

UNIVERSITÄT LEIPZIG

Medizinische Fakultät

Vorlesung 6: Punkt- und Intervallschätzung

Dr. rer. med. habil. Marita Ziepert /
Dr. rer. nat. Meinhard Mende,
IMISE, Universität Leipzig

imise.

Lernziele Konfidenzintervalle I

Am Ende dieser Vorlesung

- kennen Sie die Bedeutung des statistischen Terminus Schätzen,
- erkennen Sie, ob eine Stichprobe repräsentativ ist,
- wissen Sie um die Bedeutung des Zentralen Grenzwertsatzes,
- können Sie den Standardfehler des Mittelwertes (S.E.M.) von der Standardabweichung (SD) unterscheiden,
- können Sie Konfidenzintervalle für Mittelwerte und Anteile berechnen und interpretieren

V 6

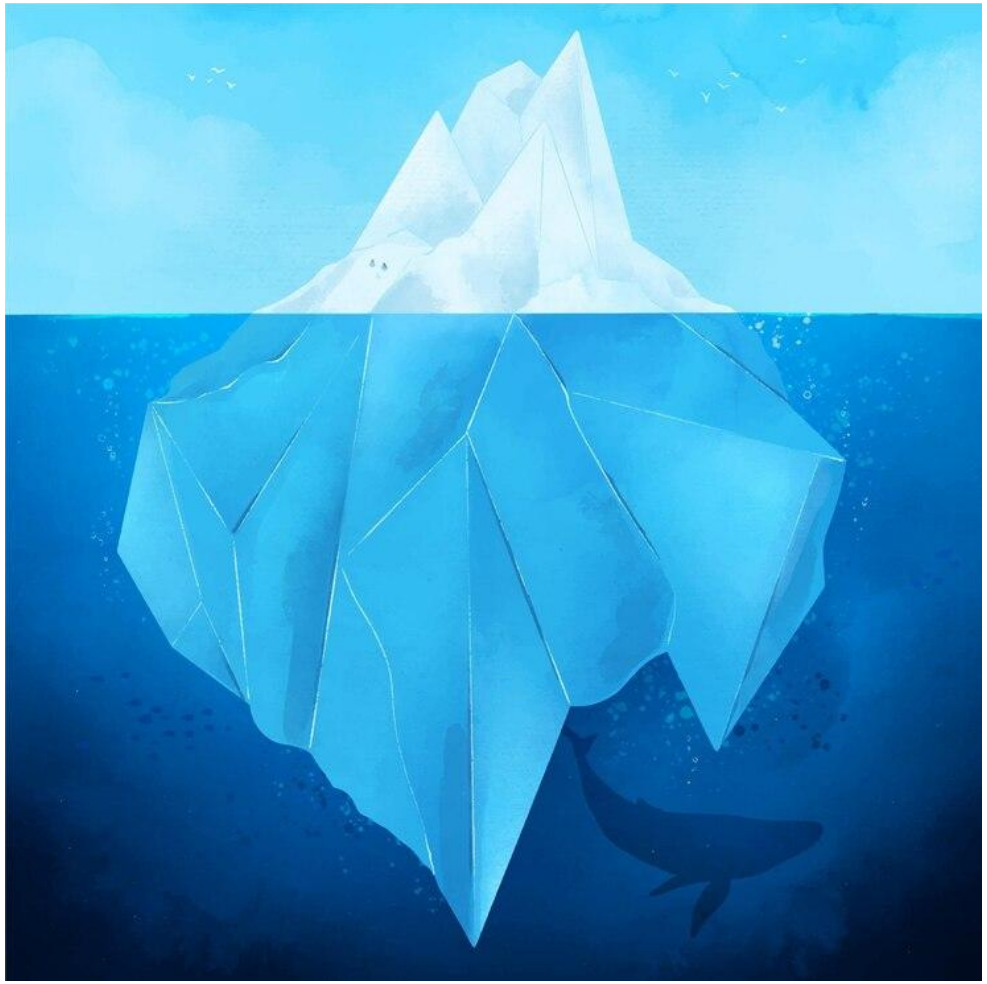
6.1 Schätzen / Stichproben

Biometrie – Leben messen?

Mögliches Motto (Martin 1868):

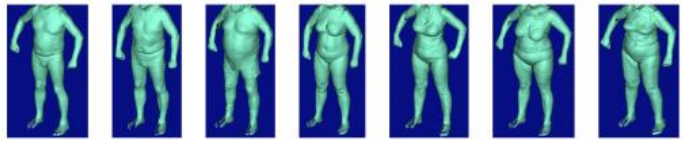
> *Man muss messen, was messbar ist, und messbar machen, was zunächst nicht messbar ist.*

- Zahlen drücken Strukturen des Kosmos aus.
- Wissenschaft erforscht das mit den Sinnen erfassbare.
  - Mess- und Labortechnik
  - Bildgebende Verfahren
  - genetische Untersuchungen



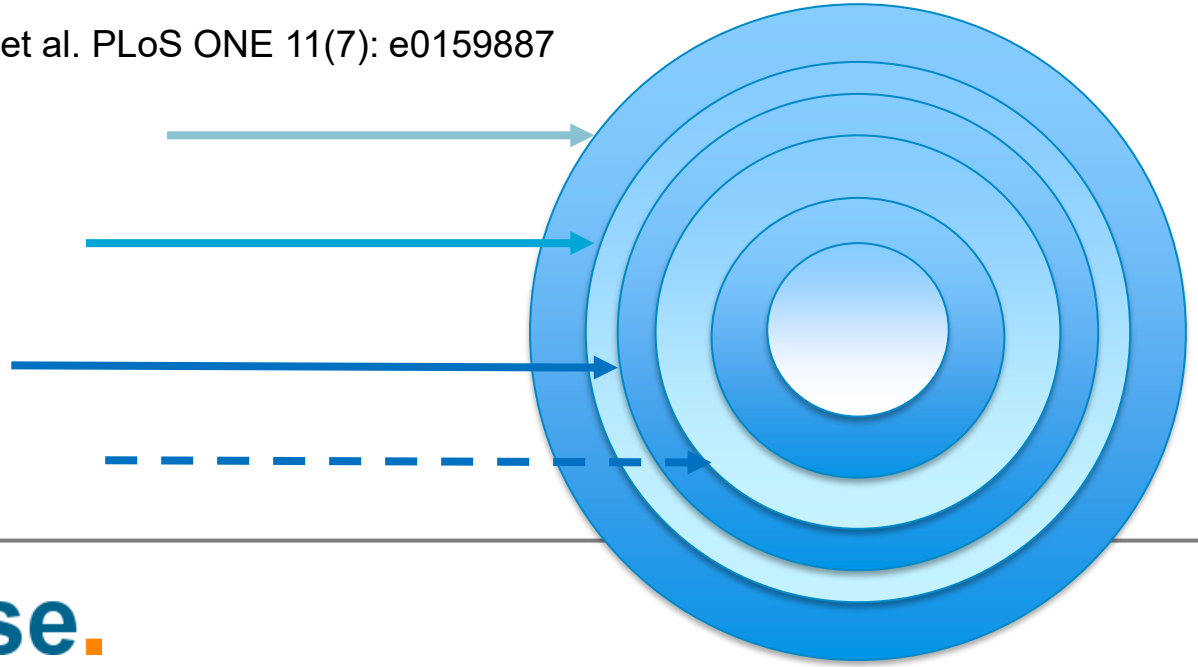
Freie Clipart von pikisuperstar

Biometrie = Leben messen



- Beispiel 1.2 LIFE-Projekt:
  - Ziel: Ursachen für Zivilisationskrankheiten erforschen
  - Studie:
    - von ca. 10 000 Leipzigern wurden viele Merkmale erfasst (soziodemographische, anthropometrische Merkmale, Labor, Echo und EKG, Fragebögen, Ganzkörperscanner, …)

Löffler-Wirth H et al. PLoS ONE 11(7): e0159887



Messen - Schätzen

Ziel

Eine Grundgesamtheit (GGS) soll anhand von Kenngrößen charakterisiert werden.

