Vorlesung Statistik II

Florian Wickelmaier

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Sommersemester 2026

Intervallschätzung

Zielsetzung

- Im Rahmen der Punktschätzung wird ein Schätzwert $\hat{\vartheta}$ ermittelt, der im Allgemeinen nicht mit dem Wert des Parameters ϑ identisch ist
- Für eine sinnvolle Interpretation ist es daher notwendig, neben dem Schätzwert $\hat{\vartheta}$ selbst die Präzision des Schätzverfahrens mit anzugeben
- Für erwartungstreue Schätzfunktionen ist der Standardfehler, d. h. die Standardabweichung der Schätzfunktion, ein sinnvolles Maß für die Präzision des Schätzverfahrens
- Die Intervallschätzung ist ein anderer Weg, die Unsicherheit bzgl. des Parameters zu beschreiben

Grundlagen

Das Ergebnis des Schätzverfahrens ist ein Intervall

Es wird dabei versucht die Wahrscheinlichkeit, dass das Verfahren ein Intervall liefert, das den Parameterwert ϑ *nicht* enthält, durch eine vorgegebene *Irrtumswahrscheinlichkeit* α zu kontrollieren

Üblicherweise werden hierfür in der Statistik die Werte $\alpha = 0.10$, $\alpha = 0.05$ (Fisher: "one in twenty") und $\alpha = 0.01$ verwendet

Entsprechend ergibt sich die Wahrscheinlichkeit, dass das Verfahren ein Intervall liefert, das den Parameterwert ϑ enthält, durch die Gegenwahrscheinlichkeit $1 - \alpha$

Diese *Überdeckungswahrscheinlichkeit* nimmt dann die Werte 0.90, 0.95 und 0.99 an

Konstruktion von Konfidenzintervallen

Sei $X_1, \ldots, X_n$ eine einfache Zufallsstichprobe mit den Realisierungen $x_1, \ldots, x_n$

Die unterliegende Verteilung in der Population hänge von einem Parameter ϑ ab

Der Schätzwert $\hat{\vartheta} = t_n(x_1, \ldots, x_n)$ für den unbekannten Parameter ϑ ist eine Realisierung der Zufallsvariable $T_n = t_n(X_1, \ldots, X_n)$

Da die beiden Intervallgrenzen auf Grund der Stichprobenziehung zufallsabhängig sind, betrachtet man zwei weitere Zufallsvariablen

$$G_u = g_u(X_1, \ldots, X_n) \quad \text{und} \quad G_o = g_o(X_1, \ldots, X_n)$$

als die untere bzw. die obere Intervallgrenze

Konstruktion von Konfidenzintervallen

Für eine gegebene Irrtumswahrscheinlichkeit α hat man dann ein zufälliges Intervall

$$[G_u, G_o] = \{x \in I\!R \mid G_u \leq x \leq G_o\},$$

das ein Ereignis ist

Für die Wahrscheinlichkeit, dass der Parameter ϑ in diesem Intervall liegt, gilt mit $G_u \leq G_o$

$$P(G_u \leq \vartheta \leq G_o) = 1 - \alpha$$

Die Breite des Intervalls $[G_u, G_o]$, die gegeben ist durch die Differenz $G_o - G_u$, kennzeichnet dann die Präzision der Parameterschätzung

Konstruktion von Konfidenzintervallen

In die Bestimmung der Intervallgrenzen gehen die Annahmen über die gemeinsame Verteilung der $X_1, \ldots, X_n$ ein, sodass sich unter verschiedenen Verteilungsannahmen auch unterschiedliche Intervalle ergeben

Beobachtet man nun die Realisierungen $x_1, \ldots, x_n$ der Zufallsstichprobe $X_1, \ldots, X_n$, dann erhält man auch zu den Intervallgrenzen G_u, G_o Realisierungen

$$g_u = g_u(x_1, \ldots, x_n) \quad \text{und} \quad g_o = g_o(x_1, \ldots, x_n)$$

und damit das Intervall

$$[g_u, g_o] = \{x \in \mathbb{R} \mid g_u \leq x \leq g_o\}$$

Zur Interpretation des Intervalls $[g_u, g_o]$ ist Folgendes zu beachten!

Interpretation eines Konfidenzintervalls

Mit g_u, g_o sowie durch ϑ werden reelle Zahlen bezeichnet

Daher ist mit der Aussage

$$g_u \leq \vartheta \leq g_o$$

keinerlei Unsicherheit verbunden

- Entweder ist die Aussage wahr oder sie ist falsch!

