Wirtschaftsmathematik

Einführung in einige Teilbereiche der Wirtschaftsmathematik

„X ist binomialverteilt mit n und p"

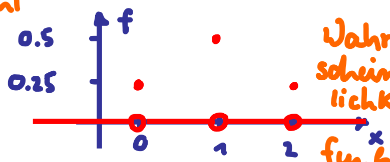
$X \sim B(n; p)$

$f(x) = \boxed{P(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}}$ $\qquad x \in 0, 1, \ldots, n$

(Funktion) Zufallsvariable

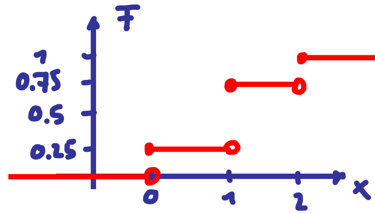
Zahl

z.B. $n = 2$, $p = 0.5$



Wahrscheinlichkeitsfunktion

x	0	1	2
$f(x)$	0.25		



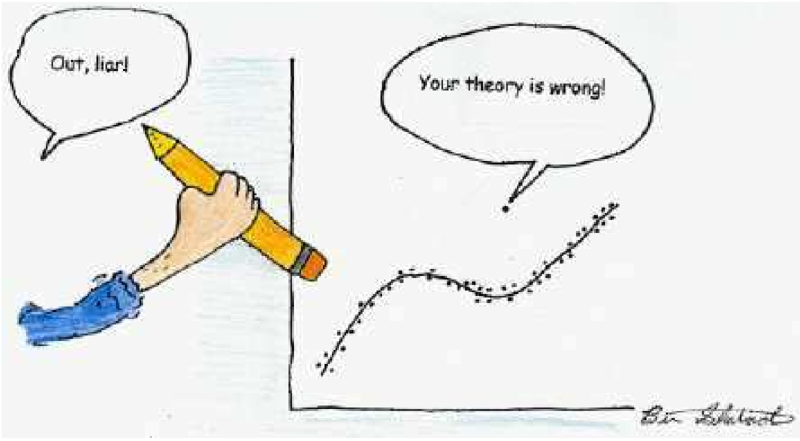
Verteilungsfunktion

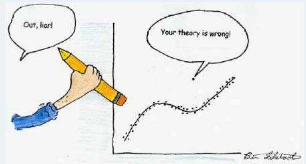
Wintersemester 2016

HSA Wing Sessionlist WS 2016				
			52	
Datum	N.	Zeit	UE	Themen
Dienstag, 20. September 2016	1	18.00-21.15	4	Einführung, Zinsen, Renten
Dienstag, 27. September 2016	2	18.00-21.15	4	Tilgung, Festverz. Wertpapiere
Dienstag, 4. Oktober 2016	3	18.00-21.15	4	Lineare Optimierung: Einführung, Lösungsmethoden
Samstag, 8. Oktober 2016	4	08.00-11.45	4	Lineare Optimierung: Standardmaximumproblem, Simplex
Dienstag, 11. Oktober 2016	5	18.00-21.15	4	Gewöhnliche Differentialgleichungen
Samstag, 15. Oktober 2016	6	08.00-11.45	4	Analytische Lösung linearer DGLs
Dienstag, 18. Oktober 2016	7	18.00-21.15	4	Einführung, univ. Statistik, Konzentration
Samstag, 22. Oktober 2016	8	11.45-15.00	4	Korrelation, Regression, Preisindizes
Dienstag, 25. Oktober 2016	9	18.00-21.15	4	Kombinatorik, Wahrscheinlichkeiten; Binomial, Hypergeo, Poisson
Samstag, 29. Oktober 2016	10	08.00-11.15	4	Zufallsvariablen, Lage- und Streuung, Stetige ZV, Gleich-vtlg.
Dienstag, 15. November 2016	11	18.00-21.15	4	Normalvtlg., Schätzen und Eigenschaften von Punktschätzern
Samstag, 26. November 2016	12	08.00-11.15	4	Konfidenzintervalle, t-Test
Dienstag, 29. November 2016	13	18.00-21.15	4	Puffer, Wiederholung Besprechung Probeklausur
Samstag, 3. Dezember 2016		09.30-11.00		Klausur (regulärer Termin 90 Min., mit Aufsicht)

Prof. Dr. Stefan Etschberger
HSA

Wirtschaftsmathematik: Table of Contents





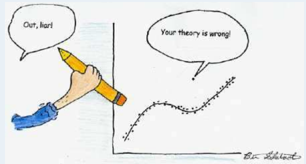
- Vollerhebung of unmöglich,
- Deshalb: Beobachte Teilgesamtheit und schließe auf Grundgesamtheit

Beispiel

Warensendung von 1000 Stück; darunter M Stück Ausschuss.
M ist unbekannt.
$\rightarrow$ Zufällige Entnahme von $n = 30$ Stück („Stichprobe").
Darunter 2 Stück Ausschuss.
Denkbare Zielsetzungen:

- Schätze M durch eine Zahl (z.B. $\frac{2}{30} \cdot 1000 = 66{,}67$)
- Schätze ein Intervall für M (z.B. $M \in [58; 84]$)
- Teste die Hypothese, dass $M > 50$ ist.

Grundbegriffe

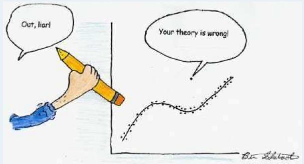


- **Grundgesamtheit (G):** Menge aller relevanten Merkmalsträger.
- **Verteilung von G:** $F(x) = P(X \leqq x) =$ Wahrscheinlichkeit, dass ein Merkmalsträger ausgewählt wird, der beim untersuchten Merkmal maximal die Ausprägung x aufweist.
- **Uneingeschränkte (reine) Zufallsauswahl:**
  Jedes Element von G hat die selbe Chance, ausgewählt zu werden.
- **Stichprobenumfang (n):** Anzahl der Merkmalsträger in der Stichprobe.
- **Einfache Stichprobe:**
  Uneingeschränkte Zufallsauswahl und unabhängige Ziehung.
  $\to$ Alle **Stichprobenvariablen** $X_1, \ldots, X_n$ sind iid.
  identical
  independ.
  distr.
- **Stichprobenergebnis:**
  n-Tupel der Realisationen der Stichprobenvariablen, $(x_1, \ldots, x_n)$.

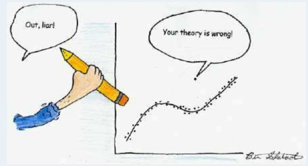
Wichtige Stichprobenfunktionen

- Gegeben: Einfache Stichprobe $X_1, \ldots, X_n$, Beliebige Verteilung,
  mit $E(X_i) = \mu$, $\mathrm{Var}(X_i) = \sigma^2$

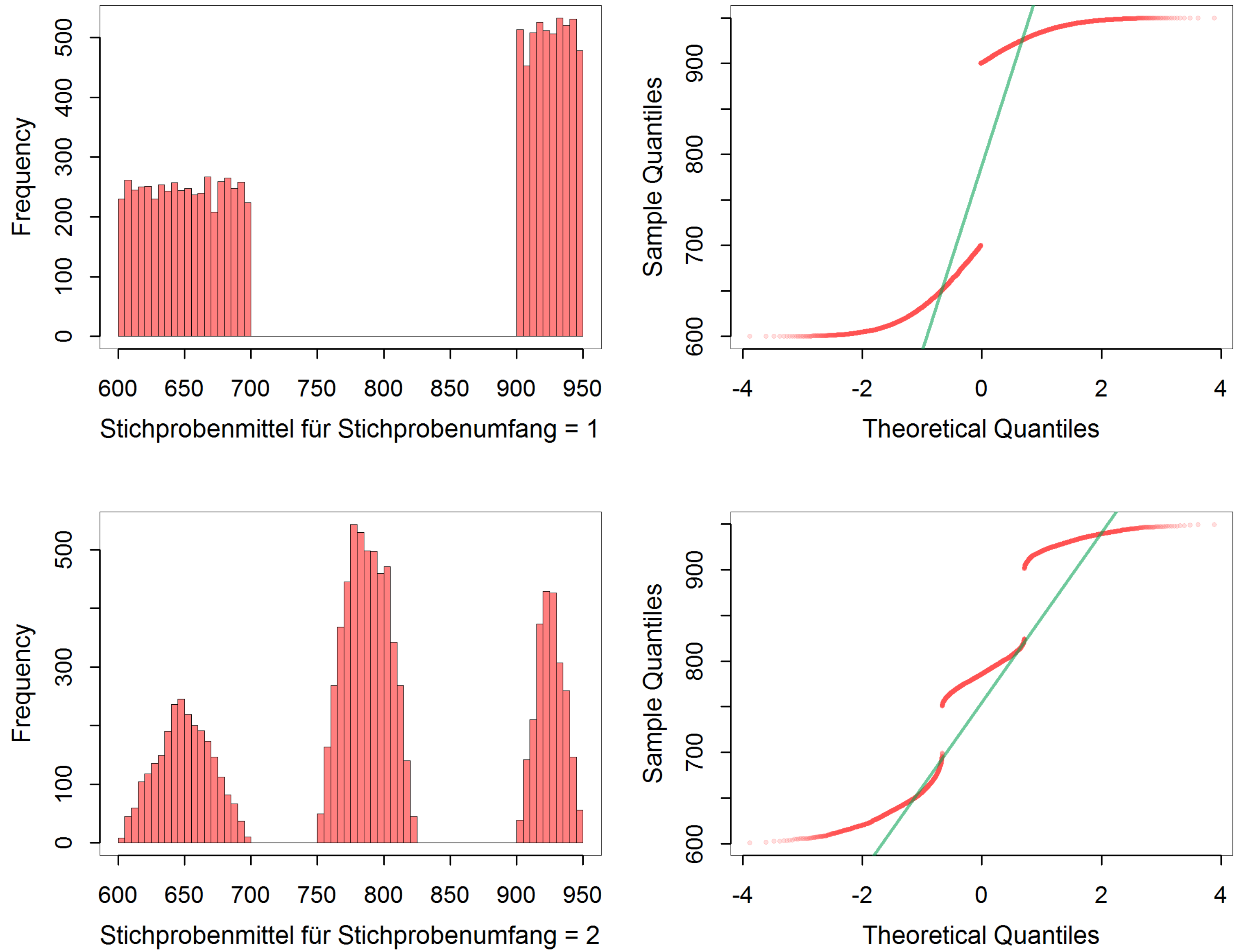
Stichprobenfunktion V	Bezeichnung	$E(V)$	$\mathrm{Var}(V)$
$\sum\limits_{i=1}^{n} X_i$	Merkmalssumme	$n\mu$	$n\sigma^2$
$\bar{X} = \frac{1}{n} \sum\limits_{i=1}^{n} X_i$	Stichprobenmittel	μ	$\frac{\sigma^2}{n}$
$\frac{\bar{X} - \mu}{\sigma} \sqrt{n}$	Gauß-Statistik	0	1
$\frac{1}{n} \sum\limits_{i=1}^{n} (X_i - \mu)^2$	mittlere quadratische Abweichung bezüglich μ	σ^2	
$\frac{1}{n} \sum\limits_{i=1}^{n} (X_i - \bar{X})^2$	mittlere quadratische Abweichung	$\frac{n-1}{n} \sigma^2$	
$S^2 = \frac{1}{n-1} \sum\limits_{i=1}^{n} (X_i - \bar{X})^2$	Stichprobenvarianz	σ^2	
$S = \sqrt{S^2}$	Stichproben-Standardabweichung		
$\frac{\bar{X} - \mu}{S} \sqrt{n}$	t-Statistik		



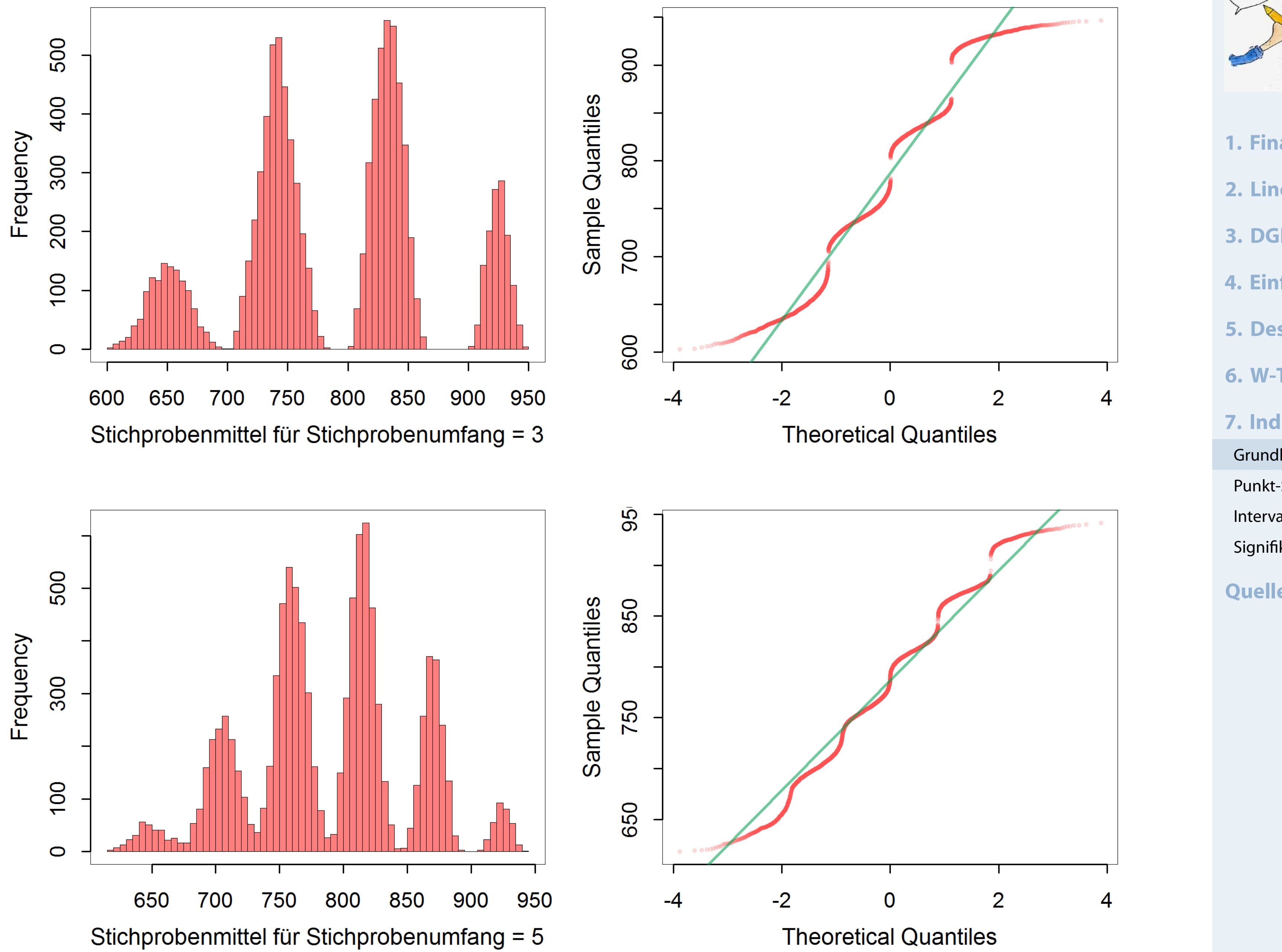
Auswirkungen der Stichprobengröße

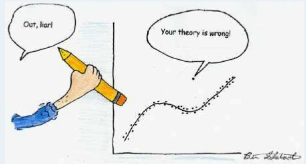


Ziehen von 10.000 Stichproben (jeweils vom Umfang n) und Berechnung der Stichprobenmittel (Verteilung: zwei überlagerte Gleichverteilungen):