Herausforderungen:

1. Gesamterhebung meist nicht möglich

   Lösung: repräsentative Stichprobe

2. Messwerte variieren zwischen den Probanden

   Lösung: Parameter als typische Kenngröße ermitteln

3. Wie sicher / genau ist ein möglicher Schätzwert?

   Lösung: Konfidenzintervall

Schätzen I

Schätzen (engl.: estimation):

> Parameter (z. B. einer Verteilung) auf Basis von Daten einer Stichprobe mit Hilfe einer Schätzfunktion (Schätzer) optimal bestimmen.

Dabei gibt es zwei zentrale Fragestellungen:

1. Wie ausgeprägt ist ein Effekt?

   Schätzen des *wahren* Wertes aus einer Stichprobe

= Punktschätzung

2. Wie präzise ist der Schätzer?

   Konfidenzintervall des Schätzers

= Intervallschätzung

Schätzen II

Die Grundgesamtheit ist die Gesamtheit aller Beobachtungseinheiten, über die man eine Aussage treffen möchte.

Optimalitätskriterien für Schätzer:

- erwartungstreu / unverfälscht (engl. unbiased)
- höchste Präzision

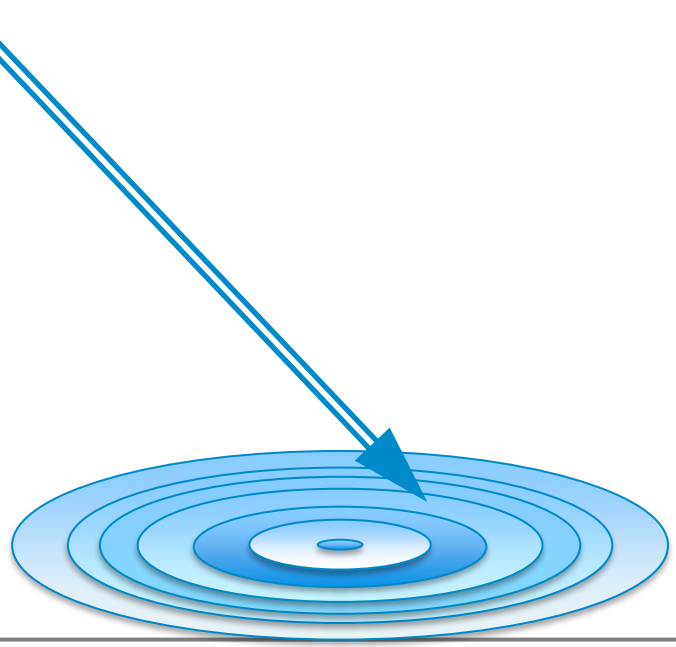
Schätzen kann man zum Beispiel

- Wahrscheinlichkeiten
- Mittelwerte
- Differenzen zweier Mittelwerte
- die Steilheit einer Regressionsgeraden oder
- die Stärke einer Korrelation

Schätzen III

– Eine Punktschätzung ist die Schätzung eines unbekannten Parameters aufgrund der Beobachtungen an einer zufällig gezogenen möglichst unverzerrten Stichprobe

wahrer Wert	unbekannt, aber gesucht
durch Stichprobe geschätzter Wert	empirisch ermittelter Wert



Wiederholung: Stichproben

- Eine einfache Zufallsstichprobe ist ein Verfahren, bei dem alle möglichen Stichproben vom Umfang *n* gleichwahrscheinlich aus der Grundgesamtheit gezogen werden können.
- Angestrebte Eigenschaften:

1. Jede Beobachtungseinheit hat dieselbe Wahrscheinlichkeit ausgewählt zu werden

→ unverzerrt (unbiased)

2. Die Auswahl der einen Beobachtungseinheit beeinflusst die Auswahl der anderen nicht

→ unabhängig

Ziel: repräsentative Stichprobe,
d. h., die Stichprobe setzt sich genauso zusammen wie die GGS.



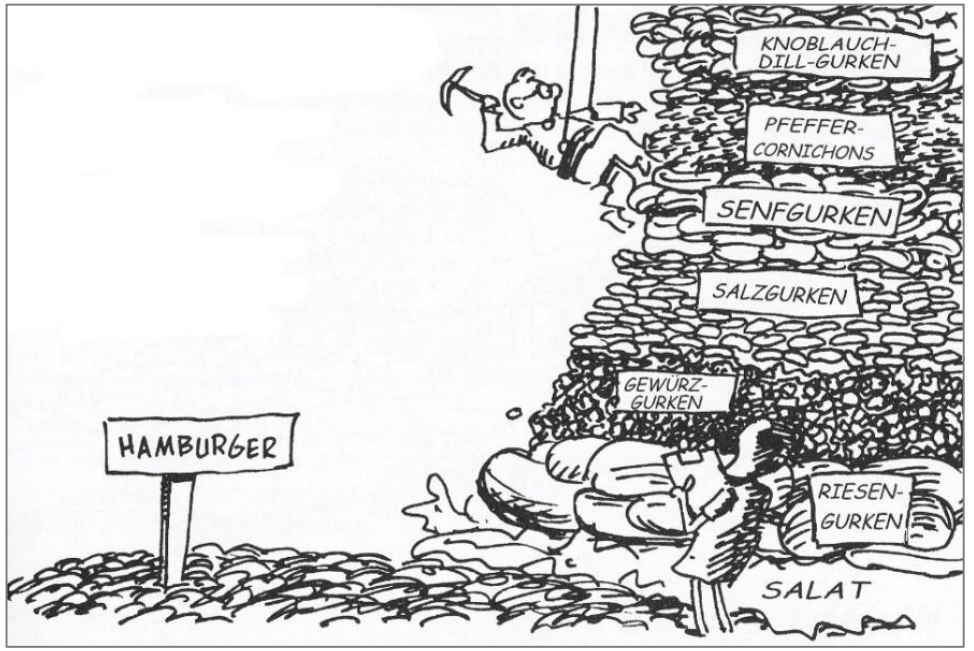
Gonick & Smith: Statistik in Cartoons

Stichproben II

- Bei einer geschichteten Stichprobe teilt man die Grundgesamtheit in homogene Gruppen (Schichten) ein und zieht für jede Gruppe eine einfache Zufallsstichprobe.

