7 Schliessende Statistik – Parameter- und Intervallschätzung

7.1 Zufallsstichproben

Repetition

Der *Ergebnisraum* $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$ ist die Menge aller sich gegenseitig ausschliessenden Ergebnisse $\omega_1, \omega_2, \omega_3, \dots$ eines *Zufallsexperiments*.
Unter einer *Zufallsvariablen* versteht man eine Funktion X, die jedem Ergebnis ω aus dem Ergebnisraum Ω eine reelle Zahl $X(\omega)$ zuordnet: $X: \Omega \to \mathbb{R}$.

Eine grundlegende Aufgabe der Statistik besteht darin, Kenntnisse und Informationen über die Eigenschaften oder Merkmale einer bestimmten Menge von Objekten zu gewinnen, ohne dass dabei alle Objekte in die Untersuchung einbezogen werden müssen. Dies ist nämlich aus den folgenden Gründen meist auch gar nicht möglich:

- Der Zeit- und Kostenaufwand ist zu hoch.
- Die Anzahl Objekte, die untersucht werden muss, ist zu gross.
- Die Objekte werden u.U. bei der Untersuchung zerstört (z.B. Bestimmung der Lebensdauer).

Die Menge von gleichartigen Objekten oder Elementen heisst *Grundgesamtheit*. Sie kann endlich oder unendlich viele Objekte enthalten. Wir entnehmen eine *Stichprobe* vom Umfang n, um daraus Informationen über die Grundgesamtheit zu gewinnen. Die Auswahl der Objekte für die Stichprobe kann gezielt oder durch Zufall geschehen. Wir betrachten hier nur den zweiten Fall, wo alle Elemente der Grundgesamtheit die gleiche Chance haben, ausgewählt zu werden. Das ist eine wichtige Voraussetzung; andernfalls stimmt die Mathematik nicht!

Definition

Eine *einfache Zufallsstichprobe* vom Umfang n ist eine Folge von Zufallsvariablen $X_1, X_2, \dots, X_n$, den sogenannten *Stichprobenvariablen*. Dabei bezeichnet X_i die Merkmalsausprägung des i-ten Elements in der Stichprobe. Die beobachteten Merkmalswerte $x_1, x_2, \dots, x_n$ der n Elemente sind Realisierungen der Zufallsvariablen $X_1, X_2, \dots, X_n$ und heissen *Stichprobenwerte*.
Die Elemente der Stichprobe werden zufällig ausgewählt, und zwar so, dass jedes Element der Grundgesamtheit mit derselben Wahrscheinlichkeit in die Stichprobe gelangt und die Ziehungen aus der Grundgesamtheit unabhängig voneinander erfolgen. Die Stichprobenvariablen sind dann
stochastisch unabhängig
…………………………………………………………………………………………………
Die X_i folgen alle derselben Verteilung $F(x)$, nämlich der Verteilung der Grundgesamtheit.

In der *deskriptiven Statistik* geht es darum, die Werte einer Stichprobe (oder auch der Grund-
übersichtlich darzustellen
gesamtheit) ……………………………………………… (mit Hilfe von Tabellen und/oder Graphiken) und durch geeignete Masszahlen (Lagemasse wie arithmetisches Mittel oder
zusammenzufassen
Median, Streuungsmasse wie Varianz) …………………………………….. .
aufgrund der Werte einer
In der *schliessenden Statistik* versucht man, ………………………………………………
Stichprobe, Aussagen über die Grundgesamtheit zu machen.
…………………………………………………………………………………………………

7.2 Parameterschätzungen

7.2.1 Einführung

Die Grundgesamtheit ist hinsichtlich eines bestimmten Merkmals durch die Verteilungsfunktion (CDF) $F(x)$ der entsprechenden Zufallsvariablen X vollständig charakterisiert.

Das betrachtete Merkmal legt den Typ der Verteilung fest – aber welche Werte soll man für die Parameter wählen?

Beispiel 1: Hochrechnung bei einer Abstimmung

Fragestellung	Wie gross ist der Anteil Ja-Stimmen?
Merkmal	Ja oder Nein
Typ der Verteilung	Bernoulli-Verteilung
Parameter	p: Wahrscheinlichkeit für Ja

Beispiel 2: Qualitätskontrolle

Fragestellung	Wie genau wird der Sollwert eingehalten?
Merkmal	Abweichung vom Sollwert
Typ der Verteilung	Normalverteilung
Parameter	μ : Erwartungswert σ : Standardabweichung

Die Werte der Parameter sollen auf der Basis einer konkreten Stichprobe geschätzt werden.

1. Frage: Wie erhält man auf der Basis einer konkreten Stichprobe Schätzwerte für die unbekannten Parameter?
→ *Schätzfunktionen* (Punktschätzung)

2. Frage: Wie genau und wie sicher sind solche Schätzwerte?
→ *Vertrauensintervalle* (Intervallschätzung)

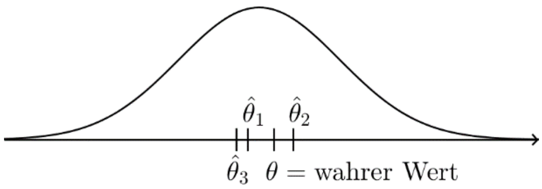
7.2.2 Schätzfunktionen

Ausgangslage: Der Parameter θ („theta“, z.B. der Mittelwert oder die Standardabweichung) charakterisiert die Verteilung der Grundgesamtheit. Der Wert von θ soll geschätzt werden.

