Statistik und Wahrscheinlichkeitsrechnung

– Wintersemester 2019/20 –

Kapitel 06: Schätzverfahren

Prof. Dr. Adrian Ulges

Angewandte Informatik (B.Sc.) /
Informatik - Technische Systeme (B.Sc.) /
Wirtschaftsinformatik (B.Sc.)

Fachbereich DCSM
Hochschule RheinMain

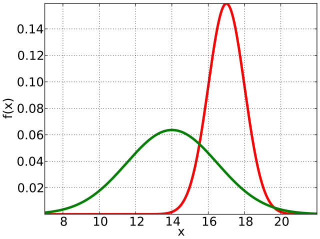
Outline

Warum Parameterschätzung?

- Wir haben im letzten Kapitel verschiedene **Verteilungstypen** kennengelernt. Diese Verteilungstypen definieren eine grobe Form der Dichtefunktion/Verteilungsfunktion.
- *Aber: Wie genau die Verteilung aussieht, hängt von* **Parametern** Θ *ab.*

Beispiel: Die Normalverteilung

- **Parameter**: Erwartungswert μ und Varianz σ^2.
- Der Parametervektor lautet $\Theta := (\mu, \sigma^2)$.



Warum Parameterschätzung?

Vorgehen in der Praxis

- Wähle einen Verteilungstyp (✓, Kapitel 5)
- Schätze die Parameter Θ (in diesem Kapitel)
- Wende die Verteilung an (✓, Kapitel 3-4)
  (und berechne so Wahrscheinlichkeiten für den Ausgang zukünftiger Experimente)

Beispiel Bild: [2]

Wir modellieren die Warteschlange eines Servers und berechnen die Anzahl eingehender Jobs pro Minute mit der Poisson-Verteilung. *Hierzu müssen wir den Parameter λ ermitteln.*

Definition: Punkt- vs. Intervallschätzung

Wir ermitteln Θ mit Methoden der **Parameterschätzung**.
Von diesen unterscheiden wir zwei Arten:

Definition (Punktschätzung)

Gegeben eine Stichprobe, bestimme **Näherungswerte $\hat{\boldsymbol{\Theta}}$** *für die Parameter Θ der zugehörigen Verteilung.*

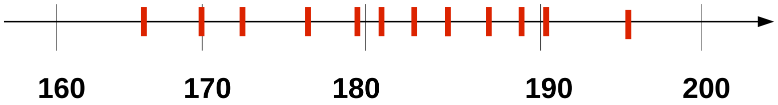
Definition (Intervallschätzung)

Gegeben eine Stichprobe, bestimme sogenannte **Konfidenz- oder Vertrauensintervalle**, *in denen die Parameter "höchstwahrscheinlich" liegen.*

Outline

Punktschätzung: Einfaches Beispiel

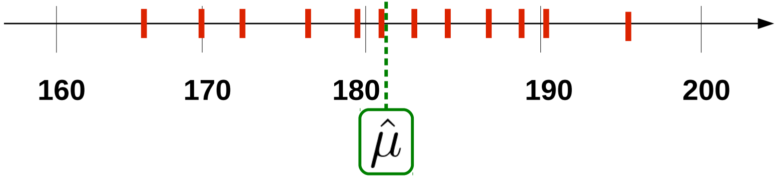
Beispiel "Körpergröße": Schätzer für den Erwartungswert μ



(Intuitive Lösung)

Unsere Schätzung für den Erwartungswert $\hat{\mu}$ entspricht gerade dem empirischen Mittelwert der Stichprobe:

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



Punktschätzung: Komplizierteres Beispiel

Das "German Tank Problem" Bild: [3]

- Historisches Problem (2. Weltkrieg)
- **Stichprobe**: Seriennummern zerstörter deutscher Panzer, z.B.

  $$x_1, .., x_5 = 135, 785, 425, 1267, 669$$

- **Annahme**: Seriennummern sind gleichverteilt im Bereich $1, 2, ..., N$
- **Ziel**: Schätze N, also die Anzahl deutscher Panzer insgesamt

month[1]	statistical estimate	intelligence estimate	German records
Jun 1940	169	1,000	122
Jun 1941	244	1,550	271
Aug 1942	327	1,550	342