Beispiel B 6.1: LIFE-Studie,
Gesundheitszustand der Leipziger Bevölkerung

- Einteilung in Altersgruppen;
- daraus Zufallsstichproben



Gonick & Smith: Statistik in Cartoons

Stichproben III

- Bei einer Cluster-Stichprobe (engl. cluster = Gruppe, Verbund, …) zieht man nicht einzelne Probanden, sondern natürlich definierte Gruppen und aus diesen einige oder alle Probanden
- Probanden eines Clusters hängen voneinander ab, sind einander ähnlicher als Probanden verschiedener Cluster.


2348 Originalarbeit | Original article

Einfluss eines täglichen Schulsportunterrichts auf das kardiovaskuläre Risiko – 2-Jahres Ergebnisse einer Cluster-randomisierten Studie

Effects of daily physical exercise at school on cardiovascular risk – Results of a 2-year cluster-randomized study

Autoren C. Walther[1,2] M. Mende[3] L. Gaede[2] U. Müller[1] K. Machalica[1] G. Schuler[1]

Beispiel 6.2:

Leipziger Schülersportstudie

(z. B. Dtsch Med Wochenschr 2011;136: 2348–2354)

- 5. / 6. Schulklassen als Cluster wurden randomisiert

Stichproben IV

- Bei einer systematischen Stichprobe startet man mit einer zufällig ausgewählten Beobachtungseinheit und wählt danach jede k-te Beobachtungseinheit für die Untersuchung aus.

- Z. B.: Zufriedenheit der Patienten in einer Ambulanz

Stichproben V

Auswahlverfahren:

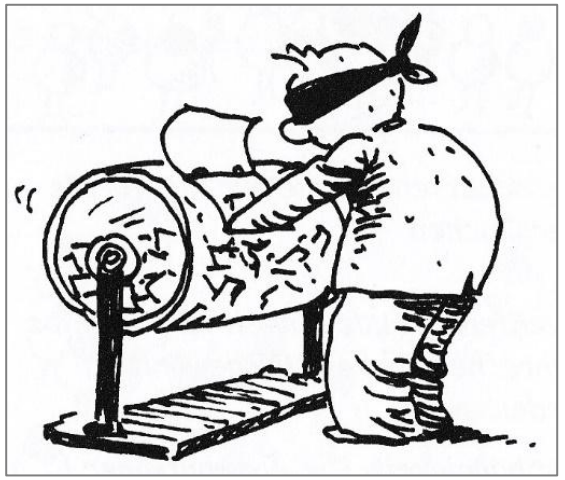
- Lostrommel
- Datenbank mit Zufallszahlengenerator

Vorteil von Stichproben:

- sparen Zeit, Aufwand und Geld

Herausforderung:

- Um zu validen Erkenntnissen zu kommen, benötigt man eine Fallzahlplanung.



Gonick & Smith: Statistik in Cartoons



Stichprobenziehung – Beispiel 6.3

- Frage:

Wie hoch ist die Prävalenz der schweren Sepsis auf Intensivstationen in Deutschland?

→ Vorlesung 14: Epidemiologie

- Sepsis-Studie

  Engel C et al. for the Competence Network Sepsis (SepNet)

  Epidemiology of sepsis in Germany: Results from a national prospective multicenter study

  Intensive Care Medicine 2007;33(4):606-618

- Stichprobenziehung?

B 6.3 Sepsis-Studie

- Sepsis:
  - Die Sepsis ist eine lebensbedrohliche Organfunktionsstörung aufgrund einer fehlregulierten Reaktion des Körpers auf eine Infektion.
    (Reinhard Larsen, Anästhesie und Intensivmedizin für die Fachpflege. 2016 Jun 14:958–963)
- Septischer Schock
  - eine Unterform der Sepsis …, wenn folgende zwei Kriterien erfüllt sind:
    - anhaltende Hypotension …
    - Serumlaktatwerte über 2 mmol/l …
- Behandlung überwiegend auf Intensivstationen
- 28-Tage-Letalität: mehr als 40%

B 6.3 Sepsis-Studie

Stichprobenziehung:

- 20 % aller Krankenhäuser in Deutschland zufällig ausgewählt
- Stichtage zufällig ausgewählt

$$\text{Schätzer für die Prävalenz} = \frac{\text{Anzahl der Sepsis-Patienten}}{\text{Anzahl aller Patienten auf den Intensivstationen an den Stichtagen}}$$

Konfidenzintervalle I

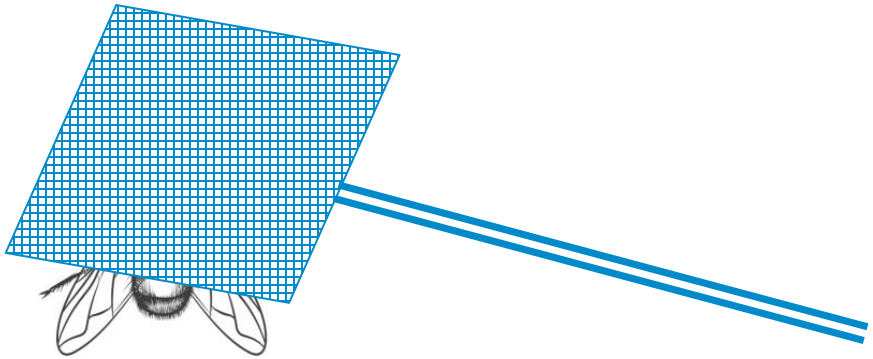
- Anhand einer Formel berechnet man einen Wert für die Kenngröße aus den Daten der Stichprobe (Punktschätzer).
- Ziel: ein Intervall bestimmen, das mit einem abschätzbaren Maß an Vertrauen (confidence) den wahren Wert enthält.

Methode:

- Schätzung (Konstruktion) von Konfidenzintervallen (C.I.)

(Theoretische) Modellannahme:

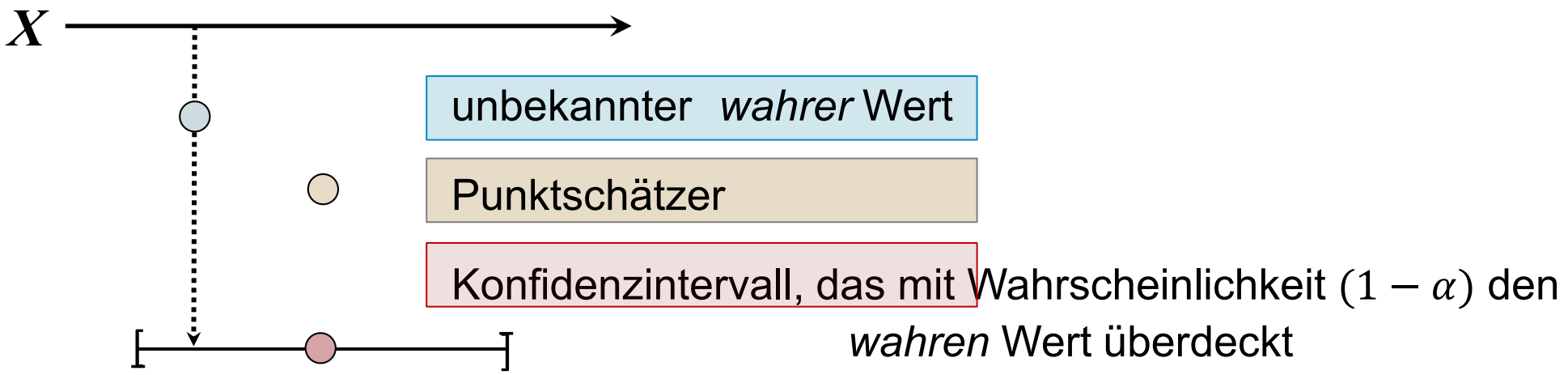
- Beliebige Wiederholbarkeit des Experiments



https://de.freepik.com/vektoren-kostenlos/hand-gezeichnete-fliegen-insekt-mit-fluegel-biologie-und-skizze_10602594.htm#fromView=keyword&page=1&position=16&uuid=7f6e1d5f-dd62-46bf-81df-cd141ae315d3&query=Fliege+clipart

