser Zufallsvektoren ist.

Beachte Die Komponenten X_{1i} und X_{2i} von X_i müssen jedoch im allgemeinen *weder* unabhängig *noch* identisch verteilt sind.

Im Rahmen dieser Vorlesung betrachten wir lediglich den Fall, dass

- ein (nichtvektorieller) Funktionswert $g(\theta) \in \mathbb{R}$ des Parametervektors $\theta = (\theta_1, \dots, \theta_m)$ aus den beobachteten Daten $(x_{11}, \dots, x_{1n_1})$ bzw. $(x_{21}, \dots, x_{2n_2})$ geschätzt werden soll, wobei $g : \Theta \rightarrow \mathbb{R}$ eine vorgegebene Borel-messbare Funktion sei.
- Dabei diskutieren wir insbesondere die Frage, wie das Modell des Konfidenzintervalls verallgemeinert werden kann, um zu Konfidenzintervallen ausgehend von (vektoriellen) Stichprobenvariablen $X_i = (X_{1i}, X_{2i})$ zu gelangen.

Zur Vereinfachung der Schreibweise setzen wir $n = \max\{n_1, n_2\}$ und betrachten

- zwei Stichprobenfunktionen $\underline{\theta} : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ und $\bar{\theta} : \mathbb{R}^{2n} \rightarrow \mathbb{R}$, so dass

$$\underline{\theta}(x_1, \dots, x_n) \leq \bar{\theta}(x_1, \dots, x_n) \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^{2n}, \quad (61)$$

- wobei die Funktionswerte $\underline{\theta}(x_1, \dots, x_n)$ und $\bar{\theta}(x_1, \dots, x_n)$ jedoch nur von den (beobachteten) Komponenten $(x_{11}, \dots, x_{1n_1})$ und $(x_{21}, \dots, x_{2n_2})$ des Vektors (x_1, \dots, x_n) abhängen mögen.

Definition 3.7 Sei $\gamma \in (0, 1)$ eine beliebige, jedoch fest vorgegebene Zahl. Dann heißt das zufällige Intervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ *Konfidenzintervall* für $g(\theta)$ zum Niveau γ , falls

$$P_\theta(\underline{\theta}(X_1, \dots, X_n) \leq g(\theta) \leq \bar{\theta}(X_1, \dots, X_n)) \geq \gamma \quad (62)$$

für jedes $\theta \in \Theta$.

Beachte

- Bisher hatten wir in Abschnitt 3 stets den Fall betrachtet, dass $g(\theta) = \theta_j$ für ein $j \in \{1, \dots, m\}$.
- In den folgenden Beispielen hat $g(\theta)$ die Form $g(\theta) = \theta_j - \theta_k$ bzw. $g(\theta) = \theta_j/\theta_k$ für ein vorgegebenes Paar $j, k \in \{1, \dots, m\}$ von Indizes.

Die Stichprobenvariablen $X_i = (X_{1i}, X_{2i})$ seien nun normalverteilte Zufallsvektoren, wobei wir in den folgenden zwei Beispielen außerdem voraussetzen, dass die Komponenten X_{1i} und X_{2i} von X_i unabhängige (jedoch im allgemeinen nicht identisch verteilte) Zufallsvariablen sind.

Beispiel *Konfidenzintervall für die Differenz zweier Erwartungswerte (bei bekannten Varianzen)*

- Für zwei beliebige, jedoch vorgegebene Zahlen $n_1, n_2 \in \mathbb{N}$ betrachten wir zwei unabhängige Zufallsstichproben $(X_{11}, \dots, X_{1n_1})$ und $(X_{21}, \dots, X_{2n_2})$.
- Dabei nehmen wir an, dass $X_{1i} \sim N(\mu_1, \sigma_1^2)$ und $X_{2i} \sim N(\mu_2, \sigma_2^2)$ für (unbekannte) $\mu_1, \mu_2 \in \mathbb{R}$ und (bekannte) $\sigma_1^2, \sigma_2^2 > 0$.
- Dann sind die Stichprobenmittel \bar{X}_{1n_1} und \bar{X}_{2n_2} unabhängige Zufallsvariable mit

$$\bar{X}_{1n_1} \sim N(\mu_1, \sigma_1^2/n_1), \quad \bar{X}_{2n_2} \sim N(\mu_2, \sigma_2^2/n_2).$$

- Hieraus folgt, dass

$$\bar{X}_{1n_1} - \bar{X}_{2n_2} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

bzw.

$$\frac{\bar{X}_{1n_1} - \bar{X}_{2n_2} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1).$$

- Also gilt für jedes $\alpha \in (0, 1)$

$$P\left(z_{\frac{\alpha}{2}} \leq \frac{\bar{X}_{1n_1} - \bar{X}_{2n_2} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \leq z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

bzw.

$$P\left(\underbrace{\bar{X}_{1n_1} - \bar{X}_{2n_2} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}_{=\underline{\theta}} \leq \mu_1 - \mu_2 \leq \underbrace{\bar{X}_{1n_1} - \bar{X}_{2n_2} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}_{=\bar{\theta}}\right) = 1 - \alpha.$$

- Die auf diese Weise bestimmten Stichprobenfunktionen $\underline{\theta}(X_1, \dots, X_n)$ und $\bar{\theta}(X_1, \dots, X_n)$ ergeben ein Konfidenzintervall für $\mu_1 - \mu_2$ zum Niveau $\gamma = 1 - \alpha$.

Beispiel *Konfidenzintervall für den Quotienten zweier Varianzen (bei unbekanntem Erwartungswerten)*

- Ähnlich wie in dem vorhergehenden Beispiel betrachten wir für zwei beliebige, jedoch vorgegebene Zahlen $n_1, n_2 \in \mathbb{N}$ zwei unabhängige Zufallsstichproben $(X_{11}, \dots, X_{1n_1})$ und $(X_{21}, \dots, X_{2n_2})$.
- Dabei nehmen wir an, dass $X_{1i} \sim N(\mu_1, \sigma_1^2)$ und $X_{2i} \sim N(\mu_2, \sigma_2^2)$ für (unbekannte) $\mu_1, \mu_2 \in \mathbb{R}$ und (unbekannte) $\sigma_1^2, \sigma_2^2 > 0$.
- Dann sind die Stichprobenvarianzen $S_{1n_1}^2$ und $S_{2n_2}^2$ unabhängige Zufallsvariable und
- gemäß Theorem 3.5 gilt

$$\frac{(n_1 - 1)S_{1n_1}^2}{\sigma_1^2} \sim \chi_{n_1-1}^2 \quad \text{bzw.} \quad \frac{(n_2 - 1)S_{2n_2}^2}{\sigma_2^2} \sim \chi_{n_2-1}^2.$$

- Hieraus und aus der Definition der F-Verteilung ergibt sich

$$\frac{S_{2n_2}^2/\sigma_2^2}{S_{1n_1}^2/\sigma_1^2} \sim F_{n_2-1, n_1-1}.$$

- Also gilt für jedes $\alpha \in (0, 1)$

$$P\left(F_{n_2-1, n_1-1, \frac{\alpha}{2}} \leq \frac{S_{2n_2}^2/\sigma_2^2}{S_{1n_1}^2/\sigma_1^2} \leq F_{n_2-1, n_1-1, 1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

bzw.

$$P\left(\underbrace{\frac{S_{1n_1}^2}{S_{2n_2}^2} F_{n_2-1, n_1-1, \frac{\alpha}{2}}}_{=\underline{\theta}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \underbrace{\frac{S_{1n_1}^2}{S_{2n_2}^2} F_{n_2-1, n_1-1, 1-\frac{\alpha}{2}}}_{=\bar{\theta}}\right) = 1 - \alpha,$$

wobei $F_{n_2-1, n_1-1, \gamma}$ das γ -Quantil der F-Verteilung mit $(n_2 - 1, n_1 - 1)$ Freiheitsgraden bezeichnet, vgl. Tabellen 4a–4f.

- Die auf diese Weise bestimmten Stichprobenfunktionen $\underline{\theta}(X_1, \dots, X_n)$ und $\bar{\theta}(X_1, \dots, X_n)$ ergeben ein Konfidenzintervall für σ_1^2/σ_2^2 zum Niveau $\gamma = 1 - \alpha$.

3.6 Asymptotische Konfidenzintervalle

Wir zeigen für zwei Beispiele von Familien parametrischer Verteilungen, wie man mit Hilfe des zentralen Grenzwertsatzes und des starken Gesetzes der großen Zahlen asymptotische Konfidenzintervalle auf einfache Weise herleiten kann.

Dabei kehren wir erneut zur Betrachtung von *Ein-Stichproben-Problemen* zurück. D.h., wir nehmen an, dass *ein* Datensatz (x_1, \dots, x_n) beobachtet wird, den wir als Realisierung *einer* Zufallsstichprobe (X_1, \dots, X_n) auffassen.

Beispiel *Konfidenzintervall für den Erwartungswert λ bei Poisson-Verteilung*

- Wir nehmen an, dass $X_i \sim \text{Poi}(\lambda)$ für ein (unbekanntes) $\lambda > 0$.
- Weil dann $\mathbb{E} X_i = \lambda$ und $\text{Var} X_i = \lambda$ gilt, ergibt sich aus dem zentralen Grenzwertsatz, dass für jedes $\alpha \in (0, 1)$

$$\lim_{n \rightarrow \infty} P\left(-z_{1-\alpha/2} \leq \sqrt{n} \frac{\bar{X}_n - \lambda}{\sqrt{\lambda}} \leq z_{1-\alpha/2}\right) = 1 - \alpha. \quad (63)$$

- Die Ungleichung

$$-z_{1-\alpha/2} \leq \sqrt{n} \frac{\bar{X}_n - \lambda}{\sqrt{\lambda}} \leq z_{1-\alpha/2}$$

wird nun so umgeformt, dass ein Konfidenzintervall für λ hergeleitet werden kann:

$$\begin{aligned}
 -z_{1-\alpha/2} &\leq \sqrt{n} \frac{\bar{X}_n - \lambda}{\sqrt{\lambda}} \leq z_{1-\alpha/2} \\
 \Leftrightarrow n \frac{(\bar{X}_n - \lambda)^2}{\lambda} &\leq z_{1-\alpha/2}^2 \\
 \Leftrightarrow \bar{X}_n^2 - 2\lambda\bar{X}_n + \lambda^2 &\leq \frac{z_{1-\alpha/2}^2}{n} \lambda \\
 \Leftrightarrow \left(\lambda - \left(\bar{X}_n + \frac{z_{1-\alpha/2}^2}{2n} \right) \right)^2 &+ \bar{X}_n^2 - \left(\bar{X}_n + \frac{z_{1-\alpha/2}^2}{2n} \right)^2 \leq 0 \\
 \Leftrightarrow \left| \lambda - \left(\bar{X}_n + \frac{z_{1-\alpha/2}^2}{2n} \right) \right| &\leq \sqrt{\frac{\bar{X}_n z_{1-\alpha/2}^2}{n} + \frac{z_{1-\alpha/2}^4}{4n^2}}
 \end{aligned}$$

Durch diese Quadrierung der Ungleichungen in (63) und anschließende Auflösung nach λ ergibt sich also, dass (63) äquivalent ist mit

$$\lim_{n \rightarrow \infty} P \left(\left| \lambda - \left(\bar{X}_n + \frac{z_{1-\alpha/2}^2}{2n} \right) \right| \leq \sqrt{\frac{\bar{X}_n z_{1-\alpha/2}^2}{n} + \frac{z_{1-\alpha/2}^4}{4n^2}} \right) = 1 - \alpha.$$

- Also ist mit

$$\underline{\theta}(X_1, \dots, X_n) = \bar{X}_n + \frac{z_{1-\alpha/2}^2}{2n} - \sqrt{\frac{\bar{X}_n z_{1-\alpha/2}^2}{n} + \frac{z_{1-\alpha/2}^4}{4n^2}},$$

$$\bar{\theta}(X_1, \dots, X_n) = \bar{X}_n + \frac{z_{1-\alpha/2}^2}{2n} + \sqrt{\frac{\bar{X}_n z_{1-\alpha/2}^2}{n} + \frac{z_{1-\alpha/2}^4}{4n^2}}$$

ein asymptotisches Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für λ zum Niveau $\gamma = 1 - \alpha$ geben.

- Für die Länge dieses Konfidenzintervalls gilt

$$\ell(X_1, \dots, X_n) = \bar{\theta}(X_1, \dots, X_n) - \underline{\theta}(X_1, \dots, X_n) = 2 \sqrt{\frac{\bar{X}_n z_{1-\alpha/2}^2}{n} + \frac{z_{1-\alpha/2}^4}{4n^2}}.$$

Beispiel *Alternatives Konfidenzintervall für den Erwartungswert λ bei Poisson-Verteilung*

- Ein weiteres asymptotisches Konfidenzintervall für λ ergibt sich, wenn die Größe $\sqrt{\lambda}$ im Nenner von (63) ersetzt wird durch $\sqrt{\bar{X}_n}$.
- Aus dem starken Gesetz der großen Zahlen ergibt sich nämlich, dass

$$\bar{X}_n \xrightarrow{\text{f.s.}} \lambda.$$

- Man kann zeigen, dass hieraus und aus (63) die Gültigkeit von

$$\lim_{n \rightarrow \infty} P \left(-z_{1-\alpha/2} \leq \sqrt{n} \frac{\bar{X}_n - \lambda}{\sqrt{\bar{X}_n}} \leq z_{1-\alpha/2} \right) = 1 - \alpha \quad (64)$$

für jedes $\alpha \in (0, 1)$ folgt, wobei diese Art von Aussagen in der Literatur der *Satz von Slutski* genannt werden. Als ihr „Entdecker“ gilt Jewgeni Jewgenjewitsch Slutski (1880–1948).

