

2.3 Intervallschätzung

2.3.1 Motivation und Hinführung

Bsp. 2.12. [Wahlumfrage]

Der wahre Anteil der „rot-grün Wähler“ unter allen Wählern war 2009 –auf eine Nachkommastelle gerundet– genau 33.7%. Wie groß ist die Wahrscheinlichkeit, in einer Zufallsstichprobe von 1000 Personen aus allen Wählern genau einen relativen Anteil von 33.7% von „rot-grün Wählern“ erhalten zu haben?

D.h., mit Wahrscheinlichkeit von etwa 97.3%, verfehlt der Schätzer den wahren Wert.

Hat man ein stetiges Merkmal, so ist sogar $P(\bar{X} = w) = 0$ für jeden Wert w , d.h. der wahre Wert wird mit Wahrscheinlichkeit 1 verfehlt.

Konsequenzen:

- Insbesondere Vorsicht bei der Interpretation „knapper Ergebnisse“ (z.B. Anteil 50.2%)
- Suche Schätzer mit möglichst kleiner Varianz, um „im Durchschnitt möglichst nahe am wahren Wert dran zu sein“
- Es ist häufig auch gar nicht nötig, sich genau auf einen Wert als Schätzung festzulegen. Oft reicht die Angabe eines Intervalls, von dem man hofft, dass es den wahren Wert überdeckt: *Intervallschätzung*

Symmetrische Intervallschätzung basierend auf einer Schätzfunktion $T = g(X_1, \dots, X_n)$:

$$I(T) = [T - a, T + a]$$

„**Trade off**“ bei der Wahl von a :

⇒ Wie soll man a wählen?

Typisches Vorgehen:

- Man gebe sich durch inhaltliche Überlegungen einen Sicherheitsgrad (*Konfidenzniveau*) γ vor.
- Dann konstruiert man das Intervall so, dass es mindestens mit der Wahrscheinlichkeit γ den wahren Parameter überdeckt ⇒ „Konfidenzintervall“ („Vertrauensintervall“)

2.3.2 Definition von Konfidenzintervallen:

Gegeben sei eine i.i.d. Stichprobe X_1, \dots, X_n zur Schätzung eines Parameters ϑ und eine Zahl $\gamma \in (0; 1)$. Ein zufälliges Intervall $\mathcal{C}(X_1, \dots, X_n)$ heißt *Konfidenzintervall* zum *Sicherheitsgrad* γ (Konfidenzniveau γ), falls für jedes ϑ gilt:

$$P_{\vartheta}(\vartheta \in \underbrace{\mathcal{C}(X_1, \dots, X_n)}) \geq \gamma.$$

Bem. 2.13.

- Die Wahrscheinlichkeitsaussage bezieht sich auf das Ereignis, dass das zufällige Intervall den festen, wahren Parameter überdeckt. Streng genommen darf man im objektivistischen Verständnis von Wahrscheinlichkeit nicht von der *Wahrscheinlichkeit* sprechen, „dass ϑ in dem Intervall liegt“, da ϑ nicht zufällig ist und somit keine Wahrscheinlichkeitsverteilung besitzt. (Ein Zurückziehen auf subjektive Wahrscheinlichkeiten führt hier aber auch auf Widersprüche.)

- Typischerweise konstruiert man Konfidenzintervalle symmetrisch um einen Schätzer T . Es sind aber auch manchmal z.B. einseitige Konfidenzintervalle, etwa der Form

$$[\bar{X}, \bar{X} + b],$$

sinnvoll

2.3.3 Konstruktion von Konfidenzintervallen

Für die Konstruktion sehr praktische Vorgehensweise: Suche Zufallsvariable Z_{ϑ} , die

- den gesuchten Parameter ϑ enthält und
- deren Verteilung aber nicht mehr von dem Parameter abhängt, („*Pivotgröße*“, dt. Angelpunkt).
- Dann wähle den Bereich C_Z so, dass $P_{\vartheta}(Z_{\vartheta} \in C_Z) = \gamma$ und
- löse nach ϑ auf.

Konfidenzintervall für den Mittelwert eines normalverteilten Merkmals bei bekannter Varianz:

X_1, \dots, X_n i.i.d. Stichprobe gemäß $X_i \sim \mathcal{N}(\mu, \sigma^2)$, wobei σ^2 bekannt sei.

$$\left[\bar{X} - \frac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\sqrt{n}}, \bar{X} + \frac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\sqrt{n}} \right] = \left[\bar{X} \pm \frac{z_{\frac{1+\gamma}{2}} \cdot \sigma}{\sqrt{n}} \right] \quad (2.7)$$

Bem. 2.14.