Konfidenzintervalle II

– Konfidenzintervall (C.I.): aus der Stichprobe rechnerisch konstruiertes Intervall um den Punktschätzer einer Kenngröße herum, das mit einem quantifizierbaren Maß (= Wahrscheinlichkeit) an Vertrauen den unbekannten (hypothetischen) *wahren* Wert überdeckt.



– Hierbei bezeichnet α die Fehlerquote (→ vgl. Vorlesung 9)

Bausteine zur Entwicklung eines C.I.

1. Wir starten mit einem Punktschätzer:
   Wir haben z. B. den Mittelwert zur Charakterisierung der Lage einer Verteilung kennengelernt.
   - Der Mittelwert hat mathematisch gute Eigenschaften.

2. Wir nutzen aus, dass der Mittelwert für große Stichproben annähernd normalverteilt ist (→ Zentraler Grenzwertsatz).

3. Um zu einem $(1-\alpha)$-Konfidenzintervall zu gelangen, schneiden wir an beiden Enden der empirischen Verteilung einen Wertebereich ab, der insgesamt nur mit Wsk. α vorkommt.

Zentraler Grenzwertsatz



Die Normalverteilung spielt eine hervorragende Rolle als Grenzverteilung für viele Verteilungen, wenn die Fallzahl wächst.

So gilt z. B. der Zentrale Grenzwertsatz von Lindeberg / Lévy (sinngemäß):

- Ist S_n die Summe von n unabhängigen, identisch verteilten Zufallsgrößen mit gemeinsamem Erwartungswert μ und Streuung σ^2.
- Dann konvergiert die standardisierte Zufallsgröße $Z = \frac{S_n - n \cdot \mu}{\sqrt{n} \cdot \sigma}$ für wachsende n gegen die Standardnormalverteilung.

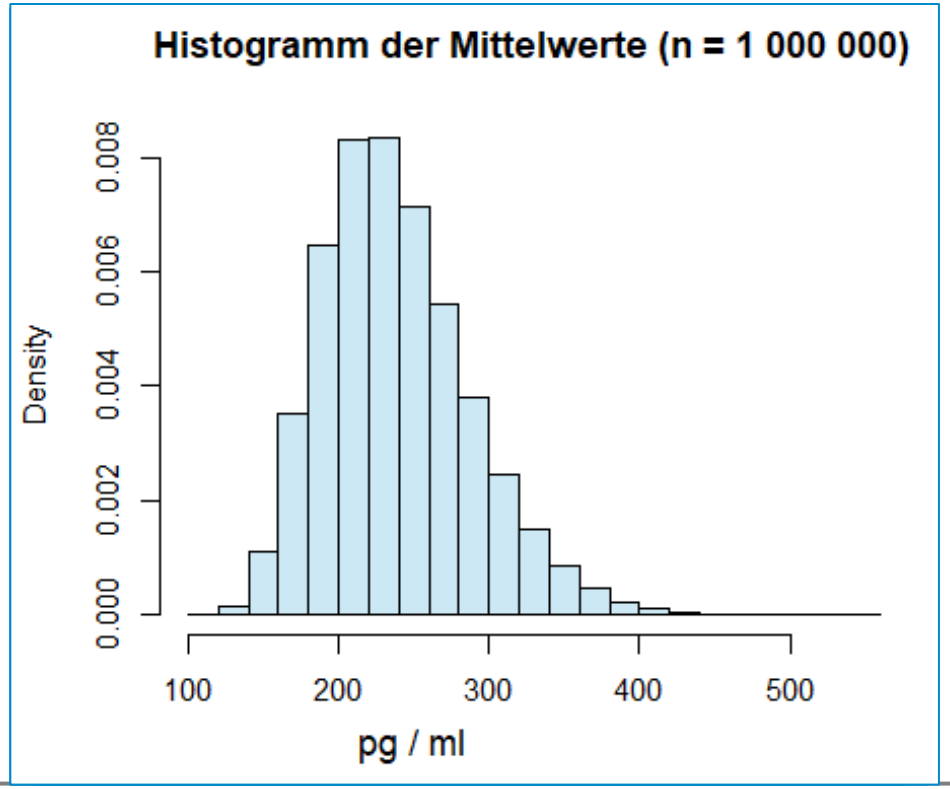
P. P. Eckstein, Repetitorium Statistik. Betriebswirtschaftlicher Verlag Dr. Th. Gabler GmbH, Wiesbaden 2001

Zentraler Grenzwertsatz II

Beispiel: NT-proBNP ist ein Herzmarker mit extrem rechtsschiefer Verteilung; d. h.

- die Verteilung erstreckt sich über mehrere Zehnerpotenzen
- und es gibt wenige sehr extreme Werte

- Bildet man jetzt n = 100, 1 000, …, 1 000 000 Stichproben dieser Werte
- und stellt das Histogramm der Mittelwerte dar,
- gehorchen diese immer besser einer Gauß-Verteilung



ZGWS: Folgerung für die Praxis

Für viele Verteilungen gilt, dass sich Summe bzw. Mittelwert für $n \to \infty$ der Normalverteilung annähert. Der theoretische Physiker E.T. Jaynes nennt etliche gewichtige Gründe, die Normalverteilung als Normalfall anzusehen.

E.T.Jaynes: Probability Theory. The Logic of Science

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj75pWj9-fxAhWTPewKHcXmCjUQFnoECAUQAA&url=https%3A%2F%2Fbayes.wustl.edu%2Fetj%2Fprob%2Fbook.pdf&usg=AOvVaw1gMptGqGmy5WUtU-LKRU1A (2026-10-02)

Insbesondere ist beim Mittelwertvergleich ein vorheriger Test auf Normalverteilung wenig sinnvoll, ja irreführend (Rasch D et al. Stat Papers 2011;52:219-231)