Es gilt also

$$P(g_u \leq \vartheta \leq g_o) = \begin{cases} 1 & \text{für } \vartheta \in [g_u, g_o] \\ 0 & \text{sonst} \end{cases}$$

Die Wahrscheinlichkeit $1 - \alpha$, mit der der Parameter ϑ in einem konstruierten Intervall liegt, ist eine Eigenschaft des Schätz- bzw. Konstruktionsverfahrens und nicht des einzelnen Intervalls

Interpretation eines Konfidenzintervalls

Die Schätzung des Intervalls ist so angelegt, dass dessen Konstruktion mit Wahrscheinlichkeit $1 - \alpha$ dazu führt, dass der Parameter ϑ in diesem Intervall tatsächlich enthalten ist

Im Sinne einer frequentistischen Interpretation bedeutet das für eine Irrtumswahrscheinlichkeit $\alpha = 0.05$

- Von 100 Intervallen, die abhängig von den jeweiligen Realisierungen $x_1, \ldots, x_n$ der Stichprobe $X_1, \ldots, X_n$ konstruiert werden, liegt der unterliegende Populationsparameter ϑ in 95 Fällen im konstruierten Intervall
- Lediglich in 5% der Fälle irrt man sich und der Parameter liegt nicht im angegebenen Intervall

$(1-\alpha)$-Konfidenzintervall

Zu vorgegebener Irrtumswahrscheinlichkeit α liefern die aus der Stichprobe $X_1, \ldots, X_n$ gebildeten Schätzfunktionen

$$G_u = g_u(X_1, \ldots, X_n) \quad \text{und} \quad G_o = g_o(X_1, \ldots, X_n)$$

ein $(1-\alpha)$-*Konfidenzintervall* (Vertrauensintervall), wenn gilt

$$P(G_u \leq G_o) = 1$$
$$P(G_u \leq \vartheta \leq G_o) = 1 - \alpha$$

$1-\alpha$ wird als *Überdeckungswahrscheinlichkeit* (auch Sicherheits- oder Konfidenzwahrscheinlichkeit) oder *Konfidenzniveau* bezeichnet

$(1-\alpha)$-Konfidenzintervall

Sind sowohl G_u als auch G_o Schätzfunktionen mit echter Variablilität, so erhält man damit ein *zweiseitiges Konfidenzintervall*

- Es ergibt sich dabei eine Realisierung $[g_u, g_o]$ des Intervalls mit im Allgemeinen nicht entarteten Grenzen $g_u \neq -\infty$, $g_o \neq \infty$

Im Folgenden beschränken wir uns auf den Fall *symmetrischer Konfidenzintervalle*

- Die Wahrscheinlichkeit, dass ϑ über der oberen Intervallgrenze liegt (d. h. $\vartheta > G_o$) ist dabei gleich der Wahrscheinlichkeit, dass ϑ unter der unteren Intervallgrenze liegt (d. h. $\vartheta < G_u$)

Will man die Abweichung der Schätzung vom Parameter nur in eine Richtung kontrollieren, so kann man die Intervallgrenzen prinzipiell gleich $-\infty$ bzw. ∞ setzen

Einseitiges $(1-\alpha)$-Konfidenzintervall

Setzt man $G_u = -\infty$ für alle Stichproben $X_1, \ldots, X_n$, so erhält man ein *einseitiges Konfidenzintervall*

$$P(\vartheta \leq G_o) = 1 - \alpha$$

mit der oberen Intervallgrenze G_o

Setzt man $G_o = \infty$ für alle Stichproben $X_1, \ldots, X_n$, so erhält man ein einseitiges Konfidenzintervall

$$P(G_u \leq \vartheta) = 1 - \alpha$$

mit der unteren Intervallgrenze G_u

Einseitiges $(1-\alpha)$-Konfidenzintervall

Einseitige Konfidenzintervalle kann man verwenden, wenn nur Abweichungen in eine Richtung erwartet werden bzw. für die untersuchte Fragestellung relevant sind

Beispiel

- Benötigt man für das Bestehen einer Klausur eine bestimmte Punktzahl und will deren Erreichen aus der Bearbeitung von Probeklausuren (den Stichprobenvariablen) ableiten, so interessiert man sich für ein einseitiges Konfidenzintervall der Form $[G_u, \infty)$
- Interessiert man sich für das Bestehen der Klausur, dann will man nur absichern, dass eine bestimmte Punktzahl überschritten wird, die konkrete Abweichung nach oben ist nicht relevant

Konfidenzintervalle für den Erwartungswert μ einer normalverteilten Zufallsvariable