[1] Quelle: en.wikipedia.org

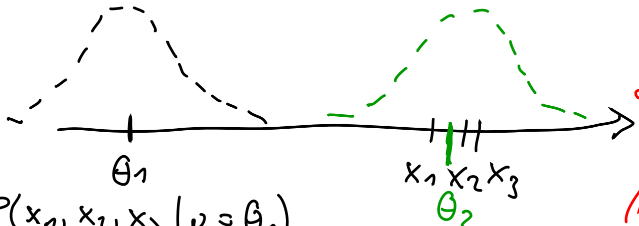
Outline

Maximum-Likelihood (ML)-Schätzung

Definition (Idee der ML-Schätzung)

Gegeben sei eine Stichprobe $x_1, ..., x_n$ und ein zu schätzender Parameter θ. Die Idee der ML-Schätzung ist, für den Parameter den Schätzwert $\hat{\theta}$ zu wählen, **der die Wahrscheinlichkeit, die Stichprobe zu beobachten, maximiert**.

Illustration $\theta = \mu$



"Likelihood" $L(\theta)$

$P(x_1, x_2, x_3 \mid \mu = \theta_1)$ niedrig $\Rightarrow$ schlecht!

$P(x_1, x_2, x_3 \mid \mu = \theta_2)$ hoch :)

ML-Schätzung: Herleitung

Wir definieren die sogenannte **Likelihood-Funktion** L und finden ihr Maximum (deshalb *"maximum-likelihood", oder kurz "ML"*):

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmax}} \; L(\Theta)$$

$$= \quad \text{''} \quad P(x_1, x_2, \ldots, x_n \mid \Theta)$$

unabhängig!

$$= \quad \text{''} \quad P(x_1 \mid \Theta) \cdot P(x_2 \mid \Theta) \cdot \ldots \cdot P(x_n \mid \Theta)$$

Für stetige Zufallsvariablen:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmax}} \; f(x_1; \Theta) \cdot f(x_2; \Theta) \cdot \ldots \cdot f(x_n; \Theta)$$

ML-Schätzung: Herleitung

Beispiel 1: Bernoulli-Verteilung

- Wir ermitteln den Parameter p *(Erfolgswahrscheinlichkeit)*
- **Stichprobe**: Wir werfen die Münze zehn mal

$$x_1, \dots, x_{10} = 1, 0, 0, 1, 1, 1, 0, 1, 1, 0$$

$$\begin{aligned}
\hat{p} &= \underset{p \in [0,1]}{\arg\max}\ P(x_1 \mid p) \cdot P(x_2 \mid p) \cdot \ldots \cdot P(x_n \mid p) \\
&= \quad " \quad P(X=1 \mid p) \cdot P(X=0 \mid p) \cdot \ldots \cdot P(X=0 \mid p) \\
&= \quad " \quad p \cdot (1-p) \cdot \ldots \cdot (1-p) \\
&= \quad " \quad p^6 \cdot (1-p)^4 \qquad \downarrow \log \\
&= \quad " \quad \log\left(p^6 \cdot (1-p)^4\right) \\
&= \quad " \quad 6 \cdot \log(p) + 4 \cdot \log(1-p)
\end{aligned}$$

Beispiel 1: Bernoulli-Verteilung

$$\hat{p} = \arg\max_{p} \; 6\cdot\log(p) + 4\cdot\log(1-p)$$

Ableiten:

$$6\cdot\frac{1}{p} + 4\cdot\frac{1}{1-p}\cdot(-1) \overset{!}{=} 0$$

$$6\cdot(1-p) - 4\cdot p = 0$$

$$6 - 10p = 0$$

$$p = 6/10.$$

Generell lautet der ML-Schätzer:

$$\hat{p} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{\#\text{Erfolge}}{\#\text{Versuche}}$$

Schätzfunktion

Definition: "Schätzfunktion"

Von der Stichprobe zur Schätzfunktion

- Die Werte einer Stichprobe $x_1, x_2, ..., x_n$ lassen sich als Realisierungen von n Zufallsvariablen $X_1, X_2, ..., X_n$ auffassen.
- Üblicher Weise nehmen wir an, die Zufallsvariablen gehören zur **gleichen Verteilung** und sind **unabhängig** (engl. *"independent and identically distributed"*, oder kurz *"i.i.d."*).
- Eine **Schätzfunktion** ist eine Funktion $g(X_1, X_2, ..., X_n)$, die den n Zufallsvariablen einen Näherungswert eines Parameters θ zuordnet.