Die Wirkung der einzelnen Größen im Konfidenzintervall erkennt man am besten durch eine „*Ceteris paribus-Analyse*“: Alle Größen bis auf eine werden festgehalten, diese wird variiert:

- Je größer σ , desto größer das Intervall!
- Je größer γ , desto größer $z_{\frac{1+\gamma}{2}}$
- Je größer n und damit \sqrt{n} , desto schmaler ist das Intervall

Konfidenzintervall für den Mittelwert eines normalverteilten Merkmals bei unbekannter Varianz: Neben dem Erwartungswert sei nun auch σ^2 unbekannt. σ^2 wird entsprechend durch den UMVU-Schätzer

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2,$$

(mit $S = \sqrt{S^2}$) geschätzt. Allerdings ist

$$Z = \frac{\bar{X} - \mu}{S} \cdot \sqrt{n}$$

jetzt nicht mehr normalverteilt, denn S ist zufällig.

Wir führen deshalb ein neues Verteilungsmodell ein.

Bem. 2.15. *t-Verteilung*

Gegeben sei eine i.i.d. Stichprobe X_1, \dots, X_n mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$. Dann heißt die Verteilung von

$$Z = \frac{\bar{X} - \mu}{S} \cdot \sqrt{n}$$

t-Verteilung (oder Student-Verteilung) mit $\nu = n - 1$ *Freiheitsgraden*. In Zeichen: $Z \sim t(\nu)$.

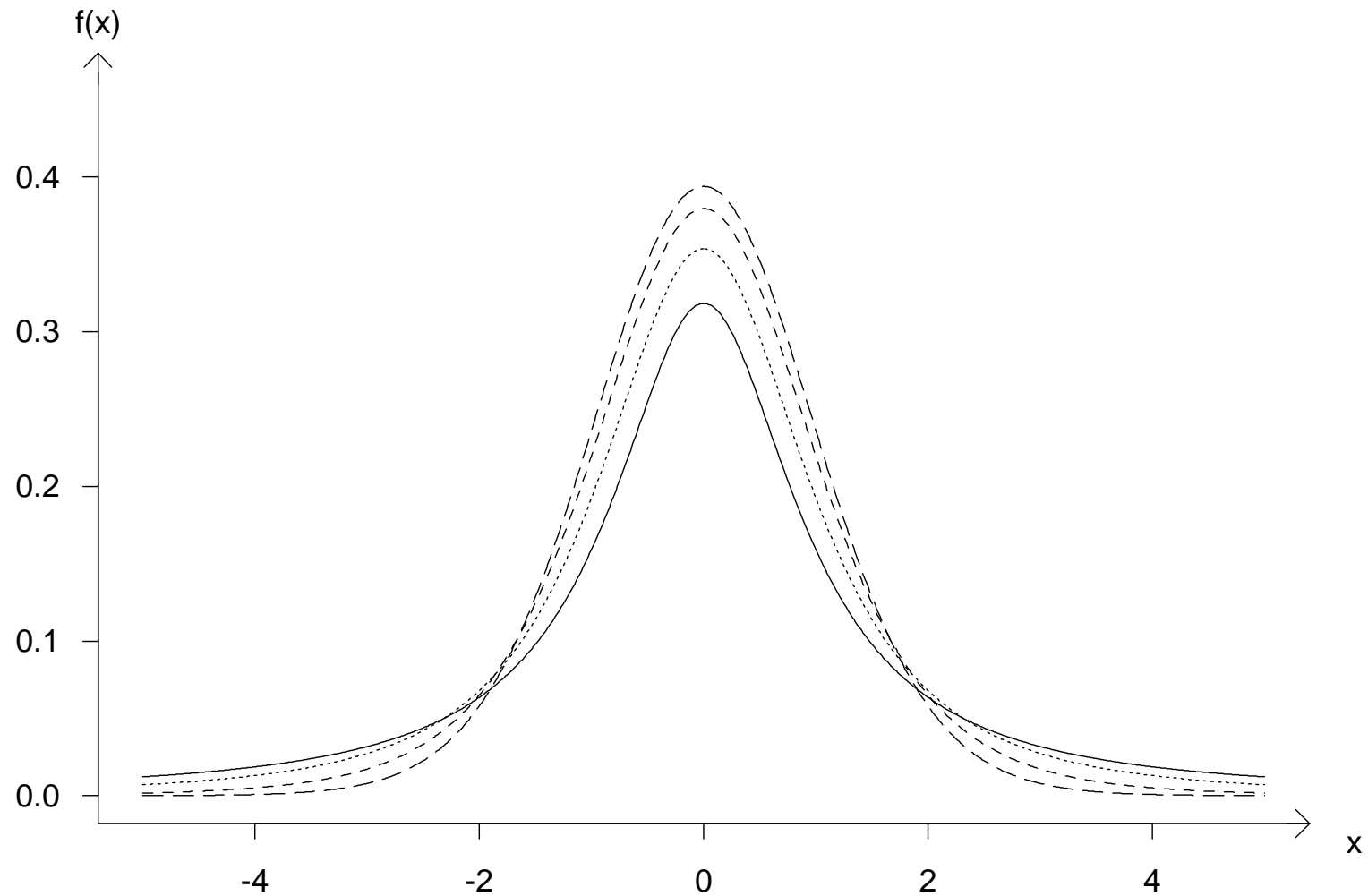
Wichtige Werte der t-Verteilung sind tabelliert. Angegeben sind, für verschiedene α , die Lösung t_α der Gleichung