V 6

6.2 Konfidenzintervalle für den Mittelwert

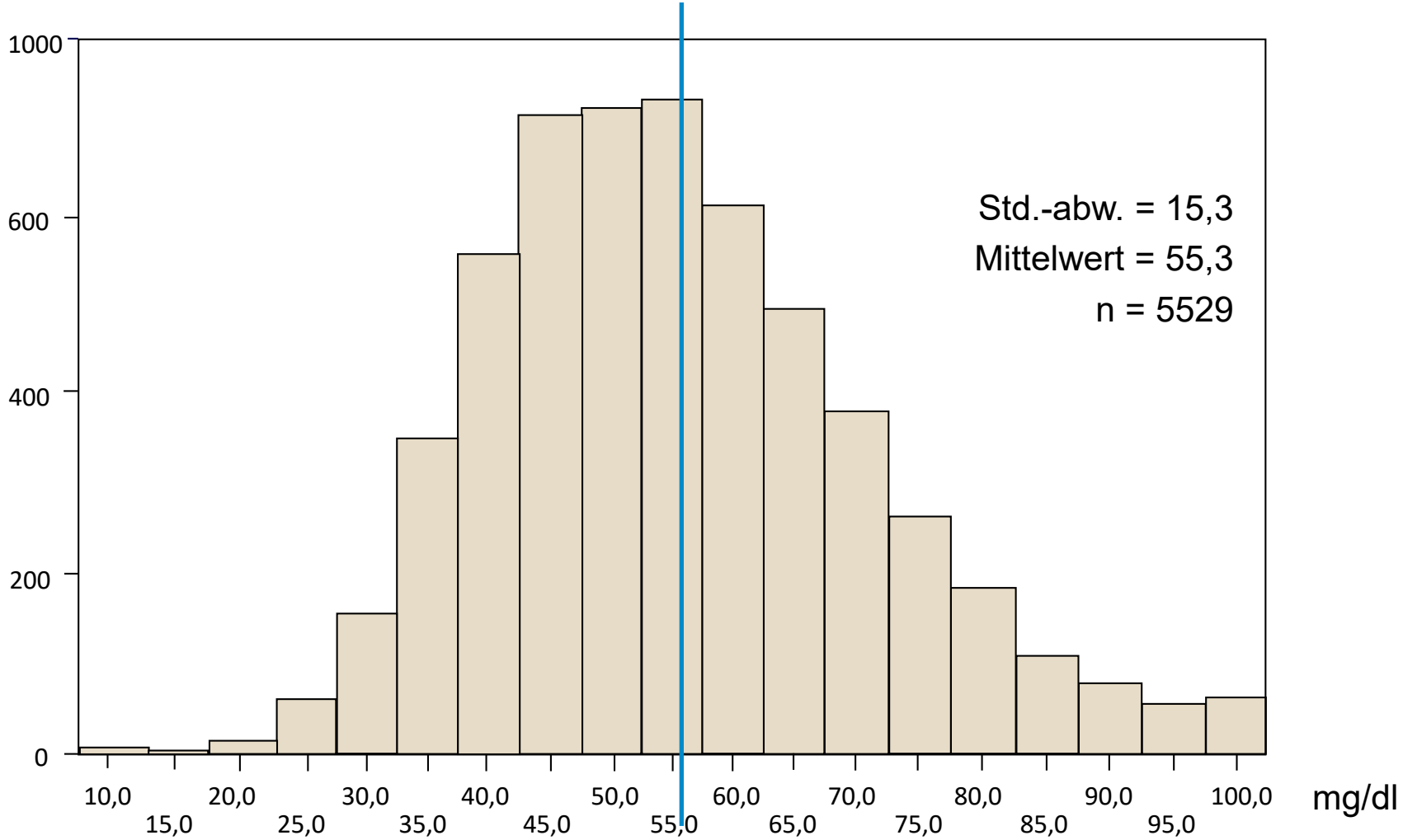
Beispiel 6.4 – Lipid-Screening

- Der mittlere HDL-Cholesterolwert von Frauen soll ermittelt (geschätzt) werden (Lipid-Screening)
- Es liegen Messungen von n = 5529 Frauen einer Grundgesamtheit vor. (s. Richter V et al. Clin Cardiol 2007;12:51-53)

- Um die Eigenschaften von Schätzwerten zu untersuchen, ziehen wir später wiederholt Stichproben verschiedener Größe aus diesen 5529 Werten.

Lipid-Screening - HDL – Frauen

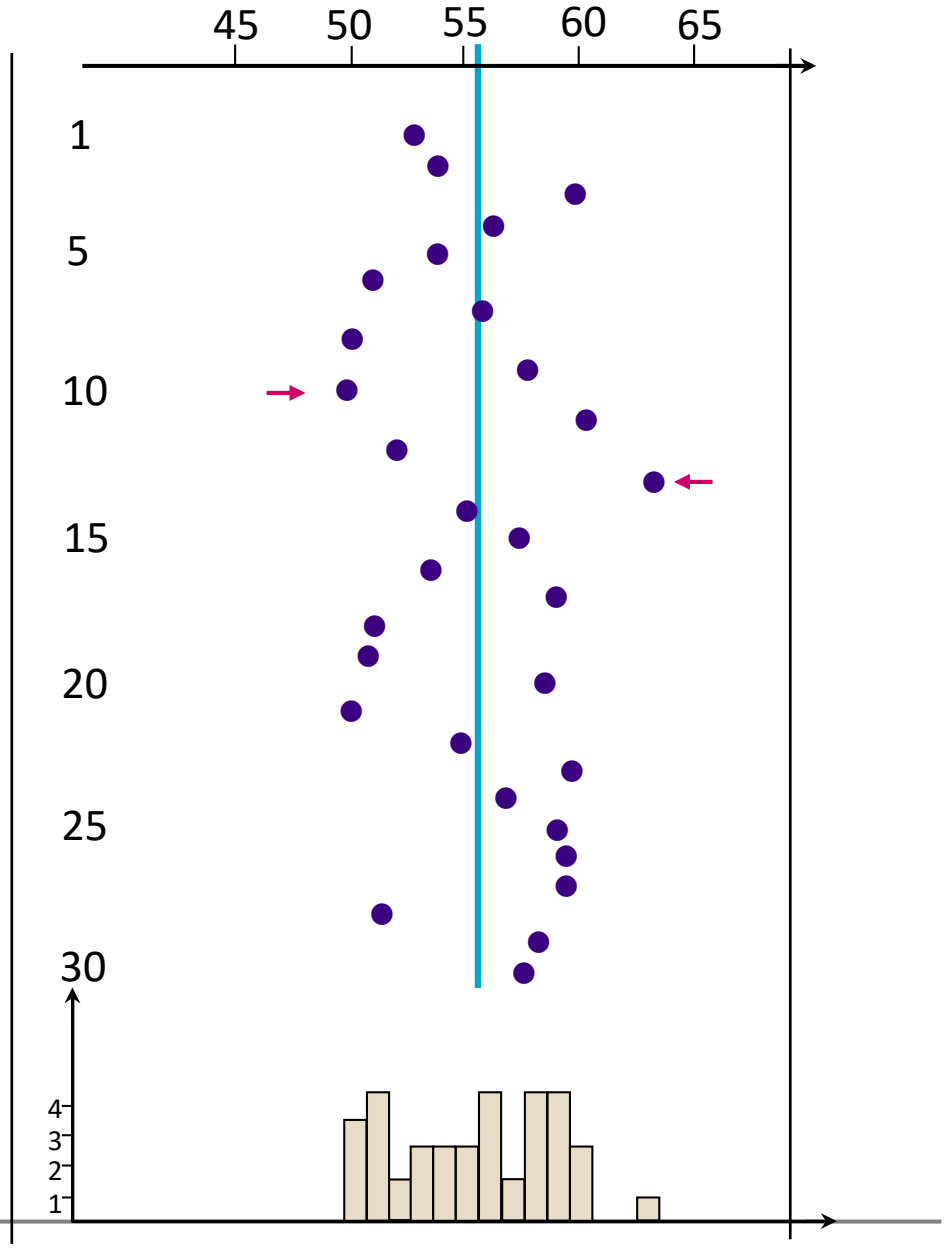
- Punktschätzung:
  $\hat{\mu} = 55{,}3\ mg/dl$