Die Stichprobe $X_1, \ldots, X_n$ bestehe aus unabhängig und identisch $N(\mu, \sigma^2)$-verteilten Zufallsvariablen

Es werden zwei Fälle unterschieden

- Der Erwartungswert μ ist unbekannt, die Varianz σ^2 ist aber bekannt
- Der Erwartungswert μ und die Varianz σ^2 sind unbekannt

Es wird sich zeigen, dass in diesen beiden Fällen der Konstruktion der Konfidenzintervalle unterschiedliche Verteilungen zugrunde zu legen sind

Konfidenzintervalle für den Erwartungswert μ einer normalverteilten Zufallsvariable bei bekannter Varianz σ^2

Den Ausgangspunkt der Konstruktion des Konfidenzintervalls bildet das arithmetische Mittel

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

als Punktschätzer für den unbekannten Erwartungswert μ

Wir haben bereits gezeigt, dass $\bar{X}$ unter den gegebenen Voraussetzungen ebenfalls normalverteilt ist gemäß $N(\mu, \sigma^2/n)$

$(1-\alpha)$-Konfidenzintervall

Die Schätzfunktion $\bar{X}$ lässt sich standardisieren, sodass gilt

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Damit hat man eine Schätzfunktion,

- die neben dem unbekannten Parameter μ nur bekannte Größen enthält
- deren Verteilung einerseits bekannt ist und andererseits nicht von μ abhängt

Für die Standardisierte lässt sich unmittelbar ein zweiseitig beschränkter Bereich angeben, in dem sie mit Wahrscheinlichkeit $1-\alpha$ liegt

$(1-\alpha)$-Konfidenzintervall

Es gilt

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

wobei $z_{1-\alpha/2}$ das $(1-\alpha/2)$-Quantil der Standardnormalverteilung ist

Dann hat man

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

$(1-\alpha)$-Konfidenzintervall

In der letzten Form sind die Bedingungen für ein $(1-\alpha)$-Konfidenzintervall erfüllt, das sich ergibt durch

$$\left[\bar{X} - z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}},\ \bar{X} + z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

Die Breite des Konfidenzintervalls ist

$$b_I = 2 \cdot z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

und hängt neben σ auch ab vom Stichprobenumfang n und der Irrtumswahrscheinlichkeit α

Breite des Konfidenzintervalls

Die Breite des Konfidenzintervalls, die die Unsicherheit der Schätzung charakterisiert,

- nimmt mit steigendem Stichprobenumfang n ab und geht für $n \to \infty$ gegen Null
- nimmt mit wachsender Sicherheitswahrscheinlichkeit $1 - \alpha$ (also abnehmender Irrtumswahrscheinlichkeit α) zu und geht für $\alpha \to 0$ gegen unendlich

Damit ergeben sich verschiedene Möglichkeiten eines Trade-Offs zwischen α und n

- Beispielweise kann man zu fest vorgegebener Irrtumswahrscheinlichkeit α den für eine bestimmte Maximalbreite des Konfidenzintervalls erforderlichen Stichprobenumfang berechnen durch $n \geq (2 \cdot z_{1-\alpha/2} \cdot \sigma / b_I)^2$ (Tippfehler in Fahrmeir et al., 2016, S. 359) bzw. durch Simulation bestimmen

Einseitiges $(1-\alpha)$-Konfidenzintervall

Entsprechend lassen sich aus der Standardisierten des arithmetischen Mittels in analoger Weise einseitige Konfidenzintervalle ableiten durch

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

Entsprechendes lässt sich für die inverse Ungleichung zeigen

Einseitiges $(1-\alpha)$-Konfidenzintervall

Damit ergeben sich die einseitigen Konfidenzintervalle zum Konfidenzniveau $1-\alpha$ durch

$$\left[\bar{X} - z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}},\ \infty\right)$$

für die Begrenzung nach unten und

$$\left(-\infty,\ \bar{X} + z_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}\right]$$

für die Begrenzung nach oben

Beispiel

In einem Experiment zur Tonhöhenwahrnehmung wird der Punkt der subjektiven Gleichheit zu einem Sinuston von 1000 Hz mit Hilfe der Herstellungsmethode bei sukzessiver Reizdarbietung ermittelt

- In jedem von 30 Durchgängen stellt die Versuchsperson nach der Vorgabe des Standardreizes einen Vergleichsreiz so ein, dass er gleich hoch erscheint
- Aus der Gültigkeit des Weberschen Gesetzes sei bekannt, dass für die Standardabweichung gilt $\sigma = 7.5\,\text{Hz}$
- Als Mittelwert der Einstellungen resultiere $\bar{x} = 996\,\text{Hz}$