Definition
Allgemein ist eine *Stichprobenfunktion* eine Funktion, die von den Stichprobenvariablen $X_1, X_2, \ldots, X_n$ abhängt. Eine *Schätzfunktion* $\Theta = g(X_1, X_2, \ldots, X_n)$ ist eine spezielle Stichprobenfunktion, nämlich eine „Formel“, mit der man den Wert eines Parameters θ der Grundgesamtheit schätzen kann: Setzt man eine konkrete Stichprobe $x_1, x_2, \ldots, x_n$ ein, so erhält man einen *Schätzwert* $\hat{\theta} = g(x_1, x_2, \ldots, x_n)$ für den Parameter θ.

Illustration

Es gibt einen wahren Wert für den gesuchten Parameter.
Der Schätzwert $\hat{\theta}$ für diesen Parameter hängt aber von der konkreten Stichprobe ab. Zieht man hintereinander viele Stichproben und bestimmt jeweils den Schätzwert $\hat{\theta}$, dann findet man in der Regel viele verschiedene Schätzwerte für θ.



Zufallsvariable
θ ist also selber eine ……………………………………, die einer gewissen Verteilung folgt.

arithmetisches Mittel, Median
Schätzfunktionen für μ : ....

empirische (korrigierte) Standardabweichung
Schätzfunktionen für σ : ....

relative Häufigkeit
Schätzfunktionen für p: ....

Beispiel 1:

Wir entnehmen eine Stichprobe von 8 Marroni aus dem Angebot eines Händlers und möchten daraus Rückschlüsse auf die Verteilung des Gewichts und den Anteil wurmstichiger Marroni des gesamten Angebots des Händlers machen.

Die Stichprobenvariablen sind:

- X_i = Gewicht in Gramm der i-ten Marroni (normalverteilt)

  $\hat{\mu}$ und $\hat{\sigma}$
  Schätzfunktionen für die Schätzwerte ……………….. aus der Stichprobe

  μ und σ
  für die Parameter ………………….. des gesamten Angebots gesucht.
- Y_i = 1, falls i-te Marroni wurmstichig, sonst 0 (bernoulliverteilt)

  $\hat{p}$ p
  Schätzfunktion für den Schätzwert ……….. aus der Stichprobe für den Parameter………..
  des gesamten Angebots gesucht.

7.2.3 Kriterien für eine optimale Schätzfunktion

Wir haben bereits zu Beginn des Semesters gesehen, dass es für denselben Parameter verschiedene Schätzfunktionen gibt. Wir brauchen also Kriterien, um zu bestimmen, welches jeweils die ‚beste' Schätzfunktion ist.

Definition

Eine Schätzfunktion Θ eines Parameters θ heisst *erwartungstreu*, wenn gilt:

$$E(\Theta) = \theta$$

Gegeben sind zwei erwartungstreue Schätzfunktionen Θ_1 und Θ_2 desselben Parameters θ. Man nennt Θ_1 *effizienter* als Θ_2, falls gilt:

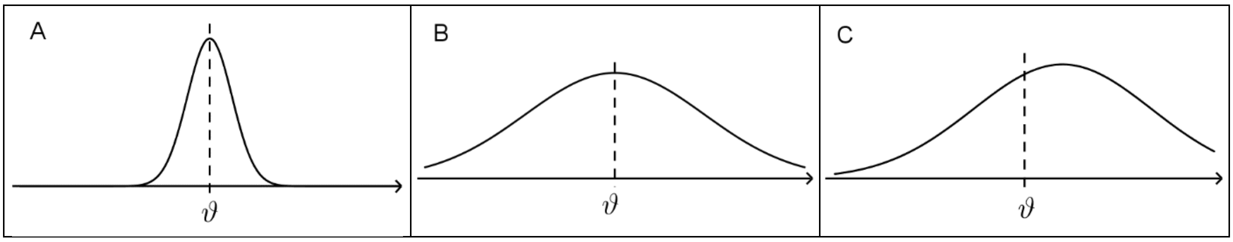
$$V(\Theta_1) < V(\Theta_2)$$

Eine Schätzfunktion Θ heisst *konsistent*, wenn gilt:

$$E(\Theta) \to \theta \;\text{ und }\; V(\Theta) \to 0 \;\text{ für }\; n \to \infty$$

Beispiel 2:
Lesen Sie noch einmal die obigen Definitionen und ordnen Sie dann die Beschreibungen 1-3 den Illustrationen A-C zu.

1. Dichtefunktion einer nicht erwartungstreuen Schätzfunktion.
2. Dichtefunktion einer erwartungstreuen und effizienten Schätzfunktion.
3. Dichtefunktion einer erwartungstreuen, aber weniger effizienten Schätzfunktion. θ



7.2.4 Schätzfunktionen für die wichtigsten statistischen Parameter

Erwartungswert

Schätzfunktion: $\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$	Schätzwert: $\hat{\mu} = \bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i$

Wegen der Linearität des Erwartungswerts und der stochastischen Unabhängigkeit der Stichprobenvariablen gilt:

$E(\bar{X}) = E(X_i)$ somit ist $\bar{X}$ .... (erwartungstreu)

$V(\bar{X}) = \frac{V(X_i)}{n} \xrightarrow{n\to\infty} 0$ somit ist $\bar{X}$ .... (konsistent)

Varianz

Schätzfunktion: $S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$	Schätzwert: $\hat{\sigma}^2 = s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2$

Es gilt (ohne Herleitung):

$E(S^2) = V(X_i)$ somit ist S^2 .... (erw.tr.)

$V(S^2) = \frac{\mu_4}{n} - \frac{(n-3)\mu_2^2}{n(n-1)} \xrightarrow{n\to\infty} 0$ somit ist $\bar{X}$ .... (kons.)

mit $\mu_4 = E((X_i - \mu)^4)$ und $\mu_2 = V(X_i)$

Die Schätzfunktion $\tilde{S}^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$ wird auch verwendet, ist aber nicht erwartungstreu!