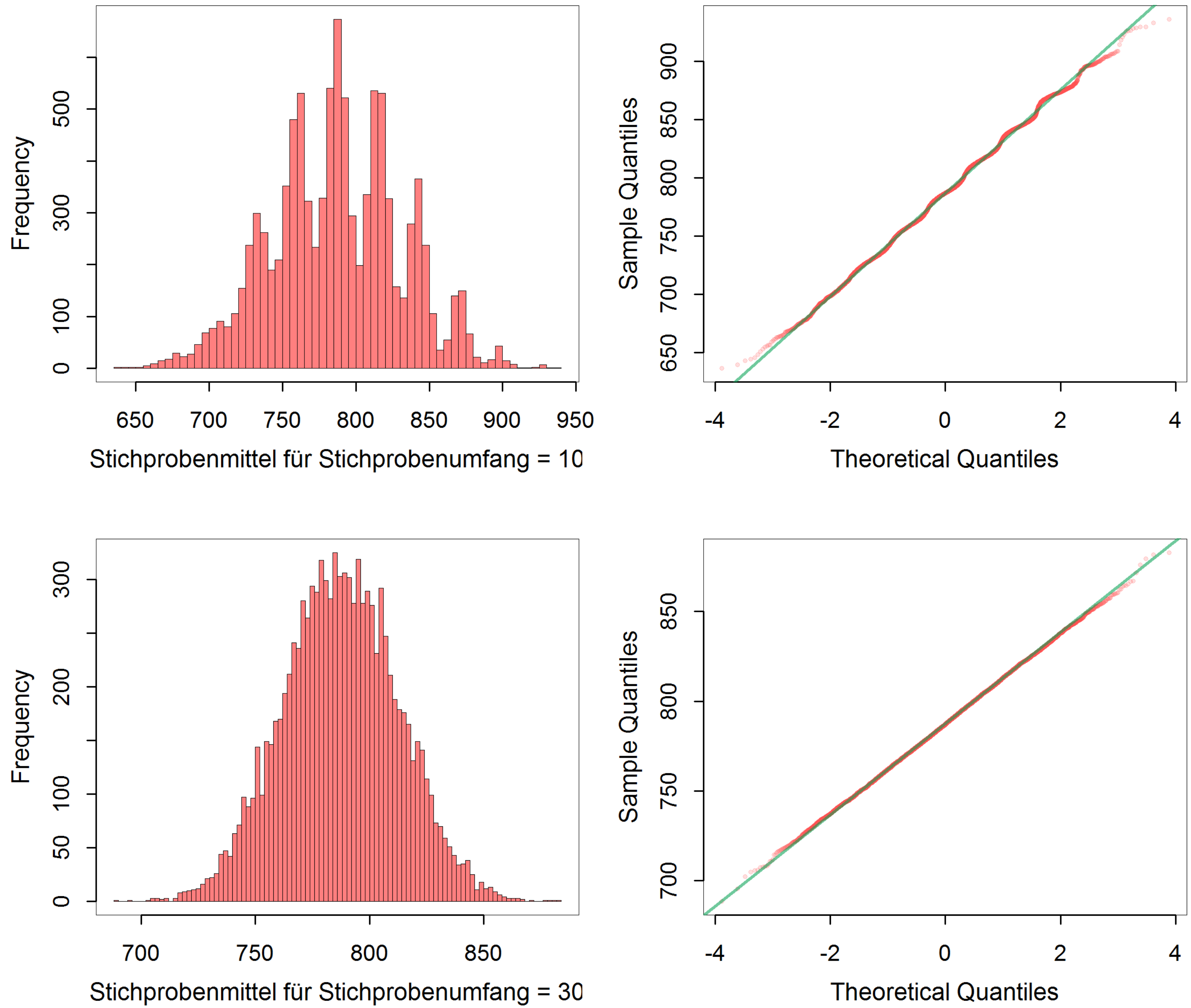


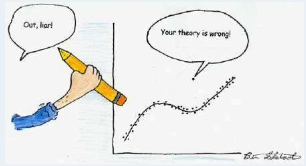
Auswirkungen der Stichprobengröße

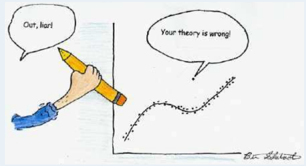




Auswirkungen der Stichprobengröße





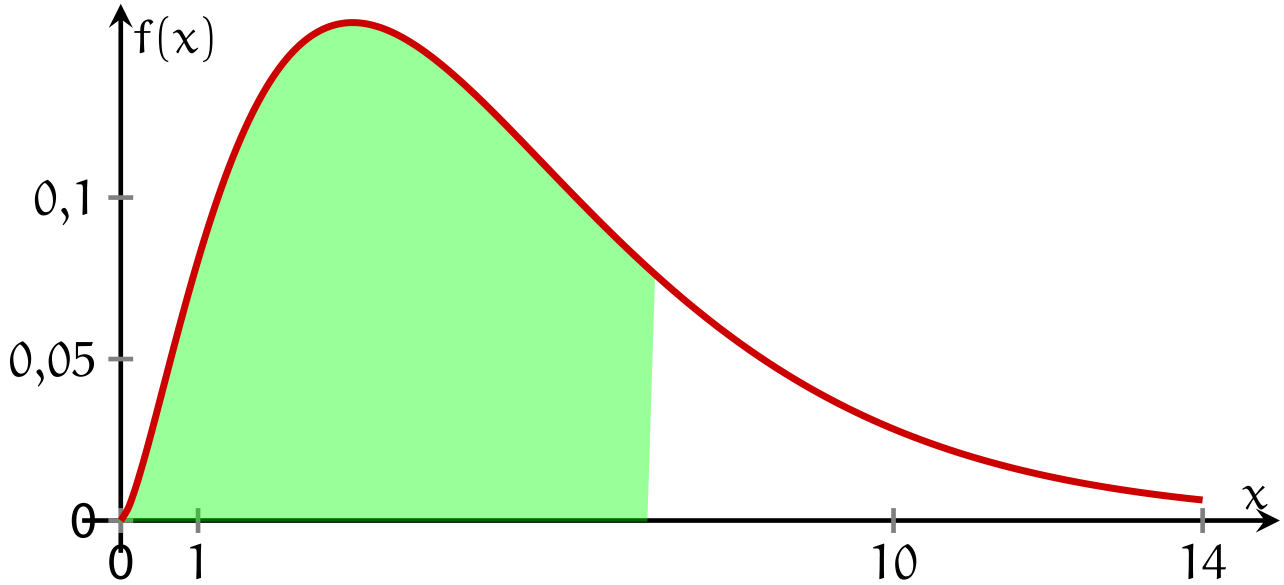


Chi-Quadrat-Verteilung

- Sind $X_1, \ldots, X_n$ iid $N(0;1)$-verteilte Zufallsvariablen, so wird die Verteilung von

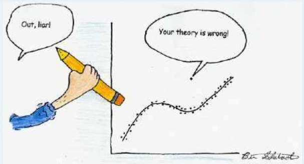
$$Z = \sum_{i=1}^{n} X_i^2$$

als **Chi-Quadrat-Verteilung mit n Freiheitsgraden** bezeichnet.



- Kurzschreibweise: $Z \sim \chi^2(n)$
- **Beispiel:** $\chi^2(30)$: $x_{0,975} = 46{,}98$

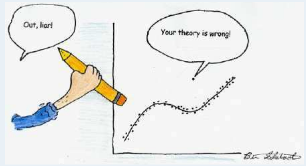
Quantilstabelle der χ^2-Verteilung mit n Freiheitsgraden



$\alpha \setminus n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.005	0.00	0.01	0.07	0.21	0.41	0.68	0.99	1.34	1.73	2.16	2.60	3.07	3.56	4.07	4.60
0.01	0.00	0.02	0.11	0.30	0.55	0.87	1.24	1.65	2.09	2.56	3.05	3.57	4.11	4.66	5.23
0.025	0.00	0.05	0.22	0.48	0.83	1.24	1.69	2.18	2.70	3.25	3.82	4.40	5.01	5.63	6.26
0.05	0.00	0.10	0.35	0.71	1.15	1.64	2.17	2.73	3.33	3.94	4.57	5.23	5.89	6.57	7.26
0.1	0.02	0.21	0.58	1.06	1.61	2.20	2.83	3.49	4.17	4.87	5.58	6.30	7.04	7.79	8.55
0.2	0.06	0.45	1.01	1.65	2.34	3.07	3.82	4.59	5.38	6.18	6.99	7.81	8.63	9.47	10.31
0.25	0.10	0.58	1.21	1.92	2.67	3.45	4.25	5.07	5.90	6.74	7.58	8.44	9.30	10.17	11.04
0.4	0.28	1.02	1.87	2.75	3.66	4.57	5.49	6.42	7.36	8.30	9.24	10.18	11.13	12.08	13.03
0.5	0.45	1.39	2.37	3.36	4.35	5.35	6.35	7.34	8.34	9.34	10.34	11.34	12.34	13.34	14.34
0.6	0.71	1.83	2.95	4.04	5.13	6.21	7.28	8.35	9.41	10.47	11.53	12.58	13.64	14.69	15.73
0.75	1.32	2.77	4.11	5.39	6.63	7.84	9.04	10.22	11.39	12.55	13.70	14.85	15.98	17.12	18.25
0.8	1.64	3.22	4.64	5.99	7.29	8.56	9.80	11.03	12.24	13.44	14.63	15.81	16.98	18.15	19.31
0.9	2.71	4.61	6.25	7.78	9.24	10.64	12.02	13.36	14.68	15.99	17.27	18.55	19.81	21.06	22.31
0.95	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92	18.31	19.68	21.03	22.36	23.68	25.00
0.975	5.02	7.38	9.35	11.14	12.83	14.45	16.01	17.53	19.02	20.48	21.92	23.34	24.74	26.12	27.49
0.99	6.63	9.21	11.34	13.28	15.09	16.81	18.48	20.09	21.67	23.21	24.73	26.22	27.69	29.14	30.58
0.995	7.88	10.60	12.84	14.86	16.75	18.55	20.28	21.95	23.59	25.19	26.76	28.30	29.82	31.32	32.80

$\alpha \setminus n$	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
0.005	5.14	5.70	6.26	6.84	7.43	8.03	8.64	9.26	9.89	10.52	11.16	11.81	12.46	13.12	13.79
0.01	5.81	6.41	7.01	7.63	8.26	8.90	9.54	10.20	10.86	11.52	12.20	12.88	13.56	14.26	14.95
0.025	6.91	7.56	8.23	8.91	9.59	10.28	10.98	11.69	12.40	13.12	13.84	14.57	15.31	16.05	16.79
0.05	7.96	8.67	9.39	10.12	10.85	11.59	12.34	13.09	13.85	14.61	15.38	16.15	16.93	17.71	18.49
0.1	9.31	10.09	10.86	11.65	12.44	13.24	14.04	14.85	15.66	16.47	17.29	18.11	18.94	19.77	20.60
0.2	11.15	12.00	12.86	13.72	14.58	15.44	16.31	17.19	18.06	18.94	19.82	20.70	21.59	22.48	23.36
0.25	11.91	12.79	13.68	14.56	15.45	16.34	17.24	18.14	19.04	19.94	20.84	21.75	22.66	23.57	24.48
0.4	13.98	14.94	15.89	16.85	17.81	18.77	19.73	20.69	21.65	22.62	23.58	24.54	25.51	26.48	27.44
0.5	15.34	16.34	17.34	18.34	19.34	20.34	21.34	22.34	23.34	24.34	25.34	26.34	27.34	28.34	29.34
0.6	16.78	17.82	18.87	19.91	20.95	21.99	23.03	24.07	25.11	26.14	27.18	28.21	29.25	30.28	31.32
0.75	19.37	20.49	21.60	22.72	23.83	24.93	26.04	27.14	28.24	29.34	30.43	31.53	32.62	33.71	34.80
0.8	20.47	21.61	22.76	23.90	25.04	26.17	27.30	28.43	29.55	30.68	31.79	32.91	34.03	35.14	36.25
0.9	23.54	24.77	25.99	27.20	28.41	29.62	30.81	32.01	33.20	34.38	35.56	36.74	37.92	39.09	40.26
0.95	26.30	27.59	28.87	30.14	31.41	32.67	33.92	35.17	36.41	37.65	38.89	40.11	41.34	42.56	43.77
0.975	28.85	30.19	31.53	32.85	34.17	35.48	36.78	38.08	39.36	40.65	41.92	43.19	44.46	45.72	46.98
0.99	32.00	33.41	34.81	36.19	37.57	38.93	40.29	41.64	42.98	44.31	45.64	46.96	48.28	49.59	50.89
0.995	34.27	35.72	37.16	38.58	40.00	41.40	42.80	44.18	45.56	46.93	48.29	49.64	50.99	52.34	53.67