$$P(Z \leq t_\alpha(\nu)) = \alpha,$$

wobei $t_\alpha(\nu)$ von der Anzahl ν der Freiheitsgrade abhängt. t_α ist das α -Quantil der entsprechenden t-Verteilung (analog zu z_α als Quantil der Standardnormalverteilung).

Die Dichte einer t-Verteilung ist der Dichte der Standardnormalverteilung sehr ähnlich: Sie ist auch symmetrisch um 0, besitzt aber etwas höhere Dichte für extreme Werte („schwerere Enden“).

n	0.8	0.9	0.95	0.975	0.99	0.995	0.999	0.9995
1	1.3764	3.0777	6.3138	12.706	31.821	63.657	318.31	636.62
2	1.0607	1.8856	2.9200	4.3027	6.9646	9.9248	22.327	31.599
3	0.9785	1.6377	2.3534	3.1824	4.5407	5.8409	10.215	12.924
4	0.9410	1.5332	2.1318	2.7764	3.7469	4.6041	7.1732	8.6103
5	0.9195	1.4759	2.0150	2.5706	3.3649	4.0321	5.8934	6.8688
6	0.9057	1.4398	1.9432	2.4469	3.1427	3.7074	5.2076	5.9588
7	0.8960	1.4149	1.8946	2.3646	2.9980	3.4995	4.7853	5.4079
8	0.8889	1.3968	1.8595	2.3060	2.8965	3.3554	4.5008	5.0413
9	0.8834	1.3830	1.8331	2.2622	2.8214	3.2498	4.2968	4.7809
10	0.8791	1.3722	1.8125	2.2281	2.7638	3.1693	4.1437	4.5869
11	0.8755	1.3634	1.7959	2.2010	2.7181	3.1058	4.0247	4.4370
12	0.8726	1.3562	1.7823	2.1788	2.6810	3.0545	3.9296	4.3178
13	0.8702	1.3502	1.7709	2.1604	2.6503	3.0123	3.8520	4.2208
14	0.8681	1.3450	1.7613	2.1448	2.6245	2.9768	3.7874	4.1405
15	0.8662	1.3406	1.7531	2.1314	2.6025	2.9467	3.7328	4.0728
16	0.8647	1.3368	1.7459	2.1199	2.5835	2.9208	3.6862	4.0150
17	0.8633	1.3334	1.7396	2.1098	2.5669	2.8982	3.6458	3.9651
18	0.8620	1.3304	1.7341	2.1009	2.5524	2.8784	3.6105	3.9216
19	0.8610	1.3277	1.7291	2.0930	2.5395	2.8609	3.5794	3.8834
20	0.8600	1.3253	1.7247	2.0860	2.5280	2.8453	3.5518	3.8495
∞	0.8416	1.2816	1.6449	1.9600	2.3263	2.5758	3.0903	3.2906



Dichten von t -Verteilungen für $\nu = 1$ (—), $= 2$ (···), $= 5$ (- - -) und $= 20$ (—) Freiheitsgrade.