Beispiel: Münzwurf

- Unsere Schätzfunktion für die Erfolgswahrscheinlichkeit p:

$$g(X_1, ..., X_n) = \frac{X_1 + ... + X_n}{n} \quad \left(= \frac{\#Erfolge}{\#Versuche} \right)$$

Definition: “Schätzfunktion”

Anmerkungen

- Ist eine konkrete Stichprobe $x_1, ..., x_n$ gegeben, nennen wir den Funktionswert der Schätzfunktion den **Schätzwert**:

$$\hat{p} := \frac{x_1 + ... + x_n}{n}$$

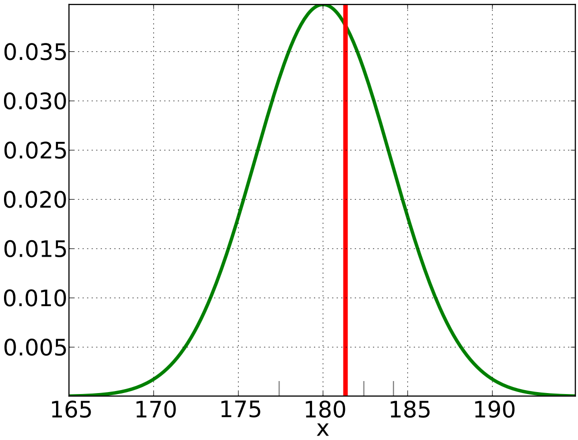
Experiment

Mit einem kleinen Experiment *(folgende Slides)* zeigen wir:

- Die Schätzfunktion $g(X_1, ..., X_2)$ ist - weil Sie von Zufallsvariablen abhängt - **ebenfalls eine Zufallsvariable**.
- Eine gute Schätzfunktion g hat als **Erwartungswert** den echten Wert θ des zu schätzenden Parameters.
- Eine gute Schätzfunktion g hat **geringe Varianz** *(die Varianz nimmt üblicher Weise auch mit der Stichprobengröße ab)*.

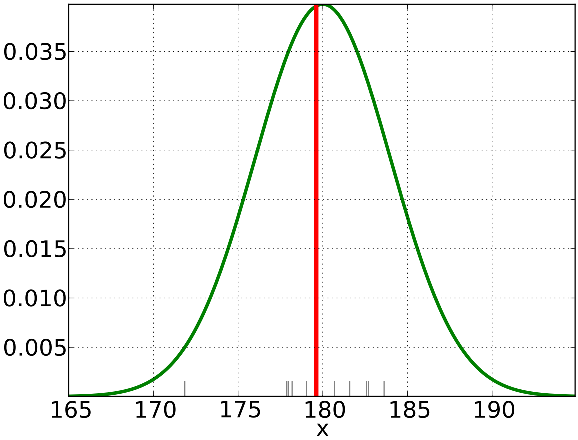
Beispiel "Körpergröße": Experiment

- Die Körpergröße sei normalverteilt mit $\mu = 180$.
- Wir generieren eine **Stichprobe** x_1, x_2, x_3 und schätzen μ:



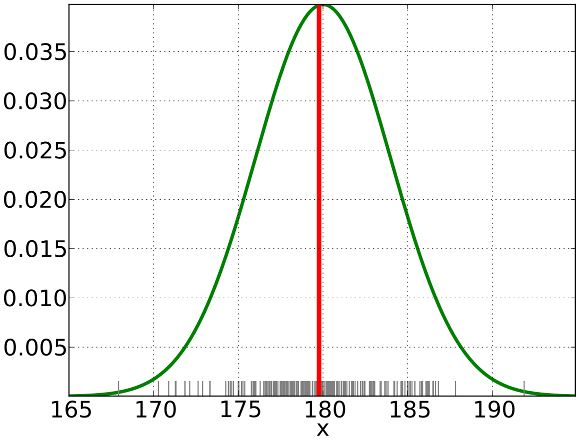
Beispiel "Körpergröße": Experiment

- Die Körpergröße ist normalverteilt mit $\mu = 180$
- Wir erhöhen die **Stichprobengröße** auf 10



Beispiel "Körpergröße": Experiment

- Die Körpergröße ist normalverteilt mit $\mu = 180$
- Wir erhöhen die **Stichprobengröße** auf 200



Was macht eine gute Schätzfunktion aus?