- Also ist mit

$$\underline{\theta}(X_1, \dots, X_n) = \bar{X}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n}, \quad \bar{\theta}(X_1, \dots, X_n) = \bar{X}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n} \quad (65)$$

ein weiteres asymptotisches Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für λ zum Niveau $\gamma = 1 - \alpha$ gegeben.

- Dieses Konfidenzintervall hat allerdings den Nachteil, dass die in (65) gegebene Zufallsvariable $\underline{\theta}(X_1, \dots, X_n)$ auch negative Werte annehmen kann, obwohl bei der Poisson-Verteilung vorausgesetzt wird, dass $\lambda > 0$.
- Deshalb kann man anstelle von (65) die Zufallsvariablen

$$\underline{\theta}(X_1, \dots, X_n) = \max\left\{0, \bar{X}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n}\right\}, \quad \bar{\theta}(X_1, \dots, X_n) = \bar{X}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n}$$

als Endpunkte des Konfidenzintervalls $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für λ betrachten.

Beispiel *Konfidenzintervall für die „Erfolgswahrscheinlichkeit“ p bei Bernoulli-Verteilung*

- Wir nehmen an, dass $X_i \sim \text{Bin}(1, p)$ für ein (unbekanntes) $p \in (0, 1)$.
- Weil dann $\mathbb{E} X_i = p$ und $\text{Var} X_i = p(1 - p)$ gilt, ergibt sich aus dem zentralen Grenzwertsatz, dass für jedes $\alpha \in (0, 1)$

$$\lim_{n \rightarrow \infty} P\left(-z_{1-\alpha/2} \leq \sqrt{n} \frac{\bar{X}_n - p}{\sqrt{p(1-p)}} \leq z_{1-\alpha/2}\right) = 1 - \alpha. \quad (66)$$

- Durch Quadrierung der Ungleichungen in (66) und anschließende Auflösung nach p ergibt sich, dass (66) äquivalent ist mit

$$\lim_{n \rightarrow \infty} P(\underline{\theta}(X_1, \dots, X_n) \leq p \leq \bar{\theta}(X_1, \dots, X_n)) = 1 - \alpha,$$

wobei

$$\underline{\theta}(X_1, \dots, X_n) = \frac{1}{n + z_{1-\alpha/2}^2} \left(n\bar{X}_n + \frac{z_{1-\alpha/2}^2}{2} - z_{1-\alpha/2} \sqrt{n\bar{X}_n(1 - \bar{X}_n) + \frac{z_{1-\alpha/2}^2}{4}} \right)$$

bzw.

$$\bar{\theta}(X_1, \dots, X_n) = \frac{1}{n + z_{1-\alpha/2}^2} \left(n\bar{X}_n + \frac{z_{1-\alpha/2}^2}{2} + z_{1-\alpha/2} \sqrt{n\bar{X}_n(1 - \bar{X}_n) + \frac{z_{1-\alpha/2}^2}{4}} \right).$$

- Auf diese Weise ergibt sich ein asymptotisches Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für p zum Niveau $\gamma = 1 - \alpha$.

Beachte

- Ähnlich wie in dem vorhergehenden Beispiel erhält man ein einfacheres asymptotisches Konfidenzintervall für p , wenn man berücksichtigt, dass

$$\bar{X}_n \xrightarrow{\text{f.s.}} p$$

wegen des starken Gesetzes der großen Zahlen.

- Hieraus und aus (66) ergibt sich (wegen des Satzes von Slutski), dass

$$\lim_{n \rightarrow \infty} P\left(-z_{1-\alpha/2} \leq \sqrt{n} \frac{\bar{X}_n - p}{\sqrt{\bar{X}_n(1 - \bar{X}_n)}} \leq z_{1-\alpha/2}\right) = 1 - \alpha. \quad (67)$$

- Also ist mit

$$\underline{\theta}(X_1, \dots, X_n) = \bar{X}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n(1 - \bar{X}_n)}$$

bzw.

$$\bar{\theta}(X_1, \dots, X_n) = \bar{X}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n(1 - \bar{X}_n)}$$

ein weiteres asymptotisches Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für p zum Niveau $\gamma = 1 - \alpha$ gegeben.

- Dieses Konfidenzintervall hat jedoch den Nachteil, dass die Zufallsvariable $\underline{\theta}(X_1, \dots, X_n)$ negative Werte und $\bar{\theta}(X_1, \dots, X_n)$ Werte größer als 1 annehmen kann, obwohl bei der Bernoulli-Verteilung vorausgesetzt wird, dass $0 < p < 1$.
- Deshalb betrachtet man die modifizierten Schätzer

$$\underline{\theta}(X_1, \dots, X_n) = \max\left\{0, \bar{X}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n(1 - \bar{X}_n)}\right\}$$

bzw.

$$\bar{\theta}(X_1, \dots, X_n) = \min\left\{1, \bar{X}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \sqrt{\bar{X}_n(1 - \bar{X}_n)}\right\}$$

als Endpunkte des Konfidenzintervalls $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für p .

4 Tests statistischer Hypothesen

4.1 Problemstellung und Modellbeschreibung

In diesem Abschnitt prüfen wir Hypothesen (d.h. Vermutungen) über die Beschaffenheit der unbekanntem Verteilung bzw. Verteilungsfunktion der Stichprobenvariablen X_1, \dots, X_n .

- Wir betrachten lediglich sogenannte *nichtrandomisierte Tests*, d.h., die untersuchte Hypothese wird entweder verworfen, oder sie wird nicht verworfen.
- Die Entscheidung, ob eine Hypothese akzeptabel ist bzw. verworfen wird, hängt von den beobachteten Daten x_1, \dots, x_n ab, d.h. von der beobachteten Realisierung (x_1, \dots, x_n) der Zufallsstichprobe (X_1, \dots, X_n) .

- Die Entscheidungsregel wird so konstruiert, dass die Wahrscheinlichkeiten möglicher Fehlentscheidungen möglichst klein sind (bzw. vorgegebene Schwellenwerte nicht überschreiten).
- Dabei kann man ähnlich wie bei der Konstruktion von Konfidenzintervallen zu einem (vorgegebenen) Konfidenzniveau $\gamma = 1 - \alpha$ vorgehen, die in Abschnitt 3 diskutiert worden sind,
- denn bei einem Konfidenzintervall $(\underline{\theta}(X_1, \dots, X_n), \bar{\theta}(X_1, \dots, X_n))$ für einen unbekanntem Parameterwert $g(\theta)$ kann man α als die ‘Fehlerwahrscheinlichkeit’ interpretieren, dass $g(\theta)$ nicht innerhalb der Schranken $\underline{\theta}(X_1, \dots, X_n)$ bzw. $\bar{\theta}(X_1, \dots, X_n)$ liegt.

Wie in den Kapiteln 1–3 nehmen wir an, dass

- die Stichprobenvariablen X_1, \dots, X_n unabhängig und identisch verteilt sind und dass
- der Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , über den die Stichprobenvariablen X_1, X_2, \dots definiert sind, der kanonische Wahrscheinlichkeitsraum ist

Sei Δ die Familie der insgesamt in Betracht gezogenen (d.h. potentiell möglichen) Verteilungsfunktionen F der Stichprobenvariablen X_1, \dots, X_n .

- Betrachte eine Zerlegung von Δ in zwei (nichtleere) Teilmengen Δ_0 und Δ_1 , d.h.

$$\Delta = \Delta_0 \cup \Delta_1, \quad \text{wobei} \quad \Delta_0 \cap \Delta_1 = \emptyset.$$

- Es werden die Hypothesen

$$H_0 : F \in \Delta_0 \quad \text{bzw.} \quad H_1 : F \in \Delta_1,$$

betrachtet, dass die unbekannte Verteilungsfunktion F der Stichprobenvariablen X_i zu der Teilmenge $\Delta_0 \subset \Delta$ bzw. $\Delta_1 \subset \Delta$ innerhalb der Familie Δ der insgesamt in Betracht gezogenen Verteilungsfunktionen gehört.

Dabei ist die folgende Sprechweise üblich:

- Die Hypothese $H_0 : F \in \Delta_0$ heißt *Nullhypothese*, während $H_1 : F \in \Delta_1$ *Alternativhypothese* genannt wird.
- Soll eine Entscheidung zwischen $H_0 : F \in \Delta_0$ bzw. $H_1 : F \in \Delta_1$ getroffen werden, so ist folgende Formulierung üblich:
‘Man testet die Nullhypothese $H_0 : F \in \Delta_0$ gegen die Alternative $H_1 : F \in \Delta_1$ ’.
- Die Nullhypothese H_0 bzw. die Alternativhypothese H_1 heißen *einfach*, falls die Teilmenge Δ_0 bzw. Δ_1 nur aus einem Element besteht. Ansonsten sagt man, dass H_0 bzw. H_1 *zusammengesetzte Hypothesen* sind.

Um zwischen der Nullhypothese H_0 und der Alternativhypothese H_1 abwägen zu können, wird ein *Test*, d.h. eine *Entscheidungsregel* nach dem folgenden Prinzip konstruiert:

- Der Stichprobenraum \mathbb{R}^n wird in zwei Borel-Mengen K und $K^c = \mathbb{R}^n \setminus K$ zerlegt.
- Dabei heißt $K \subset \mathbb{R}^n$ der *kritische Bereich* (d.h. der Ablehnungsbereich der Nullhypothese H_0).
- Die Nullhypothese H_0 wird verworfen (d.h. abgelehnt), falls $(x_1, \dots, x_n) \in K$.
- Ansonsten, d.h. falls $(x_1, \dots, x_n) \notin K$, wird H_0 nicht verworfen.
- Man sagt bewusst ' H_0 nicht verworfen' wird statt ' H_0 wird akzeptiert'. Wenn H_0 nicht verworfen wird, bedeutet dies nicht, dass H_0 gültig ist, sondern nur, dass nicht genügend Hinweise für die Gültigkeit von H_1 vorliegen.

Mit anderen Worten:

- Betrachte die Stichprobenfunktion $\varphi : \mathbb{R}^n \rightarrow \{0, 1\}$ mit

$$\varphi(x_1, \dots, x_n) = \begin{cases} 1, & \text{falls } (x_1, \dots, x_n) \in K, \\ 0, & \text{falls } (x_1, \dots, x_n) \notin K. \end{cases} \quad (68)$$

- Die Nullhypothese H_0 wird also verworfen, falls $\varphi(x_1, \dots, x_n) = 1$.
- Ansonsten, d.h. falls $\varphi(x_1, \dots, x_n) = 0$, wird H_0 nicht verworfen.

Dies führt zu der folgenden

Definition 4.1 Sei $\alpha \in (0, 1)$ eine beliebige (vorgegebene) Zahl. Man sagt, dass die in (68) eingeführte Stichprobenfunktion $\varphi : \mathbb{R}^n \rightarrow \{0, 1\}$ ein *Test zum Niveau α* ist, falls

$$P_F(\varphi(X_1, \dots, X_n) = 1) \leq \alpha \quad (69)$$

für jedes $F \in \Delta_0$. Falls außerdem noch

$$P_F(\varphi(X_1, \dots, X_n) = 1) \geq \alpha \quad (70)$$

für jedes $F \in \Delta_1$, dann heißt $\varphi : \mathbb{R}^n \rightarrow \{0, 1\}$ ein *unverfälschter Test* zum Niveau α .

Beim Testen von Hypothesen können zwei Arten von Fehlern auftreten:

- Die Nullhypothese kann abgelehnt werden, ob sie richtig ist.
- Die Nullhypothese kann nicht abgelehnt werden, obwohl sie falsch ist.

Dies führt zur folgenden Definition.

	H_0 richtig	H_0 falsch
H_0 verwerfen	Fehler 1. Art; Wahrscheinlichkeit $\alpha_n(F) \leq \alpha$	richtige Entscheidung
H_0 nicht verwerfen	richtige Entscheidung	Fehler 2. Art; Wahrscheinlichkeit $\beta_n(F)$

Definition 4.2 Für jedes $F \in \Delta_0$ heißt

$$\alpha_n(F) = P(\varphi(X_1, \dots, X_n) = 1) \quad \left(= P_F((X_1, \dots, X_n) \in K) \right)$$

die *Wahrscheinlichkeit des Fehlers 1. Art*, und

für jedes $F \in \Delta_1$ heißt

$$\beta_n(F) = P_F(\varphi(X_1, \dots, X_n) = 0) \quad \left(= P_F((X_1, \dots, X_n) \notin K) \right)$$

die *Wahrscheinlichkeit des Fehlers 2. Art*.