Beispiel

Ein Konfidenzintervall zum Niveau 0.95 erhält man dann auf folgende Weise

- Es ist $\alpha = 0.05$ und somit $1 - \alpha/2 = 0.975$
- Mit Hilfe einer Tabelle der Standardnormalverteilung oder mit R durch Eingabe von `qnorm(0.975, 0, 1)` ermittelt man für das 0.975-Quantil den Wert $z_{0.975} = 1.96$
- Mit $z_{0.975} \cdot \frac{\sigma}{\sqrt{n}} = 1.96 \cdot \frac{7.5}{\sqrt{30}} = 2.7$ ergibt sich das 0.95-Konfidenzintervall

$$\left[\bar{x} - z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right] = [996 - 2.7, 996 + 2.7]$$
$$= [993.3, 998.7]$$

Beispiel

Was kann man aus dem Ergebnis für den Punkt der subjektiven Gleichheit (PSE) ableiten?

- Der PSE ist mit einer Wahrscheinlichkeit von 0.95 in dem aus dem Verfahren resultierenden Konfidenzintervall enthalten
- Offensichtlich liegt der Standardreiz nicht im angegebenen Konfidenzintervall
- Es gilt also mit hoher Wahrscheinlichkeit (mindestens 0.95), dass der PSE vom Standardreiz verschieden ist
- Bei großer Unsicherheit (Breite des Intervalls) liegt allerdings eine starke Über- oder Unterschätzung des PSEs vor, wenn die Abweichung vom Standardreiz ("Effekt") klein ist
- Deshalb ist die Festlegung einer inhaltlich maximal zulässigen Breite vor der Datenerhebung nötig und durch die entsprechende Stichprobengröße einzuhalten!

Beispiel

Was kann man aus dem Ergebnis für den Punkt der subjektiven Gleichheit (PSE) ableiten?

- Da man zum Niveau $1 - \alpha = 0.99$ das Konfidenzintervall

$$[992.5, 999.5]$$

  erhält, ist mit der Aussage "Der PSE unterscheidet sich vom Standardreiz" eine Irrtumswahrscheinlichkeit von weniger als 0.01 (also weniger als 1 Prozent) verbunden
- Die Breite beträgt nun allerdings 7 Hz und die Überschätzung der Abweichung vom Standardreiz ist damit noch größer

Beispiel

Wie viele Durchgänge werden benötigt für ein 0.95-Konfidenzintervall von höchstens der Breite $b_l = 2\,\text{Hz}$?

Mit der oben abgeleiteten Formel berechnet man

$$n \geq \left(\frac{2 \cdot z_{1-\alpha/2} \cdot \sigma}{b_l}\right)^2 = \left(\frac{2 \cdot 1.96 \cdot 7.5}{2}\right)^2 = 216.09$$

Die Anzahl der Durchgänge ist also auf mindestens 217 zu erhöhen

Konfidenzintervall und statistisches Schließen

Gültige Schlussfolgerungen bezüglich der Lage des Parameters sind also nur bei engen Konfidenzintervallen möglich

Dafür muss bei der Studienplanung die folgende Reihenfolge eingehalten werden

1. Festlegen der inhaltlich maximal zulässigen Breite des Intervalls
2. Bestimmung des Stichprobenumfangs n, mit dem die festgelegte Breite erreicht werden kann, entweder analytisch oder per Simulation (s. Beispielsammlung)
3. Erhebung der geplanten Stichprobe und Berechnung des Konfidenzintervalls