Spezialfall: Anteilswert *p* einer Bernoulli-Verteilung

Schätzfunktion: $\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$	Schätzwert: $\hat{p} = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

Dabei sind $X_1, X_2, \ldots, X_n$ Bernoulli-Zufallsvariablen
mit $P(X_i = 1) = p$ und $P(X_i = 0) = 1 - p$.

Nach dem Zentralen Grenzwertsatz und unseren Überlegungen zur Bernoulli-Verteilung gilt:

$E(\bar{X}) = E(X_i) = p$ somit ist $\bar{X}$ .... (erw.tr.)

$V(\bar{X}) = \dfrac{V(X_i)}{n} = \dfrac{p \cdot (1-p)}{n} \xrightarrow{n\to\infty} 0$ somit ist $\bar{X}$ .... (kons.)

Beispiel 1 (Marroni) – Fortsetzung:
Die Stichprobe von 8 Marroni aus dem Angebot des Händlers ergab die folgenden Gewichte; davon waren 3 wurmstichig.

i	1	2	3	4	5	6	7	8
x_i in g	15	20	21	18	16	16	17	21
y_i	0	1	0	0	0	1	1	0

Wir interessieren uns also für zwei Merkmale des Marroni Angebots des Händlers: Das Gewicht der Marroni und das Merkmal wurmstichig.

Man erhält daraus die Schätzwerte für den Mittelwert und die Standardabweichung des Gewichts

$$\hat{\mu} = \bar{x} = \frac{1}{8} \cdot (15 + 20 + 21 + 18 + 16 + 16 + 17 + 21) = 18\text{g}$$

$$\hat{\sigma}^2 = s^2 = \frac{1}{8-1} \cdot (9 + 4 + 9 + 0 + 4 + 4 + 1 + 9) = \frac{40}{7} \approx 5.714\text{g}^2$$

$$\hat{\sigma} = s = \sqrt{5.714} = 2.390\text{g}$$

und den Schätzwert für den Anteil wurmstichiger Maroni:

$$\hat{p} = \frac{3}{8} = 0.375 = 37.5\%$$

Bemerkung zur Konstruktion von Schätzern
Sucht man nach geeigneten Schätzfunktionen, so werden meist Schätzfunktionen verwendet, denen man eine anschauliche Bedeutung zuordnen kann. Man kann diese Schätzfunktionen dann durch Simulation oder auch mathematisch daraufhin untersuchen, ob sie den obigen Kriterien genügen.

Es ist aber auch möglich, Schätzfunktionen anhand bestimmter Kriterien zu “konstruieren”.
Es gibt mehrere bewährte Methoden zur Konstruktion von Punktschätzern:

- die Maximum-Likelihood-Methode,
- die Methode der kleinsten Quadrate oder
- die Momentenmethode.

Wir werden uns im Folgenden die erste der Methoden genauer anschauen. Die Methode der kleinsten Quadrate haben wir bereits in einem vorigen Kapitel kennengelernt.

7.2.5 Maximum-Likelihood-Schätzung

Die Maximum-Likelihood-Methode ist eine dieser Methoden. Dabei wird das stochastische Modell so gewählt, dass die gemachte Beobachtung die maximale Wahrscheinlichkeit erhält

Wir betrachten eine Zufallsvariable X und ihre Dichte (PDF)

$$f_X(x|\theta)$$

welche von x und einem oder mehreren Parametern θ abhängig ist. Für eine Stichprobe vom Umfang n mit den Werten $x_1, x_2, \dots, x_n$ nennen wir die vom Parameter θ abhängige Funktion

$$L(\theta) = f_X(x_1|\theta) \cdot f_X(x_2|\theta) \cdot \dots \cdot f_X(x_n|\theta)$$

die Likelihood-Funktion der Stichprobe.

Da die Stichprobenwerte stochastisch unabhängig voneinander ermittelt wurden, entspricht der Wert der Likelihood-Funktion dem Wert der gemeinsamen Verteilungsdichte der n Zufallsvariablen, wenn man die gemachten Beobachtungen $x_1, x_2, \dots, x_n$ dort einsetzt.
Bei der Maximum-Likelihood-Schätzung wird derjenige Wert $\hat{\theta}_{ML}$ für den Parameter θ ermittelt, welcher die Likelihood-Funktion am grössten macht, bei dem also die Funktion $L(\theta)$ ihr globales Maximum annimmt.
Mit anderen Worten: Mit der ML-Schätzung ermitteln wir das stochastische Modell, welches die beobachteten Daten am wahrscheinlichsten macht.

Beispiel 3
Wir betrachten eine normalverteilte Zufallsvariable X, bei der Erwartungswert μ und Varianz σ^2 unbekannt sind. Wir möchten die Parameter μ und σ^2 mithilfe der Maximum-Likelihood-Methode schätzen. Mit einer Stichprobe $x_1, x_2, \dots, x_n$ erhalten wir die Likelihood-Funktion

$$L(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-(x_1-\mu)^2}{2\sigma^2}} \cdot \dots \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-(x_n-\mu)^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{n/2}} \cdot \prod_{i=1}^{n} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}$$

Nun suchen wir diejenigen Werte für die Parameter μ und σ^2, welche diese Likelihood-Funktion maximieren. Da der Logarithmus eine strikt monoton wachsende Funktion ist, können wir stattdessen auch den Logarithmus der Likelihood-Funktion maximieren:

$$ln(\,L(\mu, \sigma^2)) = -\frac{n}{2} \cdot ln(\,2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