Ein Überblick über Fehler 1. Art und Fehler 2. Art wird in der obigen Tabelle gegeben.

Von besonderem Interesse sind Tests, deren Wahrscheinlichkeiten für Fehler 1. Art eine vorgegebene ‘Irrtumswahrscheinlichkeit’ α nicht überschreiten und für die gleichzeitig die Wahrscheinlichkeiten für Fehler 2. Art möglichst klein sind.

4.2 Parametrische Tests

Falls die Familie Δ der insgesamt in Betracht gezogenen Verteilungsfunktionen F der Stichprobenvariablen X_1, \dots, X_n eine parametrische Familie $\Delta = \{F_\theta, \theta \in \Theta\}$ von Verteilungsfunktionen ist, dann zerlegen wir Δ auf die folgende Weise in zwei Teilmengen Δ_0 und Δ_1 :

- Betrachte eine Zerlegung

$$\Theta = \Theta_0 \cup \Theta_1, \quad \Theta_0 \cap \Theta_1 = \emptyset$$

des Parameterraumes Θ in zwei disjunkte Teilmengen Θ_0, Θ_1 , so dass

$$\Delta_0 = \{F_\theta, \theta \in \Theta_0\} \quad \text{bzw.} \quad \Delta_1 = \{F_\theta, \theta \in \Theta_1\},$$

- wobei die Hypothesen H_0 und H_1 dann wie folgt formuliert werden:

$$H_0 : \theta \in \Theta_0 \quad \text{vs.} \quad H_1 : \theta \in \Theta_1.$$

Beispiel Sei M_1 ein etabliertes Medikament mit bekannter Heilungswahrscheinlichkeit $p_1 = 0.8$. M_2 sei ein neues Medikament mit unbekannter Heilungswahrscheinlichkeit p_2 . Es soll überprüft werden, ob M_2 besser ist als M_1 . Dies führt zur folgenden Hypothesenbildung:

$$H_0 : p_2 \leq 0.8 \quad \text{vs.} \quad H_1 : p_2 > 0.8.$$

Das Medikament M_2 wird an $n = 50$ Patienten getestet. Es seien X_1, \dots, X_n die entsprechenden Stichprobenvariablen mit $X_i \sim \text{Bin}(1, p_2)$, d. h. $\{X_i = 1\}$ bedeutet, dass der i -te Patient geheilt wurde. Sollte man H_0 verwerfen, falls z. B. 42 Heilungen beobachtet wurden (d. h. $\hat{p}_2 = \bar{X}_n = 0.84$)?

Definition 4.3 Sei $\Theta = \mathbb{R}$ und $\theta_0 \in \Theta$. Folgende Hypothesen seien gegeben:

1. $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$
2. $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$
3. $H_0 : \theta \geq \theta_0$ vs. $H_1 : \theta < \theta_0$.

Im Fall (1) heißt ein entsprechender Test *zweiseitig*, bei den Fällen (2)-(3) *einseitig*.

Für parametrische Verteilungsfamilien können die Begriffe, die in Definition 4.1 eingeführt wurden, wie folgt spezifiziert bzw. durch weitere Begriffe ergänzt werden.

Definition 4.4

- Im Fall einer parametrischen Familie $\Delta = \{F_\theta, \theta \in \Theta\}$ von Verteilungsfunktionen heißt die in (68) eingeführte Stichprobenfunktion $\varphi : \mathbb{R}^n \rightarrow \{0, 1\}$ ein *Parametertest zum Niveau α* , falls

$$P_\theta(\varphi(X_1, \dots, X_n) = 1) \leq \alpha \tag{71}$$

für jedes $\theta \in \Theta_0$ gilt.

- Falls außerdem noch

$$P_\theta(\varphi(X_1, \dots, X_n) = 1) \geq \alpha \tag{72}$$

für jedes $\theta \in \Theta_1$, dann heißt $\varphi : \mathbb{R}^n \rightarrow \{0, 1\}$ ein *unverfälschter Test* zum Niveau α .

- Die Abbildung $\alpha_n : \Theta \rightarrow [0, 1]$ mit $\alpha_n(\theta) = P_\theta((\varphi(X_1, \dots, X_n) = 1))$ heißt *Gütefunktion* des Tests φ .
- Die Einschränkung $\alpha_n : \Theta_1 \rightarrow [0, 1]$ dieser Abbildung auf die Teilmenge Θ_1 des Parameterraumes heißt die *Macht* (power) von φ .
- Der Test φ zum Niveau α heißt *konsistent*, falls $\lim_{n \rightarrow \infty} \alpha_n(\theta) = 1$ für jedes $\theta \in \Theta_1$ gilt.
- Seien $\varphi : \mathbb{R}^n \rightarrow \{0, 1\}$ und $\varphi' : \mathbb{R}^n \rightarrow \{0, 1\}$ zwei Parametertests zum Niveau α . Falls die Macht von φ' größer ist als die Macht von φ , d.h.

$$P_\theta((\varphi(X_1, \dots, X_n) = 1) \leq P_\theta((\varphi'(X_1, \dots, X_n) = 1) \tag{73}$$

für jedes $\theta \in \Theta_1$, dann sagt man, dass der Test φ' *besser* als der Test φ ist.

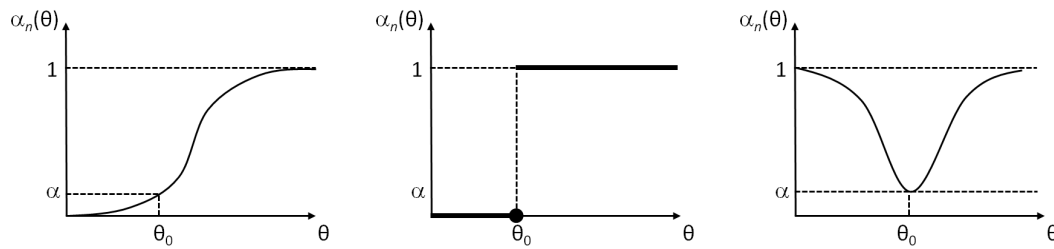


Figure 4: Links: Gütefunktion eines einseitigen Tests zum Niveau α für die Hypothesen $H_0 : \theta \leq \theta_0$ vs. $H_1 : \theta > \theta_0$. Mitte: optimale Gütefunktion. Rechts: Gütefunktion eines zweiseitigen Tests zum Niveau α für die Hypothesen $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$.

Beachte Beispiele für Gütefunktionen bei $\Theta \subset \mathbb{R}$ sind in Abbildung 4 gegeben.

Bei der praktischen Konstruktion des kritischen Bereichs K eines Parametertests zum Niveau α kann oft wie folgt vorgegangen werden:

- Bestimme eine Stichprobenfunktion $T : \mathbb{R}^n \rightarrow \mathbb{R}$ (genannt *Teststatistik*), so dass
- die Zufallsvariable $T(X_1, \dots, X_n)$ für jedes $\theta \in \Theta_0$ die gleiche (bekannte) Verteilung hat, und
- bestimme einen *Schwellenwert* $c > 0$, so dass $P_\theta(|T(X_1, \dots, X_n)| > c) \leq \alpha$ für jedes $\theta \in \Theta_0$ gilt.
- Dann ist mit $K = \{(x_1, \dots, x_n) : |T(x_1, \dots, x_n)| > c\}$ der kritische Bereich eines Parametertests zum Niveau α gegeben.

Um einen Test zum Niveau α mit einer möglichst großen Macht zu erhalten, ist es jedoch manchmal zweckmäßiger,

- zwei Schwellenwerte $c_1, c_2 \in \mathbb{R} \cup \{\pm\infty\}$ mit $c_1 < c_2$ zu betrachten, so dass für jedes $\theta \in \Theta_0$

$$P_\theta(c_1 \leq T(X_1, \dots, X_n) \leq c_2) \geq 1 - \alpha.$$

- Dann ist mit $K = \mathbb{R}^n \setminus \{(x_1, \dots, x_n) : c_1 \leq T(x_1, \dots, x_n) \leq c_2\}$ der kritische Bereich eines (asymmetrischen) Parametertests zum Niveau α gegeben.

Definition 4.5 Sei (x_1, \dots, x_n) die (konkrete) Stichprobe, die als Realisierung von (X_1, \dots, X_n) interpretiert wird. Weiterhin sei $T(X_1, \dots, X_n)$ die Teststatistik, mit der der Test φ konstruiert wurde. Der *p-Wert* des statistischen Tests φ ist das kleinste Niveau, bei dem der Wert $t = T(x_1, \dots, x_n)$ zur Verwerfung der von H_0 führt.

In den folgenden Abschnitten wird das oben beschriebene Konstruktionsprinzip anhand einer Reihe von Beispielen näher diskutiert.

Beispiel *einseitiger Binomialtest*

- Wir nehmen an, dass $X_i \sim \text{Bin}(1, p)$ für eine unbekannte ‘Erfolgswahrscheinlichkeit’ $p \in (0, 1)$.
- Für einen vorgegebenen Wert p_0 soll eine Entscheidung zwischen folgenden Hypothesen getroffen werden:

$$H_0 : p \leq p_0 \quad \text{vs.} \quad H_1 : p \geq p_0.$$

- Idee: Das Stichprobenmittel \bar{X}_n ist ein guter Schätzer für p . Also sollte H_0 für ‘grosse’ Werte von \bar{X}_n abgelehnt werden.
- Dies führt zu folgendem Ansatz:

$$\varphi(X_1, \dots, X_n) = \begin{cases} 1, & \text{falls } \sum_{i=1}^n X_i > k, \\ 0, & \text{sonst.} \end{cases}$$

- Dabei ist $k \in \mathbb{N}$ so zu wählen, dass der Test Niveau $\alpha \in (0, 1)$ hat. Sei also $p \leq p_0$. Dann gilt:

$$\begin{aligned} \alpha_n(p) &= P_p(\varphi(X_1, \dots, X_n) = 1) \\ &= P_p\left(\sum_{i=1}^n X_i > k\right) \\ &= \sum_{j=k+1}^n \binom{n}{j} p^j (1-p)^{n-j} \end{aligned}$$

- Dabei ist $\alpha_n(p)$ ist monoton wachsend in p . Daher gilt:

$$\alpha_n(p) \leq \alpha_n(p_0) \quad \text{für alle } p \leq p_0.$$

- Mit

$$k = \min \left\{ \ell \in \{-1, 0, \dots, n\} : \sum_{j=\ell+1}^n \binom{n}{j} p_0^j (1-p_0)^{n-j} \leq \alpha \right\}$$

hat der Test φ das Niveau α .

- Gleichzeitig wird der Fehler 2. Art $\beta_n(p) = 1 - \alpha_n(p) = \sum_{j=0}^k \binom{n}{j} p^j (1-p)^{n-j}$ für $p > p_0$ möglichst gering gehalten.

Der einseitige Binomialtest kann nun auf das obige Beispiel mit den zwei Medikamenten M_1 bzw. M_2 angewandt werden. Zur Erinnerung: Das Medikament M_1 hat die bekannte Heilungswahrscheinlichkeit $p_1 = 0.8$, das Medikament M_2 die unbekannte Heilungswahrscheinlichkeit p_2 . Folgende Hypothesen sollen getestet werden:

$$H_0 : p_2 \leq 0.8 \quad \text{vs.} \quad H_1 : p_2 > 0.8.$$

Das Medikament M_2 wurde an $n = 50$ Patienten getestet, wobei 42 Heilungen beobachtet wurden. Nach dem einseitigen Binomialtest wird $H_0 : p_2 \leq 0.8$ abgelehnt, falls $\sum_{i=1}^n X_i > k$. Das Niveau sei $\alpha = 0.05$. Dann ist $k = 43$, da

$$\sum_{j=\ell+1}^{50} \binom{50}{j} 0.8^j (0.2)^{50-j} = \begin{cases} 0.045 & \text{für } \ell = 43, \\ 0.1034 & \text{für } \ell = 42. \end{cases}$$

Da 42 Heilungen beobachtet wurden, wird H_0 zum Niveau $\alpha = 0.05$ nicht verworfen.

4.2.1 Parametertests bei Normalverteilung

Wir nehmen in diesem Abschnitt an, dass die Stichprobenvariablen X_1, \dots, X_n normalverteilt sind, d.h., es gelte $\Delta \subset \{N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\}$. Wir betrachten also die Familie der (eindimensionalen) Normalverteilungen bzw. Teilmengen hiervon.

Wir konstruieren zunächst Tests für den Erwartungswert μ . Dabei gibt es Ähnlichkeiten zur Konstruktion der Konfidenzintervalle für μ , die in Abschnitt 3.4 diskutiert werden.