Definition (Erwartungstreue)

Wir nennen eine Schätzfunktion g_n für einen Parameter θ **erwartungstreu**, *wenn gilt:*

$$E(\overbrace{g(X_1, ..., X_n)}^{\hat{p}}) = \theta$$

Ist der ML-Schätzer $\hat{p}$ **erwartungstreu**?

$$^{*}\, E(X_i) = 1 \cdot p + 0 \cdot \cancel{(1-p)} = p$$

$$E(\hat{p}) = p$$

$$E(\hat{p}) = E\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right)$$

$$= \frac{1}{n}\Big(\underbrace{E(X_1)}_{*\; p} + \underbrace{E(X_2)}_{p} + \ldots + \underbrace{E(X_n)}_{p}\Big)$$

$$= \frac{1}{\cancel{n}} \cancel{n} \cdot p = p \quad \checkmark$$

ML-Schätzung

Anmerkungen

- Wir wir gerade gesehen haben, können wir alternativ das Maximum der **logarithmierten** Likelihood-Funktion *(engl. "log-likelihood")* ermitteln:

$$\hat{\theta} = \arg\max_{\theta} L(\theta) = \arg\max_{\theta} \log(L(\theta))$$

  Dies **vereinfacht oft die Berechnung** *(aus einem Produkt von Wahrscheinlichkeiten wird durch das Logarithmieren eine Summe)*.

- Das ML-Verfahren lässt sich für **mehrere Parameter** verallgemeinern: Für unbekannte Parameter $\theta_1, ..., \theta_r$ erhalten wir r Bedingungen

$$\frac{\partial L}{\partial \theta_1} = \frac{\partial L}{\partial \theta_2} = ... = \frac{\partial L}{\partial \theta_r} = 0$$

Beispiel 2: Normalverteilung

- **Gegeben**: Eine Normalverteilung $\mathcal{N}(\mu, \sigma^2)$
- **Gesucht**: Erwartungswert μ und Varianz σ^2
- Gegeben ist außerdem eine **Stichprobe** $x_1, \ldots, x_n$.
- Wir leiten die ML-Schätzer $\hat{\mu}$ und $\hat{\sigma}^2$ her.

$$e^a \cdot e^b = e^{a+b}$$

$$L(\mu,\sigma) = \underbrace{f(x_1;\mu,\sigma)}_{\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-\frac{(x_1-\mu)^2}{2\sigma^2}}} \cdot f(x_2;\mu,\sigma) \cdots \underbrace{f(x_n;\mu,\sigma)}_{\frac{1}{\sqrt{2\pi}\,\sigma}\cdot e^{-\frac{(x_n-\mu)^2}{2\sigma^2}}}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \cdot \left(\frac{1}{\sigma}\right)^n \cdot e^{-\frac{1}{2\sigma^2}\cdot\left((x_1-\mu)^2 + \ldots + (x_n-\mu)^2\right)}$$

$$\log(L(\mu,\sigma)) = n\cdot\log\left(\frac{1}{\sqrt{2\pi}}\right) + n\cdot\log\left(\frac{1}{\sigma}\right) - \frac{1}{2\sigma^2}\cdot\sum_i (x_i-\mu)^2$$

Beispiel 2: Normalverteilung

$$\log L(\mu,\sigma) = n \cdot \log\left(\frac{1}{\sqrt{2\pi}}\right) + n \cdot \log\left(\frac{1}{\sigma}\right) - \frac{1}{2\sigma^2} \cdot \sum_i (x_i-\mu)^2$$

$$\log L_\mu(\mu,\sigma) = 0 + 0 \cancel{-} \frac{1}{\cancel{2}\sigma^2} \cdot \sum_i \cancel{2} \cdot (x_i-\mu) \cdot \cancel{(-1)}$$

(f_x)

$$\cancel{\frac{1}{\sigma^2}} \sum_i (x_i-\mu) \overset{!}{=} 0 \quad | \cdot \sigma^2$$

$$x_1+x_2+\ldots+x_n \underbrace{-\mu-\mu-\mu-\mu\ldots-\mu}_{n\times} = 0$$

$$\sum_i x_i = n \cdot \mu \quad | : n$$

$$\mu = \frac{1}{n} \sum_i x_i$$

Beispiel 2: Normalverteilung

Beispiel 2: Normalverteilung

Beispiel 2: Normalverteilung

Die Maximum-Likelihood-Schätzwerte für Erwartungswert μ und Varianz σ^2 einer Normalverteilung lauten also:

$$\hat{\mu} = \frac{1}{n} \sum_i x_i \quad \left(= \overline{x} \right)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_i (x_i - \overline{x})^2 \quad \left(= s^2 \right)$$