Konfidenzintervalle III

Gedankenexperiment:

- Wir ziehen 30 Stichproben von je 20 Probandinnen
- und tragen die Mittelwerte auf:
- Jede Stichprobe liefert einen anderen Schätzwert $\hat{\mu}$ für den Mittelwert

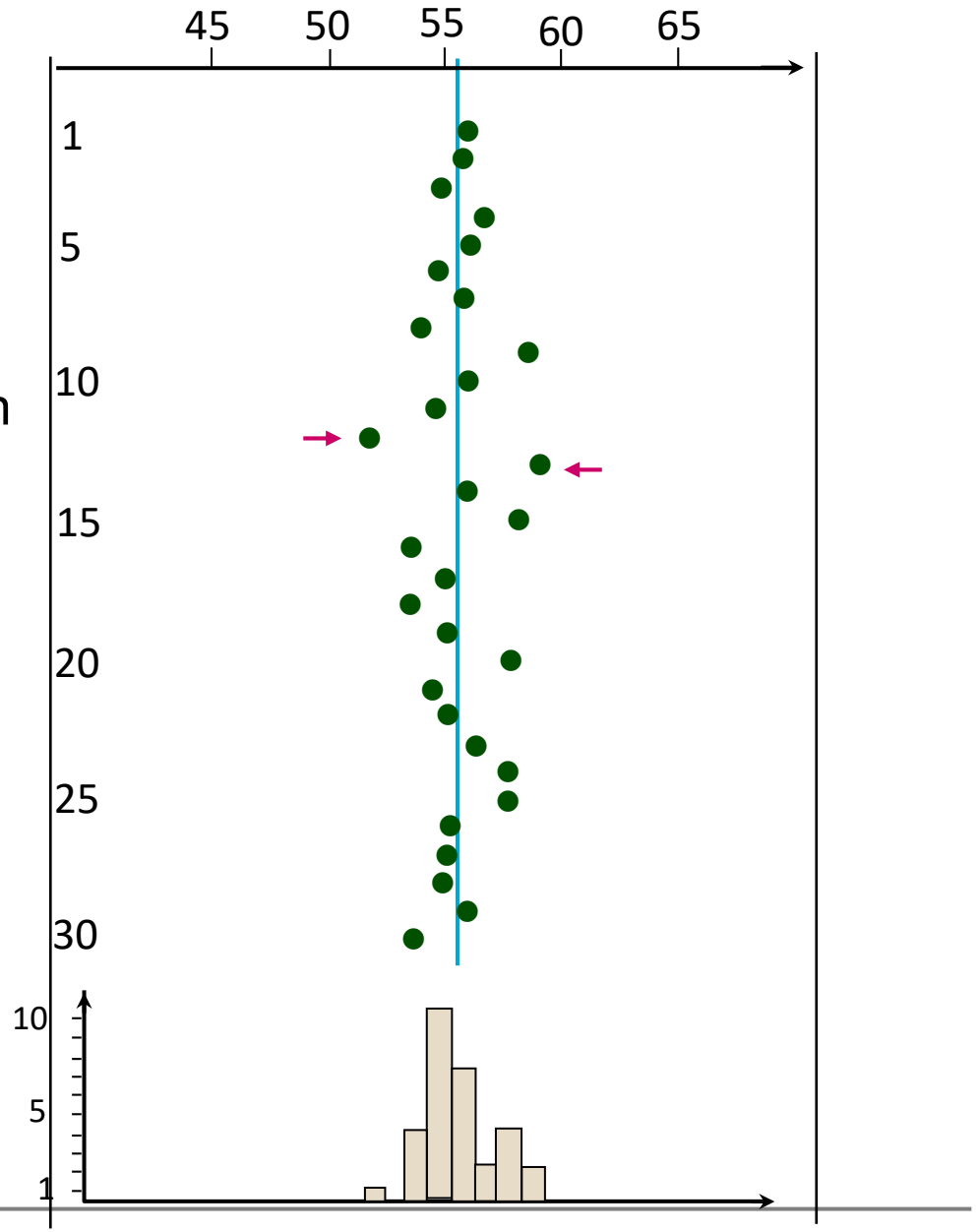


Konfidenzintervalle IV

Gedankenexperiment:

- Wir ziehen 30 Stichproben von je 100 Probandinnen
- und tragen die Mittelwerte auf:
- Jede Stichprobe liefert einen anderen Schätzwert $\hat{\mu}$ für den Mittelwert

- Je größer die Stichprobe, desto enger dürfen wir den Bereich ziehen, in dem 95% aller Werte liegen werden.



Standardfehler des Mittelwertes

– Der Standardfehler des Mittelwertes (standard error of the mean, S.E.M.) beschreibt die Genauigkeit der Schätzung eines Mittelwertes durch eine Stichprobe des Umfanges *n*.

$$s.e.m. = \sqrt{\frac{s^2}{n}} = \frac{s}{\sqrt{n}}$$

Der S. E. M.

– wird aus der empirischen Standardabweichung der Einzelwerte in der Stichprobe geschätzt und

– sinkt für wachsende n mit dem Faktor $\frac{1}{\sqrt{n}}$

– und beschreibt den erwarteten Messfehler.

Unterscheide: S. E. M. und Standardabweichung

– Die Standardabweichung beschreibt die Streuung der Einzelwerte in der Stichprobe,
d. h., sie wird als Kenngröße zur deskriptiven Statistik verwendet:

$$\bar{x} \pm SD$$

Standard Deviation

– Der Standardfehler beschreibt den erwarteten Fehler bei der Schätzung des Mittelwerts:

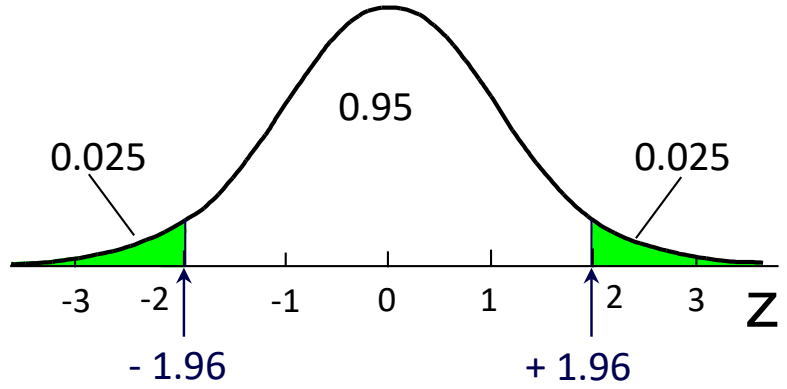
$$\bar{x} \pm S.E.M.$$

Standard Error of the Mean

Der S.E.M. wird zur Konstruktion von Konfidenzintervallen angewandt.
Ein Haupteinsatzgebiet ist die Schätzung von Effekten in klinischen Studien.