Je größer ν ist, umso ähnlicher sind sich die $t(\nu)$ -Verteilung und die Standardnormalverteilung. Für $\nu \rightarrow \infty$ sind sie gleich, ab $\nu = 30$ gilt der Unterschied als vernachlässigbar.

Konfidenzintervall zum Konfidenzniveau γ :

Bem. 2.16.

- Es gelten analoge Aussagen zur Wirkung des Stichprobenumfangs und Konfidenzniveaus wie bei bekannter Varianz. (vgl. Bem. 2.14)
- Für jedes γ (und jedes ν) gilt

$$t_{\frac{1+\gamma}{2}}(n) > z_{\frac{1+\gamma}{2}} \quad (2.8)$$

also ist das t-Verteilungs-Konfidenzintervall (etwas) breiter.

- Da die Unterschiede zur Normalverteilung für $n \geq 30$ als vernachlässigbar gelten (s.o.), rechnet man für $n \geq 30$ bei der t-Verteilung auch einfach mit $z_{\frac{1+\gamma}{2}}$.

Bsp. 2.17.

Eine Maschine füllt Gummibärchen in Tüten ab, die laut Aufdruck 250g Füllgewicht versprechen. Wir nehmen im Folgenden an, dass das Füllgewicht normalverteilt ist. Bei 16 zufällig aus der Produktion herausgegriffenen Tüten wird ein mittleres Füllgewicht von 245g und eine Stichprobenstreuung (Standardabweichung) von 10g festgestellt.

- a) Berechnen Sie ein Konfidenzintervall für das mittlere Füllgewicht zum Sicherheitsniveau von 95%.
- b) Wenn Ihnen zusätzlich bekannt würde, dass die Stichprobenstreuung gleich der tatsächlichen Streuung ist, wäre dann das unter a) zu berechnende Konfidenzintervall für das mittlere Füllgewicht breiter oder schmaler? Begründen Sie ihre Antwort ohne Rechnung.

Bsp. 2.18. [Klausurergebnisse (Fiktives Beispiel)]

i	punkte	i	punkte	i	punkte	i	punkte	i	punkte	i	punkte
1	16.5	7	27.5	13	23	19	14	25	19.5	31	35.5
2	16.5	8	22	14	15	20	21	26	18	32	12.5
3	25.5	9	16.5	15	31	21	19.5	27	37.5	33	25
4	25	10	8	16	26	22	17.5	28	15.5	34	25.5
5	20.5	11	33.5	17	13.5	23	36	29	7.5	35	18.5
6	27.5	12	19.5	18	24	24	31.5	30	18		

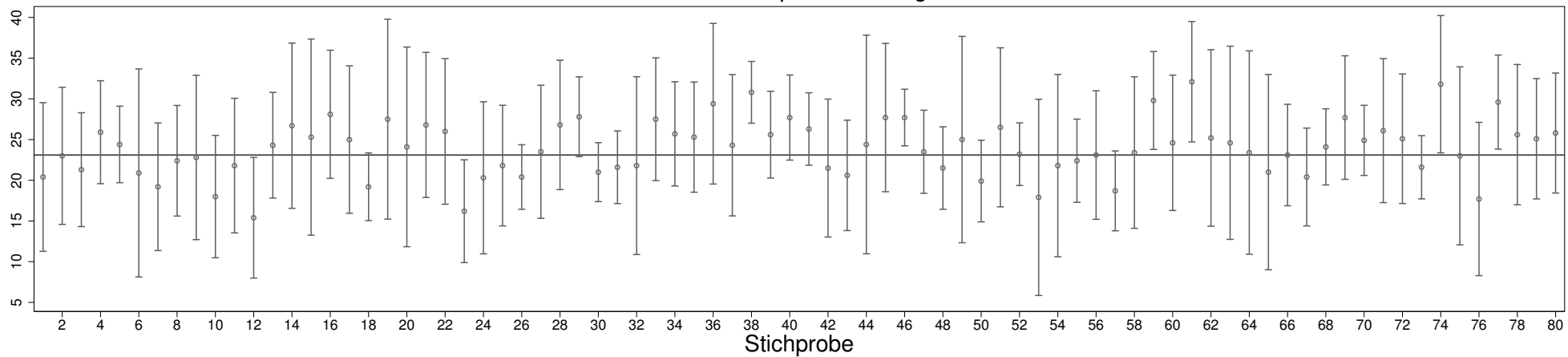
- Mittelwert und Varianz der Grundgesamtheit (alle 35 Klausuren):

