



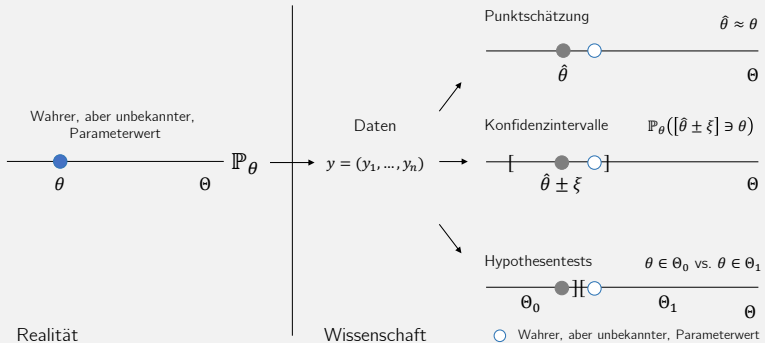
Wahrscheinlichkeitstheorie und Frequentistische Inferenz

BSc Psychologie WiSe 2023/24

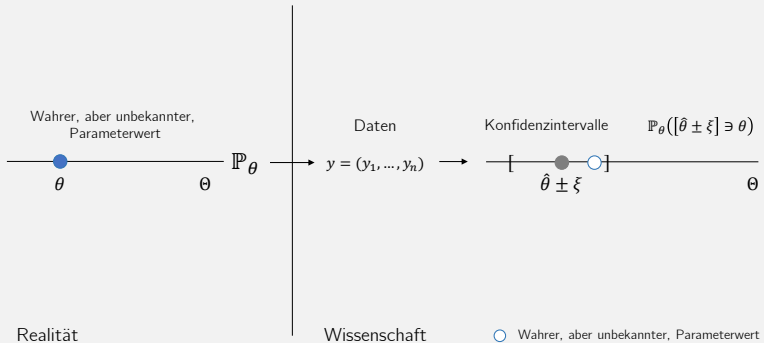
Prof. Dr. Dirk Ostwald

(10) Konfidenzintervalle

Standardproblemstellungen und Standardannahme Frequentistischer Inferenz



Standardproblemstellungen und Standardannahme Frequentistischer Inferenz



Standardannahme Frequentistischer Inferenz

$\mathcal{M}$ sei ein Frequentistisches Inferenzmodell mit Zufallsvektor v . Es wird angenommen, dass ein Datensatz $y \in \mathbb{R}^n$ eine der möglichen Realisierungen von v ist.

Aus Frequentistischer Sicht kann man eine Studie unendlich oft wiederholen und zu jedem Datensatz Schätzer oder Statistiken auswerten, z.B. das Stichprobenmittel:

$$\text{Datensatz (1)} : y^{(1)} = (y_1^{(1)}, y_2^{(1)}, \dots, y_n^{(1)}) \text{ mit } \bar{y}^{(1)} = \frac{1}{n} \sum_{i=1}^n y_i^{(1)}$$

$$\text{Datensatz (2)} : y^{(2)} = (y_1^{(2)}, y_2^{(2)}, \dots, y_n^{(2)}) \text{ mit } \bar{y}^{(2)} = \frac{1}{n} \sum_{i=1}^n y_i^{(2)}$$

$$\text{Datensatz (3)} : y^{(3)} = (y_1^{(3)}, y_2^{(3)}, \dots, y_n^{(3)}) \text{ mit } \bar{y}^{(3)} = \frac{1}{n} \sum_{i=1}^n y_i^{(3)}$$

$$\text{Datensatz (4)} : y^{(4)} = (y_1^{(4)}, y_2^{(4)}, \dots, y_n^{(4)}) \text{ mit } \bar{y}^{(4)} = \frac{1}{n} \sum_{i=1}^n y_i^{(4)}$$

$$\text{Datensatz (5)} : y^{(5)} = \dots$$

Um die Qualität ihrer Methoden zu beurteilen betrachtet die Frequentistische Inferenz deshalb die Wahrscheinlichkeitsverteilungen von Schätzern und Statistiken. Was zum Beispiel ist die Verteilung der $\bar{y}^{(1)}, \bar{y}^{(2)}, \bar{y}^{(3)}, \bar{y}^{(4)}, \dots$ also die Verteilung der Zufallsvariable $\bar{v}_n$?

Wenn eine Methode im Sinne der Frequentistischen Standardannahme "gut" ist, dann heißt das also, dass sie bei häufiger Anwendung "im Mittel gut" ist. Im Einzelfall, also im Normalfall nur eines vorliegenden Datensatzes, kann sie auch "schlecht" sein.