Beispiel *Test des Erwartungswertes μ (bei bekannter Varianz σ^2)*

- Wir nehmen an, dass $X_i \sim N(\mu, \sigma^2)$ für ein (unbekanntes) $\mu \in \mathbb{R}$ und ein (bekanntes) $\sigma^2 > 0$.
- Für einen vorgegebenen (hypothetischen) Zahlenwert $\mu_0 \in \mathbb{R}$ soll die Hypothese $H_0 : \mu = \mu_0$ (gegen die Alternative $H_1 : \mu \neq \mu_0$) getestet werden, d.h. $\Theta = \{\mu : \mu \in \mathbb{R}\}$ mit

$$\Theta_0 = \{\mu_0\} \quad \text{bzw.} \quad \Theta_1 = \{\mu : \mu \in \mathbb{R}, \mu \neq \mu_0\}$$

- Wir betrachten die Teststatistik $T : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$T(x_1, \dots, x_n) = \sqrt{n} \frac{\bar{x}_n - \mu_0}{\sigma}. \quad (74)$$

- Unter H_0 (d.h. falls H_0 gültig ist) gilt:

$$T(X_1, \dots, X_n) \sim N(0, 1).$$

- Als Schwellenwert c wird $c = z_{1-\alpha/2}$ verwendet, d.h. das $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung.
- Dann gilt $P_{\mu_0}(|T(X_1, \dots, X_n)| > z_{1-\alpha/2}) = \alpha$.
- Die Hypothese $H_0 : \mu = \mu_0$ wird abgelehnt, falls $|T(x_1, \dots, x_n)| > z_{1-\alpha/2}$.
- Dieser Test wird als *zweiseitiger Gaußtest* bezeichnet.

Für $\alpha = 0.05$ und $\sigma = 0.10$ wollen wir nun prüfen, ob

- die Hypothese $H_0 : \mu = 42.0$ mit der bereits in Abschnitt 3.2 betrachteten (konkreten) Stichprobe

$$(x_1, \dots, x_{10}) = (41.60, 41.48, 42.34, 41.95, 41.86, 42.18, 41.72, 42.26, 41.81, 42.04)$$

vereinbar ist.

- In diesem Fall gilt $\bar{x}_{10} = 41.924$, und aus (74) ergibt sich somit

$$\begin{aligned} |T(x_1, \dots, x_{10})| &= \left| \sqrt{10} \frac{\bar{x}_{10} - \mu_0}{\sigma} \right| \\ &= \left| \sqrt{10} \frac{41.924 - 42.0}{0.10} \right| \\ &= 2.40 > z_{0.975} = 1.96, \end{aligned}$$

wobei das Quantil $z_{0.975} = 1.96$ aus Tabelle 1 entnommen wurde.

- Die Hypothese $H_0 : \mu = 42.0$ wird also verworfen.

Beachte

1. Man kann sich leicht überlegen, dass

- für die Gütefunktion $\alpha_n : \mathbb{R} \rightarrow [0, 1]$ dieses Tests

$$\begin{aligned} \alpha_n(\mu) &= P_\mu(|T(X_1, \dots, X_n)| > z_{1-\alpha/2}) \\ &= 1 - \Phi\left(z_{1-\alpha/2} + \frac{\mu_0 - \mu}{\sigma} \sqrt{n}\right) + \Phi\left(-z_{1-\alpha/2} + \frac{\mu_0 - \mu}{\sigma} \sqrt{n}\right) \end{aligned}$$

für beliebige $n \in \mathbb{N}$ und $\mu \in \mathbb{R}$ gilt.

- Hieraus ergibt sich insbesondere, dass $\alpha_n(\mu) \geq \alpha$ für beliebige $n \in \mathbb{N}$ und $\mu \in \mathbb{R}$, und

$$\lim_{n \rightarrow \infty} \alpha_n(\mu) = 1$$

für jedes $\mu \neq \mu_0$.

- Der Test ist also unverfälscht und konsistent.

2. Falls die Hypothese $H_0 : \mu = \mu_0$ gegen die (einseitige) Alternative $H_1 : \mu > \mu_0$ getestet werden soll, d.h. $\Theta = \{\mu : \mu \in \mathbb{R}, \mu \geq \mu_0\}$ mit

$$\Theta_0 = \{\mu_0\} \quad \text{bzw.} \quad \Theta_1 = \{\mu : \mu \in \mathbb{R}, \mu > \mu_0\},$$

dann könnte man zwar so wie bisher vorgehen und für die in (74) definierte Teststatistik T

- den kritischen Bereich $K = \{(x_1, \dots, x_n) : |T(x_1, \dots, x_n)| > z_{1-\alpha/2}\}$ betrachten.
- Ein besserer Test ergibt sich jedoch, wenn der folgende (einseitige) kritische Bereich

$$K' = \{(x_1, \dots, x_n) : T(x_1, \dots, x_n) > z_{1-\alpha}\}$$

betrachtet wird, denn für jedes $\mu > \mu_0$ gilt

$$1 - \Phi\left(z_{1-\alpha/2} + \frac{\mu_0 - \mu}{\sigma} \sqrt{n}\right) + \Phi\left(-z_{1-\alpha/2} + \frac{\mu_0 - \mu}{\sigma} \sqrt{n}\right) < 1 - \Phi\left(z_{1-\alpha} + \frac{\mu_0 - \mu}{\sigma} \sqrt{n}\right).$$

3. Falls die Hypothese $H_0 : \mu = \mu_0$ gegen die Alternative $H_1 : \mu < \mu_0$ getestet werden soll, dann wird der kritische Bereich $K'' = \{(x_1, \dots, x_n) : T(x_1, \dots, x_n) < -z_{1-\alpha}\}$ betrachtet.

Wir diskutieren nun einen Test des Erwartungswertes μ , wobei angenommen wird, dass die Varianz σ^2 ebenfalls unbekannt ist.

Beispiel Test des Erwartungswertes μ (bei unbekannter Varianz σ^2)

- Im Unterschied zu dem vorhergehenden Beispiel nehmen wir nun an, dass $X_i \sim N(\mu, \sigma^2)$ für ein (unbekanntes) $\mu \in \mathbb{R}$ und ein (unbekanntes) $\sigma^2 > 0$.
- Für einen vorgegebenen (hypothetischen) Zahlenwert $\mu_0 \in \mathbb{R}$ soll erneut die Hypothese $H_0 : \mu = \mu_0$ (gegen die zweiseitige Alternative $H_1 : \mu \neq \mu_0$) getestet werden.
- Dies bedeutet, dass jetzt $\Theta = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$ mit

$$\Theta_0 = \{(\mu, \sigma^2) : \mu = \mu_0, \sigma^2 > 0\} \quad \text{bzw.} \quad \Theta_1 = \{(\mu, \sigma^2) : \mu \neq \mu_0, \sigma^2 > 0\},$$

d.h., sowohl H_0 als auch H_1 sind jetzt zusammengesetzte Hypothesen.

- Damit hängt auch zusammen, dass die in (74) betrachtete Teststatistik T nun nicht mehr geeignet ist, weil in (74) die unbekannte Größe σ vorkommt.
- Wir betrachten vielmehr die Teststatistik $T : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$T(x_1, \dots, x_n) = \sqrt{n} \frac{\bar{x}_n - \mu_0}{s_n}. \quad (75)$$

- Unter H_0 gilt (vgl. Korollar 3.6):

$$T(X_1, \dots, X_n) \sim t_{n-1}.$$

- Als Schwellenwert wird $c = t_{n-1, 1-\alpha/2}$ verwendet, d.h. das $(1 - \alpha/2)$ -Quantil der t -Verteilung mit $n - 1$ Freiheitsgraden.
- Es gilt $P_{(\mu_0, \sigma^2)}(|T(X_1, \dots, X_n)| > t_{n-1, 1-\alpha/2}) = \alpha$ für jedes $\sigma^2 > 0$.
- Die Hypothese $H_0 : \mu = \mu_0$ wird abgelehnt, falls $|T(x_1, \dots, x_n)| > t_{n-1, 1-\alpha/2}$.
- Der Test wird als *Einstichproben-t-Test* bezeichnet.

Für $\alpha = 0.05$ wollen wir nun prüfen, ob

- die Hypothese $H_0 : \mu = 42.0$ mit der schon in dem vorhergehenden Beispiel betrachteten (konkreten) Stichprobe

$$(x_1, \dots, x_{10}) = (41.60, 41.48, 42.34, 41.95, 41.86, 42.18, 41.72, 42.26, 41.81, 42.04)$$

vereinbar ist, wobei jetzt jedoch angenommen wird, dass die Varianz σ^2 unbekannt ist.

- Es gilt $\bar{x}_{10} = 41.924$ und $s_{10}^2 = 0.0807$ bzw. $s_{10} = 0.284$.
- Aus (75) ergibt sich somit, dass

$$\begin{aligned} |T(x_1, \dots, x_{10})| &= \left| \sqrt{10} \frac{\bar{x}_{10} - \mu_0}{s_{10}} \right| \\ &= \left| \sqrt{10} \frac{41.924 - 42.0}{0.284} \right| \\ &= 0.846 < t_{9, 0.975} = 2.262, \end{aligned}$$

wobei das Quantil $t_{9, 0.975} = 2.262$ aus Tabelle 3 entnommen wurde.

- Unter den jetzt gemachten Modellannahmen wird also die Hypothese $H_0 : \mu = 42.0$ nicht verworfen.
- Der Hauptgrund hierfür ist, dass die Stichprobenvarianz $s_{10}^2 = 0.0807$ deutlich größer ist als der bei dem vorhergehenden Beispiel betrachtete Wert $\sigma^2 = 0.01$ der (z.B. aufgrund von Vorinformationen) als bekannt angenommenen Varianz σ^2 .
- Mit anderen Worten: Wegen der relativ großen Stichprobenvarianz s_{10}^2 wird jetzt die Abweichung $0.076 = |41.924 - 42.0|$ des Stichprobenmittels $\bar{x}_{10} = 41.924$ von dem hypothetischen Erwartungswert $\mu_0 = 42.0$ toleriert, während sie bei dem vorhergehenden Beispiel (vor allem) wegen der kleineren Varianz $\sigma^2 = 0.01$ nicht toleriert wurde.

Beachte

1. Falls die Hypothese $H_0 : \mu = \mu_0$ gegen die Alternative $H_1 : \mu > \mu_0$ getestet werden soll, dann wird (ähnlich wie in dem vorhergehenden Beispiel) für die in (75) definierte Teststatistik T der kritische Bereich K' betrachtet mit

$$K' = \{(x_1, \dots, x_n) : T(x_1, \dots, x_n) > t_{n-1, 1-\alpha}\}.$$

2. Analog wird für den Test der Hypothese $H_0 : \mu = \mu_0$ gegen die Alternative $H_1 : \mu < \mu_0$ der kritische Bereich K'' betrachtet mit

$$K'' = \{(x_1, \dots, x_n) : T(x_1, \dots, x_n) < -t_{n-1, 1-\alpha}\}.$$

Beispiel Test der Varianz σ^2 (bei bekanntem Erwartungswert μ)

- Wir nehmen an, dass $X_i \sim N(\mu, \sigma^2)$ für ein (bekanntes) $\mu \in \mathbb{R}$ und ein (unbekanntes) $\sigma^2 > 0$.
- Für einen vorgegebenen (hypothetischen) Zahlenwert $\sigma_0^2 > 0$ soll die Hypothese $H_0 : \sigma^2 = \sigma_0^2$ (gegen die Alternative $H_1 : \sigma^2 \neq \sigma_0^2$) getestet werden, d.h. $\Theta = \{\sigma^2 : \sigma^2 > 0\}$ mit

$$\Theta_0 = \{\sigma_0^2\} \quad \text{bzw.} \quad \Theta_1 = \{\sigma^2 : \sigma^2 \neq \sigma_0^2\}.$$

- Wir betrachten die Teststatistik $T : \mathbb{R}^n \rightarrow (0, \infty)$ mit

$$T(x_1, \dots, x_n) = \frac{n\tilde{s}_n^2}{\sigma_0^2}. \quad (76)$$

- Unter H_0 gilt:

$$T(X_1, \dots, X_n) \sim \chi_n^2.$$

- Als Schwellenwerte werden $c_1 = \chi_{n, \alpha/2}^2$ bzw. $c_2 = \chi_{n, 1-\alpha/2}^2$ verwendet. Dabei ist \tilde{S}_n^2 die in Abschnitt 3.4 eingeführte Größe

$$\tilde{S}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2.$$

- Dann gilt $P_{\sigma_0^2}(\chi_{n,\alpha/2}^2 \leq T(X_1, \dots, X_n) \leq \chi_{n,1-\alpha/2}^2) = 1 - \alpha$.
- Die Hypothese $H_0 : \sigma^2 = \sigma_0^2$ wird abgelehnt, falls $T(x_1, \dots, x_n) < \chi_{n,\alpha/2}^2$ oder $T(x_1, \dots, x_n) > \chi_{n,1-\alpha/2}^2$.