Um die Maximalstelle von $ln(\,L(\mu, \sigma^2))$ zu finden, bestimmen wir deren partielle Ableitungen nach μ und σ^2 und setzen diese gleich Null:

$$0 = \frac{d}{d\mu} ln(\,L(\mu, \sigma^2)) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu)$$

und

$$0 = \frac{d}{d\sigma^2} ln(\,L(\mu, \sigma^2)) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

Daraus folgern wir

$$0 = \sum_{i=1}^{n}(x_i - \mu) = \left(\sum_{i=1}^{n} x_i\right) - n\mu \qquad \text{also } \mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

und

$$0 = -n + \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n}(x_i - \mu)^2 \qquad \text{also } \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Wir finden also nur eine einzige kritische Stelle. Da die Likelihood-Funktion nach oben beschränkt ist, muss dort das globale Maximum von $L(\mu, \sigma^2)$ sein.

Die Maximum-Likelihood-Schätzwerte für den Erwartungswert und die Varianz sind also:

$\hat{\mu}_{ML} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$	$\hat{\sigma}^2{}_{ML} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = s^{*2}$

Die im Beispiel angewendete Technik des Logarithmierens der Likelihood-Funktion wird sehr häufig verwendet, da dies wie gezeigt die Maximierung vereinfacht.

Beispiel 4
Herr Maier hat während einer Dienstreise einen längeren Aufenthalt auf einem Flughafen.
Um sich die Zeit zu vertreiben, stoppt er die Zeit zwischen zwei landenden Flugzeugen auf derselben Rollbahn.
Er notiert folgende Stichprobenwerte x_i (in Minuten):

$$3, 6, 6, 4, 8, 2, 4, 5, 9, 3.$$

Die Zufallsvariable X_i bezeichne das Intervall zwischen zwei aufeinanderfolgenden Landungen und wird als exponentialverteilt mit dem unbekannten Parameter λ angenommen mit der Dichtefunktion

$$f(t; \lambda) = \lambda \cdot e^{-\lambda t} \text{ für } t \geq 0$$

Die Likelihood-Funktion für die realisierte Stichprobe $(x_1, x_2, \dots, x_{10})$ ist gegeben mit

$$L(\lambda, 3, 6, 6, 4, 8, 2, 4, 5, 9, 3) = \lambda e^{-3\lambda} \cdot \lambda e^{-6\lambda} \cdot \dots \cdot \lambda e^{-3\lambda} = \lambda^{10} \cdot e^{-50\lambda}$$

und logarithmiert mit

$$\ln L(\lambda) = 10 \ln \lambda - 50\lambda$$

Ableiten nach λ und Nullsetzen führt zu

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = \frac{10}{\lambda} - 50 = 0$$

Für die ML-Schätzung $\hat{\lambda}$ für λ der exponentialverteilten Grundgesamtheit resultiert:

$$\hat{\lambda} = \frac{10}{50} = 0.2 = \frac{1}{\bar{x}}$$

Die zweite Ableitung ist kleiner 0, womit die Bedingung für ein Maximum erfüllt ist.

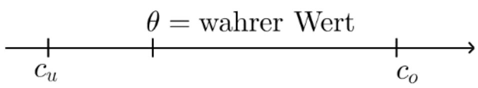
7.3 Vertrauensintervalle

7.3.1 Einführung

Problem: Punktschätzungen erlauben keine Aussage über die Genauigkeit und Sicherheit der Schätzung – der aus einer Zufallsstichprobe gewonnene Schätzwert kann nämlich erheblich vom tatsächlichen Wert des Parameters abweichen (insbesondere bei kleinem Stichprobenumfang).

Fragestellung: Wie genau und wie sicher sind die aus den Stichprobenuntersuchungen mit Hilfe von Schätzfunktionen bestimmten Schätzwerte für die unbekannten Parameter?

Idee: Anhand einer konkreten Stichprobe ein Intervall $[c_u; c_o]$ bestimmen, das den unbekannten Parameter θ mit Sicherheit enthält.



Aber: ……………………………………………………

Ein solches Intervall kann nur trivial existieren, d.h. alle möglichen Werte beinhalten, da eine Stichprobe keine absolut sicheren Rückschlüsse auf die Grundgesamtheit zulässt.

……………………………………………………

Neue Idee: Anhand einer konkreten Stichprobe ein Intervall $[c_u; c_o]$ bestimmen, das den unbekannten Parameter θ mit einer gegebenen grossen Wahrscheinlichkeit γ enthält.

Aber: ……………………………………………………

Der Parameter θ ist zwar unbekannt, aber fix – er hängt nicht vom Zufall ab!

c_u und c_o hängen vom Zufall, d.h. von der Stichprobe ab.

Erwartungsgemäss liegt in γ*100 Prozent der Intervalle $[c_u, c_o]$ der gesuchte Parameter θ.

……………………………………………………

Vorgehen

Wir legen eine beliebige grosse Wahrscheinlichkeit γ fest (z.B. $\gamma = 95\%$). γ heisst *statistische Sicherheit* oder *Vertrauensniveau*. $\alpha = 1 - \gamma$ ist die sog. *Irrtumswahrscheinlichkeit*.