Simulation der Überdeckungswahrscheinlichkeit

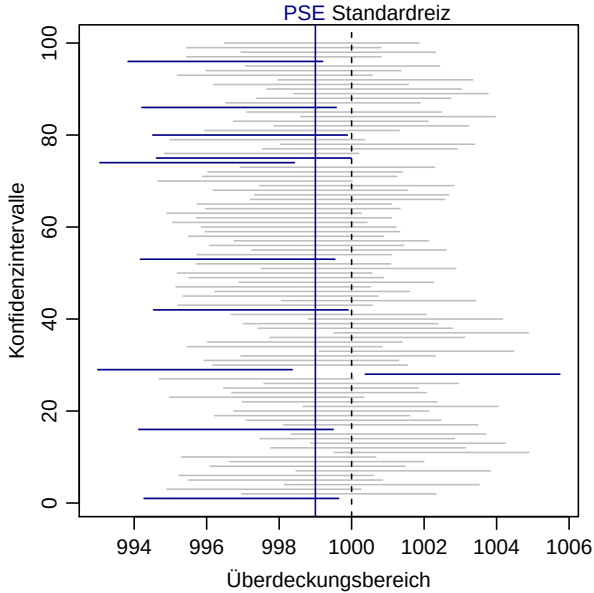
Grafische Darstellung der Konfidenzintervalle von 100 Stichprobenziehungen vom Umfang $n = 30$ aus einer $N(999, 7.5^2)$-Verteilung

```
alpha <- 0.05; n <- 30
pse <- 999; sigma <- 7.5
z <- qnorm(c(alpha/2, 1 - alpha/2))

ci <- replicate(100, {
  x <- rnorm(n, pse, sigma)     # ziehen der Stichprobe
  mean(x) + z * sigma/sqrt(n)  # berechnen des Intervalls
})

matplot(ci, rbind(1:100, 1:100), type = "l",
        lty = 1, xlab = "Ueberdeckungsbereich",
        ylab = "Konfidenzintervalle")
abline(v = c(pse, 1000), lty = 1:2)
```

Simulation der Überdeckungswahrscheinlichkeit



Simulation der Überdeckungswahrscheinlichkeit

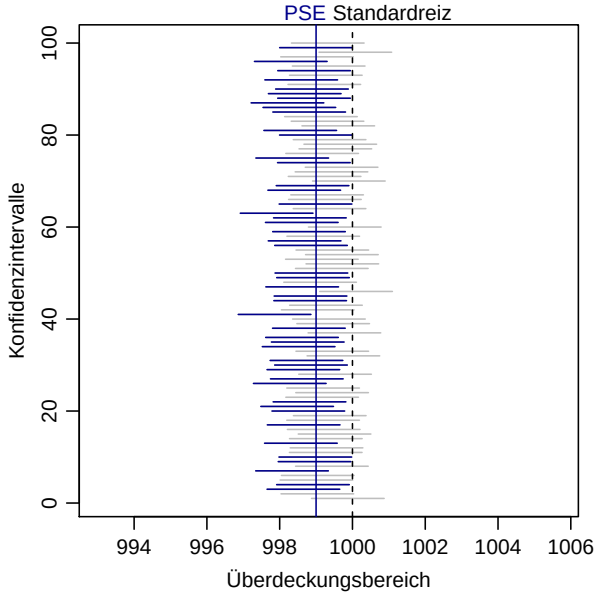
Grafische Darstellung der Konfidenzintervalle von 100 Stichprobenziehungen vom Umfang $n = 217$ aus einer $N(999, 7.5^2)$-Verteilung

```
alpha <- 0.05; n <- 217
pse <- 999; sigma <- 7.5
z <- qnorm(c(alpha/2, 1 - alpha/2))

ci <- replicate(100, {
  x <- rnorm(n, pse, sigma)
  mean(x) + z * sigma/sqrt(n)
})

matplot(ci, rbind(1:100, 1:100), type = "l",
        lty = 1, xlab = "Ueberdeckungsbereich",
        ylab = "Konfidenzintervalle")
abline(v = c(pse, 1000), lty = 1:2)
```

Simulation der Überdeckungswahrscheinlichkeit



Simulation der Überdeckungswahrscheinlichkeit

Simulation zur Überdeckungswahrscheinlichkeit der Konfidenzintervalle von 2000 Stichprobenziehungen vom Umfang $n = 30$ aus einer $N(999, 7.5^2)$-Verteilung

```
alpha <- 0.05; n <- 30
pse <- 999; sigma <- 7.5
z <- qnorm(c(alpha/2, 1 - alpha/2))

ci <- replicate(2000, {
  x <- rnorm(n, pse, sigma)
  mean(x) + z * sigma/sqrt(n)
})

## Ueberdeckungswahrscheinlichkeit
mean(ci[1, ] <= pse & pse <= ci[2, ])
[1] 0.944
```

Breite der Konfidenzintervalle

Grafische Darstellung in Abhängigkeit von der Irrtumswahrscheinlichkeit α und dem Stichprobenumfang *n*

```
sigma <- 7.5
alpha <- seq(0.01, 0.5, length.out = 50)
n     <- seq(1, 80, length.out = 50)

## Linien gleicher Breite
contour(n, alpha,
  outer(n, alpha,
    function(x, y)
      2*qnorm(1 - y/2) * sigma/sqrt(x)
  ),
  nlevels = 15, xlab = "n",
  ylab = expression(alpha))
```