Beispiel Test der Varianz σ^2 (bei unbekanntem Erwartungswert μ)

- Wir nehmen nun an, dass $X_i \sim N(\mu, \sigma^2)$ für ein (unbekanntes) $\mu \in \mathbb{R}$ und ein (unbekanntes) $\sigma^2 > 0$.
- Für einen vorgegebenen (hypothetischen) Zahlenwert $\sigma_0^2 \in \mathbb{R}$ soll die Hypothese $H_0 : \sigma^2 = \sigma_0^2$ (gegen die Alternative $H_1 : \sigma^2 \neq \sigma_0^2$) getestet werden, d.h. $\Theta = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$ mit

$$\Theta_0 = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 = \sigma_0^2\} \quad \text{bzw.} \quad \Theta_1 = \{(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 \neq \sigma_0^2\}.$$

- Wir betrachten die Teststatistik $T : \mathbb{R}^n \rightarrow (0, \infty)$ mit

$$T(x_1, \dots, x_n) = \frac{(n-1)s_n^2}{\sigma_0^2}. \quad (77)$$

- Unter H_0 gilt (vgl. Theorem 3.5):

$$T(X_1, \dots, X_n) \sim \chi_{n-1}^2.$$

Als Schwellenwerte werden $c_1 = \chi_{n-1,\alpha/2}^2$ bzw. $c_2 = \chi_{n-1,1-\alpha/2}^2$ verwendet.

- Es gilt

$$P_{(\mu, \sigma_0^2)}(\chi_{n-1,\alpha/2}^2 \leq T(X_1, \dots, X_n) \leq \chi_{n-1,1-\alpha/2}^2) = 1 - \alpha$$

für jedes $\mu \in \mathbb{R}$.

- Die Hypothese $H_0 : \sigma^2 = \sigma_0^2$ wird abgelehnt, falls

$$T(x_1, \dots, x_n) < \chi_{n-1,\alpha/2}^2 \quad \text{oder} \quad T(x_1, \dots, x_n) > \chi_{n-1,1-\alpha/2}^2.$$

4.2.2 Zwei-Stichproben-Tests

So wie in Abschnitt 3.5 nehmen wir nun an, dass die Stichprobenvariablen X_1, \dots, X_n (zweidimensionale) normalverteilte Zufallsvektoren sind mit $X_i = (X_{1i}, X_{2i})$. Dabei setzen wir voraus, dass die Komponenten X_{1i} und X_{2i} von X_i unabhängige (jedoch im allgemeinen nicht identisch verteilte) Zufallsvariablen sind.

Für zwei beliebige, jedoch vorgegebene Zahlen $n_1, n_2 \in \mathbb{N}$ betrachten wir also zwei unabhängige Zufallsstichproben $(X_{11}, \dots, X_{1n_1})$ und $(X_{21}, \dots, X_{2n_2})$, wobei wir

- die beobachteten Daten $(x_{11}, \dots, x_{1n_1})$ und $(x_{21}, \dots, x_{2n_2})$ als Realisierungen von $(X_{11}, \dots, X_{1n_1})$ bzw. $(X_{21}, \dots, X_{2n_2})$ auffassen und
- zur Vereinfachung der Schreibweise $n = \max\{n_1, n_2\}$ bzw. $x_i = (x_{1i}, x_{2i})$ für $i = 1, \dots, n$ setzen.

Beispiel *Test der Gleichheit zweier Erwartungswerte (bei bekannten Varianzen)*

- Wir nehmen an, dass $X_{1i} \sim N(\mu_1, \sigma_1^2)$ und $X_{2i} \sim N(\mu_2, \sigma_2^2)$ für (unbekannte) $\mu_1, \mu_2 \in \mathbb{R}$ und (bekannte) $\sigma_1^2, \sigma_2^2 > 0$.
- Dabei soll die Hypothese $H_0 : \mu_1 = \mu_2$ (gegen die Alternative $H_1 : \mu_1 \neq \mu_2$) getestet werden, d.h. $\Theta = \{(\mu_1, \mu_2) : \mu_1, \mu_2 \in \mathbb{R}\}$ mit

$$\Theta_0 = \{(\mu_1, \mu_2) : \mu_1 = \mu_2\} \quad \text{bzw.} \quad \Theta_1 = \{(\mu_1, \mu_2) : \mu_1 \neq \mu_2\}$$

- Wir betrachten die Teststatistik $T : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ mit

$$T(x_1, \dots, x_n) = \frac{\bar{x}_{1n_1} - \bar{x}_{2n_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (78)$$

und den Schwellenwert $c = z_{1-\alpha/2}$.

- Weil $\bar{X}_{1n_1} \sim N(\mu_1, \sigma_1^2/n_1)$ bzw. $\bar{X}_{2n_2} \sim N(\mu_2, \sigma_2^2/n_2)$, gilt dann

$$\bar{X}_{1n_1} - \bar{X}_{2n_2} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

- Unter H_0 gilt also:

$$T(X_1, \dots, X_n) \sim N(0, 1)$$

und daher

$$P_{(\mu_1, \mu_2)}(|T(X_1, \dots, X_n)| > z_{1-\alpha/2}) = \alpha$$

für alle $\mu_1, \mu_2 \in \mathbb{R}$ mit $\mu_1 = \mu_2$.

- Die Hypothese $H_0 : \mu_1 = \mu_2$ wird abgelehnt, falls $|T(x_1, \dots, x_n)| > z_{1-\alpha/2}$.

Beispiel *Test der Gleichheit zweier Varianzen (bei unbekanntem Erwartungswerten)*

- Wir nehmen wir an, dass $X_{1i} \sim N(\mu_1, \sigma_1^2)$ und $X_{2i} \sim N(\mu_2, \sigma_2^2)$ für (unbekannte) $\mu_1, \mu_2 \in \mathbb{R}$ und (unbekannte) $\sigma_1^2, \sigma_2^2 > 0$.
- Dabei soll die Hypothese $H_0 : \sigma_1^2 = \sigma_2^2$ (gegen die Alternative $H_1 : \sigma_1^2 \neq \sigma_2^2$) getestet werden, d.h. $\Theta = \{(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) : \mu_1, \mu_2 \in \mathbb{R}, \sigma_1^2, \sigma_2^2 > 0\}$ mit

$$\Theta_0 = \{(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) : \sigma_1^2 = \sigma_2^2\} \quad \text{bzw.} \quad \Theta_1 = \{(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) : \sigma_1^2 \neq \sigma_2^2\}$$

- Wir betrachten die Teststatistik $T : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ mit

$$T(x_1, \dots, x_n) = \frac{s_{2n_2}^2}{s_{1n_1}^2}, \quad (79)$$

wobei

$$s_{in_i}^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{in_i})^2 \quad i = 1, 2.$$

- Für $\sigma_1^2 = \sigma_2^2$ gilt dann

$$T(X_1, \dots, X_n) \sim F_{n_2-1, n_1-1}.$$

- Mit den Schwellenwerten $c_1 = F_{n_2-1, n_1-1, \alpha/2}$ und $c_2 = F_{n_2-1, n_1-1, 1-\alpha/2}$ gilt also

$$P_{(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)}(F_{n_2-1, n_1-1, \alpha/2} \leq T(X_1, \dots, X_n) \leq F_{n_2-1, n_1-1, 1-\alpha/2}) = 1 - \alpha$$

für alle $\mu_1, \mu_2 \in \mathbb{R}$ und für alle $\sigma_1^2, \sigma_2^2 > 0$ mit $\sigma_1^2 = \sigma_2^2$.

- Die Hypothese $H_0 : \sigma_1^2 = \sigma_2^2$ wird somit abgelehnt, falls

$$T(x_1, \dots, x_n) < F_{n_2-1, n_1-1, \alpha/2} \quad \text{oder} \quad T(x_1, \dots, x_n) > F_{n_2-1, n_1-1, 1-\alpha/2}.$$

4.2.3 Asymptotische Parametertests

Wir setzen jetzt nicht mehr voraus, dass die Stichprobenvariablen X_1, X_2, \dots normalverteilt sind, sondern lediglich, dass die Verteilungsfunktion F von X_i zu einer (beliebigen) parametrischen Familie $\Delta = \{F_\theta, \theta \in \Theta\}$ von Verteilungsfunktionen gehört.

So wie bisher in Abschnitt 4 betrachten wir

- eine Zerlegung $\Theta = \Theta_0 \cup \Theta_1$ des Parameterraumes in zwei disjunkte Teilmengen Θ_0, Θ_1 und
- die zugehörigen Hypothesen $H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1$.
- Außerdem betrachten wir für jedes $n = 1, 2, \dots$ eine Borel-Menge $K_n \subset \mathbb{R}^n$ und
- eine Stichprobenfunktion $\varphi_n : \mathbb{R}^n \rightarrow \{0, 1\}$ mit

$$\varphi_n(x_1, \dots, x_n) = \begin{cases} 1, & \text{falls } (x_1, \dots, x_n) \in K_n, \\ 0, & \text{falls } (x_1, \dots, x_n) \notin K_n. \end{cases}$$

Definition 4.6 Sei $\alpha \in (0, 1)$ ein beliebige (vorgegebene) Zahl. Dann sagt man, dass durch die Folge von Stichprobenfunktionen $\{\varphi_n\}$ ein *asymptotischer Parametertest* zum Niveau α gegeben ist, falls

$$\lim_{n \rightarrow \infty} P_\theta(\varphi_n(X_1, \dots, X_n) = 1) \leq \alpha$$

für jedes $\theta \in \Theta_0$ gilt.

Beispiel *Test des Erwartungswertes λ bei Poisson-Verteilung*

- Wir nehmen an, dass $X_i \sim \text{Poi}(\lambda)$ für ein (unbekanntes) $\lambda > 0$.
- Für einen vorgegebenen (hypothetischen) Zahlenwert $\lambda_0 > 0$ soll die Hypothese $H_0 : \lambda = \lambda_0$ (gegen die Alternative $H_1 : \lambda \neq \lambda_0$) getestet werden, d.h. $\Theta = \{\lambda : \lambda > 0\}$ mit

$$\Theta_0 = \{\lambda_0\} \quad \text{bzw.} \quad \Theta_1 = \{\lambda : \lambda \neq \lambda_0\}$$

- Wir betrachten die Teststatistik $T_n : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$T_n(x_1, \dots, x_n) = \begin{cases} \sqrt{n} \frac{\bar{x}_n - \lambda_0}{\sqrt{\bar{x}_n}} & \text{falls } \bar{x}_n > 0 \\ 0, & \text{falls } \bar{x}_n = 0 \end{cases} \quad (80)$$

und den Schwellenwert $c = z_{1-\alpha/2}$.

- Dann gilt (vgl. die Überlegungen im ersten Beispiel des Abschnittes 3.6)

$$\lim_{n \rightarrow \infty} P_{\lambda_0}(|T_n(X_1, \dots, X_n)| > z_{1-\alpha/2}) = \alpha.$$

- Durch die in (80) eingeführte Folge $\{T_n\}$ von Testfunktionen ist also ein asymptotischer Test des Erwartungswertes λ zum Niveau α gegeben.
- Dabei wird bei der praktischen Anwendung dieses Tests (bei großem Stichprobenumfang n) die Hypothese $H_0 : \lambda = \lambda_0$ abgelehnt, falls $|T_n(x_1, \dots, x_n)| > z_{1-\alpha/2}$.

4.3 Kolmogorow-Smirnow-Test; χ^2 -Anpassungstest

In diesem Abschnitt setzen wir *nicht* voraus, dass die Familie Δ der insgesamt in Betracht gezogenen Verteilungsfunktionen F der Stichprobenvariablen X_i eine parametrische Familie von Verteilungsfunktionen ist.

- Dabei soll die Hypothese $H_0 : F = F_0$ getestet werden, dass die Verteilungsfunktion F der Stichprobenvariablen X_i gleich einer vorgegebenen (hypothetischen) Verteilungsfunktion F_0 ist (gegen die Alternativhypothese $H_1 : F \neq F_0$).
- Zur Überprüfung der Hypothesen

$$H_0 : F = F_0 \quad \text{vs.} \quad H_1 : F \neq F_0 \quad (81)$$

werden in der Literatur verschiedene (asymptotische) Tests vorgeschlagen.

Beispiel *Kolmogorow-Smirnow-Test*

- Es soll getestet werden, ob die unbekannte Verteilungsfunktion F gleich einer vorgegebenen (hypothetischen) Verteilungsfunktion F_0 ist, wobei F_0 stetig sei. Falls F_0 nicht stetig ist, so ist der Test konservativ, hält also länger an der Nullhypothese fest als notwendig.
- Man testet also die Hypothese

$$H_0 : F = F_0$$

gegen die Alternative

$$H_1 : F \neq F_0.$$

- Der Kolmogorow-Smirnow-Test basiert auf der Untersuchung der in (16) eingeführten empirischen Verteilungsfunktion $\hat{F}_n : \mathbb{R} \times \mathbb{R}^n \rightarrow [0, 1]$, wobei

$$\hat{F}_n(x; x_1, \dots, x_n) = \frac{1}{n} |\{i : 1 \leq i \leq n, x_i \leq x\}|.$$

- Dabei wird die Teststatistik $\sqrt{n}T_n : \mathbb{R}^n \rightarrow [0, \infty)$ betrachtet mit

$$T_n(x_1, \dots, x_n) = \sup_{x \in \mathbb{R}} |\hat{F}_n(x; x_1, \dots, x_n) - F_0(x)|.$$

- Unter H_0 gilt nach dem Satz von Gliwenko/Cantelli (vgl. Theorem 1.11), dass

$$P_{F_0} \left(\lim_{n \rightarrow \infty} T_n(X_1, \dots, X_n) = 0 \right) = 1,$$

d. h. unter H_0 konvergiert $T_n(X_1, \dots, X_n)$ fast sicher gegen null.