Testverteilungen: t-Verteilung

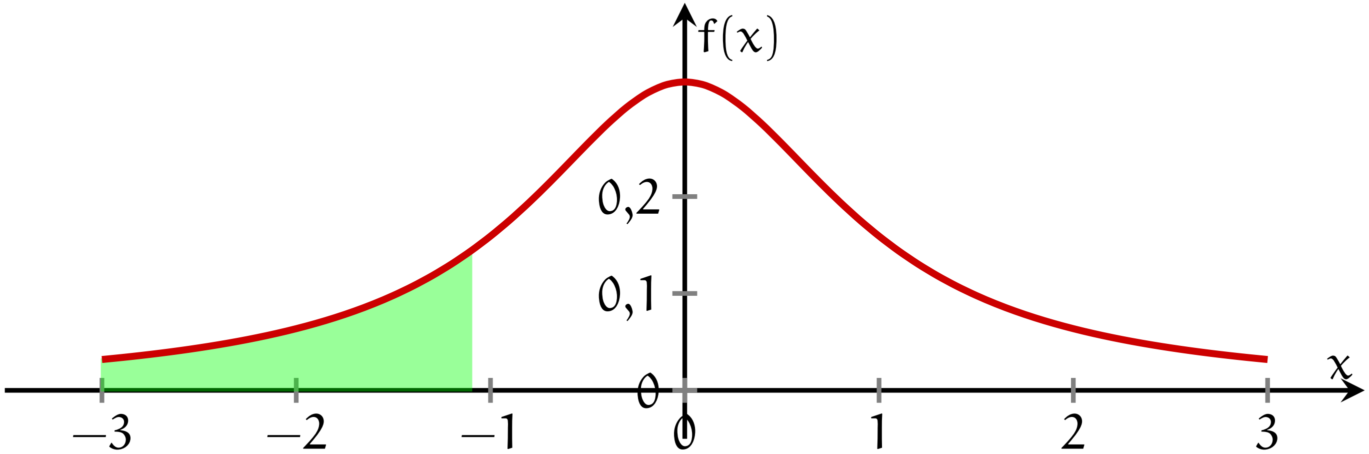


- Ist $X \sim N(0;1)$, $Z \sim \chi^2(n)$, X, Z unabhängig, so wird die Verteilung von

$$T = \frac{X}{\sqrt{\frac{1}{n} Z}}$$

als t-Verteilung mit n Freiheitsgraden bezeichnet.

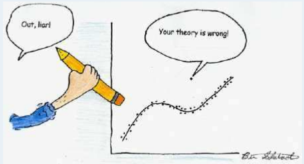
William Sealy Gosset
1876 – 1937



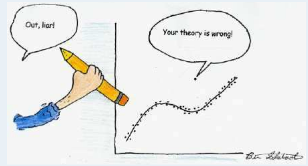
- Kurzschreibweise: $T \sim t(n)$
- **Beispiel:** $t(10)$ $x_{0,6} = 0{,}260$, $x_{0,5} = 0$, $x_{0,1} = -x_{0,9} = -1{,}372$

Quantilstabelle der t-Verteilung mit n Freiheitsgraden

$\alpha \setminus n$	0.6	0.75	0.8	0.9	0.95	0.975	0.99	0.995
1	0.325	1.000	1.376	3.078	6.314	12.706	31.820	63.657
2	0.289	0.816	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.765	0.979	1.638	2.353	3.183	4.541	5.841
4	0.271	0.741	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.718	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.711	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.706	0.889	1.397	1.860	2.306	2.897	3.355
9	0.261	0.703	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.700	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.698	0.875	1.363	1.796	2.201	2.718	3.106
12	0.259	0.696	0.873	1.356	1.782	2.179	2.681	3.054
13	0.259	0.694	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.692	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.691	0.866	1.341	1.753	2.131	2.603	2.947
16	0.258	0.690	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.689	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.688	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.688	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.687	0.860	1.325	1.725	2.086	2.528	2.845
21	0.257	0.686	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.686	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.685	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.685	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.684	0.856	1.316	1.708	2.059	2.485	2.787
26	0.256	0.684	0.856	1.315	1.706	2.055	2.479	2.779
27	0.256	0.684	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.683	0.855	1.312	1.701	2.048	2.467	2.763
29	0.256	0.683	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.683	0.854	1.310	1.697	2.042	2.457	2.750

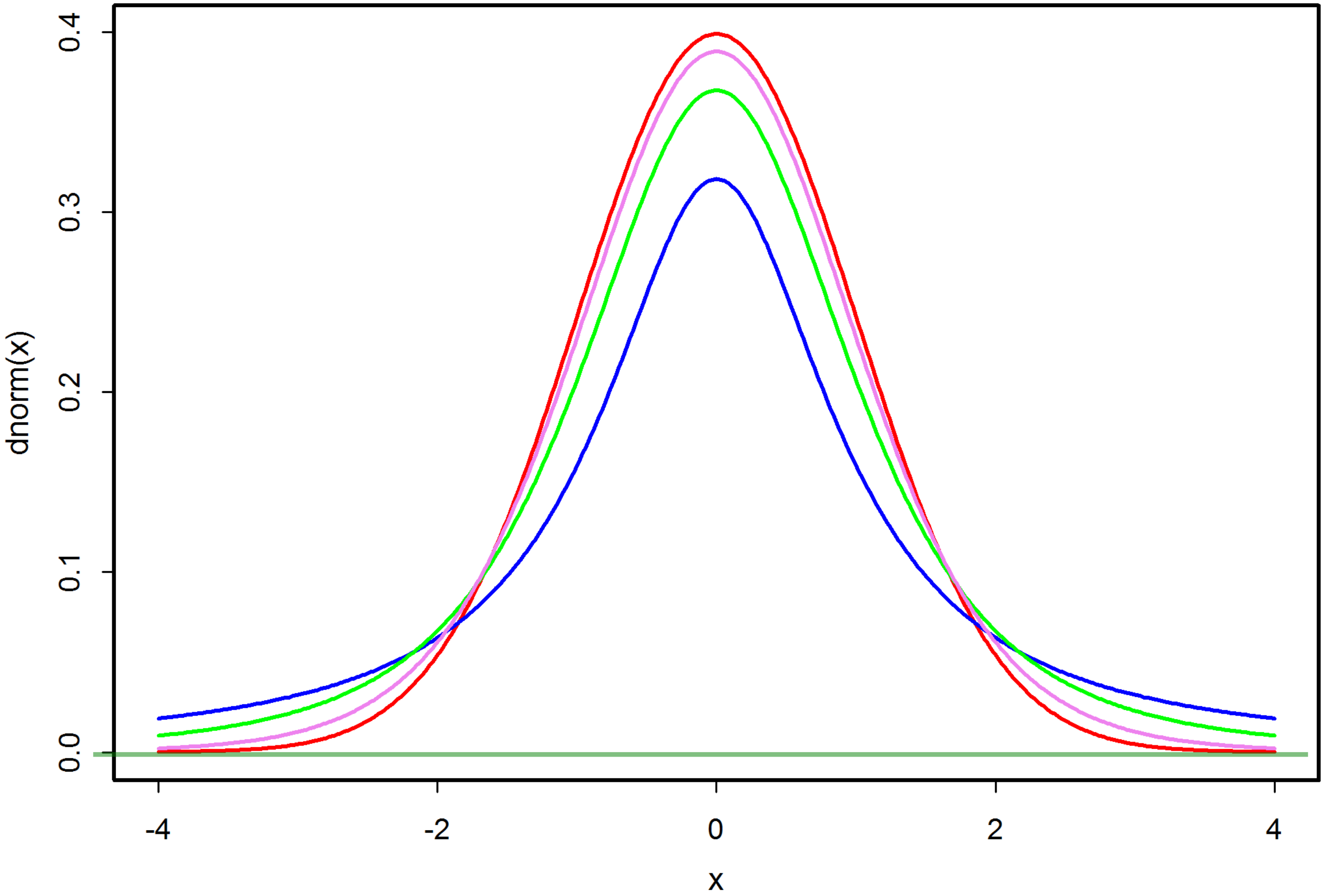


t-Verteilung vs. Normalverteilung

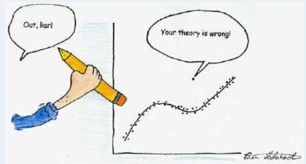


Dichtefunktion

- t-Verteilung mit 1 (blau), 3 (grün) und 10 (lila) Freiheitsgraden
- Standardnormalverteilung (rot)



Punkt-Schätzung



- Ein unbekannter Parameter ϑ der Verteilung von G soll auf Basis einer Stichprobe geschätzt werden.
- Zum Beispiel: σ von $N(10; \sigma)$
- Schätzwert: $\hat{\vartheta}$
- Vorgehen: Verwendung einer Schätzfunktion

$$\hat{\Theta} = g(X_1, \ldots, X_n)$$

Beachte: Der Schätzwert $\hat{\vartheta}$ ist die Realisierung der ZV (!) $\hat{\Theta}$.

- Frage: Welche Stichprobenfunktion ist zur Schätzung geeignet?

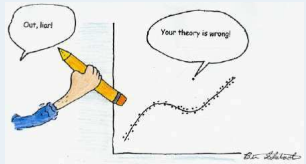
➠ Kriterien für die Beurteilung/Konstruktion von Schätzfunktionen!

- Im Folgenden: Vorliegen einer einfachen Stichprobe, d.h. $X_1, \ldots, X_n$ iid.

Beispiel

- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

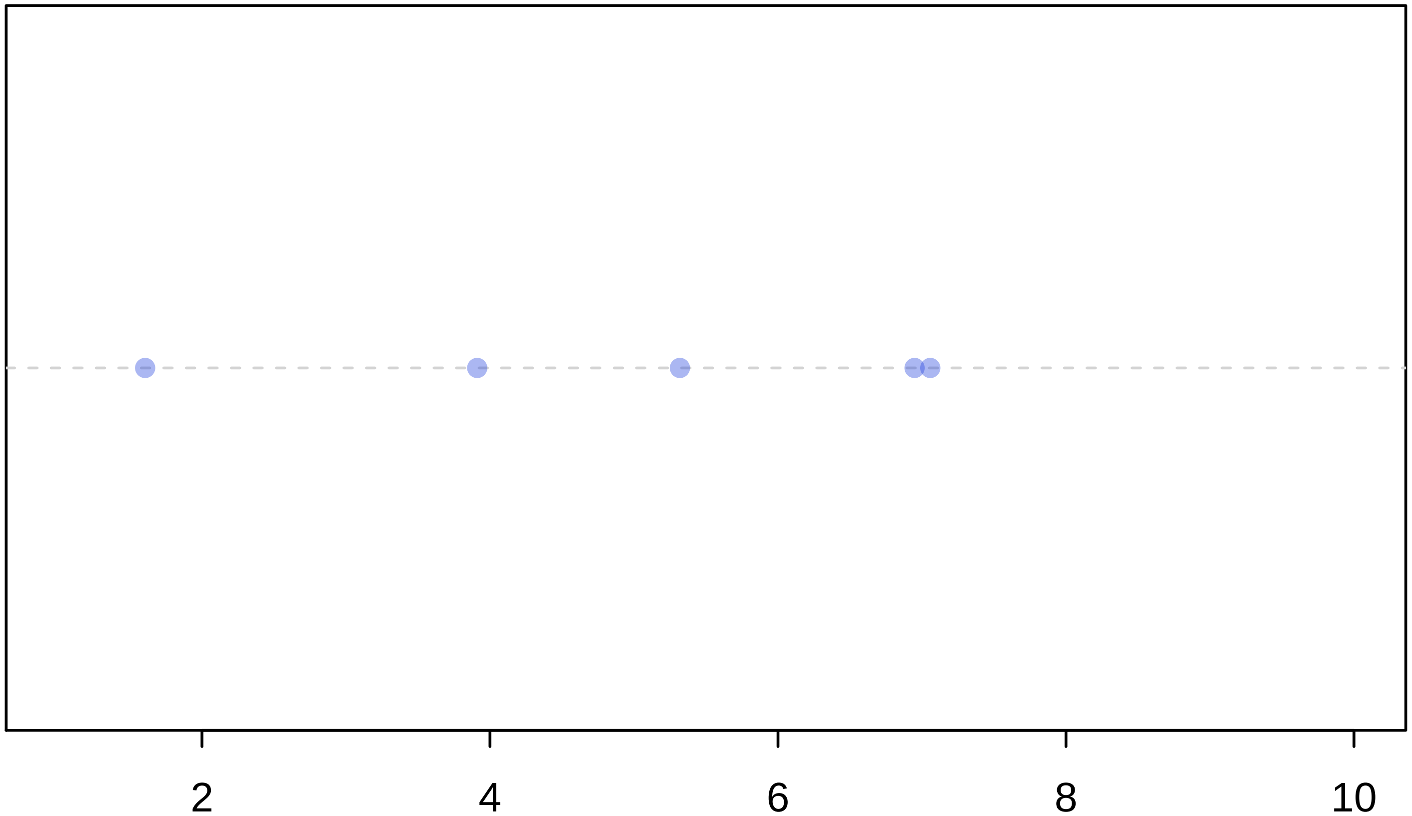
$$\hat{\Theta}_1 = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1} \sum_{i=1}^{n} X_i$$