$$\hat{\mu} = \overline{\text{punkte}} = 21.81$$

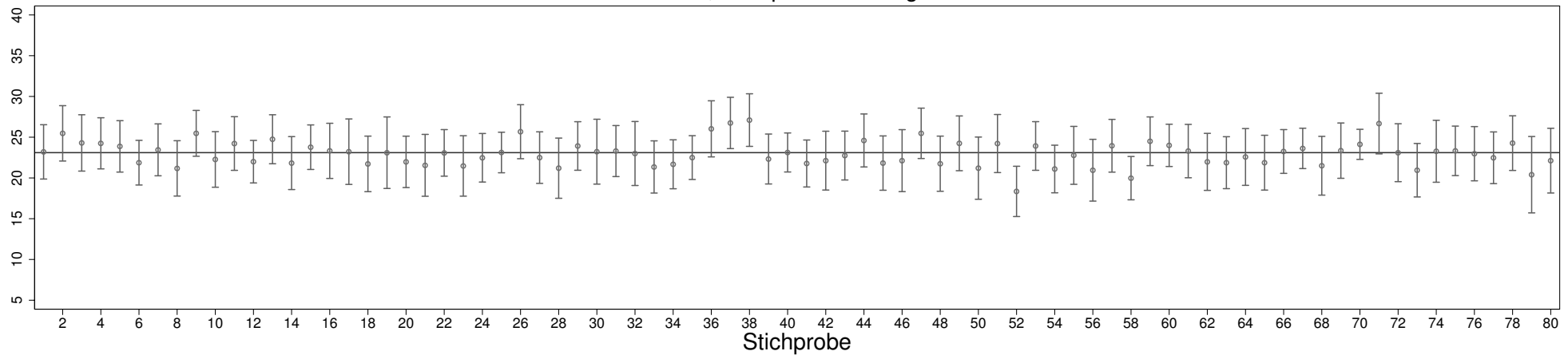
$$\hat{\sigma}^2 = s_{\text{punkte}}^2 = 7.56^2$$

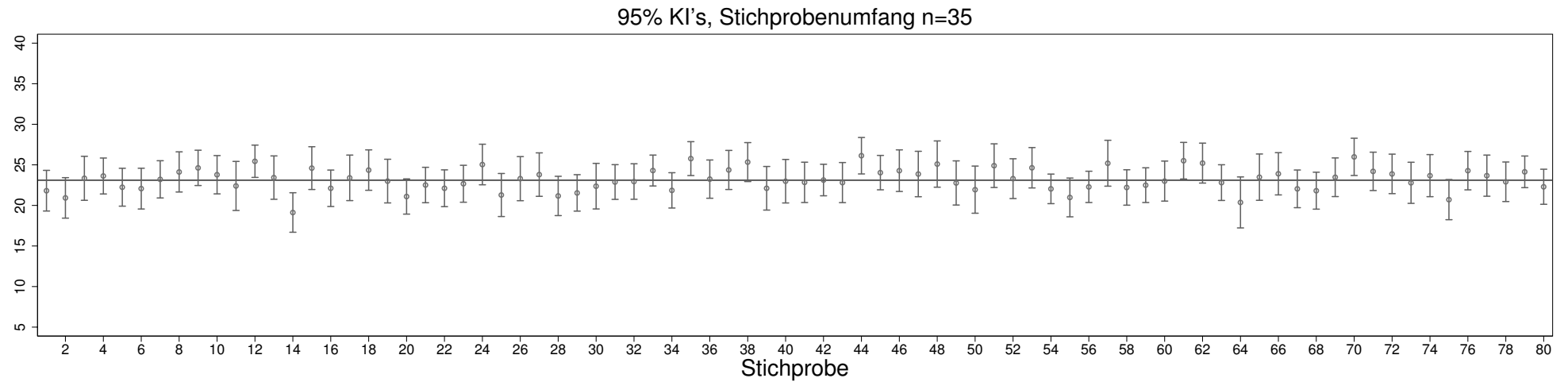
- Wir ziehen insgesamt 80 Stichproben vom Umfang $n = 5$, $n = 10$, $n = 20$, $n = 35$ und bestimmen Konfidenzintervalle zum Niveau 95%.