Definition

Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Anwendungsbeispiel

Selbstkontrollfragen

Definition

Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Anwendungsbeispiel

Selbstkontrollfragen

Definition (δ -Konfidenzintervall)

Es sei v die Stichprobe eines Frequentistischen Inferenzmodells mit wahren, aber unbekannten Parameter, $\theta \in \Theta$, es sei $\delta \in]0, 1[$ und es seien $G_u(v)$ und $G_o(v)$. Dann heißt ein Intervall der Form

$$\kappa(v) := [G_u(v), G_o(v)], \quad (1)$$

so dass

$$\mathbb{P}_\theta(\kappa(v) \ni \theta) = \mathbb{P}_\theta(G_u(v) \leq \theta \leq G_o(v)) = \delta \text{ für alle } \theta \in \Theta \text{ gilt} \quad (2)$$

ein δ -Konfidenzintervall für θ . δ ist die Überdeckungswahrscheinlichkeit von $\kappa(v)$ für θ und wird *Konfidenzlevel* genannt. Die Statistiken $G_u(v)$ und $G_o(v)$ heißen die *unteren* und *oberen Grenzen* des Konfidenzintervalls, respektive.

Bemerkungen

- θ ist fest, nicht zufällig, und unbekannt.
- κ ist ein zufälliges Intervall, weil $G_u(v)$ und $G_o(v)$ Zufallsvariablen sind.
- $\kappa \ni \theta$ bedeutet $\theta \in \kappa$, aber κ ist zufällig und steht deshalb vorn (cf. $\mathbb{P}(\xi = x)$).
- Ein δ -Konfidenzintervall überdeckt den wahren Wert θ mit Wahrscheinlichkeit δ .
- Oft wird $\delta = 0.95$ gewählt, also *95%-Konfidenzintervalle* betrachtet.

Zwei Interpretationen von δ -Konfidenzintervallen

- (1) Wird ein Zufallsvorgang unter gleichen Umständen häufig wiederholt, so überdeckt das zugehörige δ -Konfidenzintervall den wahren, aber unbekannten Parameterwert im langfristigen Mittel in $\delta \cdot 100\%$ der Fälle. Technischer ausgedrückt, für unabhängig und identisch realisierte Stichproben einer Verteilung mit wahren, aber unbekannten, Parameter θ überdeckt im langfristigen Mittel ein entsprechendes δ -Konfidenzintervall θ in $\delta \cdot 100\%$ aller Fälle.
- (2) Gegeben sei eine Menge von Zufallsvorgängen mit wahren, aber unbekannten, Parametern $\theta_1, \theta_2, \dots$ und realisierte δ -Konfidenzintervalle für eben jene Menge von wahren, aber unbekannten Parametern $\theta_1, \theta_2, \dots$. Dann überdecken im langfristigen Mittel $\delta \cdot 100\%$ der Konfidenzintervalle den wahren, aber unbekannten, Wert θ_i für $i = 1, 2, \dots$. Technischer ausgedrückt, für unabhängig realisierte Stichproben von Verteilungen mit wahren, aber unbekannten, Parametern $\theta_1, \theta_2, \dots$ überdecken im langfristigen Mittel entsprechende δ -Konfidenzintervalle θ_i für $i = 1, 2, \dots$ in $\delta \cdot 100\%$ aller Fälle.

Wir demonstrieren im Folgenden beide Interpretationen mithilfe von Simulationen.

Allgemeine Konstruktion von Konfidenzintervallen

- (1) Definition des Frequentistischen Inferenzmodells
- (2) Definition der Konfidenzintervallstatistik
- (3) Analyse der Verteilung der Konfidenzintervallstatistik
- (4) Etablierung der Konfidenzbedingung

Beispiel

- (1) Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Definition

Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Anwendungsbeispiel

Selbstkontrollfragen

(1) Definition des Frequentistischen Inferenzmodells

$v := (v_1, \dots, v_n)$ sei die Stichprobe eines Normalverteilungsmodells mit unbekannten Parametern μ und $\sigma^2 > 0$. Wir entwickeln ein δ -Konfidenzintervall für den Erwartungswertparameter μ .