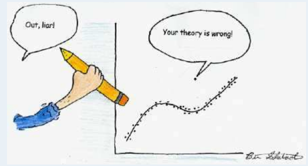


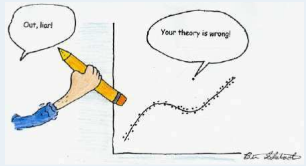
Beispiel

- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\hat{\Theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$



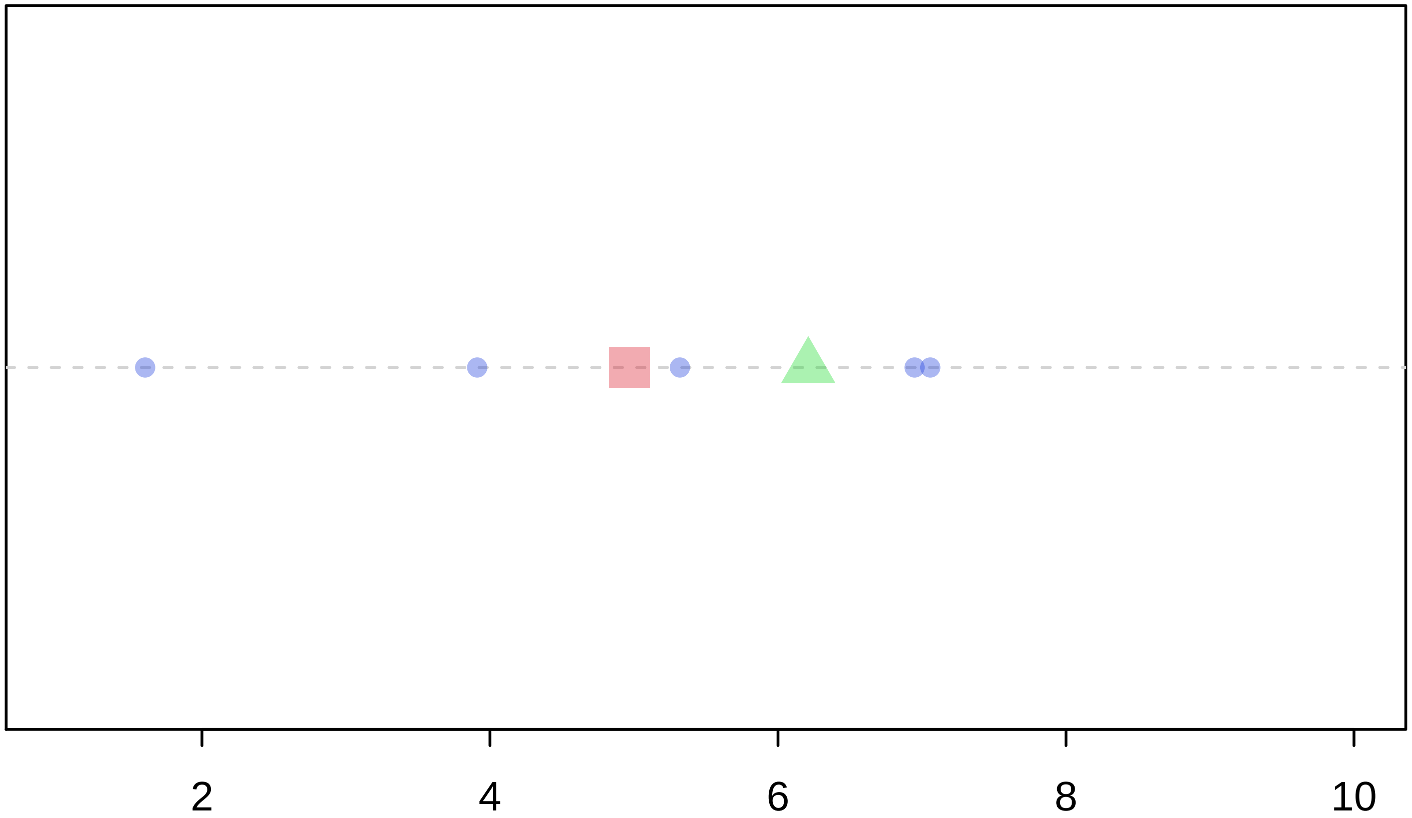




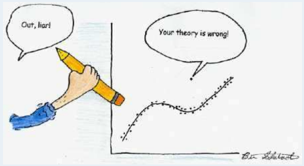
Beispiel

- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\hat{\Theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$

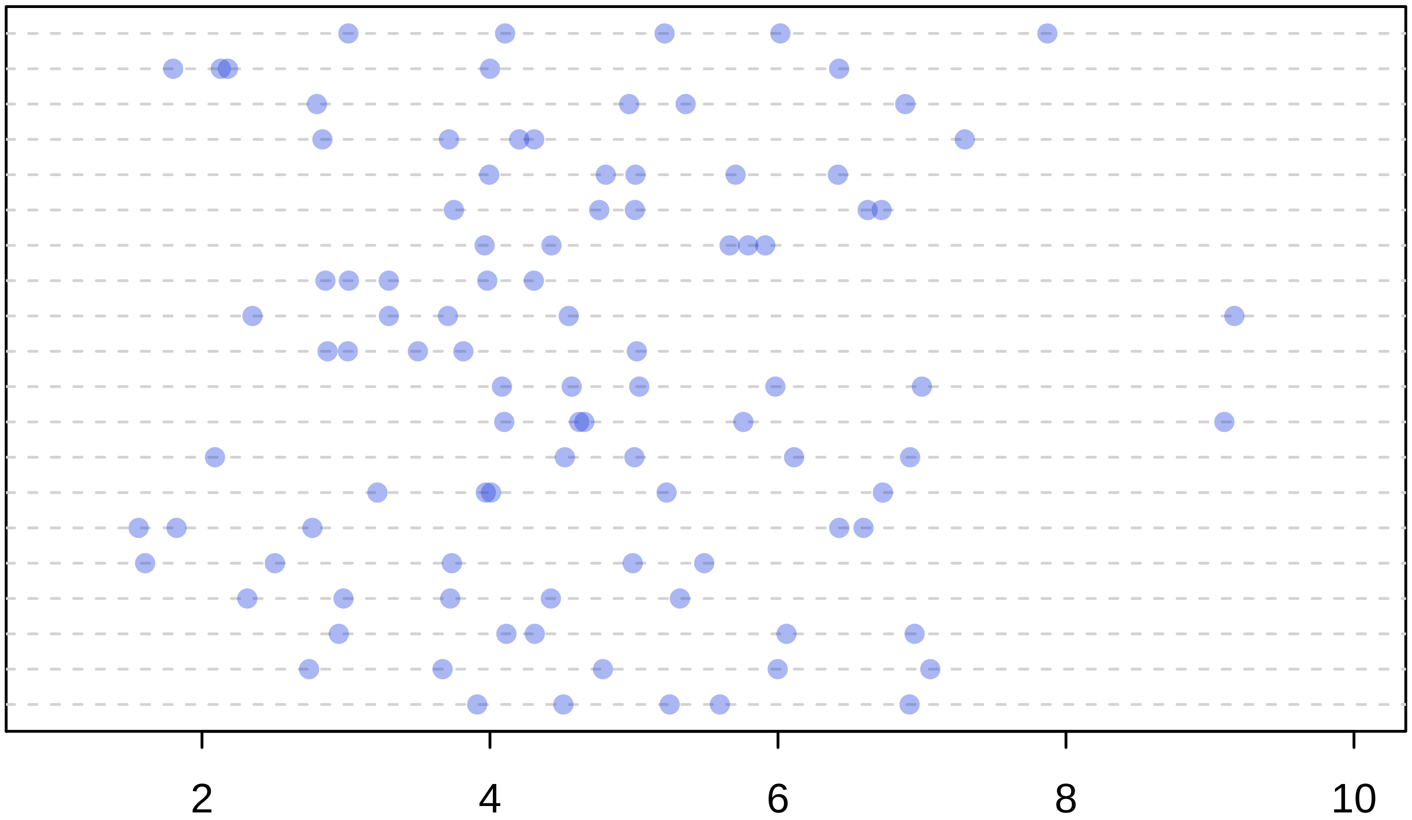


Beispiel

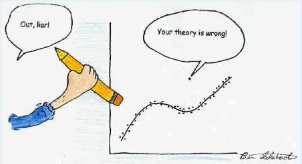


- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\hat{\Theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$

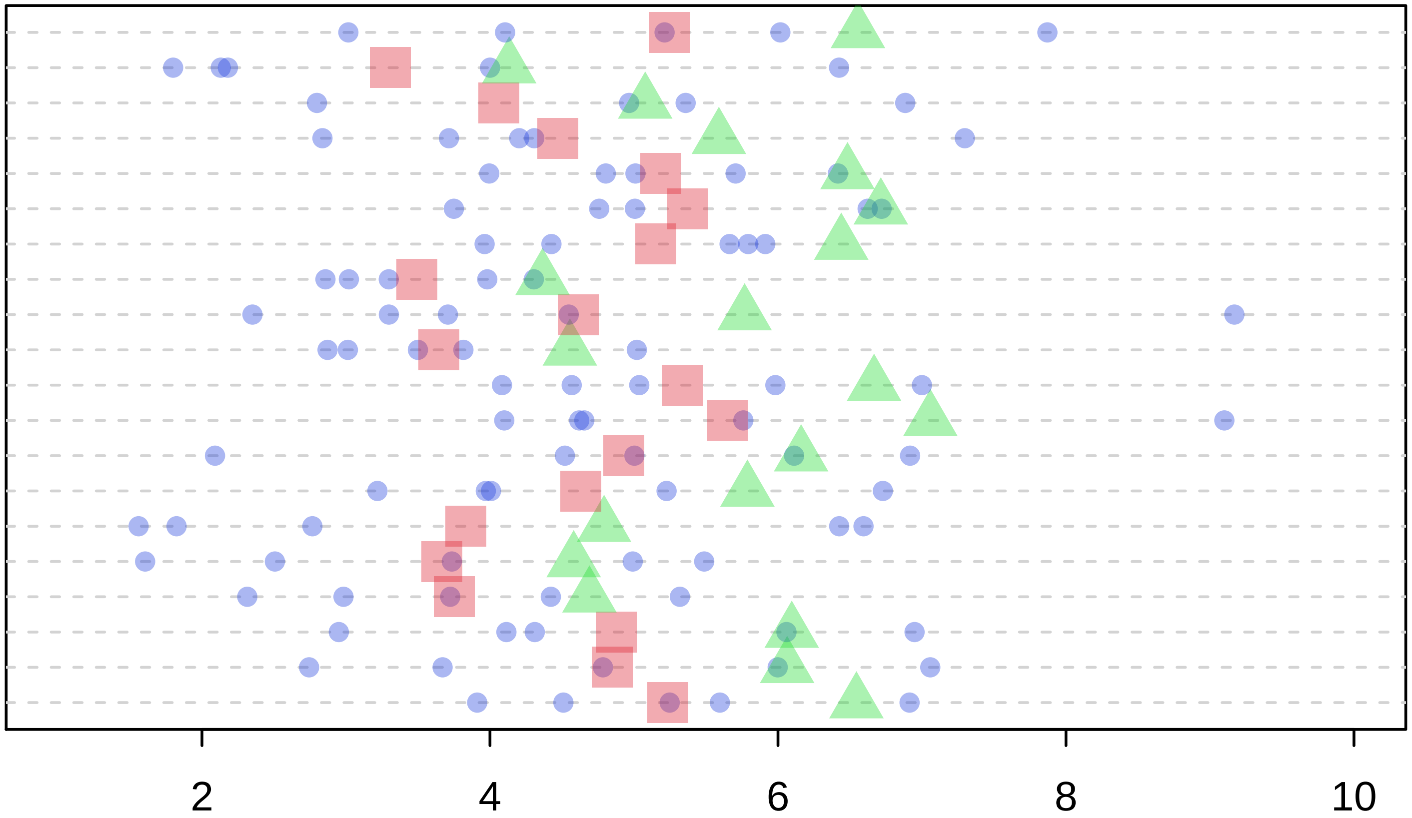


Beispiel

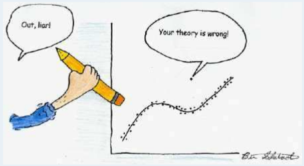


- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\hat{\Theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$