Dann bestimmen wir zwei Zufallsvariablen Θ_u und Θ_o so, dass sie den wahren Parameterwert θ mit der Wahrscheinlichkeit γ einschliessen:

$$P(\Theta_u \leq \theta \leq \Theta_o) = \gamma$$

Dabei hängen Θ_u und Θ_o von den Stichprobenwerten ab, d.h. sie sind Stichprobenfunktionen der n unabhängigen Zufallsvariablen $X_1, X_2, \ldots, X_n$:

$$\Theta_u = g_u(X_1, X_2, \ldots, X_n) \;\text{ und }\; \Theta_o = g_o(X_1, X_2, \ldots, X_n)$$

Setzt man nun die Werte $x_1, x_2, \ldots, x_n$ einer konkreten Stichprobe ein, so erhält man c_u und c_o. Sie bilden die Grenzen eines Intervalls, das als *Vertrauensintervall* für den unbekannten Parameter θ bezeichnet wird.

Bemerkungen

Um das Vertrauensintervall zu bestimmen, muss die Verteilung der verwendeten Schätzfunktion Θ für θ bekannt sein.
Die Länge $l = c_o - c_u$ des Vertrauensintervalls ist ein Mass für die Genauigkeit der Parameterschätzung. Sie hängt nicht nur von den Stichprobenwerten $x_1, x_2, \dots, x_n$ ab.
Welche anderen Grössen beeinflussen die Länge l des Vertrauensintervalls? Wie?

Je grösser γ, desto grösser das Vertrauensintervall.

....

Je grösser der Stichprobenumfang n, desto kürzer das Vertrauensintervall – bei konsistenen Schätzern.

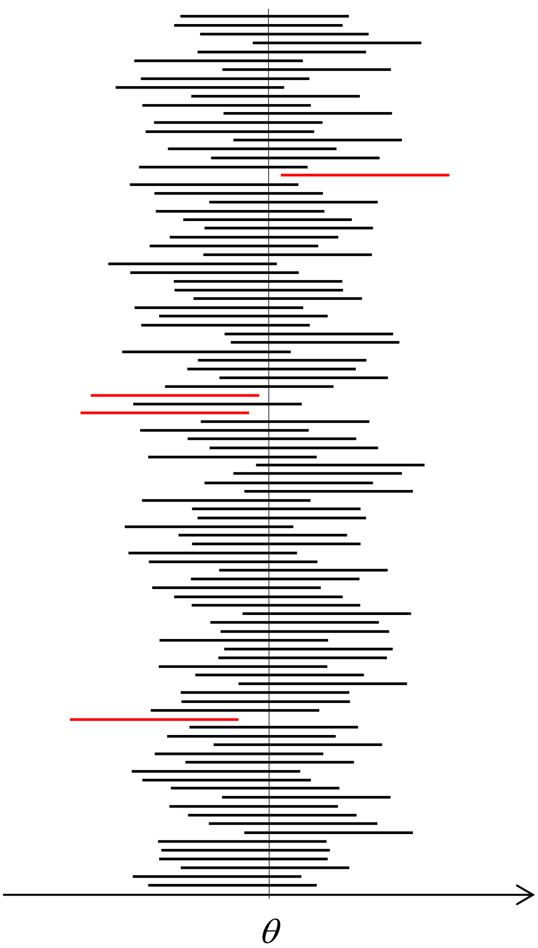
....

....

Illustration

Aus einer Grundgesamtheit werden 100 verschiedene Stichproben der Grösse n entnommen. Für jede dieser Stichproben wird das zugehörige 95%-Vertrauensintervall berechnet. Etwa 95 dieser 100 Vertrauensintervalle werden den wahren Parameterwert θ enthalten.

Die zufällige Stichprobe bestimmt das Vertrauensintervall. Der wahre Parameterwert kann entweder darin liegen oder nicht!



7.3.2 Konstruktion eines Vertrauensintervalls

Gegeben ist eine *normalverteilte* Zufallsvariable X mit unbekanntem μ und *bekanntem* σ^2.
Es soll ein Vertrauensintervall für den Mittelwert μ konstruiert werden.
(Bei vielen Messinstrumenten, Maschinen und Automaten wird die Varianz σ^2 bereits vom Hersteller als eine Art Gerätekonstante angegeben.)

Wir wählen das Vertrauensniveau $\gamma = 95\%$ (ebenfalls üblich ist 99%). Welcher Wert für γ gewählt werden soll, ist keine mathematische Frage, sondern hängt davon ab, wie sicher das Ergebnis sein soll; dabei bezahlt man höhere Sicherheit mit geringerer Genauigkeit.

Die Schätzfunktion für μ: $\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$ liefert einen Ausgangspunkt für unser Vertrauensintervall; ausgehend von diesem Punkt berechnen wir die Grenzen.

Nun wollen wir die Grenzen für das Vertrauensintervall bestimmen; diese sollen symmetrisch um $\bar{X}$ liegen. Wir suchen also eine Schranke e, so dass gilt;

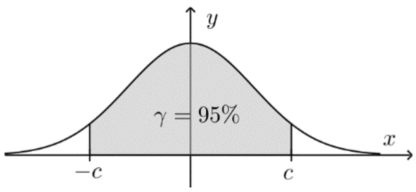
$P(\bar{X} - e \leq \mu \leq \bar{X} + e) = \gamma$ oder anders ausgedrückt: $P(|\bar{X} - \mu| \leq e) = \gamma$ (*)

Nach dem Zentralen Grenzwertsatz ist $\bar{X}$ normalverteilt mit $E(\bar{X}) = \mu$ und $\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$. Damit wir Wahrscheinlichkeiten bestimmen können, müssen wir statt $\bar{X}$ die standardisierte Zufallsvariable $U = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ verwenden; U ist standardnormalverteilt.
Die Gleichung (*) lässt sich umformen zu:

$$P\left(|U| \leq \frac{e}{\sigma/\sqrt{n}}\right) = \gamma$$

Die Illustration rechts zeigt die Situation; dabei ist $c = \frac{e}{\sigma/\sqrt{n}}$.
Wir suchen also das c mit $\phi(c) = \frac{1+\gamma}{2} = 0.975$.
Bestimmen Sie aus der Tabelle 1 diese Schranke c:

$c =$ 1.9600



Aus diesem c können wir nun e berechnen:

$$e = c \cdot \frac{\sigma}{\sqrt{n}}$$

Also lauten die Formeln für die Stichprobenfunktionen Θ_u und Θ_o für die Grenzen:

$$\Theta_u = \bar{X} - c \cdot \frac{\sigma}{\sqrt{n}} \qquad \Theta_o = \bar{X} + c \cdot \frac{\sigma}{\sqrt{n}}$$

Beispiel 5
In einer Fabrik werden Schokoladentafeln hergestellt. Zur Qualitätskontrolle wird eine Stichprobe entnommen und jeweils das Gewicht der Tafel ermittelt:

i	1	2	3	4	5	6	7	8	9	10
x_i	100	97	101	96	98	102	96	100	101	99

Bestimmen Sie ein 95%-Vertrauensintervall für den Mittelwert μ, wenn die Varianz der Grundgesamtheit bekannt ist: $\sigma^2 = 4.667$.

$$\boldsymbol{\bar{x} = \frac{1}{10} \cdot (100 + 97 + 101 + 96 + 98 + 102 + 96 + 100 + 101 + 99) = 99}$$

$$\boldsymbol{e = c \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{2.160}{\sqrt{10}} = 1.339}$$

Vertrauensintervall für $\boldsymbol{\mu}$: $[\boldsymbol{\bar{x} - e; \bar{x} + e}] = [\boldsymbol{97.661; 100.339}]$

Bemerkung

Wir haben c oben so gewählt, dass gilt: $\phi(c) = \frac{1+\gamma}{2}$. Man nennt c deshalb auch das $\frac{1+\gamma}{2}$-Quantil der Standardnormalverteilung.

Für gängige Werte von p können die p -Quantile einfacher mit Tabelle 2 bestimmt werden:

$$p = \phi(c) = 0.95 \quad \Rightarrow \quad c = \overset{1.645}{\ldots\ldots\ldots\ldots\ldots}$$

$$p = \phi(c) = 0.975 \quad \Rightarrow \quad c = \overset{1.960}{\ldots\ldots\ldots\ldots\ldots}$$

$$p = \phi(c) = 0.995 \quad \Rightarrow \quad c = \overset{2.576}{\ldots\ldots\ldots\ldots\ldots}$$

Weitere Stichprobenverteilungen

Zusätzlich zu den Quantilen der Standardnormalverteilung werden bei der Bestimmung von Vertrauensintervallen bei Stichprobenvariablen die Quantile weiterer Wahrscheinlichkeits-verteilungen benötigt, wie die der

- t-Verteilung (Student-t-Verteilung) und
- Chi-Quadrat-Verteilung (χ^2-Verteilung)

Die standardisierte Schätzfunktion des Stichproben-Mittelwerts normalverteilter Daten ist nicht mehr normalverteilt, sondern t-verteilt, wenn die zur Standardisierung des Mittelwerts benötigte Varianz des Merkmals unbekannt ist und mit der Stichprobenvarianz geschätzt werden muss.

Die χ^2-Verteilung findet Anwendung, wenn man die empirische Varianz bestimmt hat und die Schätzung des Vertrauensintervalls ermitteln möchte, das den (unbekannten) Wert der Varianz der Grundgesamtheit mit einer gewissen Wahrscheinlichkeit einschliesst.

Wir verzichten darauf, diese beiden stetigen Verteilungsfunktionen im Detail zu betrachten, sondern arbeiten direkt mit den tabellierten Quantilen dieser Funktionen (Tabelle 3 und 4).

Das Vorgehen je nach Fragestellung ist in der folgenden Tabelle zusammengefasst.

7.3.3 Übersicht über verschiedene Vertrauensintervalle zum Niveau γ

	(1) Verteilung der Grundgesamtheit	(2) Param.	(3) Schätzfunktionen	(4) zugehörige standardisierte Zufallsvariable	(5) Verteilung und benötigte Quantile	(6) Zufallsvariablen für Intervallgrenzen
1	Normalverteilung (Varianz σ^2 bekannt)	μ	$\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$	$U = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	**Standardnormalverteilung** (Tabelle 2) $c = u_p$ mit $p = \frac{1+\gamma}{2}$	$\Theta_u = \bar{X} - c \cdot \frac{\sigma}{\sqrt{n}}$ $\Theta_o = \bar{X} + c \cdot \frac{\sigma}{\sqrt{n}}$
2	Normalverteilung (Varianz σ^2 unbekannt und $n \leq 30$; sonst Fall 1 mit s als Schätzwert für σ)	μ	$\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$ $S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$	$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$	***t*-Verteilung** (Tabelle 4) mit $f = n - 1$ $c = t_{(p;f)}$ mit $p = \frac{1+\gamma}{2}$	$\Theta_u = \bar{X} - c \cdot \frac{S}{\sqrt{n}}$ $\Theta_o = \bar{X} + c \cdot \frac{S}{\sqrt{n}}$
3	Normalverteilung	σ^2	$\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$ $S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$	$Z = (n-1)\frac{S^2}{\sigma^2}$	**Chi-Quadrat-Verteilung** (Tabelle 3) mit $f = n - 1$ $c_1 = z_{(p_1;f)}$ mit $p_1 = \frac{1-\gamma}{2}$ $c_2 = z_{(p_2;f)}$ mit $p_2 = \frac{1+\gamma}{2}$	$\Theta_u = \frac{(n-1) \cdot S^2}{c_2}$ $\Theta_o = \frac{(n-1) \cdot S^2}{c_1}$
4	Bernoulli-Verteilung Anteilsschätzung (mit $n\hat{p}(1-\hat{p}) > 9$)	p	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ X_i 0/1-wertig mit $P(X_i = 1) = p$	$U = \frac{\bar{X} - p}{\sqrt{p(1-p)/n}}$	**Standardnormalverteilung** (näherungsweise), Tabelle 2 $c = u_q$ mit $q = \frac{1+\gamma}{2}$	$\Theta_u = \bar{X} - c \cdot \sqrt{\frac{\bar{X} \cdot (1-\bar{X})}{n}}$ $\Theta_o = \bar{X} + c \cdot \sqrt{\frac{\bar{X} \cdot (1-\bar{X})}{n}}$
5	beliebig mit $n > 30$	μ, σ^2	wie im Fall 1 (gegebenenfalls mit s als Schätzwert für σ) bzw. im Fall 3			