95% KI's, Stichprobenumfang n=5

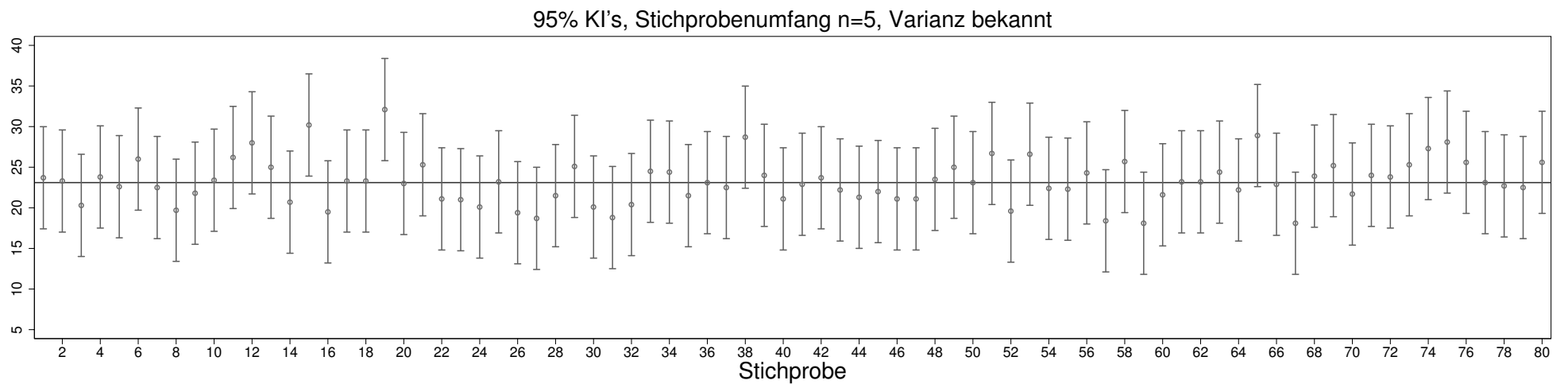
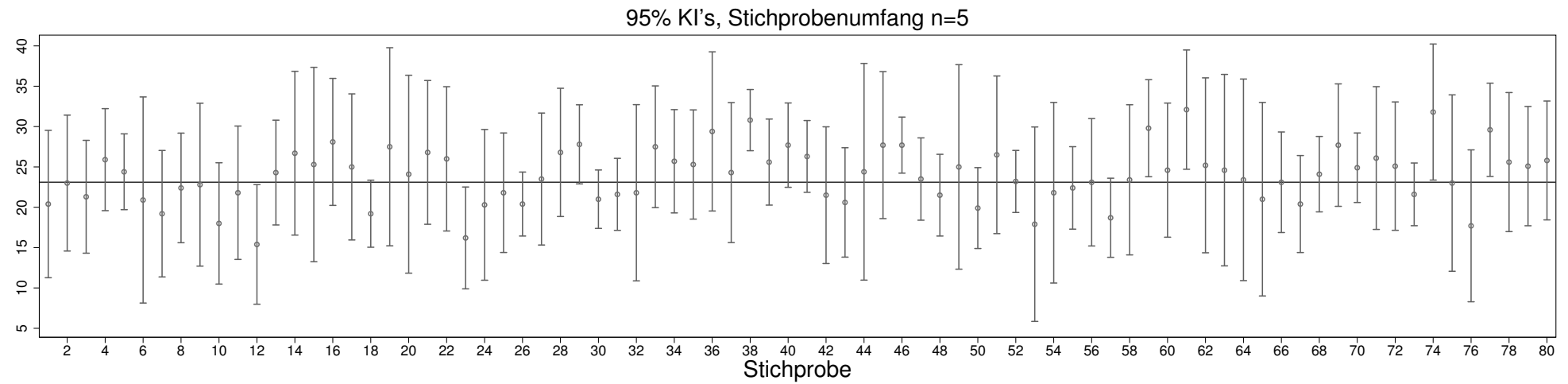


95% KI's, Stichprobenumfang n=20





Punktzahl der Klausur Statistik I: Konfidenzintervalle für 80 Stichproben mit jeweiligem Stichprobenumfang $n = 5$ (oben), $n = 20$ (Mitte), $n = 35$ (unten).



Punktzahl der Klausur Statistik I: Konfidenzintervalle für 80 Stichproben mit Stichprobenumfang $n = 5$. Oben ist die Varianz unbekannt, unten als bekannt vorausgesetzt.