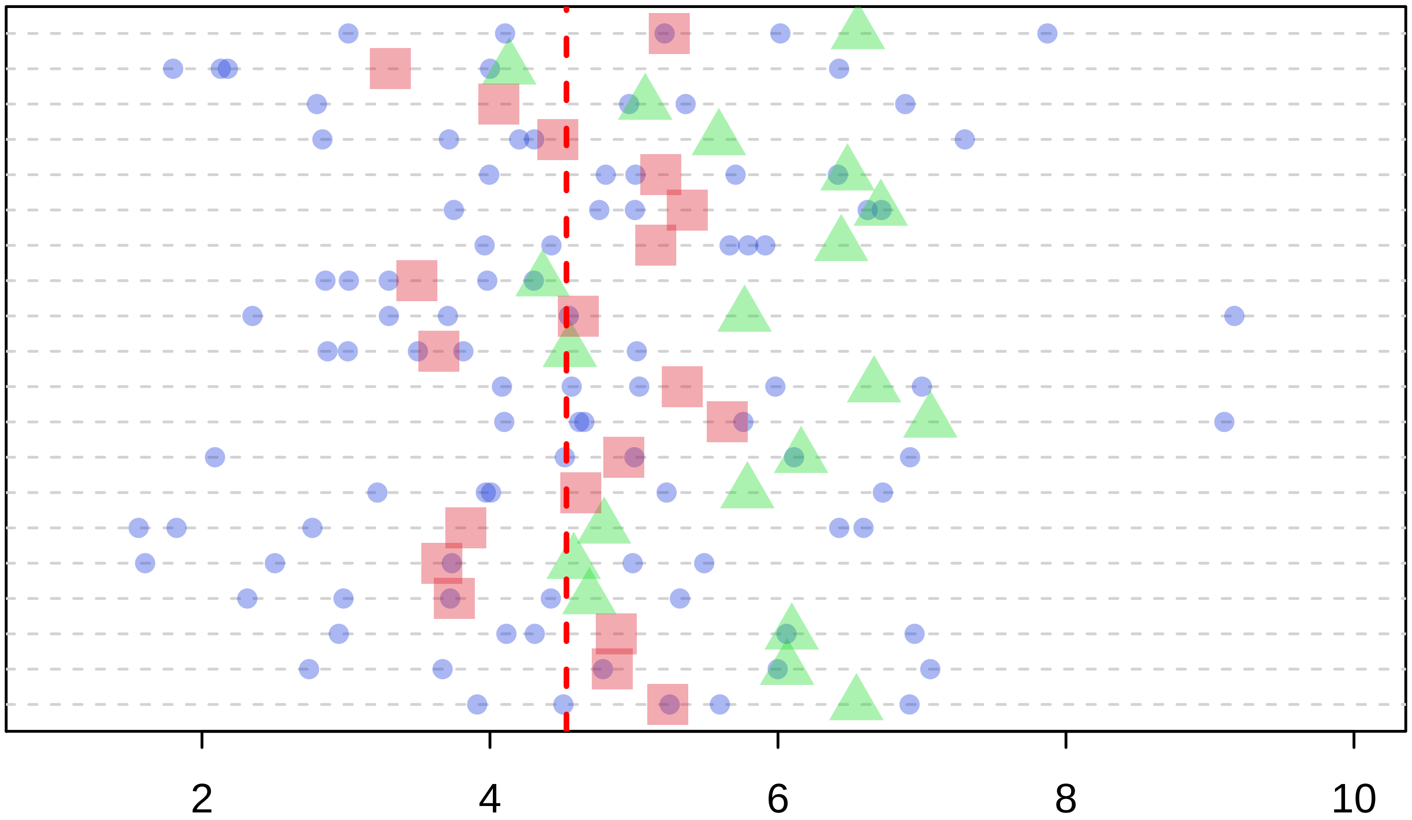


Beispiel



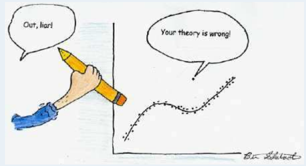
- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\hat{\Theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$



Mittelwert Grundgesamtheit = 4.53

Erwartungstreue und Wirksamkeit



- Eine Schätzfunktion $\hat{\Theta} = g(X_1, \ldots, X_n)$ heißt erwartungstreu oder unverzerrt für ϑ, wenn unabhängig vom numerischen Wert von ϑ gilt:

$$E(\hat{\Theta}) = \vartheta$$

Beispiel

Sind $\hat{\Theta}_1 = \bar{X}$, $\hat{\Theta}_2 = \frac{X_1+X_n}{2}$, $\hat{\Theta}_3 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$ erwartungstreu für μ?

a) $\hat{\Theta}_1$: $\quad E(\bar{X}) = \mu$

$\Rightarrow \hat{\Theta}_1$ ist erwartungstreu.

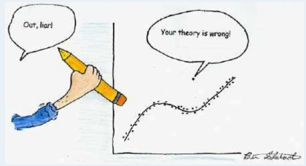
b) $\hat{\Theta}_2$: $\quad E\left(\frac{X_1+X_n}{2}\right) = \frac{1}{2}[E(X_1) + E(X_n)] = \frac{1}{2}(\mu + \mu) = \mu$

$\Rightarrow \hat{\Theta}_2$ ist erwartungstreu.

c) $\hat{\Theta}_3$: $E\left(\frac{1}{n-1}\sum_{i=1}^{n} X_i\right) = \frac{1}{n-1}\sum_{i=1}^{n} E(X_i) = \frac{1}{n-1}\sum_{i=1}^{n} \mu = \frac{n}{n-1}\mu \neq \mu$

$\Rightarrow \hat{\Theta}_3$ ist nicht erwartungstreu

Erwartungstreue und Wirksamkeit



- Welche der erwartungstreuen Schätzfunktionen $\widehat{\Theta}_1, \widehat{\Theta}_2$ ist „besser"?
- Von zwei erwartungstreuen Schätzfunktionen $\widehat{\Theta}_1$, $\widehat{\Theta}_2$ für ϑ heißt $\widehat{\Theta}_1$ wirksamer als $\widehat{\Theta}_2$, wenn unabhängig vom numerischen Wert von ϑ gilt:

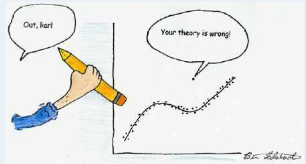
$$\operatorname{Var}(\widehat{\Theta}_1) < \operatorname{Var}(\widehat{\Theta}_2)$$

Beispiel: $(\widehat{\Theta}_1 = \bar{X},\ \widehat{\Theta}_2 = \frac{X_1 + X_n}{2})$
Wegen

$$\left.\begin{aligned} \operatorname{Var}(\widehat{\Theta}_1) &= \operatorname{Var}(\bar{X}) && = \frac{\sigma^2}{n} \\ \operatorname{Var}(\widehat{\Theta}_2) &= \operatorname{Var}\left(\frac{X_1 + X_n}{2}\right) = \frac{1}{4}(\sigma^2 + \sigma^2) && = \frac{\sigma^2}{2} \end{aligned}\right\} \Rightarrow \operatorname{Var}(\widehat{\Theta}_1) < \operatorname{Var}(\widehat{\Theta}_2)$$

(falls $n > 2$) ist $\widehat{\Theta}_1$ wirksamer als $\widehat{\Theta}_2$.

Intervall-Schätzung



- Für einen unbekannten Verteilungsparameter ϑ soll auf Basis einer Stichprobe ein Intervall geschätzt werden.
- Verwendung der Stichprobenfunktionen V_u, V_o, so dass $V_u \leqq V_o$ und

$$P(V_u \leqq \vartheta \leqq V_o) = 1 - \alpha$$

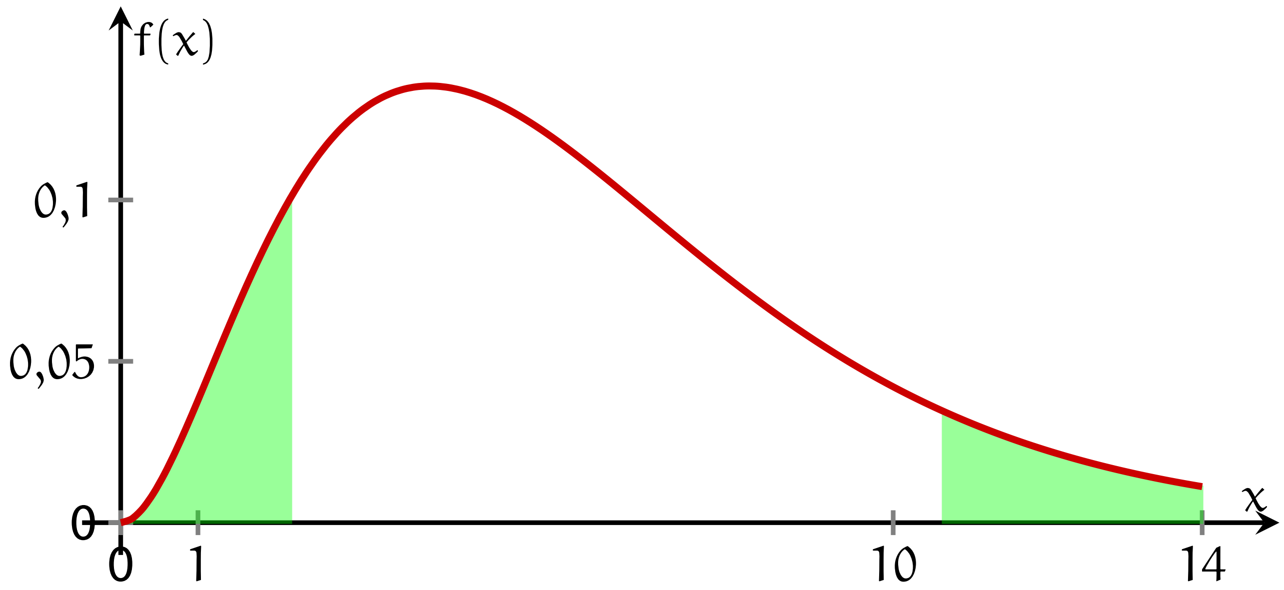
  stets gelten.
  $[V_u; V_o]$ heißt Konfidenzintervall (KI) für ϑ zum Konfidenzniveau $1 - \alpha$.
- Beachte: Das Schätzintervall $[v_u; v_o]$ ist Realisierung der Zufallsvariablen (!) V_u, V_o.
  ➠ Irrtumswahrscheinlichkeit α (klein, i.d.R. $\alpha \leqq 0{,}1$)
- Frage: Welche Konfidenzintervalle sind zur Schätzung geeignet?
  ➠ Hängt von Verteilung von G sowie vom unbekannten Parameter (μ, σ^2) ab!
- Im Folgenden: Einfache Stichprobe $X_1, \ldots, X_n$ mit $E(X_i) = \mu$, $Var(X_i) = \sigma^2$

Intervall-Schätzung

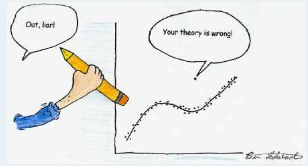
Wichtiger Spezialfall: Symmetrische Konfidenzintervalle

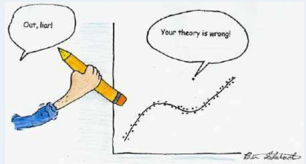
- Symmetrisch heißt **nicht**, dass die Dichte symmetrisch ist, sondern
- übereinstimmende Wahrscheinlichkeiten für Über-/Unterschreiten des Konfidenzintervalls, d.h.

$$P(V_u > \vartheta) = P(V_o < \vartheta) = \frac{\alpha}{2}$$



- **Wichtig:** Eine Verkleinerung von α bewirkt eine Vergrößerung des Konfidenzintervalls.





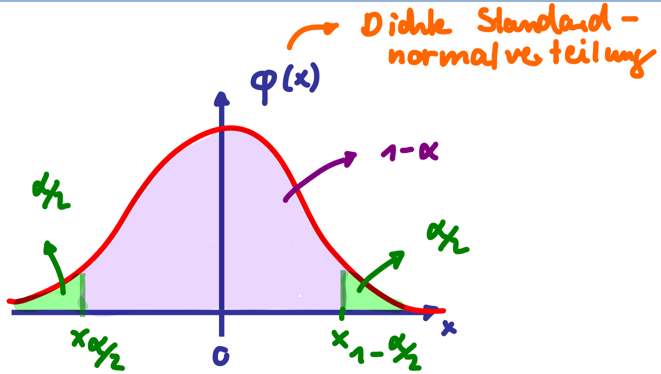
$\bar X = \frac{1}{n} \sum_i X_i$

Stichprobenmittel

gegeben: Standardabweichung $\mathrm{Std}[X_i] = \sigma$

Varianz: $\mathrm{Var}[X_i] = \sigma^2$

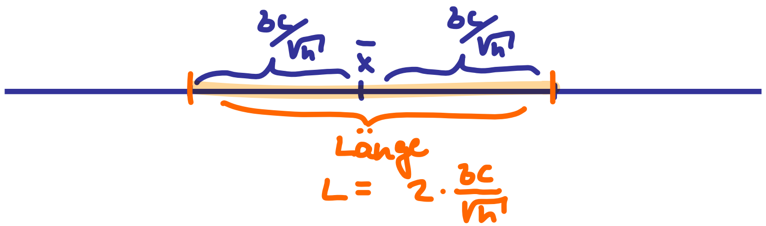
$\Rightarrow \mathrm{Var}[\bar X] = \sigma^2/n \quad \Rightarrow \mathrm{Std}[\bar X] = \sigma/\sqrt{n}$

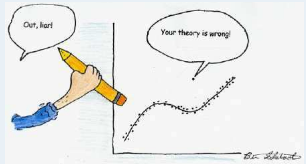


Vorgehensweise:

1. Festlegen des Konfidenzniveaus $1-\alpha$
2. Bestimmung des $\left(1-\frac{\alpha}{2}\right)$-Fraktils c der $N(0,1)$-Verteilung
3. Berechnen des Stichprobenmittels $\bar x$
4. Berechnen des Wertes $\frac{\sigma c}{\sqrt{n}}$
5. Ergebnis der Intervall-Schätzung:

$$\left[\bar x - \frac{\sigma c}{\sqrt{n}}\,;\quad \bar x + \frac{\sigma c}{\sqrt{n}}\right]$$





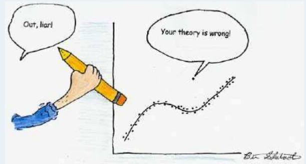
Beispiel

Normalverteilung mit $\sigma = 2{,}4$
$(x_1, \ldots, x_9)$ = (184.2, 182.6, 185.3, 184.5, 186.2, 183.9, 185.0, 187.1, 184.4)

Gesucht: Konfidenzintervall für μ zum Konfidenzniveau $1 - \alpha = 0{,}99$

1. $1 - \alpha = 0{,}99$
2. $N(0;1)$: $c = x_{1-\frac{\alpha}{2}} = x_{1-\frac{0,01}{2}} = x_{0,995} = 2{,}576$ (Tab. 3; Interpolation) 2.575 (auch i.O.)
3. $\bar{x} = \frac{1}{9}(184{,}2 + \cdots + 184{,}4) = 184{,}8$
4. $\frac{\sigma c}{\sqrt{n}} = \frac{2{,}4 \cdot 2{,}576}{\sqrt{9}} = 2{,}06$
5. $\text{KI} = [184{,}8 - 2{,}06; 184{,}8 + 2{,}06] = [182{,}74; 186{,}86]$

Interpretation: Mit 99 % Wahrscheinlichkeit ist $\mu \in [182{,}74; 186{,}86]$.

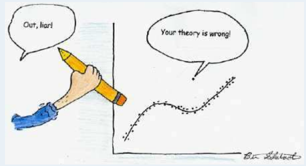


Wichtige Fraktilswerte

Wichtige $N(0;1)$-Fraktilswerte:

α	x_α
0,9	1,281552
0,95	1,644854
0,975	1,959964
0,99	2,326348
0,995	2,575829

(I.d.R. genügen drei Nachkommastellen.)