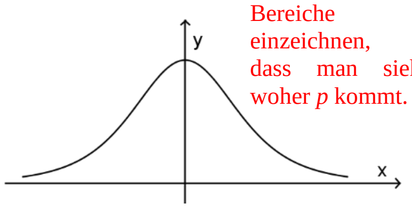
Beispiel 1 (Marroni) – Fortsetzung
Annahmen: Die Grundgesamtheit ist normalverteilt und μ und σ^2 sind unbekannt.
Bestimmen Sie ein 95%-Vertrauensintervall für den Mittelwert μ.

1. Wir betrachten Tabelle 7.3.3. Welche Zeile passt zu unserer Problemstellung? ……… 2
   Beachten Sie die ersten beiden Spalten: Welcher Verteilung folgt die Grundgesamtheit? Für welchen Parameter brauchen wir ein Vertrauensintervall?
2. Wir berechnen die benötigten Schätzwerte $\bar{x}$ und s (s. Spalte 3).
   Bereits erledigt: $\bar{x} = 18\text{g}$, $s = 2.390\text{g}$.
3. Wir bestimmen die Grenze c für die standardisierte Zufallsvariable.

   Vertrauensniveau gemäss Aufgabenstellung: $\gamma =$……… 95%
   Gemäss Spalte 5 ist die standardisierte Zufallsvariable
   t-verteilt (s. Tabelle 4).

   Anzahl Freiheitsgrade: $f =$ ………………………… $n - 1 = 7$

   $p =$……………… 0.975 $\Rightarrow$ $c =$ ……………… 2.365



Die verschiedenen Bereiche einzeichnen, so dass man sieht, woher p kommt.

4. Wir bestimmen das Vertrauensintervall.

   Hilfsgrösse: $e = c \cdot \frac{s}{\sqrt{n}} \approx 1.998\text{g}$

   Vertrauensintervall für μ: $[\bar{x} - e; \bar{x} + e] = [16.002\text{g}; 19.998\text{g}]$

Beispiel 5 (Schokoladentafel) – Fortsetzung
Bestimmen Sie ein 95%-Vertrauensintervall für den Mittelwert μ. Nehmen Sie diesmal an, die Varianz sei unbekannt. Hier nochmals die Werte:

i	1	2	3	4	5	6	7	8	9	10
x_i	100	97	101	96	98	102	96	100	101	99

Vergleichen Sie die Vertrauensintervalle für bekannte und unbekannte Varianz.

1. Wir betrachten Tabelle 7.3.3. Welche Zeile passt zu unserer Problemstellung? ……… 2
2. Wir berechnen die benötigten Schätzwerte $\bar{x}$ und s^2 (s. Spalte 3).
   Bereits erledigt: $\bar{x} = 99$.

$$s^2 = \frac{1}{9} \cdot (1^2 + (-2)^2 + 2^2 + (-3)^2 + (-1)^2 + 3^2 + (-3)^2 + 1^2 + 2^2 + 0^2) = \frac{1}{9} \cdot 42$$
$$= 4.667$$

3. Wir bestimmen die Grenze c für die standardisierte Zufallsvariable.

   Vertrauensniveau gemäss Aufgabenstellung: $\gamma =$ 95%
   Gemäss Spalte 5 ist die standardisierte Zufallsvariable t -verteilt
   $f =$ $n-1=9$, $p =$ 0.975 $\Rightarrow$ $c =$ 2.262

4. Wir bestimmen das Vertrauensintervall.

   Hilfsgrösse: $e = c \cdot \frac{s}{\sqrt{n}} = 2.262 \cdot \frac{2.160}{\sqrt{10}} = 1.545$

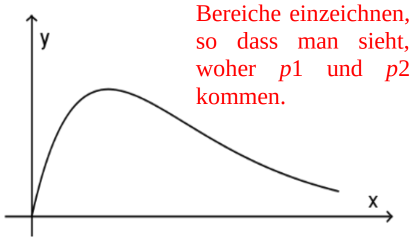
   Vertrauensintervall für μ: $[\bar{x} - e;\ \bar{x} + e] = [97.455; 100.545]$

Vergleich mit dem Vertrauensintervall bei bekannter Varianz (es gilt hier $s^2 = \sigma^2$!):
Ist die Varianz unbekannt, hat man weniger Informationen, also ist die Genauigkeit kleiner.
Entsprechend wird das Vertrauensintervall länger.