Erwartungstreue

- Achtung: Der Schätzer $\hat{\mu}$ ist erwartungstreu, aber $\hat{\sigma}^2$ nicht!
- Deshalb verwendet man häufiger die **korrigierte** Stichprobenvarianz *(siehe Kapitel 1)*. Diese ist erwartungstreu:

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_i (x_i - \overline{x})^2 \quad \left(= s^{*2} \right)$$

Normalverteilung: Erwartungstreue

Normalverteilung: Erwartungstreue

Normalverteilung: Erwartungstreue

Überblick: ML-Schätzer für wichtige Verteilungen

Verteilung	**Parameter**	**Schätzwert**	**Bemerkungen**
Binomial $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	p	$\hat{p}=\frac{k}{n}$	k = #Erfolge in der Stichprobe
Poisson $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$	λ	$\hat{\lambda}=\overline{x}$	$\overline{x}$ = Mittelwert der Stichprobe
Exponential $f(x)=\lambda\cdot e^{-\lambda x}$	λ	*Seminar*	—
Normalverteilung $f(x)=\frac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ σ^2	$\hat{\mu}=\overline{x}$ $\hat{\sigma}^2=s^2$	s^2 = Stichprobenvarianz

Outline

Outline

Motivation von Intervallschätzungen

Problem

- Punktschätzer treffen generell keine Aussage über ihre **Genauigkeit/Zuverlässigkeit**.
- Gerade bei kleinen Stichproben kann die Schätzung stark vom wahren Wert abweichen!

Deshalb Intervallschätzer

- Intervallschätzer liefern uns ein **Konfidenzintervall**, in dem der wahre Parameter mit einer gegebenen Wahrscheinlichkeit liegt (dies drückt die **Zuverlässigkeit des Schätzers** aus).

Illustration

Intervallschätzer: Genereller Ansatz

Anmerkungen

- Je *größer* wir die Stichprobe wählen...,
  desto **kleiner** das Konfidenzintervall.
- Je *kleiner* das Konfidenzintervall (bei festem γ),
  desto **genauer** der Schätzer.
- Je *größer* wir γ wählen..., desto **größer** das Konfidenzintervall.

Intervallschätzer: Vorgehen

Definition (Vorgehen bei Intervallschätzern)

Es sei X eine Zufallsvariable mit bekannter Verteilung und unbekanntem Parameter θ. Es sei $x_1, \ldots, x_n$ eine (i.i.d.) Stichprobe. Wir bestimmen das Konfidenzintervall $[c_u, c_o]$ für θ mit diesen Schritten:

1. *Wähle ein* **Konfidenzniveau** γ
   (bzw. eine **Irrtumswahrscheinlichkeit** $\alpha := 1-\gamma$*).*
2. *Bestimme* **zwei Schätzfunktionen** g_u, g_o*, die mit Wahrscheinlichkeit* γ *den wahren Wert von* θ *einschließen:*

$$P\Big(g_u(X_1, \ldots, X_n) \leq \theta \leq g_o(X_1, \ldots, X_n)\Big) = \gamma$$

3. *Bestimme* $c_u := g_u(x_1, \ldots, x_n)$ *und* $c_o := g_o(x_1, \ldots, x_n)$ *aus der gegebenen* **Stichprobe** $x_1, \ldots, x_n$.
4. *Das* **Konfidenzintervall** *lautet also* $[c_u, c_o]$.
   θ *liegt mit Wahrscheinlichkeit* γ *innerhalb dieses Intervalls.*

Outline

Erwartungswert μ: Intervallschätzer

- X sei normalverteilt, σ^2 sei bekannt, und μ gesucht.
- Aus einer Stichprobe $x_1, \ldots, x_n$ haben wir den ML-Schätzer $\hat{\mu} = \overline{x}$ berechnet.
- Wir wählen ein Konfidenzniveau $\gamma := 1 - \alpha$ und leiten das Konfidenzintervall $[c_u, c_o]$ her:

Erwartungswert μ: Intervallschätzer

Erwartungswert μ: Intervallschätzer

Erwartungswert μ: Intervallschätzer

Satz (Intervallschätzer für μ bei bekanntem σ)

Gegeben ein Konfidenzniveau $\gamma := 1-\alpha$ *und eine Stichprobe* $x_1, \dots, x_n \sim \mathcal{N}(\mu; \sigma)$, *liegt der gesuchte Erwartungswert* μ *mit Wahrscheinlichkeit* γ *in folgendem Intervall* $[c_u, c_o]$*:*

$$P(\underbrace{\overline{x} - x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}}_{c_u} \leq \mu \leq \underbrace{\overline{x} + x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}}_{c_o}) = \gamma$$

Hierbei bezeichnet $x_{1-\frac{\alpha}{2}}$ *das* $(1-\frac{\alpha}{2})$*-Quantil der Standardnormalverteilung.*

Anmerkungen

- Gegeben γ (bzw. α), lesen wir also das (1-$\frac{\alpha}{2}$)-Quantil aus einer **Wertetabelle** ab und setzen es in die Formel ein. *Beispiel:* $\gamma = 90\% \rightarrow \alpha = 10\% \rightarrow 95\%$*-Quantil.*

Beispiel: Schokolade![2] Bild: [5]

$$\boxed{P(\underbrace{\overline{x} - x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}}_{c_u} \leq \mu \leq \underbrace{\overline{x} + x_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}}_{c_o}) = \gamma}$$

- Wir stellen Schokotafeln mit einem Sollgewicht von 100 g her.
- Das Gewicht der Tafeln variiert, mit einer Standardabweichung von $\sigma = 2$ g.
- Die Tafeln dürfen weder zu leicht sein *(Kunde ist unzufrieden)* noch zu schwer *(Kunde wird dick und Firma geht pleite)*.
- Wir wiegen 10 Tafeln. Der Mittelwert beträgt $\overline{x} = 98.9$ g.
- **Ist das eine zufällige Abweichung, oder ist unsere Maschine falsch eingestellt?**
- ...
- Wie lautet das **Konfidenzintervall für das Gewicht** μ?

[2] Beispiel siehe Teschl/Teschl, Kapitel 30, S. 336f

Beispiel: Schokolade! Bild: [5]

- Wie lautet das **Konfidenzintervall für das Sollgewicht** μ?
- Wir legen ein Konfidenzniveau von $\gamma = 95\%$ fest.

Beispiel: Schokolade! Bild: [5]

Beispiel: Schokolade! Bild: [5]

Wie sieht das Konfidenzintervall aus, wenn wir (statt 10 Tafeln) **20 Tafeln** wiegen und das Konfidenzniveau auf $\gamma = 0.90$ senken (wir nehmen an, der gemessene Mittelwert $\overline{x}$ ändert sich nicht)?

α	x_α
0.001	-3.090
0.005	-2.576
0.01	-2.326
0.025	-1.960
0.05	-1.645
0.1	-1.282
0.9	1.282
0.95	1.645
0.975	1.960
0.99	2.326
0.995	2.576
0.999	3.090

Outline

Erwartungswert μ bei unbekanntem σ Bilder: [1] [4]

- Im obigen Beispiel haben wir angenommen, σ sei **bekannt**. Dies ist in der Praxis aber meist nicht der Fall.
- Bei unbekanntem σ gehen wir analog vor. Wir ersetzen lediglich die (nun unbekannte) Standardabweichung σ gegen die zugehörige **Schätzung s^* (die korrigierte Standardabweichung der Stichprobe)**.
- Statt der standardnormalverteilten Variable **Y** betrachten wir eine (neue) Zufallsvariable **T**:

$$Y := \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} \qquad \Rightarrow \qquad T := \frac{\overline{x} - \mu}{s^*/\sqrt{n}}$$

- Die Variable T folgt einer neuen Verteilung, der **t-Verteilung von Student** *(nach William Sealy Gosset)*.