- Ergebnis:
 - * Die Breite der Konfidenzintervalle nimmt mit wachsendem n ab.
 - * Nicht alle Konfidenzintervalle enthalten den wahren Mittelwert $\mu = 23.1$ (per Konstruktion tritt dieser Effekt mit Wahrscheinlichkeit 5% ein).
 - * Die Intervalle mit bekannter Varianz sind im Mittel enger.
 - * Die Intervalle mit bekannter Varianz sind immer gleich lang.

2.3.4 Approximative Konfidenzintervalle

Ist der Stichprobenumfang groß genug, so kann wegen des zentralen Grenzwertsatzes das Normalverteilungs-Konfidenzintervall auf den Erwartungswert beliebiger Merkmale (mit existierender Varianz) angewendet werden. Man erhält approximative Konfidenzintervalle, die meist auch der Berechnung mit Software zugrundeliegen.

Beispiel: Konfidenzintervall für den Anteil π : Seien X_1, \dots, X_n i.i.d. mit

$$X_i = \begin{cases} 1 \\ 0 \end{cases}, \quad P(X_i = 1) = \pi.$$

Bsp. 2.19. [Wahlumfrage]

Seien $n = 500$, $\bar{X} = 46.5\%$ und $\gamma = 95\%$.

Inhaltliche Bemerkungen Eine große praktische Bedeutung haben Konfidenzintervalle auch für die Stichprobenplanung. Sie werden herangezogen, um den benötigten Stichprobenumfang zu ermitteln.

2.3.5 Bestimmung des Stichprobenumfangs

- Je genauer und sicherer, desto größer muss der Stichprobenumfang sein
- Genauigkeit: Halbe Länge des Konfidenzintervalls
- Gib Konfidenzniveau (oft 95%) und eine Schranke b_{\max} an (inhaltliche Entscheidungen!), also de facto halbe Breite für die minimal zulässige Genauigkeit im Sinne von $\frac{1}{b_{\max}}$
- Bestimme den Mindeststichprobenumfang so, dass γ und b_{\max} eingehalten werden

Konkrete Umsetzung für die Anteilsschätzung

γ : Konfidenzniveau
g: Genauigkeit

$$z_{\frac{1+\gamma}{2}} \cdot \sqrt{\frac{\bar{X}(1 - \bar{X})}{n}} \leq b_{\max}$$

Auflösen nach n :

$$n \geq \frac{1}{b_{\max}^2} z_{\frac{1+\gamma}{2}}^2 \cdot \bar{X}(1 - \bar{X})$$

Da \bar{X} unbekannt ist, verwendet man zur Stichprobenplanung z.B. Schätzwerte früherer Untersuchungen oder als konservative Schranke $\bar{X}(1 - \bar{X}) \leq 0.25$.

Bsp. 2.20.

Gegeben:

- Konfidenzniveau: $\gamma = 95\%$
- Genauigkeit: $b_{\max} = 0.1$

Bestimmung von n :

$$n \geq \frac{1}{g^2} z_{\frac{1+\gamma}{2}}^2 \cdot \bar{X}(1 - \bar{X}) = \frac{1}{0.1^2} 1.96^2 \cdot 0.25 = 96.04$$

wobei konservativ angestzt wurde $\bar{X}(1 - \bar{X}) \leq 0.25$

Also sollten ca. 100 Personen befragt werden.

Bei $g = 5\%$ ergibt sich $n = 385$, bei $g = 1\%$ ergibt sich $n = 9604$

2.3.6 Ausblick

Viele weitere Konfidenzintervalle, auch bezüglich der Unterschiede bei zwei Stichproben

- Differenz von Mittelwerten bei unabhängigen Stichproben
- Differenz von Anteilen bei unabhängigen Stichproben
- Differenz von Mittelwerten bei „verbundenen Stichproben“

Diese gehorchen immer dem gleichen Konstruktionsprinzip (vgl. von Kap. 2.3.3) und sind analog zu den genauer besprochenen Beispielen zu interpretieren. Man vergleiche auch die Konstruktion von Tests im Kapitel 2.4, die auf einem ähnlichen Ausgangspunkt rekurrieren. In der Tat gibt es einen engen Zusammenhang zwischen Hypothesentests und Konfidenzintervallen (vgl. „Dualität“ am Ende von Kap. 2.4).