- Bei bekannter Standardabweichung gilt offenkundig

$$L = V_o - V_u = \frac{2\sigma c}{\sqrt{n}}$$

$\frac{\sigma c}{\sqrt{n}}$ $\bar{x}$ $\frac{\sigma c}{\sqrt{n}}$ x

Länge

$L = 2 \cdot \frac{\sigma c}{\sqrt{n}}$

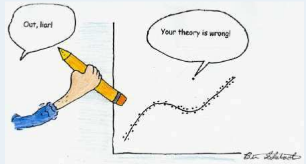
- Welcher Stichprobenumfang n sichert eine vorgegebene (Maximal-)Länge L? $\Rightarrow$ Nach n auflösen! $\Rightarrow$

$$n \geqq \left(\frac{2\sigma c}{L}\right)^2$$

- Eine Halbierung von L erfordert eine Vervierfachung von n!
- Angewendet auf letztes **Beispiel**:

$$L = 4 \Rightarrow n \geqq \left(\tfrac{2\cdot 2{,}4\cdot 2{,}576}{4}\right)^2 = 9{,}556 \Rightarrow n \geqq 10$$

$$L = 2 \Rightarrow n \geqq \left(\tfrac{2\cdot 2{,}4\cdot 2{,}576}{2}\right)^2 = 38{,}222 \Rightarrow n \geqq 39$$



Konfidenzintervall für μ bei Normalverteilung mit unbekanntem σ^2

- Vorgehensweise:
  1. Festlegen des Konfidenzniveaus $1-\alpha$
  2. Bestimmung des $\left(1-\frac{\alpha}{2}\right)$-Fraktils c der $t(\overbrace{n-1}^{\text{Anz. Freiheitsgrade}})$-Verteilung
  3. Berechnen des Stichprobenmittels $\bar{x}$ und der Stichproben-Standardabweichung s
  4. Berechnen des Wertes $\frac{sc}{\sqrt{n}}$
  5. Ergebnis der Intervall-Schätzung:

$$\left[\bar{x}-\frac{sc}{\sqrt{n}}\,;\quad \bar{x}+\frac{sc}{\sqrt{n}}\right]$$

- Zu Schritt 2: Falls $n-1>30$ wird die $N(0;1)$-Verteilung verwendet.

Konfidenzintervalllänge

α-Fraktile der t-Verteilung mit n Freiheitsgraden

$\downarrow n \setminus \alpha \rightarrow$	0,6	0,75	0,8	0,9	0,95	0,975	0,99	0,995
1	0,325	1,000	1,376	3,078	6,314	12,706	31,820	63,657
2	0,289	0,816	1,061	1,886	2,920	4,303	6,965	9,925
3	0,277	0,765	0,979	1,638	2,353	3,183	4,541	5,841
4	0,271	0,741	0,941	1,533	2,132	2,776	3,747	4,604
5	0,267	0,727	0,920	1,476	2,015	2,571	3,365	4,032
6	0,265	0,718	0,906	1,440	1,943	2,447	3,143	3,707
7	0,263	0,711	0,896	1,415	1,895	2,365	2,998	3,499
8	0,262	0,706	0,889	1,397	1,860	2,306	2,897	3,355
9	0,261	0,703	0,883	1,383	1,833	2,262	2,821	3,250

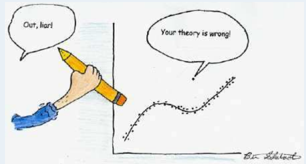
Beispiel:

Wie das letzte Beispiel, jedoch σ unbekannt.

1. $1 - \alpha = 0{,}99$
2. $t(8)$: $c = x_{1-\frac{\alpha}{2}} = x_{1-\frac{0{,}01}{2}} = x_{0{,}995} = 3{,}355$ (Tab. 4)
3. $\bar{x} = \frac{1}{9}\,(184{,}2 + \cdots + 184{,}4) = 184{,}8$

   $s = \sqrt{\frac{1}{8}\,[(184{,}2^2 + \cdots + 184{,}4^2) - 9 \cdot 184{,}8^2]} = 1{,}31$
4. $\frac{sc}{\sqrt{n}} = \frac{1{,}31 \cdot 3{,}355}{\sqrt{9}} = 1{,}47$
5. $\mathrm{KI} = [184{,}8 - 1{,}47; 184{,}8 + 1{,}47] = [183{,}33; 186{,}27]$

Interpretation: Mit 99 % Wahrscheinlichkeit ist $\mu \in [183{,}33; 186{,}27]$.



Beispiel : Geldbeuteldaten

Stichprobe: 57.59, 103.44, 26.71, 103.44, 0.72

gesucht: Konf.intervall für μ der Grundgesamtheit
mit Konf.niveau 0.95.

(Annahme: G ist normalverteilt)

① $1-\alpha = 0.95$ $\quad$ ($\alpha = 0.05$, $1-\alpha/2 = 0.975$

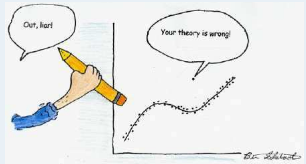
② $t(4)$ - Verteilung: $x_{0.975} \approx 2.776 = c$

③ $\bar{x} = 58.38$
$s = 45.796$

④ $KI = \left[58.38 \pm \frac{45.796 \cdot 2.776}{\sqrt{5}} \right]$

$\approx [1.53 ; 115.23]$

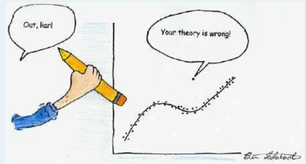
R Beispiel




```r
x <- c(184.2, 182.6, 185.3, 184.5, 186.2,
       183.9, 185.0, 187.1, 184.4)
t.test(x,conf.level=.99)

##
##  One Sample t-test
##
## data:  x
## t = 422.11, df = 8, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 99 percent confidence interval:
##  183.331 186.269
## sample estimates:
## mean of x
##     184.8
```

Konfidenzintervall für μ bei beliebiger Verteilung



- Voraussetzung: $n > 30$, bzw. falls G dichotom: $5 \leqq \sum\limits_{i=1}^{n} x_i \leqq n - 5$
- Vorgehensweise:
  1. Festlegen des Konfidenzniveaus $1 - \alpha$
  2. Bestimmung des $(1 - \frac{\alpha}{2})$-Fraktils c der Standardnormalverteilung $N(0;1)$
  3. Berechnung des Stichprobenmittels $\bar{x}$ sowe eines Schätzwertes $\hat{\sigma}$ für die Standardabweichung σ der GG mittels

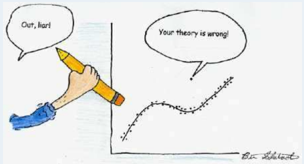
$$\hat{\sigma} = \begin{cases} \sigma, & \text{falls } \sigma \text{ bekannt} \\ \sqrt{\bar{x}(1-\bar{x})}, & \text{falls GG dichotom} \\ s, & \text{sonst} \end{cases}$$

  4. Berechnung von $\frac{\hat{\sigma} c}{\sqrt{n}}$
  5. Ergebnis der Intervallschätzung:

$$\left[\bar{x} - \frac{\hat{\sigma} c}{\sqrt{n}} ; \bar{x} + \frac{\hat{\sigma} c}{\sqrt{n}}\right]$$

- Zu Schritt 3: Manchmal kann anderer Schätzwert $\hat{\sigma}$ sinnvoller sein.

Konfidenzintervall für μ bei beliebiger Verteilung



Beispiel:

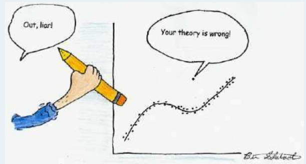
Poisson-Verteilung mit $\lambda \; (= \mu = \sigma^2)$ unbekannt.

$(x_1, \ldots, x_{40}) = (3; 8; \ldots; 6)$

Gesucht: KI für λ zum Konfidenzniveau $1 - \alpha = 0{,}9$

1. $1 - \alpha = 0{,}9$
2. $N(0;1): \; c = x_{1-\frac{\alpha}{2}} = x_{1-\frac{0,1}{2}} = x_{0,95} = 1{,}645$
3. $\bar{x} = \frac{1}{40}(3 + 8 + \cdots + 6) = 6{,}5$
   $\hat{\sigma} = \sqrt{\bar{x}} = \sqrt{6{,}5} = 2{,}55$ (da $\sigma^2 = \lambda$)
4. $\frac{\hat{\sigma} c}{\sqrt{n}} = \frac{2{,}55 \cdot 1{,}645}{\sqrt{40}} = 0{,}66$
5. $\text{KI} = [6{,}5 - 0{,}66; 6{,}5 + 0{,}66] = [5{,}84; 7{,}16]$

Konfidenzintervall für σ^2 bei Normalverteilung



Vorgehensweise

1. Festlegen eines Konfidenzniveaus $1 - \alpha$
2. Bestimmung der $\frac{\alpha}{2}$- bzw. $(1 - \frac{\alpha}{2})$-Fraktile (c_1 bzw. c_2) der $\chi^2(n-1)$-Verteilung
3. Aus der Stichprobe: Berechnung der Größe

$$(n-1)s^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 v$$

4. Berechnung des Konfidenzintervalls

$$\left[\frac{(n-1)s^2}{c_2} ; \frac{(n-1)s^2}{c_1}\right]$$

KI für σ^2 bei Normalverteilung

Beispiel:

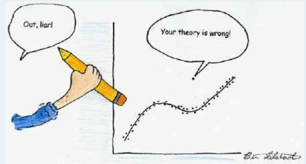
$G \sim N(\mu; \sigma)$;

$(x_1, \ldots, x_5) = (1, 1.5, 2.5, 3, 2)$

Gesucht: KI für σ^2 zum Konfidenzniveau $1 - \alpha = 0{,}99$

1. $1 - \alpha = 0{,}99$
2. $\chi^2(5-1): \; c_1 = x_{\frac{\alpha}{2}} = x_{0{,}005} = 0{,}21$
   $c_2 = x_{1-\frac{\alpha}{2}} = x_{0{,}995} = 14{,}86$
3. $\bar{x} = \frac{1}{5}(1 + 1{,}5 + 2{,}5 + 3 + 2) = 2$
   $$\sum_{i=1}^{5} x_i^2 - 5 \cdot \bar{x}^2 = 1^2 + 1{,}5^2 + 2{,}5^2 + 3^2 + 2^2 - 5 \cdot 2^2 = 2{,}5$$
4. $\text{KI} = \left[\frac{2{,}5}{14{,}86}; \frac{2{,}5}{0{,}21}\right] = [0{,}17; 11{,}9]$

(Extrem groß, da n klein.)

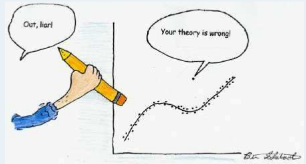


Signifikanztests

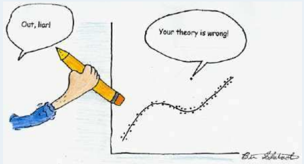
- Vorliegen einer Hypothese über die Verteilung(en) der Grundgesamtheit(en).
- Beispiele:
  - „Der Würfel ist fair.“
  - „Die Brenndauern zweier unterschiedlicher Glühbirnensorten sind gleich.“
- Hypothese soll anhand einer Stichprobe überprüft werden.
- Prinzip:
  - Hypothese verwerfen, wenn „signifikanter“ Widerspruch zur Stichprobe.
  - Ansonsten: Hypothese nicht verwerfen.
- Eine verworfene Hypothese gilt als statistisch widerlegt.
- Nicht-Verwerfung ist dagegen ein „Freispruch aus Mangel an Beweisen“.

Zu Beachten:

Nicht-Verwerfung ist **kein** „statistischer Beweis“, dass Hypothese wahr ist!
(„Trick“: Hypothese falsch $\iff$ Gegenhypothese wahr!)