- Also sollte H_0 für große Abstände $\sqrt{n}T_n(x_1, \dots, x_n)$ verworfen werden.
- Man kann außerdem zeigen, dass die Verteilung von $T_n(X_1, \dots, X_n)$ unabhängig von der hypothetischen Verteilung F_0 ist.
- Für kleine Stichprobenumfänge n können die entsprechenden Quantile, die zur Verwerfung von H_0 führen, in Tabellen nachgeschaut werden.
- Für große Stichprobenumfänge n kann folgender Grenzwertsatz angewandt werden (vgl. Theorem 1.12):

$$\sqrt{n}T_n(X_1, \dots, X_n) \xrightarrow{d} D \text{ für } n \rightarrow \infty, \quad (82)$$

wobei D eine Zufallsvariable ist, deren Verteilungsfunktion gegeben ist durch

$$P(D \leq y) = \begin{cases} \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2 y^2) & \text{für } y \geq 0, \\ 0 & \text{für } y < 0. \end{cases} \quad (83)$$

- Die Hypothese $H_0 : F = F_0$ wird abgelehnt, wenn $\sqrt{n}T_n(x_1, \dots, x_n) > \xi_{1-\alpha}$, wobei $\xi_{1-\alpha}$ das $(1 - \alpha)$ -Quantil der Verteilungsfunktion in (83) bezeichnet.
- Der Kolmogorow-Smirnow-Test hat jedoch den Nachteil, dass die Bestimmung der empirischen Verteilungsfunktion \hat{F}_n relativ aufwendig ist, wenn der Stichprobenumfang n groß ist.

Wir vereinfachen deshalb die ursprüngliche Fragestellung (81) auf die folgende Weise.

- Zerlegen den Wertebereich der Stichprobenvariablen X_i in r Klassen $(a_1, b_1], \dots, (a_r, b_r]$ mit

$$-\infty \leq a_1 < b_1 = a_2 < b_2 = \dots = a_r < b_r \leq \infty,$$

wobei r eine (hinreichend große) natürliche Zahl ist.

- Anstelle der Zufallsstichprobe (X_1, \dots, X_n) betrachten wir den Zufallsvektor (Y_1, \dots, Y_r) der 'Klassenstärken', wobei

$$Y_j = |\{i : 1 \leq i \leq n, a_j < X_i \leq b_j\}|, \quad j = 1, \dots, r. \quad (84)$$

Lemma 4.7 Der in (84) definierte Zufallsvektor (Y_1, \dots, Y_r) ist *multinomialverteilt* mit den Parametern $n \in \mathbb{N}$ und $p = (p_1, \dots, p_{r-1}) \in [0, 1]^{r-1}$, d.h., für beliebige Zahlen $k_1, \dots, k_r \in \mathbb{N}$ mit $k_1 + \dots + k_r = n$ gilt

$$P(Y_1 = k_1, \dots, Y_r = k_r) = \frac{n!}{k_1! \cdot \dots \cdot k_r!} p_1^{k_1} \cdot \dots \cdot p_r^{k_r}, \quad (85)$$

wobei $p_j = F(b_j) - F(a_j)$ für jedes $j = 1, \dots, r$.

Beweis Weil die Stichprobenvariablen X_1, \dots, X_n unabhängig und identisch verteilt sind (mit der Verteilungsfunktion F), gilt

$$P(X_1 \in (a_{i_1}, b_{i_1}], \dots, X_n \in (a_{i_n}, b_{i_n}]) = \prod_{j=1}^n F(b_{i_j}) - F(a_{i_j})$$

für jede Folge von Intervallen $(a_{i_1}, b_{i_1}], \dots, (a_{i_n}, b_{i_n}]$. Hieraus ergibt sich (85) durch Permutation der Stichprobenvariablen.

Beachte

- Die Multinomialverteilung mit den Parametern $n \in \mathbb{N}$ und $p = (p_1, \dots, p_{r-1}) \in [0, 1]^{r-1}$ bezeichnen wir mit $M_{r-1}(n, p)$.
- Für den Parametervektor $p = (p_1, \dots, p_{r-1}) \in [0, 1]^{r-1}$ der Multinomialverteilung $M_{r-1}(n, p)$ gilt

$$0 \leq p_r = 1 - \sum_{i=1}^{r-1} p_i \leq 1,$$

d.h., die Wahrscheinlichkeit p_r wird durch $p = (p_1, \dots, p_{r-1})$ eindeutig bestimmt.

- Mit anderen Worten: p_r ist *keine* unabhängig von p wählbare Parameterkomponente.
- Man kann sich leicht überlegen, dass für $r = 2$ die Multinomialverteilung $M_1(n, p_1)$ mit der Binomialverteilung $\text{Bin}(n, p_1)$ übereinstimmt.

Anstelle die ursprüngliche Fragestellung (81) zu untersuchen, testen wir nun

- die Hypothese $H_0 : p = p_0$ (gegen die Alternative $H_1 : p \neq p_0$) für einen vorgegebenen (hypothetischen) Vektor $p_0 = (p_{0,1}, \dots, p_{0,r-1}) \in (0, 1)^{r-1}$ mit $\sum_{i=1}^{r-1} p_{0,i} < 1$.
- Wir zerlegen also die Familie Δ der insgesamt in Betracht gezogenen Verteilungsfunktionen F in die Teilmengen

$$\Delta_0 = \{F : F(b_j) - F(a_j) = p_{0,j} \forall j = 1, \dots, r-1\} \quad \text{bzw.} \quad \Delta_1 = \Delta \setminus \Delta_0.$$

- Wir betrachten die Teststatistik $T_n : \mathbb{R}^n \rightarrow [0, \infty)$ mit

$$T_n(x_1, \dots, x_n) = \sum_{j=1}^r \frac{(y_j(x_1, \dots, x_n) - np_{0,j})^2}{np_{0,j}}, \quad (86)$$

wobei $y_j(x_1, \dots, x_n)$ die Anzahl derjenigen Stichprobenwerte x_1, \dots, x_n bezeichnet, die im Intervall $(a_j, b_j]$ liegen.

- Mit Hilfe von Lemma 4.7 kann man zeigen, dass

$$\lim_{n \rightarrow \infty} P_F(T_n(X_1, \dots, X_n) > \chi_{r-1, 1-\alpha}^2) = \alpha \quad (87)$$

für jede Verteilungsfunktion F der Stichprobenvariablen X_i gilt, für die $p_{0,j} = F(b_j) - F(a_j)$ für jedes $j = 1, \dots, r$.

- Bei großem n wird die Hypothese $H_0 : p = p_0$ abgelehnt, falls

$$T_n(x_1, \dots, x_n) > \chi_{r-1, 1-\alpha}^2. \quad (88)$$

Beachte

- Als der „Entdecker“ des Grenzwertsatzes (87) gilt der britische Statistiker Karl Pearson (1857–1936).
- Der asymptotische Test (88), der auf dem Grenzwertsatz (87) beruht, heißt χ^2 -Anpassungstest.

Wie das folgende Beispiel zeigt, kann der χ^2 -Anpassungstest zur Überprüfung der Güte von Zufallszahlengeneratoren bei der *Monte-Carlo-Simulation* verwendet werden.

Beispiel Test auf Gleichverteilung

- Wir nehmen an, dass die Daten x_1, \dots, x_n sogenannte *Pseudozufallszahlen* sind, die von einem Zufallszahlengenerator erzeugt worden sind.
- Dabei wollen wir prüfen, ob die Stichprobe (x_1, \dots, x_n) als Realisierung von Stichprobenvariablen X_1, \dots, X_n angesehen werden kann, die gleichverteilt im Intervall $(0, 1]$ sind.
- Wir zerlegen das Intervall $(0, 1]$ in r gleichlange Teilintervalle $(0, 1/r], \dots, ((r-1)/r, 1]$ und
- betrachten den $(r-1)$ -dimensionalen (hypothetischen) Vektor $p_0 = (1/r, \dots, 1/r)$ bzw.
- die Teststatistik $T_n : \mathbb{R}^n \rightarrow [0, \infty)$ mit

$$T_n(x_1, \dots, x_n) = \sum_{j=1}^r \frac{(y_j(x_1, \dots, x_n) - n/r)^2}{n/r},$$

wobei $y_j(x_1, \dots, x_n) = |\{i : 1 \leq i \leq n, j-1 < rx_i \leq j\}|$ die Anzahl derjenigen Pseudozufallszahlen x_1, \dots, x_n bezeichnet, die im Intervall $((j-1)/r, j/r]$ liegen.

- Bei großem n wird die Hypothese $H_0 : p = p_0$ abgelehnt, falls $T(x_1, \dots, x_n) > \chi_{r-1, 1-\alpha}^2$.

Für $\alpha = 0.05$, $n = 100\,000$ und $r = 10$ wollen wir nun prüfen, ob

- die Hypothese der Gleichverteilung mit einer (konkreten) Stichprobe $(x_1, \dots, x_{100\,000})$ von Pseudozufallszahlen vereinbar ist, für die sich der folgende Vektor der Klassenhäufigkeiten (y_1, \dots, y_{10}) ergibt:

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}
9 995	10 045	10 127	9 816	10 130	10 040	9 890	9 858	10 083	10 016

- In diesem Fall ist $T_{100\,000}(x_1, \dots, x_{100\,000}) = 10.99$, und es gilt somit

$$T_{100\,000}(x_1, \dots, x_{100\,000}) = 10.99 < \chi_{9,0.95}^2 = 16.92,$$

wobei das Quantil $\chi_{9,0.95}^2 = 16.92$ aus Tabelle 2 entnommen wurde.

- Die Hypothese der Gleichverteilung wird also nicht verworfen.

Bei der Überprüfung der Güte von Zufallszahlengeneratoren sind nicht nur Tests auf Gleichverteilung von Interesse. Ein weiteres Gütekriterium besteht darin, ob die von einem Zufallszahlengenerator erzeugten Pseudozufallszahlen x_1, x_2, \dots als Realisierungen *unabhängiger* Zufallsvariablen X_1, X_2, \dots angesehen werden können.

Der folgende Test auf Unabhängigkeit führt erneut zur Konstruktion eines χ^2 -Anpassungstests.

Beispiel *Test auf Unabhängigkeit (Run-Test)*

- Wir wollen nun prüfen, ob die Pseudozufallszahlen x_1, x_2, \dots als Realisierungen von unabhängigen Zufallsvariablen X_1, X_2, \dots angesehen werden können, die gleichverteilt im Intervall $(0, 1]$ sind.
- Hierfür definieren wir die Zufallsvariablen U_1, U_2, \dots rekursiv durch

$$U_{j+1} = \min\{i : i > U_j + 1, X_i > X_{i+1}\}, \quad (89)$$

wobei $U_1 = \min\{i : i \geq 1, X_i > X_{i+1}\}$.

- Die Zufallsvariablen V_1, V_2, \dots mit

$$V_1 = U_1, \quad V_{j+1} = U_{j+1} - U_j - 1 \quad \text{für } j = 1, 2, \dots \quad (90)$$

werden dann die *Runs* der Folge X_1, X_2, \dots genannt.

- Man kann zeigen, dass die in (90) eingeführten Zufallsvariablen V_1, V_2, \dots unabhängig und identisch verteilt sind mit

$$P(V_j = k) = \frac{k}{(k+1)!} \quad k = 1, 2, \dots, \quad (91)$$

falls die Zufallsvariablen X_1, X_2, \dots unabhängig und (identisch) gleichverteilt im Intervall $(0, 1]$ sind.

- Es mögen nun hinreichend viele Pseudozufallszahlen x_1, x_2, \dots erzeugt werden, so dass sich aus ihnen gemäß (89) und (90) die n Runs v_1, \dots, v_n ergeben.

- Wir zerlegen die positive Halbachse in r Intervalle $(a_1, b_1], \dots, (a_r, b_r]$, so dass
- die Wahrscheinlichkeiten

$$p_{0,j} = \sum_{k \in \mathbb{N} \cap (a_j, b_j]} \frac{k}{(k+1)!}, \quad j = 1, \dots, r$$

annähernd gleich sind, und

- für diese Wahrscheinlichkeiten betrachten wir den $(r-1)$ -dimensionalen (hypothetischen) Vektor $p_0 = (p_{0,1}, \dots, p_{0,r-1})$ bzw.
- die Teststatistik $T_n : \mathbb{R}^n \rightarrow [0, \infty)$ mit

$$T_n(v_1, \dots, v_n) = \sum_{j=1}^r \frac{(y_j(v_1, \dots, v_n) - np_{0,j})^2}{np_{0,j}},$$

wobei $y_j(v_1, \dots, v_n) = |\{i : 1 \leq i \leq n, j-1 < rv_i \leq j\}|$ die Anzahl derjenigen Runlängen v_1, \dots, v_n bezeichnet, die in der j -ten Klasse liegen.