(2) Definition der Statistik

Wir betrachten die T -Konfidenzintervallstatistik

$$T := \frac{\sqrt{n}}{S}(\bar{v} - \mu) \text{ mit } \bar{v} := \frac{1}{n} \sum_{i=1}^n v_i \text{ und } S := \sqrt{\frac{1}{n-1} \sum_{i=1}^n (v_i - \bar{v})^2}. \quad (3)$$

(3) Analyse der Verteilung der Konfidenzintervallstatistik

Für die T -Konfidenzintervallstatistik gilt $T \sim t(n-1)$, die T -Konfidenzintervallstatistik ist also eine t -verteilte Zufallsvariable mit Freiheitsgradparameter $n-1$. Wir verzichten auf einen Beweis. Die T -Konfidenzintervallstatistik ist eine Funktion der Stichprobe $v_1, \dots, v_n$ (via $\bar{v}$ und S), während ihre Verteilung weder von μ noch von σ^2 abhängt. Wir bezeichnen die WDF einer t -verteilten Zufallsvariable mit t , die KVF einer t -verteilten Zufallsvariable mit Ψ und die inverse KVF einer t -verteilten Zufallsvariable mit Ψ^{-1} .

(3) Analyse der Verteilung der Konfidenzintervallstatistik (fortgeführt)

```
# Modellformulierung
mu      = 10
sigsqr  = 4
n       = 12
ns      = 1e4
res     = 1e3

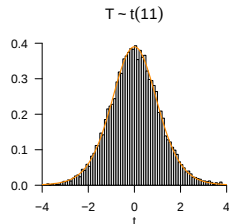
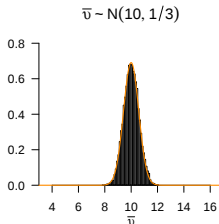
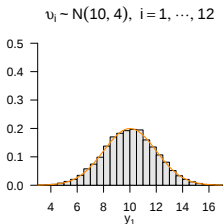
# w.a.u. Erwartungswertparameter
# wahrer bekannter Varianzparameter
# Stichprobenumfang
# Anzahl Stichprobenrealisierungen
# Ausgangsraumaufloesung

# analytische Definitionen und Resultate
y_1     = seq(3,17,len = res)
t       = seq(-4,4,len = res)
p_y_1   = dnorm(y_1,mu,sqrt(sigsqr))
p_y_bar = dnorm(y_1,mu,sqrt(sigsqr/n))
p_t     = dt(t,n-1)

# y_1 Raum
# t Raum
# y_1 WDF
# y_bar WDF
# t WDF

# Simulation
y_i      = rep(NaN,ns)
y_bar    = rep(NaN,ns)
Tee      = rep(NaN,ns)
for(s in 1:ns){
  y       = rnorm(n,mu,sqrt(sigsqr))
  y_i[s]  = y[1]
  y_bar[s] = mean(y)
  Tee[s]  = sqrt(n)*((y_bar[s] - mu)/sqrt(var(y)))
}
```

(3) Analyse der Verteilung der Konfidenzintervallstatistik (fortgeführt)



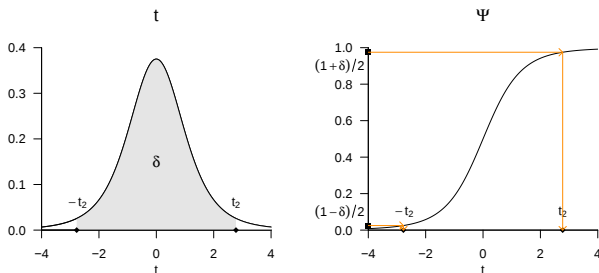
Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

(4) Etablierung der Konfidenzbedingung

Für $\delta \in]0, 1[$ seien

$$t_1 := \Psi^{-1}\left(\frac{1-\delta}{2}; n-1\right) \text{ und } t_2 := \Psi^{-1}\left(\frac{1+\delta}{2}; n-1\right) \quad (4)$$

Es gilt dann $(1+\delta)/2 - (1-\delta)/2 = \delta$ und zum Beispiel gilt für $n = 5$ und $\delta = 0.95$, $t_1 = \Psi^{-1}(0.025; 4) = -2.57$ und $t_2 = \Psi^{-1}(0.975; 4) = 2.57$. Weiterhin gilt mit der Symmetrie von $t(n-1)$, $t_1 = -t_2$. Es gilt hier also per Definition $\mathbb{P}(-t_2 \leq T \leq t_2) = \delta$.