Konstruktion eines Konfidenzintervalls

- Kennt man die Verteilung der gesuchten Kenngröße, wählt man als Intervallgrenzen Quantile so, dass der Bereich außerhalb der Grenzen nur mit Wahrscheinlichkeit α auftritt.
- Bei symmetrischen Verteilungen (z. B. Normalverteilung) sind dies logischerweise symmetrische Quantile, also z. B.
  $z_{\frac{\alpha}{2}}$ und $z_{1-\frac{\alpha}{2}}$ (vgl. Grafik für $\alpha = 0{,}05$)
- Die Werte im Inneren des Intervalls kommen dann wie gefordert mit Wahrscheinlichkeit $1 - \alpha$ vor, hier also 95%.



Anmerkung: Die Normalverteilung eignet sich wegen des Zentralen Grenzwertsatzes (für große n) zur Konstruktion von Konfidenzintervallen

Konfidenzintervalle – allgemein

$(1-\alpha)$-Konfidenzintervalle für einen Parameter p einer symmetrischen Verteilung haben allgemein die Form:

- $$\left[\hat{p} - (1-\frac{\alpha}{2})-\text{Quantil} \cdot \begin{matrix}\text{Standard}-\\ f\text{ehler}\end{matrix}, \quad \hat{p} + (1-\frac{\alpha}{2})-\text{Quantil} \cdot \begin{matrix}\text{Standard}-\\ f\text{ehler}\end{matrix}\right],$$
- wobei $\hat{p}$ der Punktschätzer für p ist.
- Die mathematische Herausforderung ist, eine Formel für den Standardfehler zu finden oder ihn anderweitig zu berechnen.
- Soll das $(1-\alpha)$-Konfidenzintervall für ein (nicht zu schief verteiltes) Merkmal berechnet werden, unterscheiden wir den Fall
  - die Varianz des Merkmals in der GGS ist bekannt,
  - die Varianz ist unbekannt.

Fall 1: C.I. – Mittelwert (bekannte Varianz) I

- Das einfachste C.I. ergibt sich, falls die Varianz σ des betrachteten Merkmals bekannt ist. (Das ist in der Praxis eher der Ausnahmefall.)
- Dann ergibt sich folgendes Intervall:

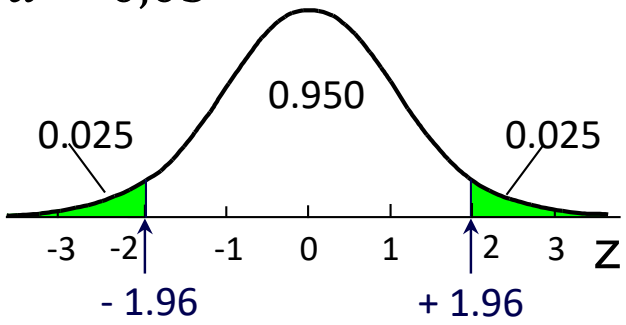
$$\bar{x} \pm z_{\left(1-\frac{\alpha}{2}\right)} \cdot S.E.M. \qquad mit \;\; S.E.M. = \frac{\sigma}{\sqrt{n}}$$

- wobei $z_{\left(1-\frac{\alpha}{2}\right)}$ das $\left(1-\frac{\alpha}{2}\right)$-Quantil der Standard-Normalverteilung ist (vgl. Folie 30)

Fall 1: bekannte Varianz – Beispiel I

z. B: Lipid- Screening

- HDL bei n = 20 Frauen erhoben:
- $\bar{x} = 52\ g/dl$
- $\sigma = 13{,}3\ g/dl$ sei bekannt
- $\alpha = 0{,}05$



$$\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

Tabelle 11.1: *Verteilungsfunktion $\phi(z)$ der standardisierten Normalverteilu*

	Werte der Standardnormalverteilung $\Phi(z)$						
z	.00	.01	.02	.03	.04	.05	.06
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750

- Es wird das $\left(1 - \frac{\alpha}{2}\right)$-Quantil (d. h., 97,5%-Quantil) benötigt.
- Man liest ab: → 1.9 → 6, also 1,96.

Fall 1: bekannte Varianz – Beispiel

$$\bar{x} - z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$n = 20; \alpha = 0{,}05$

$\bar{x} = 52 \frac{g}{dl}$

$\sigma = 13{,}3\ g/dl$ (bekannt)

Lipid-Screening

– (1-α) Konfidenzintervall (zweiseitig)

$$\bar{x} - 1{,}96 \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1{,}96 \cdot \frac{\sigma}{\sqrt{n}}$$

$$\underbrace{52 - 1{,}96 \cdot \frac{13{,}3}{\sqrt{20}}}_{= 46{,}2} \leq \mu \leq \underbrace{52 + 1{,}96 \cdot \frac{13{,}3}{\sqrt{20}}}_{= 57{,}8}$$

95%−C.I. $[46{,}2\ g/dl; 57{,}8\ g/dl]$

Fall 2: unbekannte Varianz

- Gegeben: Stichprobe des Umfangs n
- Schätzen des Mittelwertes: $\bar{x}$
- Schätzen der Varianz: $\frac{s_x^2}{n}$ Quadrat des S.E.M.

Weil uns nur die geschätzte Varianz zur Verfügung steht, dürfen wir nicht das Quantil der Standardnormalverteilung verwenden.

- Aber welche Verteilung liegt vor?

- W. Gosset (publizierte unter dem Pseudonym student) fand 1908 die Lösung:

  die t-Verteilung.