Beispiel 1 (Marroni) – Fortsetzung
Annahmen: Die Grundgesamtheit ist normalverteilt und μ und σ^2 sind unbekannt.
Bestimmen Sie ein 95%-Vertrauensintervall für die Varianz σ^2:

1. Wir betrachten Tabelle 7.3.3. Welche Zeile passt zu unserer Problemstellung? 3
2. Wir berechnen den benötigten Schätzwert s^2 (s. Spalte 3). Bereits erledigt: $s^2 = 5.714g^2$.
3. Wir bestimmen die Grenzen c_1 und c_2 für die standardisierte Zufallsvariable.

   Vertrauensniveau gemäss Aufgabenstellung: $\gamma =$ 95%.
   Gemäss Spalte 5 ist die standardisierte Zufallsvariable χ^2-verteilt. Wir verteilen die Irrtumswahrscheinlichkeit α gleichmässig auf beide Seiten.

   Die verschiedenen Bereiche einzeichnen, so dass man sieht, woher $p1$ und $p2$ kommen.

   $f =$ $n-1=7$

   $p_1 =$ 0.025 $\Rightarrow$ $c_1 =$ 1.69 $\quad p_2 =$ 0.975 $\Rightarrow$ $c_2 =$ 16.01

4. Wir bestimmen das Vertrauensintervall.

   Vertrauensgrenzen: $\theta_u = \frac{(n-1)\cdot s^2}{c_2} = \frac{(8-1)\cdot 5.714\text{g}^2}{16.01} = 2.498\text{g}^2$

   $$\theta_o = \frac{(n-1)\cdot s^2}{c_1} = \frac{(8-1)\cdot 5.714\text{g}^2}{1.69} = 23.667\text{g}^2$$

   Vertrauensintervall für σ^2: $[2.498\text{g}^2; 23.667\text{g}^2]$

Beispiel 6

Bei einem Verkehrsunternehmen werden Massnahmen zur Verbesserung der Kundenzufriedenheit durchgeführt. Um die Massnahme zu bewerten, werden 1200 Personen befragt und 473 der Befragten sprechen sich für die Massnahmen aus. Es soll der Anteil der Befürworter der Massnahmen in der Gesamtbevölkerung (Bernoulli-verteilte Grundgesamtheit) mit einer Irrtumswahrscheinlichkeit $\alpha = 1\%$ geschätzt werden.

1. Wir betrachten Tabelle 7.3.3. Welche Zeile passt zu unserer Problemstellung? ……… 4

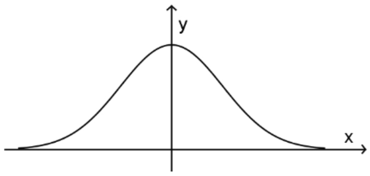
   Faustregel erfüllt? (erst Schritt 2) $n \cdot \hat{p} \cdot (1 - \hat{p}) = 1200 \cdot 0.394 \cdot 0.606 = 286.5 > 9$

2. Wir berechnen den benötigten Schätzwert $\hat{p} = \bar{x}$ (s. Spalte 3).

   Schätzwert für Anteil Befürworter: $\hat{p} = \frac{473}{1200} = 0.394$

3. Wir bestimmen die Grenze c für die standardisierte Zufallsvariable.

   Vertrauensniveau gemäss Aufgabenstellung: $\gamma =$……… 99%
   Gemäss Spalte 5 ist die standardisierte Zufallsvariable näherungsweise standardnormalverteilt (s. Tabelle 2).
   $q =$……………… 0.995 $\Rightarrow$ $c =$……………… 2.576



4. Wir bestimmen das Vertrauensintervall.

   Hilfsgrösse: $e = c \cdot \sqrt{\frac{\bar{x} \cdot (1 - \bar{x})}{n}} = 2.576 \cdot \sqrt{\frac{0.394 \cdot 0.606}{1200}} = 0.0363$

   Vertrauensintervall für p: $[\bar{x} - e; \bar{x} + e] = [0.3577; 0.4303]$

7.4 Lernziele für dieses Kapitel

- ☐ Ich kann an Beispielen erklären, was eine *einfache Zufallsstichprobe* ist.
- ☐ Ich kenne die Definitionen der wichtigsten Schätzfunktionen für den *Mittelwert*, die *Varianz* und die *Standardabweichung*.
- ☐ Ich kann anhand des Mittelwertes und der Varianz erklären, wie eine *Parameterschätzung* definiert wird.
- ☐ Ich kann Schätzwerte für den *Mittelwert*, die *Varianz* und die *Standardabweichung* in Beispielen berechnen.
- ☐ Ich kann erklären, was eine *erwartungstreue* Schätzfunktion ist.
- ☐ Ich kann erklären, was eine *konsistente* Schätzfunktion ist.
- ☐ Ich kann beschreiben, was es bedeutet, dass eine Schätzfunktion *effizienter* ist als eine andere.
- ☐ Ich kann für Beispiele die Gleichung der *Likelihood-Funktion* aufstellen.
- ☐ Ich kann durch Differenzieren der Likelihood-Funktion die *Schätzfunktion* für den/die gegebenen Parameter herleiten.
- ☐ Ich kann den Unterschied zwischen einer *Punkt*- und einer *Intervallschätzung* beschreiben.
- ☐ Ich kann an Beispielen erklären, was ein *Vertrauensintervall* ist.
- ☐ Ich kann die Begriffe der *statistischen Sicherheit* und der *Irrtumswahrscheinlichkeit* für Vertrauensintervalle erklären.
- ☐ Ich kann Tabelle 7.3.3 für Beispiele anwenden, um *Vertrauensintervalle* zu bestimmen.
- ☐ Ich kann die Quantiltabellen der Standardnormalverteilung, der Student-t-Verteilung und der Chi-Quadrat-Verteilung anwenden, um die kritischen Grenzen der Vertrauensintervalle zu berechnen.