(4) Etablierung der Konfidenzbedingung (fortgeführt)

Mit der Definition von t_2 wie oben folgt dann aber

$$\begin{aligned}\delta &= \mathbb{P}(-t_2 \leq T \leq t_2) \\&= \mathbb{P}\left(-t_2 \leq \frac{\sqrt{n}}{S}(\bar{v} - \mu) \leq t_2\right) \\&= \mathbb{P}\left(-\frac{S}{\sqrt{n}}t_2 \leq \bar{v} - \mu \leq \frac{S}{\sqrt{n}}t_2\right) \\&= \mathbb{P}\left(-\bar{v} - \frac{S}{\sqrt{n}}t_2 \leq -\mu \leq -\bar{v} + \frac{S}{\sqrt{n}}t_2\right) \\&= \mathbb{P}\left(\bar{v} + \frac{S}{\sqrt{n}}t_2 \geq \mu \geq \bar{v} - \frac{S}{\sqrt{n}}t_2\right) \\&= \mathbb{P}\left(\bar{v} - \frac{S}{\sqrt{n}}t_2 \leq \mu \leq \bar{v} + \frac{S}{\sqrt{n}}t_2\right) \\&= \mathbb{P}\left(\left[\bar{v} - \frac{S}{\sqrt{n}}t_2, \bar{v} + \frac{S}{\sqrt{n}}t_2\right] \ni \mu\right).\end{aligned}\tag{5}$$

Theorem (Konfidenzintervall für den Erwartungswertparameter)

Gegeben sei das Normalverteilungsmodell

$$v_1, \dots, v_n \sim N(\mu, \sigma^2) \quad (6)$$

mit wahren, aber unbekannten, Parametern μ und σ^2 , es sei $\delta \in]0, 1[$ und es sei

$$t_\delta := \Psi^{-1} \left(\frac{1+\delta}{2}; n-1 \right). \quad (7)$$

mit der inversen KVF Ψ^{-1} einer t -verteilten Zufallsvariable. Dann gilt für das Intervall

$$\kappa(v) := \left[\bar{v} - \frac{S}{\sqrt{n}} t_\delta, \bar{v} + \frac{S}{\sqrt{n}} t_\delta \right], \quad (8)$$

mit dem Stichprobenmittel und der Stichprobenstandardabweichung

$$\bar{v} := \frac{1}{n} \sum_{i=1}^n v_i \text{ und } S := \sqrt{\frac{1}{n-1} \sum_{i=1}^n (v_i - \bar{v})^2}, \quad (9)$$

respektive, dass

$$\mathbb{P}_\mu(\kappa(v) \ni \mu) = \delta. \quad (10)$$

Bemerkung

- Der Beweis dieses Theorems ist durch obige Konstruktion gegeben.
- κ ist ein zufälliges Intervall, weil $\bar{v}$ und S Zufallsvariablen sind.

Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Simulation der ersten Interpretation eines Konfidenzintervalls

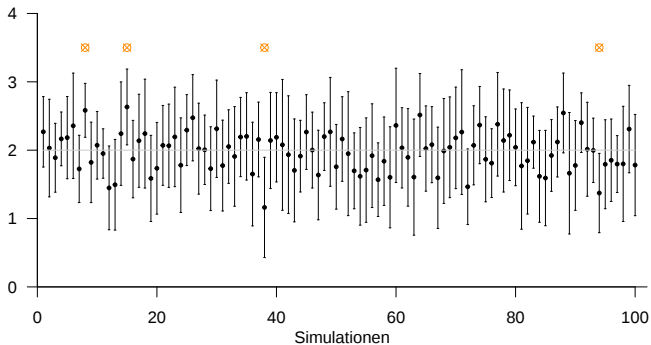
```
# Modellformulierung
set.seed(1)                                # random number generator seed
mu      = 2                                # w.a.u. Erwartungswertparameter
sigsqr  = 1                                # w.a.u. Varianzparameter
sigma   = sqrt(sigsqr)                     # w.a.u. St.Abweichungsparameter
n       = 12                               # Stichprobenumfang
delta   = 0.95                             # Konfidenzbedingung
t_delta = qt((1+delta)/2,n-1)              # \Psi^{-1}((\delta + 1)/2, n-1)

# Simulation
ns      = 1e2                               # Anzahl Simulationen
y_bar   = rep(NA,n)                         # Stichprobenmittelarray
S       = rep(NA,n)                         # St.Abweichungsarray
kappa   = matrix(rep(NA,2*ns), ncol = 2)    # Konfidenzintervallarray
for(i in 1:ns){
  y      = rnorm(n,mu,sigma)                # Stichprobenrealisierung
  y_bar[i] = mean(y)                       # Stichprobenmittel
  S[i]    = sd(y)                          # Stichprobenstandardabweichung
  kappa[i,1] = y_bar[i] - (S[i]/sqrt(n))*t_delta # untere Konfidenzintervallgrenze
  kappa[i,2] = y_bar[i] + (S[i]/sqrt(n))*t_delta # obere Konfidenzintervallgrenze
}
```

Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Simulation der ersten Interpretation eines Konfidenzintervalls

$$\mu = 2 \quad \sigma^2 = 2, \quad n = 12, \quad \delta = 0.95$$



Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Simulation der zweiten Interpretation eines Konfidenzintervalls

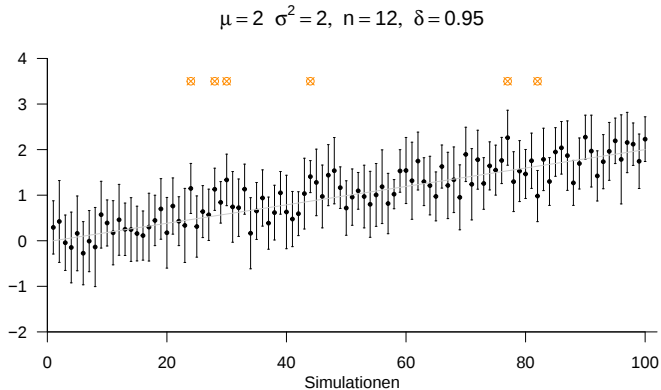
```
# Anzahl Simulationen mit \theta_1, \theta_2,...
set.seed(2)                                # random number generator seed
ns      = 1e2                              # Anzahl Simulationen

# Modellformulierung
mu       = 2 * seq(0,1,len = ns)           # w.a.u. Erwartungswertparameter
sigsqr   = 1                               # w.a.u. Varianzparameter
sigma    = sqrt(sigsqr)                   # w.a.u. St.Abweichungsparameter
n        = 12                             # Stichprobenumfang
delta    = 0.95                           # Konfidenzbedingung
t_delta  = qt((1+delta)/2,n-1)            # \Psi^{-1}((\delta + 1)/2, n-1)

# Simulation
y_bar    = rep(NaN,ns)                    # Stichprobenmittelarray
S        = rep(NaN,ns)                    # St.Abweichungsarray
kappa    = matrix(rep(NaN,2*ns), ncol = 2) # Konfidenzintervallarray
for(i in 1:ns){
  y      = rnorm(n,mu[i],sigma)           # Stichprobenrealisierung
  y_bar[i] = mean(y)                      # Stichprobenmittel
  S[i]    = sd(y)                         # Stichprobenstandardabweichung
  kappa[i,1] = y_bar[i] - (S[i]/sqrt(n))*t_delta # untere Konfidenzintervallgrenze
  kappa[i,2] = y_bar[i] + (S[i]/sqrt(n))*t_delta # obere Konfidenzintervallgrenze
}
```

Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Simulation der zweiten Interpretation eines Konfidenzintervalls



Definition

Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Anwendungsbeispiel

Selbstkontrollfragen

Definition

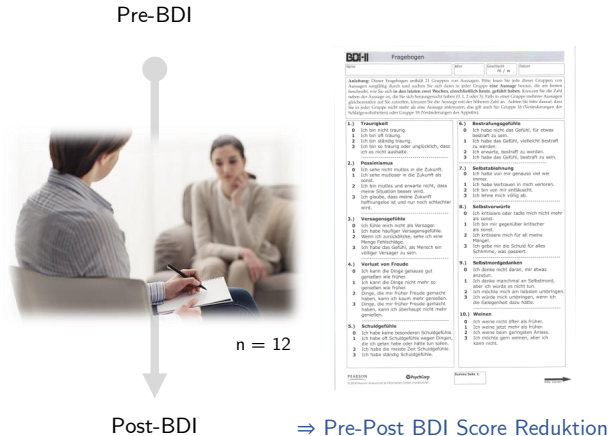
Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Anwendungsbeispiel