Die t-Verteilung von Student

Definition (Die t-Verteilung)

Die Verteilung einer stetigen Zufallsvariable T mit der Dichtefunktion

$$f(t;n) := B_n \cdot \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

ist die **(Student'sche) t-Verteilung** *mit n Freiheitsgraden.*

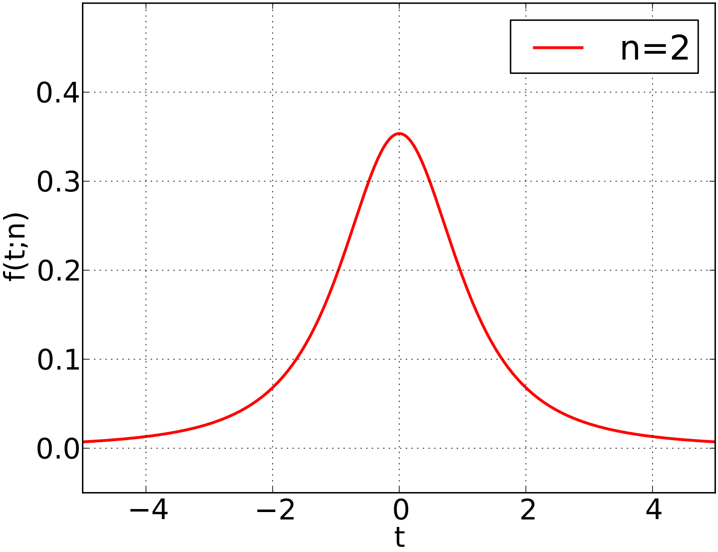
Anmerkungen

- Der Normierungsfaktor B_n ist so gewählt, dass die Dichte zu 1 integriert. Wir werden hierauf nicht näher eingehen.

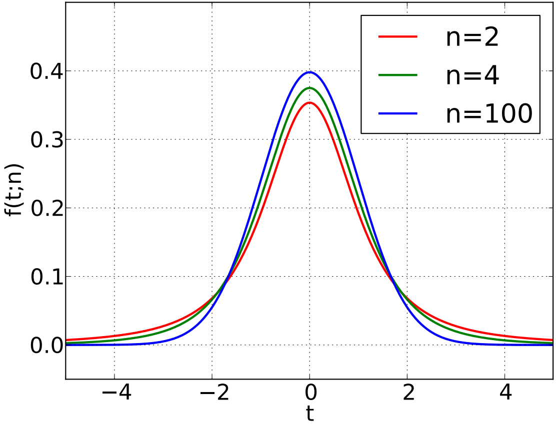
Parameter

- Einziger Parameter: $n(> 0)$, die Anzahl der Freiheitsgrade.

Die Student'sche t-Verteilung: Illustration



Die Student'sche t-Verteilung: Eigenschaften



- $f(t)$ ist symmetrisch zu $t = 0$, und $lim_{|t|\to\infty} f(t) = 0$
- Je höher der Freiheitsgrad n, desto besser wird die t-Verteilung durch die Standardnormalverteilung approximiert.
- **Faustregel**: Für $\mathbf{n} > \mathbf{30}$ können wir statt der t-Verteilung die Standardnormalverteilung verwenden.

Die Student'sche t-Verteilung: Quantile

Wir lesen die Werte der t-Verteilung üblicher Weise aus **Quantiltabellen**[3] ab:

	p				
f	0,90	0,95	0,975	0,99	0,995
1	3,078	6,314	12,707	31,820	63,654
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
22	1,321	1,717	2,074	2,508	2,819
24	1,318	1,711	2,064	2,492	2,797
26	1,315	1,706	2,056	2,479	2,779
28	1,313	1,701	2,048	2,467	2,763
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
50	1,299	1,676	2,009	2,403	2,678
60	1,296	1,671	2,000	2,390	2,660
100	1,290	1,660	1,984	2,364	2,626
200	1,286	1,653	1,972	2,345	2,601
500	1,283	1,648	1,965	2,334	2,586
⋮	⋮	⋮	⋮	⋮	⋮
∞	1,282	1,645	1,960	2,326	2,576

[3]Quelle: Papula, Mathematik für Ingenieure und Naturwissenschaftler, Band 3.