- Bei großem n , was die Erzeugung entsprechend vieler Pseudozufallszahlen x_1, x_2, \dots erfordert, wird die Hypothese $H_0 : p = p_0$ abgelehnt, falls $T(v_1, \dots, v_n) > \chi_{r-1, 1-\alpha}^2$.

5 Tabellen für Verteilungsfunktionen und Quantile

Tabelle 1 Verteilungsfunktion $\Phi(x)$ der Standardnormalverteilung

x	0	1	2	3	4	5	6	7	8	9
0,0	0,500000	0,503989	0,507978	0,511967	0,515953	0,519939	0,523922	0,527903	0,531881	0,535856
0,1	0,539828	0,543795	0,547758	0,551717	0,555670	0,559618	0,563559	0,567495	0,571424	0,575345
0,2	0,579260	0,583166	0,587064	0,590954	0,594835	0,598706	0,602568	0,606420	0,610261	0,614092
0,3	0,617911	0,621719	0,625516	0,629300	0,633072	0,636831	0,640576	0,644309	0,648027	0,651732
0,4	0,655422	0,659097	0,662757	0,666402	0,670031	0,673645	0,677242	0,680822	0,684386	0,687933
0,5	0,691462	0,694974	0,698468	0,701944	0,705402	0,708840	0,712260	0,715661	0,719043	0,722405
0,6	0,725747	0,729069	0,732371	0,735653	0,738914	0,742154	0,745373	0,748571	0,751748	0,754903
0,7	0,758036	0,761148	0,764238	0,767305	0,770350	0,773373	0,776373	0,779350	0,782305	0,785236
0,8	0,788145	0,791030	0,793892	0,796731	0,799546	0,802338	0,805106	0,807850	0,810570	0,813267
0,9	0,815940	0,818589	0,821214	0,823814	0,826391	0,828944	0,831472	0,833977	0,836457	0,838913
1,0	0,841345	0,843752	0,846136	0,848495	0,850830	0,853141	0,855428	0,857690	0,859929	0,862143
1,1	0,864334	0,866500	0,868643	0,870762	0,872857	0,874928	0,876976	0,878999	0,881000	0,882977
1,2	0,884930	0,886860	0,888767	0,890651	0,892512	0,894350	0,896165	0,897958	0,899727	0,901475
1,3	0,903199	0,904902	0,906582	0,908241	0,909877	0,911492	0,913085	0,914656	0,916207	0,917736
1,4	0,919243	0,920730	0,922196	0,923641	0,925066	0,926471	0,927855	0,929219	0,930563	0,931888
1,5	0,933193	0,934478	0,935744	0,936992	0,938220	0,939429	0,940620	0,941792	0,942947	0,944083
1,6	0,945201	0,946301	0,947384	0,948449	0,949497	0,950529	0,951543	0,952540	0,953521	0,954486
1,7	0,955435	0,956367	0,957284	0,958185	0,959071	0,959941	0,960796	0,961636	0,962462	0,963273
1,8	0,964070	0,964852	0,965621	0,966375	0,967116	0,967843	0,968557	0,969258	0,969946	0,970621
1,9	0,971284	0,971933	0,972571	0,973197	0,973810	0,974412	0,975002	0,975581	0,976148	0,976705
2,0	0,977250	0,977784	0,978308	0,978822	0,979325	0,979818	0,980301	0,980774	0,981237	0,981691
2,1	0,982136	0,982571	0,982997	0,983414	0,983823	0,984222	0,984614	0,984997	0,985371	0,985738
2,2	0,986097	0,986447	0,986791	0,987126	0,987455	0,987776	0,988089	0,988396	0,988696	0,988989
2,3	0,989276	0,989556	0,989830	0,990097	0,990358	0,990613	0,990863	0,991106	0,991344	0,991576
2,4	0,991802	0,992024	0,992240	0,992451	0,992656	0,992857	0,993053	0,993244	0,993431	0,993613
2,5	0,993790	0,993963	0,994132	0,994297	0,994457	0,994614	0,994766	0,994915	0,995060	0,995201
2,6	0,995339	0,995473	0,995603	0,995731	0,995855	0,995975	0,996093	0,996207	0,996319	0,996427
2,7	0,996533	0,996636	0,996736	0,996833	0,996928	0,997020	0,997110	0,997197	0,997282	0,997365
2,8	0,997445	0,997523	0,997599	0,997673	0,997744	0,997814	0,997882	0,997948	0,998012	0,998074
2,9	0,998134	0,998193	0,998250	0,998305	0,998359	0,998411	0,998462	0,998511	0,998559	0,998605
3,0	0,998650	0,998694	0,998736	0,998777	0,998817	0,998856	0,998893	0,998930	0,998965	0,998999
3,5	0,999767	0,999776	0,999784	0,999792	0,999800	0,999807	0,999815	0,999821	0,999828	0,999835
4,0	0,999968	0,999970	0,999971	0,999972	0,999973	0,999974	0,999975	0,999976	0,999977	0,999978

Tabelle 2 γ -Quantil $\chi_{r,\gamma}^2$ der χ^2 -Verteilung mit r Freiheitsgraden

r \ γ	0,7	0,75	0,8	0,85	0,9	0,95	0,975	0,99	0,995
1	1,07	1,32	1,64	2,07	2,71	3,84	5,02	6,63	7,88
2	2,41	2,77	3,22	3,79	4,61	5,99	7,38	9,21	10,60
3	3,66	4,11	4,64	5,32	6,25	7,81	9,35	11,34	12,84
4	4,88	5,39	5,99	6,74	7,78	9,49	11,14	13,28	14,86
5	6,06	6,63	7,29	8,12	9,24	11,07	12,83	15,09	16,75
6	7,23	7,84	8,56	9,45	10,64	12,59	14,45	16,81	18,55
7	8,38	9,04	9,80	10,75	12,02	14,07	16,01	18,48	20,28
8	9,52	10,22	11,03	12,03	13,36	15,51	17,53	20,09	21,95
9	10,66	11,39	12,24	13,29	14,68	16,92	19,02	21,67	23,59
10	11,78	12,55	13,44	14,53	15,99	18,31	20,48	23,21	25,19
11	12,90	13,70	14,63	15,77	17,28	19,68	21,92	24,73	26,76
12	14,01	14,85	15,81	16,99	18,55	21,03	23,34	26,22	28,30
13	15,12	15,98	16,98	18,20	19,81	22,36	24,74	27,69	29,82
14	16,22	17,12	18,15	19,41	21,06	23,68	26,12	29,14	31,32
15	17,32	18,25	19,31	20,60	22,31	25,00	27,49	30,58	32,80
16	18,42	19,37	20,47	21,79	23,54	26,30	28,85	32,00	34,27
17	19,51	20,49	21,61	22,98	24,77	27,59	30,19	33,41	35,72
18	20,60	21,60	22,76	24,16	25,99	28,87	31,53	34,81	37,16
19	21,69	22,72	23,90	25,33	27,20	30,14	32,85	36,19	38,58
20	22,77	23,83	25,04	26,50	28,41	31,41	34,17	37,57	40,00
21	23,86	24,93	26,17	27,66	29,62	32,67	35,48	38,93	41,40
22	24,94	26,04	27,30	28,82	30,81	33,92	36,78	40,29	42,80
23	26,02	27,14	28,43	29,98	32,01	35,17	38,08	41,64	44,18
24	27,10	28,24	29,55	31,13	33,20	36,42	39,36	42,98	45,56
25	28,17	29,34	30,68	32,28	34,38	37,65	40,65	44,31	46,93
30	33,53	34,80	36,25	37,99	40,26	43,77	46,98	50,89	53,67
40	44,16	45,62	47,27	49,24	51,81	55,76	59,34	63,69	66,77
50	54,72	56,33	58,16	60,35	63,17	67,50	71,42	76,15	79,49
60	65,23	66,98	68,97	71,34	74,40	79,08	83,30	88,38	91,95
70	75,69	77,58	79,71	82,26	85,53	90,53	95,02	100,43	104,21
80	86,12	88,13	90,41	93,11	96,58	101,88	106,63	112,33	116,32
90	96,52	98,65	101,05	103,90	107,57	113,15	118,14	124,12	128,30
100	106,91	109,14	111,67	114,66	118,50	124,34	129,56	135,81	140,17
150	158,58	161,29	164,35	167,96	172,58	179,58	185,80	193,21	198,36
200	209,99	213,10	216,61	220,74	226,02	233,99	241,06	249,45	255,26
500	516,09	520,95	526,40	532,80	540,93	553,13	563,85	576,49	585,21

Tabelle 3 γ -Quantil $t_{r,\gamma}$ der t-Verteilung mit r Freiheitsgraden

$r \setminus \gamma$	0,65	0,7	0,75	0,8	0,85	0,9	0,95	0,975	0,99	0,995
1	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,656
2	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925
3	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841
4	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604
5	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032
6	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707
7	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499
8	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355
9	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250
10	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169
11	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106
12	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055
13	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012
14	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977
15	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947
16	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921
17	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898
18	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878
19	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861
20	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845
21	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831
22	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819
23	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807
24	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797
25	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787
30	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750
40	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704
50	0,388	0,528	0,679	0,849	1,047	1,299	1,676	2,009	2,403	2,678
60	0,387	0,527	0,679	0,848	1,045	1,296	1,671	2,000	2,390	2,660
70	0,387	0,527	0,678	0,847	1,044	1,294	1,667	1,994	2,381	2,648
80	0,387	0,526	0,678	0,846	1,043	1,292	1,664	1,990	2,374	2,639
90	0,387	0,526	0,677	0,846	1,042	1,291	1,662	1,987	2,368	2,632
100	0,386	0,526	0,677	0,845	1,042	1,290	1,660	1,984	2,364	2,626
150	0,386	0,526	0,676	0,844	1,040	1,287	1,655	1,976	2,351	2,609
200	0,386	0,525	0,676	0,843	1,039	1,286	1,653	1,972	2,345	2,601
500	0,386	0,525	0,675	0,842	1,038	1,283	1,648	1,965	2,334	2,586
1000	0,385	0,525	0,675	0,842	1,037	1,282	1,646	1,962	2,330	2,581

Tabelle 4a γ -Quantil $F_{r,s,\gamma}$ der F-Verteilung mit (r, s) Freiheitsgraden

$s \setminus r$	1	2	3	4	5	6	7	8	9	10	11	12	γ
1	161	200	216	225	230	234	237	239	241	242	243	244	0,95
	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6082	6106	0,99
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,40	19,41	0,95
	98,49	99,00	99,17	99,25	99,30	99,33	99,34	99,36	99,38	99,40	99,41	99,42	0,99
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,76	8,74	0,95
	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,13	27,05	26,92	0,99
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91	0,95
	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,45	14,37	0,99
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,70	4,68	0,95
	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,96	9,89	0,99
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	0,95
	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,79	7,72	0,99
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,60	3,57	0,95
	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,54	6,47	0,99
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,31	3,28	0,95
	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,74	5,67	0,99
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,10	3,07	0,95
	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	5,18	5,11	0,99
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,94	2,91	0,95
	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,78	4,71	0,99
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,82	2,79	0,95
	9,65	7,20	6,22	5,67	5,32	5,07	4,88	4,74	4,63	4,54	4,46	4,40	0,99
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,72	2,69	0,95
	9,33	6,93	5,95	5,41	5,06	4,82	4,65	4,50	4,39	4,30	4,22	4,16	0,99

Tabelle 4b γ -Quantil $F_{r,s,\gamma}$ der F-Verteilung mit (r, s) Freiheitsgraden