Test des Erwartungswertes bei bekannter Varianz in der Grundgesamtheit



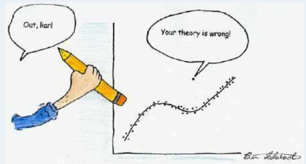
- Zunächst:
  - $G \sim N(\mu; \sigma)$ mit σ bekannt
  - Einfache Stichprobe $X_1, \ldots, X_n$
  - (Null-)Hypothese $H_0: \mu = \mu_0$
- **Beispiel:**
  $X_1, \ldots, X_{25}$ mit X_i = Füllmenge der i-ten Flasche $\sim N(\mu; 1{,}5)$
  Nullhypothese $H_0: \mu = 500$, d.h. $\mu_0 = 500$
- Je nach Interessenlage sind unterschiedliche **Gegenhypothesen** möglich:

$$
\begin{array}{ll}
\text{a)} & H_1: \mu \neq \mu_0 \\
\text{b)} & H_1: \mu < \mu_0 \\
\text{c)} & H_1: \mu > \mu_0
\end{array}
$$

- Entscheidung:

$$
\begin{array}{lll}
 & H_0: \mu = \mu_0 & \text{wird abgelehnt gegenüber} \\
\text{a)} & H_1: \mu \neq \mu_0, & \text{wenn } |\bar{x} - \mu_0| \text{ „sehr groß" ist} \\
\text{b)} & H_1: \mu < \mu_0, & \text{wenn } \bar{x} \text{ „weit kleiner" als } \mu_0 \text{ ist} \\
\text{c)} & H_1: \mu > \mu_0, & \text{wenn } \bar{x} \text{ „weit größer" als } \mu_0 \text{ ist}
\end{array}
$$

Test des Erwartungswertes bei bekannter Varianz in der Grundgesamtheit



Entscheidungskriterium aus Stichprobe:

$$v = \frac{\bar{x} - \mu_0}{\sigma} \sqrt{n}$$

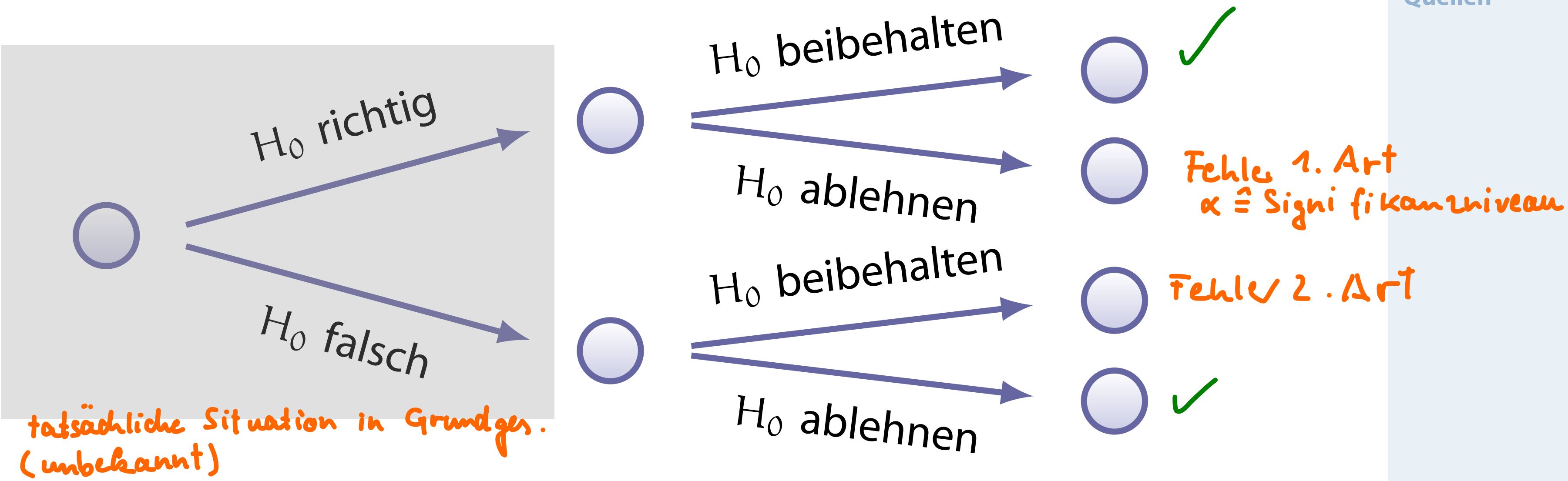
- Vorteil: Verteilung bekannt: N(0; 1)
- Dann:

$H_0 : \mu = \mu_0$ wird abgelehnt gegenüber

a) $H_1 : \mu \neq \mu_0$, wenn $|v|$ „sehr groß“ ist
b) $H_1 : \mu < \mu_0$, wenn v „sehr negativ“ ist
c) $H_1 : \mu > \mu_0$, wenn v „sehr positiv“ ist

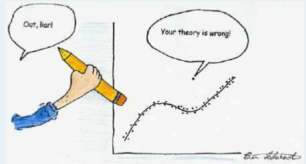
Mögliche Fehlentscheidungen

- **Ablehnung von H_0**, obwohl H_0 richtig ist: Fehler 1. Art
- **Nicht-Ablehnung von H_0**, obwohl H_0 falsch ist: Fehler 2. Art



- Signifikanzniveau — Maximal erlaubte Wahrscheinlichkeit für einen Fehler 1. Art

Test des Erwartungswertes bei bekannter Varianz in der Grundgesamtheit



- Mithilfe von α und V kann geklärt werden, was „sehr groß“ usw. heißt:
  Wahrscheinlichkeit für Fehler 1. Art im Fall a): $|v| > x$, obwohl H_0 richtig:

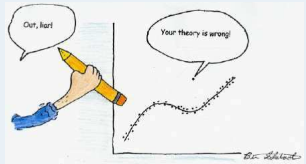
  $$\begin{aligned} P(|V| > x) &= P(V > x) + P(V < -x) \\ &= 2 \cdot P(V > x) \qquad \text{(Symmetrie der Normalverteilung)} \\ &= 2 \cdot [1 - P(V \leqq x)] = 2 \cdot [1 - \Phi(x)] \stackrel{!}{=} \alpha \end{aligned}$$

  $$\iff \Phi(x) = 1 - \tfrac{\alpha}{2}$$

  $$\iff x = x_{1-\frac{\alpha}{2}}$$

  H_0 wird demnach verworfen,
  wenn $|v| > x_{1-\frac{\alpha}{2}}$ bzw. $v \in B$ ist.
  $B = (-\infty; -x_{1-\frac{\alpha}{2}}) \cup (x_{1-\frac{\alpha}{2}}; \infty)$ heißt Verwerfungsbereich.
- Analoge Vorgehensweise für die Fälle b) und c)

Test des Erwartungswertes bei bekannter Varianz in der GG

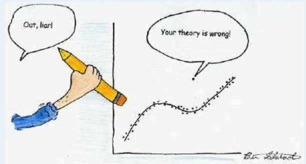


Rezept

1. Ein Signifikanzniveau α wird festgelegt.
2. Der Verwerfungsbereich

$$
\begin{aligned}
B &= \left(-\infty;\ -x_{1-\alpha/2}\right) \cup \left(x_{1-\alpha/2};\ \infty\right) && \text{im Fall a)} \\
B &= \left(-\infty;\ -x_{1-\alpha}\right) && \text{im Fall b)} \\
B &= \left(x_{1-\alpha};\ \infty\right) && \text{im Fall c)}
\end{aligned}
$$

   wird festgelegt, wobei $x_{1-\alpha/2}$ bzw. $x_{1-\alpha}$ das $(1-\alpha/2)$- bzw. das $(1-\alpha)$-Fraktil der $N(0,1)$-Verteilung ist. (Wichtig: Der Ablehnungsbereich ist also unabhängig von der Stichprobe)
3. Wichtig: Erst jetzt werden die Daten der Stichprobe erhoben/beachtet:
   Der Testfunktionswert $v = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}$ wird berechnet.
4. H_0 wird genau dann verworfen, wenn $v \in B$ gilt.



Beispiel:

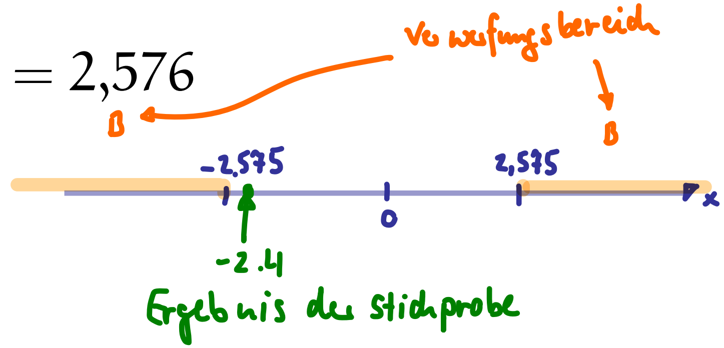
$\sigma \rightarrow$ hier bekannt

$X_1, \ldots, X_{25}$ mit $X_i \sim N(\mu; \underbrace{1{,}5}_{\sigma})$ und $\bar{x} = 499{,}28$

Prüfe $H_0: \ \mu = 500, \ H_1: \ \mu \neq 500$ zum Signifikanzniveau $\alpha = 0{,}01$

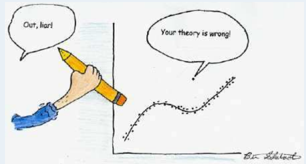
Lösung: Einstichproben-Gaußtest, Fall a)

1. $\alpha = 0{,}01$
2. $N(0;1): \ x_{1-\frac{\alpha}{2}} = x_{1-0{,}005} = x_{0{,}995} = 2{,}576$
   $\Rightarrow B = (-\infty; -2{,}576) \cup (2{,}576; \infty)$
3. $v = \frac{499{,}28 - 500}{1{,}5} \cdot \sqrt{25} = -2{,}4$
4. $v \notin B \Rightarrow H_0$ nicht verwerfen



Interpretation: Zum Signifikanzniveau 1 % kann der Brauerei keine Abweichung vom Sollwert $\mu_0 = 500$ nachgewiesen werden.

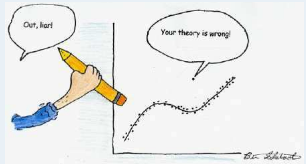
Aufbau und Klassifikation von Signifikanztests



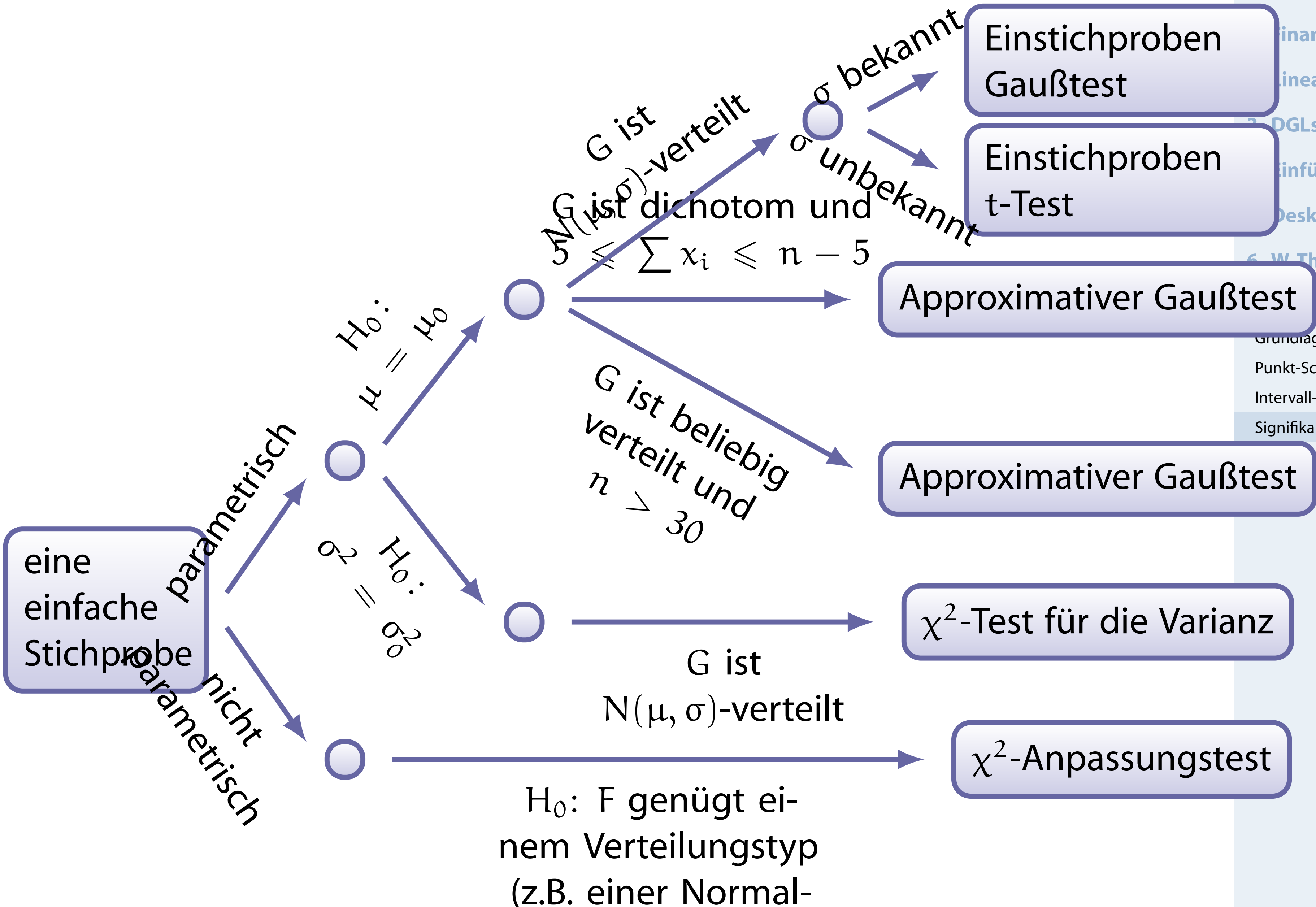
Der jeweils geeignete Test hängt ab von ...