Zurück zum Intervallschätzer des Erwartungswerts

Satz (Intervallschätzer für μ bei unbekanntem σ)

Gegeben ein Konfidenzniveau $\gamma := 1-\alpha$ und eine Stichprobe $x_1, \dots, x_n \sim \mathcal{N}(\mu; \sigma)$ mit unbekanntem σ, liegt der gesuchte Erwartungswert μ mit Wahrscheinlichkeit γ in folgendem Intervall:

$$P(\underbrace{\overline{x} - t^{n-1}_{1-\frac{\alpha}{2}} \cdot \frac{s^*}{\sqrt{n}}}_{c_u} \leq \mu \leq \underbrace{\overline{x} + t^{n-1}_{1-\frac{\alpha}{2}} \cdot \frac{s^*}{\sqrt{n}}}_{c_o}) = \gamma$$

*Hierbei bezeichnet s^* die korrigierte Standardabweichung der Stichprobe, und $t^{n-1}_{1-\frac{\alpha}{2}}$ das $(1-\frac{\alpha}{2})$-***Quantil der t-Verteilung mit n-1 Freiheitsgraden***.*

Anmerkungen

- Gegeben γ (bzw. α), lesen wir also das (1-$\frac{\alpha}{2}$)-Quantil aus einer Quantiltabelle der Student'schen t-Verteilung ab.

Beispiel: Widerstände

- Wir entnehmen einer Produktion **Bauteile (elektrische Widerstände)** und vermessen acht Widerstände:

i	1	2	3	4	5	6	7	8
x_i (Ω)	100	104	98	96	101	104	98	99

- Annahme: Der Widerstand X ist normalverteilt mit **unbekannter Varianz** σ^2.
- Wir berechnen das Vertrauensintervall für μ:

Beispiel: Widerstände

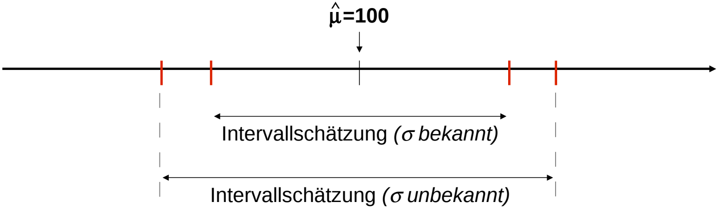
Beispiel: Widerstände

Wir nehmen zum Vergleich an, 2.875 Ω wäre nicht die *geschätzte Standardabweichung*, sondern die **echte Standardabweichung**.

- Hier gilt dieselbe Formel – nur verwenden wir *(siehe oben)* das 0.975-Quantil der **Standardnormalverteilung** x_α, nicht der **t-Verteilung** t_α^{n-1}:

α	x_α
0.001	-3.090
0.005	-2.576
0.01	-2.326
0.025	-1.960
0.05	-1.645
0.1	-1.282
0.9	1.282
0.95	1.645
0.975	1.960
0.99	2.326
0.995	2.576
0.999	3.090

Bekannte vs. unbekannte Varianz



- Die **Breite des Konfidenzintervalls** beträgt ca. 4 bei bekannter Varianz, ca. 5 bei unbekannter Varianz.
- Generell gilt: Das Konfidenzintervall ist bei unbekannter Varianz **breiter** als bei bekannter Varianz (falls $s^* = \sigma$).
- **Grund**: Die t-Verteilung verläuft *breiter* als die Normalverteilung.
- Wird die Stichprobe größer... *(Faustregel: $n \geq 30$)*
  - ... wird σ immer besser durch s^* approximixert
  - ... nähert sich die t-Verteilung der Standardnormalverteilung
  - ... sind die Ergebnisse der Intervallschätzung bei bekannter und unbekannter Varianz (nahezu) identisch.

References

[1] British statistician William Sealy Gosset, known as "Student", taken in 1908.
https://commons.wikimedia.org/wiki/File:William_Sealy_Gosset.jpg (retrieved: Jan 2017).

[2] Dylan Meconis: Parent Line.
https://flic.kr/p/5R2khv (changed to black and white, CC license, retrieved: Dec 2016).

[3] Bild 101I-635-3966-27 / Hebenstreit / CC-BY-SA 3.0 Bundesarchiv.
Panzerfabrik in Deutschland Info non-talk.svgx.
https://en.wikipedia.org/wiki/German_tank_problem#/media/File:Bundesarchiv_Bild_101I-635-3966-27,_Panzerfabrik_in_Deutschland.jpg (retrieved: Jan 2017).

[4] Chris Sansenbach (Kurisu).
Heaven (this is where the Irish go when they die).
https://flic.kr/p/31GTy (released under CC BY-ND 2.0 license, retrieved: Jan 2017).

[5] Robert Müller (robert.molinarius).
mouth-watering !
https://flic.kr/p/9g7ZPn (released under CC 2.0 license, cropped, horizontally flipped, background whitened, retrieved: Jan 2017).