$s \setminus r$	14	16	20	24	30	40	50	75	100	200	500	∞	γ
1	245	246	248	249	250	251	252	253	253	254	254	254	0,95
	6142	6169	6208	6234	6258	6286	6302	6323	6334	6352	6361	6366	0,99
2	19,42	19,43	19,44	19,45	19,46	19,47	19,47	19,48	19,49	19,49	19,50	19,50	0,95
	99,43	99,44	99,45	99,46	99,47	99,48	99,48	99,49	99,49	99,49	99,50	99,50	0,99
3	8,71	8,69	8,66	8,64	8,62	8,60	8,58	8,57	8,56	8,54	8,54	8,53	0,95
	26,83	26,69	26,60	26,50	26,41	26,35	26,27	26,23	26,18	26,14	26,12	34,12	0,99
4	5,87	5,84	5,80	5,77	5,74	5,71	5,70	5,68	5,66	5,65	5,64	5,63	0,95
	14,24	14,15	14,02	13,93	13,83	13,74	13,69	13,61	13,57	13,52	13,48	13,46	0,99
5	4,64	4,60	4,56	4,53	4,50	4,46	4,44	4,42	4,40	4,38	4,37	4,36	0,95
	9,77	9,68	9,55	9,47	9,38	9,29	9,24	9,17	9,13	9,07	9,04	9,02	0,99
6	3,96	3,92	3,87	3,84	3,81	3,77	3,75	3,72	3,71	3,69	3,68	3,67	0,95
	7,60	7,52	7,39	7,31	7,23	7,14	7,09	7,02	6,99	6,94	6,90	6,88	0,99
7	3,52	3,49	3,44	3,41	3,38	3,34	3,32	3,29	3,28	3,25	3,24	3,23	0,95
	6,35	6,27	6,15	6,07	5,98	5,90	5,85	5,78	5,75	5,70	5,67	5,65	0,99
8	3,23	3,20	3,15	3,12	3,08	3,05	3,03	3,00	2,98	2,96	2,94	2,93	0,95
	5,56	5,48	5,36	5,28	5,20	5,11	5,06	5,00	4,96	4,91	4,88	4,86	0,99
9	3,02	2,98	2,93	2,90	2,86	2,82	2,80	2,77	2,76	2,73	2,72	2,71	0,95
	5,00	4,92	4,80	4,73	4,64	4,56	4,51	4,45	4,41	4,36	4,33	4,31	0,99
10	2,86	2,82	2,77	2,74	2,70	2,67	2,64	2,61	2,59	2,56	2,55	2,54	0,95
	4,60	4,52	4,41	4,33	4,25	4,17	4,12	4,05	4,01	3,96	3,93	3,91	0,99
11	2,74	2,70	2,65	2,61	2,57	2,53	2,50	2,47	2,45	2,42	2,41	2,40	0,95
	4,29	4,21	4,10	4,02	3,94	3,86	3,80	3,74	3,70	3,66	3,62	3,60	0,99
12	2,64	2,60	2,54	2,50	2,46	2,42	2,40	2,36	2,35	2,32	2,31	2,30	0,95
	4,05	3,98	3,86	3,78	3,70	3,61	3,56	3,49	3,46	3,41	3,38	3,36	0,99

Tabelle 4c γ -Quantil $F_{r,s,\gamma}$ der F-Verteilung mit (r, s) Freiheitsgraden

$s \setminus r$	1	2	3	4	5	6	7	8	9	10	11	12	γ
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,63	2,60	0,95
	9,07	6,70	5,74	5,20	4,86	4,62	4,44	4,30	4,19	4,10	4,02	3,96	0,99
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,56	2,53	0,95
	8,86	6,51	5,56	5,03	4,69	4,46	4,28	4,14	4,03	3,94	3,86	3,80	0,99
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,51	2,48	0,95
	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,73	3,67	0,99
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42	0,95
	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,61	3,55	0,99
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,41	2,38	0,95
	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,52	3,45	0,99
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	0,95
	8,28	6,01	5,09	4,58	4,25	4,01	3,85	3,71	3,60	3,51	3,44	3,37	0,99
19	4,38	3,51	3,13	2,90	2,74	2,63	2,55	2,48	2,43	2,38	2,34	2,31	0,95
	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,36	3,30	0,99
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	2,31	2,28	0,95
	8,10	5,85	4,94	4,43	4,10	3,87	3,71	3,56	3,45	3,37	3,30	3,23	0,99
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	0,95
	8,02	5,78	4,87	4,37	4,04	3,81	3,65	3,51	3,40	3,31	3,24	3,17	0,99
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35	2,30	2,26	2,23	0,95
	7,94	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,18	3,12	0,99
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32	2,28	2,24	2,20	0,95
	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,14	3,07	0,99
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30	2,26	2,22	2,18	0,95
	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,25	3,17	3,09	3,03	0,99
25	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,28	2,24	2,20	2,16	0,95
	7,77	5,57	4,68	4,18	3,86	3,63	3,46	3,32	3,21	3,13	3,05	2,99	0,99

Tabelle 4d γ -Quantil $F_{r,s,\gamma}$ der F-Verteilung mit (r, s) Freiheitsgraden

$s \setminus r$	14	16	20	24	30	40	50	75	100	200	500	∞	γ
13	2,55	2,51	2,46	2,42	2,38	2,34	2,32	2,28	2,26	2,24	2,22	2,21	0,95
	3,85	3,78	3,67	3,59	3,51	3,42	3,37	3,30	3,27	3,21	3,18	3,16	0,99
14	2,48	2,44	2,39	2,35	2,31	2,27	2,24	2,21	2,19	2,16	2,14	2,13	0,95
	3,70	3,62	3,51	3,43	3,34	3,26	3,21	3,14	3,11	3,06	3,02	3,00	0,99
15	2,43	2,39	2,33	2,29	2,25	2,21	2,18	2,15	2,12	2,10	2,08	2,07	0,95
	3,56	3,48	3,36	3,29	3,20	3,12	3,07	3,00	2,97	2,92	2,89	2,87	0,99
16	2,37	2,33	2,28	2,24	2,20	2,16	2,13	2,09	2,07	2,04	2,02	2,01	0,95
	3,45	3,37	3,25	3,18	3,10	3,01	2,96	2,89	2,86	2,80	2,77	2,75	0,99
17	2,33	2,29	2,23	2,19	2,15	2,11	2,08	2,04	2,02	1,99	1,97	1,96	0,95
	3,35	3,27	3,16	3,08	3,00	2,92	2,86	2,79	2,76	2,70	2,67	2,65	0,99
18	2,29	2,25	2,19	2,15	2,11	2,07	2,04	2,00	1,98	1,95	1,93	1,92	0,95
	3,27	3,19	3,07	3,00	2,91	2,83	2,78	2,71	2,68	2,62	2,59	2,57	0,99
19	2,26	2,21	2,15	2,11	2,07	2,02	2,00	1,96	1,94	1,91	1,90	1,88	0,95
	3,19	3,12	3,00	2,92	2,84	2,76	2,70	2,63	2,60	2,54	2,51	2,49	0,99
20	2,23	2,18	2,12	2,08	2,04	1,99	1,96	1,92	1,90	1,87	1,85	1,84	0,95
	3,13	3,05	2,94	2,86	2,77	2,69	2,63	2,56	2,53	2,47	2,44	2,42	0,99
21	2,20	2,15	2,09	2,05	2,00	1,96	1,93	1,89	1,87	1,84	1,82	1,81	0,95
	3,07	2,99	2,88	2,80	2,72	2,63	2,58	2,51	2,47	2,42	2,38	2,36	0,99
22	2,18	2,13	2,07	2,03	1,98	1,93	1,91	1,87	1,84	1,81	1,80	1,78	0,95
	3,02	2,94	2,83	2,75	2,67	2,58	2,53	2,46	2,42	2,37	2,33	2,31	0,99
23	2,14	2,10	2,05	2,00	1,96	1,91	1,88	1,84	1,82	1,79	1,77	1,76	0,95
	2,97	2,89	2,78	2,70	2,62	2,53	2,48	2,41	2,37	2,32	2,28	2,26	0,99
24	2,13	2,09	2,02	1,98	1,94	1,89	1,86	1,82	1,80	1,76	1,74	1,73	0,95
	2,93	2,85	2,74	2,66	2,58	2,49	2,44	2,36	2,33	2,27	2,23	2,21	0,99
25	2,11	2,06	2,00	1,96	1,92	1,87	1,84	1,80	1,77	1,74	1,72	1,71	0,95
	2,89	2,81	2,70	2,62	2,54	2,45	2,40	2,32	2,29	2,23	2,19	2,17	0,99

Tabelle 4e γ -Quantil $F_{r,s,\gamma}$ der F-Verteilung mit (r, s) Freiheitsgraden

$s \setminus r$	1	2	3	4	5	6	7	8	9	10	11	12	γ
26	4,22	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15	0,95
	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,17	3,09	3,02	2,96	0,99
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25	2,20	2,16	2,13	0,95
	7,68	5,49	4,60	4,11	3,79	3,56	3,39	3,26	3,14	3,06	2,98	2,93	0,99
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24	2,19	2,15	2,12	0,95
	7,64	5,45	4,57	4,07	3,76	3,53	3,36	3,23	3,11	3,03	2,95	2,90	0,99
29	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22	2,18	2,14	2,10	0,95
	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,08	3,00	2,92	2,87	0,99
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	2,12	2,09	0,95
	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,06	2,98	2,90	2,84	0,99
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,07	2,04	2,00	0,95
	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,88	2,80	2,73	2,66	0,99
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,02	1,98	1,95	0,95
	7,17	5,06	4,20	3,72	3,41	3,18	3,02	2,88	2,78	2,70	2,62	2,56	0,99
60	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	0,95
	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,56	2,50	0,99
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,01	1,97	1,93	1,89	0,95
	7,01	4,92	4,08	3,60	3,29	3,07	2,91	2,77	2,67	2,59	2,51	2,45	0,99
80	3,96	3,11	2,72	2,48	2,33	2,21	2,12	2,05	1,99	1,95	1,91	1,88	0,95
	6,96	4,88	4,04	3,56	3,25	3,04	2,87	2,74	2,64	2,55	2,48	2,41	0,99
100	3,94	3,09	2,70	2,46	2,30	2,19	2,10	2,03	1,97	1,92	1,88	1,85	0,95
	6,90	4,82	3,98	3,51	3,20	2,99	2,82	2,69	2,59	2,51	2,43	2,36	0,99
200	3,89	3,04	2,65	2,41	2,26	2,14	2,05	1,98	1,92	1,87	1,83	1,80	0,95
	6,76	4,71	3,88	3,41	3,11	2,90	2,73	2,60	2,50	2,41	2,34	2,28	0,99
1000	3,85	3,00	2,61	2,38	2,22	2,10	2,02	1,95	1,89	1,84	1,80	1,76	0,95
	6,66	4,62	3,80	3,34	3,04	2,82	2,66	2,53	2,43	2,34	2,26	2,20	0,99

Tabelle 4f γ -Quantil $F_{r,s,\gamma}$ der F-Verteilung mit (r, s) Freiheitsgraden

$s \setminus r$	14	16	20	24	30	40	50	75	100	200	500	∞	γ
26	2,10	2,05	1,99	1,95	1,90	1,85	1,82	1,78	1,76	1,72	1,70	1,69	0,95
	2,86	2,77	2,66	2,58	2,50	2,41	2,36	2,28	2,25	2,19	2,15	2,13	0,99
27	2,08	2,03	1,97	1,93	1,88	1,84	1,80	1,76	1,74	1,71	1,68	1,67	0,95
	2,83	2,74	2,63	2,55	2,47	2,38	2,33	2,25	2,21	2,16	2,12	2,10	0,99
28	2,06	2,02	1,96	1,91	1,87	1,81	1,78	1,75	1,72	1,69	1,67	1,65	0,95
	2,80	2,71	2,60	2,52	2,44	2,35	2,30	2,22	2,18	2,13	2,09	2,06	0,99
29	2,05	2,00	1,94	1,90	1,85	1,80	1,77	1,73	1,71	1,68	1,65	1,64	0,95
	2,77	2,68	2,57	2,49	2,41	2,32	2,27	2,19	2,15	2,10	2,06	2,03	0,99
30	2,04	1,99	1,93	1,89	1,84	1,79	1,76	1,72	1,69	1,66	1,64	1,62	0,95
	2,74	2,66	2,55	2,47	2,38	2,29	2,24	2,16	2,13	2,07	2,03	2,01	0,99
40	1,95	1,90	1,84	1,79	1,74	1,69	1,66	1,61	1,59	1,55	1,53	1,51	0,95
	2,56	2,49	2,37	2,29	2,20	2,11	2,05	1,97	1,94	1,88	1,84	1,81	0,99
50	1,90	1,85	1,78	1,74	1,69	1,63	1,60	1,55	1,52	1,48	1,46	1,44	0,95
	2,46	2,39	2,26	2,18	2,10	2,00	1,94	1,86	1,82	1,76	1,71	1,68	0,99
60	1,86	1,81	1,75	1,70	1,65	1,59	1,56	1,50	1,48	1,44	1,41	1,39	0,95
	2,40	2,32	2,20	2,12	2,03	1,93	1,87	1,79	1,74	1,68	1,63	1,60	0,99
70	1,84	1,79	1,72	1,67	1,62	1,56	1,53	1,47	1,45	1,40	1,37	1,35	0,95
	2,35	2,28	2,15	2,07	1,98	1,88	1,82	1,74	1,69	1,62	1,56	1,53	0,99
80	1,82	1,77	1,70	1,65	1,60	1,54	1,51	1,45	1,42	1,38	1,35	1,32	0,95
	2,32	2,24	2,11	2,03	1,94	1,84	1,78	1,70	1,65	1,57	1,52	1,49	0,99
100	1,79	1,75	1,68	1,63	1,57	1,51	1,48	1,42	1,39	1,34	1,30	1,28	0,95
	2,26	2,19	2,06	1,98	1,89	1,79	1,73	1,64	1,59	1,51	1,46	1,43	0,99
200	1,74	1,69	1,62	1,57	1,52	1,45	1,42	1,35	1,32	1,26	1,22	1,19	0,95
	2,17	2,09	1,97	1,88	1,79	1,69	1,62	1,53	1,48	1,39	1,33	1,28	0,99
1000	1,70	1,65	1,58	1,53	1,47	1,41	1,36	1,30	1,26	1,19	1,13	1,08	0,95
	2,09	2,01	1,89	1,81	1,71	1,61	1,54	1,44	1,38	1,28	1,19	1,11	0,99