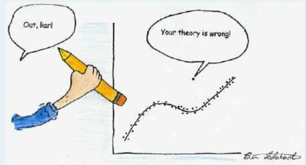
- dem zu testenden Hypothesenpaar H_0, H_1; unterscheide:
  - Parametrische Hypothesen:
    Beziehen sich auf unbekannte(n) Verteilungsparameter $(\mu, \sigma^2, \dots)$
  - Nichtparametrische Hypothesen:
    Beinhalten sonstige Aussagen, z.B. „Alter und Einkommen sind unabh.“
- den Voraussetzungen an die Verteilung/parameter (z.B. $G \sim N(\mu; \sigma)$)
- den Voraussetzungen an den Stichprobenumfang (z.B. $n > 30$)
- Art und Anzahl der Stichproben; unterscheide:
  - Signifikanztests bei einer einfachen Stichprobe
  - Signifikanztests bei mehreren unabhängigen Stichproben
  - Signifikanztests bei zwei verbundenen Stichproben

In dieser Vorlesung: Nur **einfache Stichproben**



Signifikanztests bei einer einfachen Stichprobe





Gegeben:

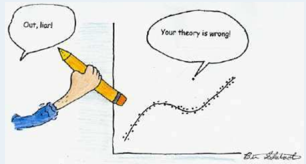
- Einfache Stichprobe $X_1, \ldots, X_n$ mit
- $E(X_i) = \mu, \; Var(X_i) = \sigma^2$

Hypothesenpaare:

$$
\begin{array}{llll}
\text{a)} & H_0: \mu = \mu_0 & & H_1: \mu \neq \mu_0 \\
\text{b)} & H_0: \mu = \mu_0 & (\text{oder } \mu \geqq \mu_0), & H_1: \mu < \mu_0 \\
\text{c)} & H_0: \mu = \mu_0 & (\text{oder } \mu \leqq \mu_0), & H_1: \mu > \mu_0
\end{array}
$$

Voraussetzungen:

1. Normalverteilung mit σ unbekannt (Einstichproben-t-Test)
   oder
2. Beliebige Verteilung
   mit $n > 30$ bzw. $5 \leqq \sum x_i \leqq n - 5$ (bei $B(1; p)$)
   (approximativer Gaußtest)



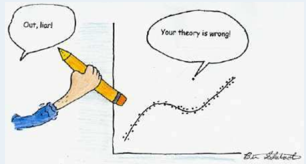
Ablauf:

1. Festlegen des Signifikanzniveaus α
2. Festlegen des Verwerfungsbereichs B:
   - Falls $H_1 : \mu \neq \mu_0$: $\quad B = (-\infty; -x_{1-\alpha/2}) \cup (x_{1-\alpha/2}; \infty)$
   - Falls $H_1 : \mu < \mu_0$: $\quad B = (-\infty; -x_{1-\alpha})$
   - Falls $H_1 : \mu > \mu_0$: $\quad B = (x_{1-\alpha}; \infty)$

   Dabei steht $x_{1-\alpha/2}$ bzw. $x_{1-\alpha}$ für das jeweilige Fraktil
   - der $t(n-1)$-Verteilung bei $n \leqslant 29$ bzw.
   - der $N(0;1)$-Verteilung bei $n \geqslant 30$.
3. Berechnen des Testfunktionswertes:

$$v = \begin{cases} \dfrac{\bar{x}-\mu_0}{s}\sqrt{n} & \text{falls Grundgesamtheit } N(\mu;\sigma)\text{-verteilt, } \sigma \text{ unbekannt} \\ & \text{oder falls Verteilung der GG beliebig, } n > 30, \ \sigma \text{ unbekannt} \\ \dfrac{\bar{x}-\mu_0}{\sigma}\sqrt{n} & \text{falls Verteilung der GG beliebig, } n > 30, \ \sigma \text{ bekannt} \\ \dfrac{\bar{x}-\mu_0}{\sqrt{\mu_0(1-\mu_0)}}\sqrt{n} & \text{falls GG gemäß } B(1;\mu)\text{-verteilt, } n > 30 \end{cases}$$

Einstichproben-t-Test: Beispiel



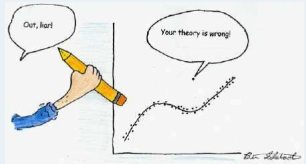
Beispiel t-Test: Energieaufnahme von Frauen

- Empfohlene täglich Energieaufnahme für Frauen: 7724 kJ (1845 kcal)
- Nehme einfache Stichprobe von 11 Frauen und teste zum Signifkanzniveau $\alpha = 0{,}05$ für
- H_0: „Der Erwartungswert der täglichen Energieaufnahme für Frauen ist 7724 kJ“ (μ_0)
- gegen H_1: $\mu \neq \mu_0$

```
daily.intake <- c(5260, 5470, 5640, 6180, 6390, 6515, 6805, 7515, 7515, 8230, 8770)
t.test(daily.intake, alternative="two.sided", mu=7724, conf.level=0.95)

##
##  One Sample t-test
##
## data:  daily.intake
## t = -2.8179, df = 10, p-value = 0.01823
## alternative hypothesis: true mean is not equal to 7724
## 95 percent confidence interval:
##  5986.348 7520.925
## sample estimates:
## mean of x
##  6753.636
```

Einstichproben-t-Test, approx. Gaußtest



Beispiel:

$X_1, \ldots, X_{2000} \sim B(1; p)$ mit

$$X_i = \begin{cases} 1, & \text{falls i-te Person Wähler einer bestimmten Partei} \\ 0, & \text{sonst} \end{cases}$$

Ergebnis der Stichprobe: $\sum\limits_{i=1}^{2000} x_i = 108$

Prüfe $H_0 : p \leqq 0{,}05$ gegen $H_1 : p > 0{,}05$ zum Signifikanzniveau 2 %

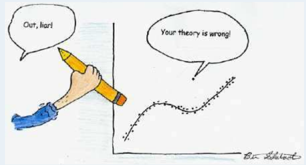
Lösung:

approximativer Gaußtest bei dichotomer (zweiwertiger) Verteilung; Voraussetzung 2 erfüllt: $5 \leqq 108 \leqq 2000 - 5$

1. $\alpha = 0{,}02$
2. $N(0;1) : x_{1-\alpha} = x_{0,98} = 2{,}05$ (Tabelle) $\Rightarrow B = (2{,}05; \infty)$
3. $v = \dfrac{\frac{108}{2000} - 0{,}05}{\sqrt{0{,}05 \cdot (1 - 0{,}05)}} \sqrt{2000} = 0{,}82$
4. $v \notin B \Rightarrow H_0$ nicht verwerfen

Zusatzfrage: Entscheidung, falls $\alpha = 0{,}01$? $\rightarrow$ Keine Änderung!

Chi-Quadrat-Test für die Varianz



- Gegeben: Einfache Stichprobe $X_1, \ldots, X_n \sim N(\mu; \sigma)$
- Hypothesenpaare:

a)	$H_0: \sigma^2 = \sigma_0^2$		$H_1: \sigma^2 \neq \sigma_0^2$
b)	$H_0: \sigma^2 = \sigma_0^2$	(oder $\sigma^2 \geqq \sigma_0^2$),	$H_1: \sigma^2 < \sigma_0^2$
c)	$H_0: \sigma^2 = \sigma_0^2$	(oder $\sigma^2 \leqq \sigma_0^2$),	$H_1: \sigma^2 > \sigma_0^2$

- **Vorgehensweise:**
  1. Festlegen des Signifikanzniveaus α.
  2. Festlegen des Verwerfungsbereichs:

$$B = [0; x_{\alpha/2}) \cup (x_{1-\alpha/2}; \infty) \qquad \text{im Fall a)}$$

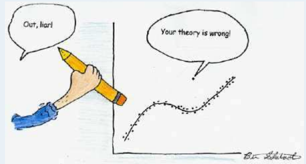
$$B = [0; x_\alpha) \qquad \text{im Fall b)}$$

$$B = (x_{1-\alpha}; \infty) \qquad \text{im Fall c)}$$

  3. Berechnung des Testfunktionswertes:

$$v = \frac{(n-1)s^2}{\sigma_0^2} = \frac{1}{\sigma_0^2} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Chi-Quadrat-Test für die Varianz



Beispiel: $G \sim N(\mu; \sigma)$

$$(x_1, \ldots, x_{10}) = (2100; 2130; 2150; 2170; 2210; 2070; 2230; 2150; 2230; 2200)$$

Prüfe $H_0: \ \sigma = 40, \ H_1: \ \sigma \neq 40$ zum Signifikanzniveau $\alpha = 0{,}1$

Lösung: χ^2-Test für die Varianz, Hypothese Fall a);
Voraussetzungen sind erfüllt

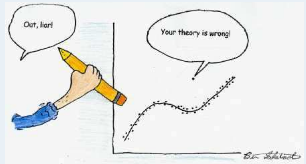
1. $\alpha = 0{,}1$
2. $\chi^2(9): \ x_{\frac{\alpha}{2}} = x_{0,05} = 3{,}33; \ x_{1-\frac{\alpha}{2}} = x_{0,95} = 16{,}92$
   (Tabelle der χ^2-Verteilung)

$$\Rightarrow B = [0; 3{,}33) \cup (16{,}92; \infty)$$

3. $\bar{x} = \frac{1}{10}(2100 + 2130 + \cdots + 2200) = 2164$

   $v = \frac{1}{40^2}\left[(2100 - 2164)^2 + \cdots + (2200 - 2164)^2\right] = 16{,}65$

   $\Rightarrow v \notin B \Rightarrow H_0$ nicht verwerfen



Zwei verbundene einfache Stichproben: Kontingenztest

- Situation: In Grundgesamtheit G: Zwei verbundene einfache Stichproben, also Beobachtung zweier Merkmale X, Y
- Hypothese:

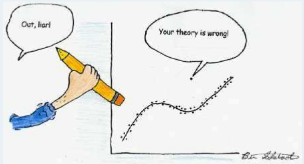
$$\begin{aligned} H_0 &: \text{ Die beiden Merkmale } X \text{ und } Y \text{ sind in } G \text{ unabhängig.} \\ H_1 &: \ X \text{ und } Y \text{ sind in } G \text{ abhängig.} \end{aligned}$$

Vorgehensweise Kontingenztest:

1. Festlegen des Signifikanzniveaus α.
2. Unterteilung der x-Achse in $k \geqslant 2$ und die y-Achse in $l \geqslant 2$ disjunkte, aneinander angrenzende Intervalle $A_1, \ldots, A_k$ bzw. $B_1, \ldots, B_l$
3. Erstellen einer Kontingenztabelle mit Randhäufigkeiten:

$x\downarrow \; y\rightarrow$	B_1	B_2	$\cdots$	B_l	
A_1	h_{11}	h_{12}	$\cdots$	h_{1l}	$h_{1\bullet}$
A_2	h_{21}	h_{22}	$\cdots$	h_{2l}	$h_{2\bullet}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	
A_k	h_{k1}	h_{k2}	$\cdots$	h_{kl}	$h_{k\bullet}$
	$h_{\bullet 1}$	$h_{\bullet 2}$	$\cdots$	$h_{\bullet l}$	n

Zwei verbundene einfache Stichproben: Kontingenztest



Vorgehensweise Kontingenztest (Fortsetzung):

4. Mit dem Fraktilswert $x_{1-\alpha}$ der χ^2-Verteilung mit $(k-1)\cdot(l-1)$ Freiheitsgraden: Berechnung des Verwerfungsbereichs

$$B = (x_{1-\alpha}; \infty)$$

5. Zu jeder Kombination aus $i = 1, \ldots, k$ und $j = 1, \ldots, l$: Berechnung der Größe

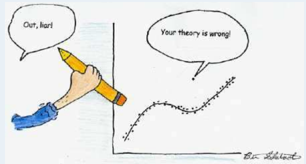
$$\tilde{h}_{ij} = \frac{h_{i\bullet} \cdot h_{\bullet j}}{n}$$

6. Berechnung des Testfunktionswerts v:

$$v = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{\left(\tilde{h}_{ij} - h_{ij}\right)^2}{\tilde{h}_{ij}} = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{h_{ij}^2}{\tilde{h}_{ij}} - n$$

7. Ablehnung von H_0 genau dann, wenn $v \in B$.

Zwei verbundene einfache Stichproben: Kontingenztest



Kontingenztest: Beispiel

- 400 Erstkandidaten einer praktischen Führerscheinprüfung schneiden abhängig von der besuchten Fahrschule folgendermaßen ab:

	Fahrschule		
	A	B	C
bestanden	130	88	62
durchgefallen	70	38	12

- Zum Signifikanzniveau von 5 % soll getestet werden, ob das Bestehen der Prüfung unabhängig von der besuchten Fahrschule ist.

Testdurchführung

1. Signifikanzniveau $\alpha = 5\,\%$
2. entfällt, da Skalenniveau nominal
3. Kontingenztabelle:

	A	B	C	$\sum$
best.	130	88	62	280
durchg.	70	38	12	120
$\sum$	200	126	74	400

4. χ^2-Verteilung mit $(3-1)\cdot(2-1) = 2$ Freiheitsgraden: $x_{1-0{,}05} = x_{0{,}95} = 5{,}99$:

$$B = (5{,}99; \infty)$$

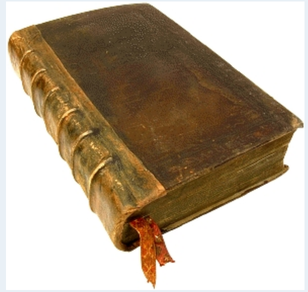
5. Berechnung der $\tilde{h}_{ij}$:

	A	B	C
best.	140	88,2	51,8
durchg.	60	37,8	22,2

6. $$v = \frac{(140-130)^2}{140} + \ldots + \frac{(22{,}2-12)^2}{22{,}2} \approx 9{,}077$$

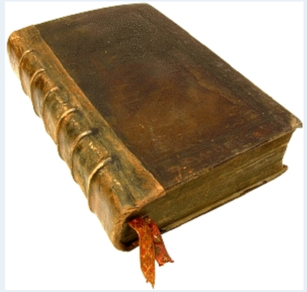
7. $v \in B$: Also wird H_0 abgelehnt, die Prüfungsergebnisse sind abhängig von der Fahrschule.

Quellenübersicht



Bücher

- Bamberg, Günter, Franz Baur und Michael Krapp (2011). **Statistik**. 16. Aufl. München: Oldenbourg Verlag. ISBN: 3486702580.
- Fahrmeir, Ludwig, Rita Künstler, Iris Pigeot und Gerhard Tutz (2009). **Statistik: Der Weg zur Datenanalyse**. 7. Aufl. Berlin, Heidelberg: Springer. ISBN: 3642019382.
- Luderer, Bernd (2003). **Starthilfe Finanzmathematik. Zinsen, Kurse, Renditen**. 2. Aufl. Stuttgart, Leipzig, Wiesbaden: Teubner.
- Opitz, Otto (2004). **Mathematik für Wirtschaftswissenschaftler**. 9. Aufl. München: Oldenbourg.



Quellen zu Bildern und Daten

- Anscombe, Francis (1973). „Graphs in Statistical Analysis“. In: **The American Statistician**, S. 195–199.
- Bach, Axel, Reinhard Brüning, Katrin Krieft, Hilmar Liebsch und Martin Rosenberg (2006). **Mit Zahlen lügen**. URL: `http://www.wdr.de/tv/quarks/sendungsbeitraege/2006/1017/000_zahlen.jsp`.
- Fahrmeir, Ludwig, Rita Künstler, Iris Pigeot und Gerhard Tutz (2009). **Statistik: Der Weg zur Datenanalyse**. 7. Aufl. Berlin, Heidelberg: Springer. ISBN: 3642019382.
- Kramer, Walter (2011). **So lügt man mit Statistik**. Piper Verlag. ISBN: 3492264131.