Selbstkontrollfragen

Anwendungsbeispiel

Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression



Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

i	BDI.Reduktion
1	-1
2	3
3	-2
4	9
5	3
6	-2
7	4
8	5
9	5
10	1
11	9
12	4

Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

Für die Pre-Post BDI Score Reduktion v_i der i ten von n Patient:innen legen wir das Modell

$$v_i = \mu + \varepsilon_i \text{ mit } \varepsilon_i \sim N(0, \sigma^2) \text{ u.i.v. für } i = 1, \dots, n \quad (11)$$

zugrunde. Dabei wird die Pre-Post BDI Reduktion v_i der i ten Patient:in also mithilfe einer über die Gruppe von Patient:innen identischen Pre-Post BDI Score Reduktion $\mu \in \mathbb{R}$ und einer Patient:innen-spezifischen normalverteilten Pre-Post BDI Score Reduktionsabweichung ε_i erklärt

Wie gezeigt ist dieses Modell äquivalent zum Normalverteilungsmodell

$$v_1, \dots, v_n \sim N(\mu, \sigma^2). \quad (12)$$

Die Standardproblemstellungen der Frequentistischen Inferenz führen dann auf folgende Fragen:

- (1) Was sind sinnvolle Tipps für die wahren, aber unbekannten, Parameterwerte μ und σ^2 ?
- (2) Wie kann im Sinne einer Intervallschätzung eine möglichst sichere Schätzung von μ gelingen?
- (3) Entscheiden wir uns sinnvollerweise für die Hypothese, dass gilt $\mu \neq 0$?

Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

Wir haben in (10) Parameterschätzung gesehen, dass unverzerzte Schätzer für den Erwartungswertparameter μ und den Varianzparameter σ^2 durch das Stichprobenmittel und die Stichprobenvarianz gegeben sind.

```
mu_hat      = mean(y)    # Stichprobenmittel als Erwartungswertparameterschätzer  
sigsqr_hat = var(y)      # Stichprobenvarianz als Varianzparameterschätzer
```

Es sind also $\hat{\mu} = 3.17$ und $\hat{\sigma}^2 = 13.8$ sinnvolle Tipps für μ und σ^2 basierend auf den vorliegenden 12 Datenpunkten.

Um eine möglichst sichere Schätzung für μ zu erlangen, geht die Frequentistische Inferenz zur Intervallschätzung mithilfe von δ -Konfidenzintervallen über, für die die assoziierte Unsicherheit dann ein prädefiniertes Level von $1 - \delta$ hat, im langfristigen Mittel überdeckt ein angegebenes δ -Konfidenzintervall den wahren, aber unbekannten Parameterwert in (nur) $1 - \delta$ 100 von 100 Fällen nicht. Für ein großes δ wie $\delta = 0.95$ ist die mit dieser Intervallschätzung assoziierte Sicherheit also eher hoch, die assoziierte Unsicherheit eher gering.

Beispiel | Evidenzbasierte Evaluation von Psychotherapie bei Depression

```
# Konfidenzintervall für den Erwartungswertparameter
delta    = 0.95                                # Konfidenzlevel
n        = length(y)                          # Anzahl Datenpunkte
t_delta  = qt((1+delta)/2,n-1)                # \psi^{-1}((\delta + 1)/2, n-1)
y_bar    = mean(y)                            # Stichprobenmittel
s        = sd(y)                              # Stichprobenstandardabweichung
mu_hat   = y_bar                              # Erwartungswertparameterschätzer
mu_hat_u = y_bar - (s/sqrt(n))*t_delta        # untere Konfidenzintervallgrenze
mu_hat_o = y_bar + (s/sqrt(n))*t_delta        # obere Konfidenzintervallgrenze
```

Das 0.95-Konfidenzintervall für den Erwartungswertparameter ist $[0.80, 5.52]$. Im langfristigen Mittel überdeckt so berechnetes Konfidenzintervall den wahren, aber unbekannten, Erwartungswertparameter in 95 von 100 Fällen.

Definition

Konfidenzintervall für den Erwartungswertparameter der Normalverteilung

Anwendungsbeispiel

Selbstkontrollfragen

1. Geben Sie die Definition des Begriffs eines δ -Konfidenzintervalls wieder.
2. Erläutern Sie die zwei Interpretationen eines δ -Konfidenzintervalls.
3. Erläutern Sie die typischen Schritte zur Konstruktion eines δ -Konfidenzintervalls.
4. Geben Sie das Theorem zum δ -Konfidenzintervall für den Erwartungswert der Normalverteilung wieder.

1. Siehe Definition (δ -Konfidenzintervall).
2. Siehe Folie 9, *Zwei Interpretationen von δ -Konfidenzintervallen*.
3. Siehe Folie 10, *Allgemeine Konstruktion von Konfidenzintervallen*.
4. Siehe Theorem (Konfidenzintervall für den Erwartungswertparameter).