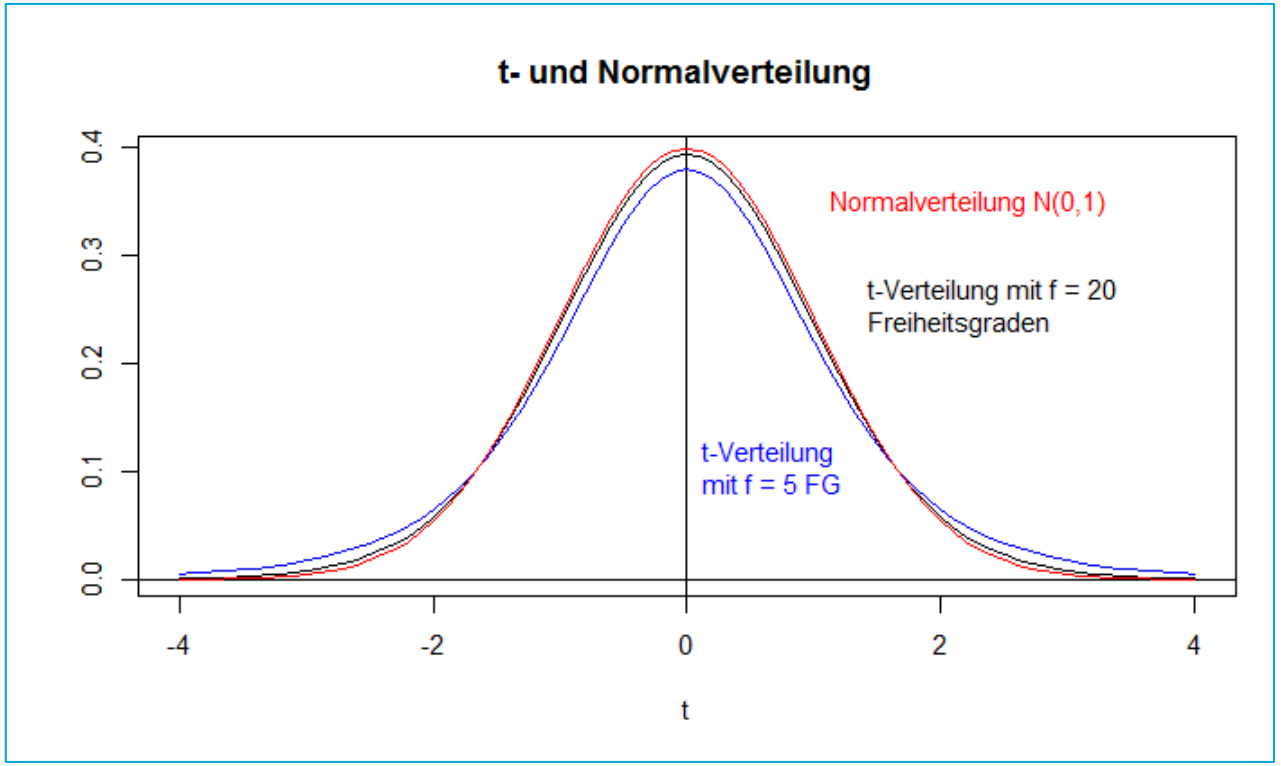
t – Verteilung I

- Die gesuchte Verteilung hat folgende Eigenschaften:
- Sie ist symmetrisch und eingipflig,
- sie soll die größere Unsicherheit bei unbekannter Varianz berücksichtigen, insbesondere bei kleinen Fallzahlen
- sie nähert sich bei guter Informationslage (d. h. *n* → groß) der Normalverteilung an.



https://de.wikipedia.org/wiki/William_Sealy_Gosset

t-Verteilung mit f FG



Die t-Verteilungsdichte

- ähnelt sehr der NV-Dichte,
- ist symmetrisch um 0,
- mit wachsender Zahl der FG f
  → sinkt die Streuung und
- nähert sich die Verteilung
  der Normalverteilung,
  - für $f = 20$ Freiheitsgrade kaum noch von der NV zu unterscheiden,
  - für $f = 100$ Freiheitsgrade fast identisch (→ ZGWS)

t – Verteilung: Quantile

Quantile der t-Verteilung

FG ... Freiheitsgrade

FG \ α	Irrtumswahrscheinlichkeit α für den zweiseitigen Test								
	0,50	0,20	0,10	0,05	0,02	0,01	0,002	0,001	0,0001
1	1,000	3,078	6,314	12,706	31,821	63,657	318,309	636,619	6366,198
2	0,816	1,886	2,920	4,303	6,965	9,925	22,327	31,598	99,992
3	0,765	1,638	2,353	3,182	4,541	5,841	10,214	12,924	28,000
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173	8,610	15,544
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893	6,869	11,178
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208	5,959	9,082
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785	5,408	7,885
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501	5,041	7,120
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297	4,781	6,594
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144	4,587	6,211
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025	4,437	5,921
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930	4,318	5,694
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852	4,221	5,513
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787	4,140	5,363
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733	4,073	5,239
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686	4,015	5,134
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646	3,965	5,044
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610	3,922	4,966
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579	3,883	4,897
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552	3,850	4,837
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527	3,819	4,784
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505	3,792	4,736
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485	3,767	4,693
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467	3,745	4,654

Skriptenheft

Tab. 11.3, S. 178



Gonick & Smith: Statistik in Cartoons

Fall 2: unbekannte Varianz

- Gegeben: $\bar{x},\ s,\ n$: aus Stichprobe Gesucht: zweiseitiges $(1-\alpha)$-Konfidenzintervall

$$\bar{x} - t_{\mathrm{f};1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\mathrm{f};1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$$

- Quantil der t-Verteilung
- f Freiheitsgrade: $f = n - 1$ Standardfehler (S.E.M.)

- Da der S.E.M. aus jeder Stichprobe neu bestimmt wird,
  unterscheiden sich die Breiten der Konfidenzintervalle verschiedener Stichproben

Fall 2: unbekannte Varianz - Beispiel

– Beispiel Lipid-Screening: HDL bei n = 20 Frauen erhoben:

$$\bar{x} = 52\ g/dl$$
$$s = 14\ g/dl$$
$$\alpha = 0{,}05$$

$$\bar{x} - t_{n-1;1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\mathrm{n}-1;1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}}$$

– $95\% - C.I.:$

– Anmerkung: Bei hinreichend großen Stichproben kann man für 95%-Konfidenzintervalle des Mittelwerts vereinfacht mit dem gerundeten Quantil $q = 2$ rechnen.

t-Quantil ablesen

Quantile der t-Verteilung

	Irrtumswahrscheinlichkeit α für den zweiseitigen Test								
FG \ α	0,50	0,20	0,10	0,05	0,02	0,01	0,002	0,001	0,0001
1	1,000	3,078	6,314	12,706	31,821	63,657	318,309	636,619	6366,198
2	0,816	1,886	2,920	4,303	6,965	9,925	22,327	31,598	99,992
3	0,765	1,638	2,353	3,182	4,541	5,841	10,214	12,924	28,000
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173	8,610	15,544
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893	6,869	11,178
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208	5,959	9,082
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785	5,408	7,885
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501	5,041	7,120
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297	4,781	6,594
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144	4,587	6,211
11	0,697	1,363	1,796	2,201	2,718	3,106	4,025	4,437	5,921
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930	4,318	5,694
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852	4,221	5,513
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787	4,140	5,363
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733	4,073	5,239
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686	4,015	5,134
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646	3,965	5,044
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610	3,922	4,966
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579	3,883	4,897
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552	3,850	4,837
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527	3,819	4,784
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505	3,792	4,736
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485	3,767	4,693
24	0,685	1,318	1,711	2,064	2,492	2,797	3,467	3,745	4,654

C.I. für den Mittelwert – Zusammenfassung

Allgemein: $(1-\alpha)$-C.I. für den Mittelwert (→ Folie 33)

$$\left[\bar{x} - (1-\tfrac{\alpha}{2})-\text{Quantil} \cdot \text{Std.}-\text{fehler} \leq \mu \leq \bar{x} + (1-\tfrac{\alpha}{2})-\text{Quantil} \cdot \text{Std.}-\text{fehler}\right]$$

1. Varianz σ^2 bekannt:

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

2. Varianz σ^2 unbekannt, d. h., mit geschätztem $\hat{\sigma}^2 = s^2$:

$$\left[\bar{x} - t_{n-1;\,1-\alpha/2} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1;1-\alpha/2} \cdot \frac{s}{\sqrt{n}}\right]$$



Gonick & Smith: Statistik in Cartoons

Anmerkung: für $n > 100$ ist: $t_{1-\frac{\alpha}{2},n-1} \approx z_{1-\frac{\alpha}{2}}$

Übungsaufgaben 6.1

- Sie haben ein 95%-Konfidenzintervall berechnet. Eine Kollegin verlangt ein 99%-Konfidenzintervall. Wird dieses
  ○ kleiner ○ größer oder ○ genauso breit wie das vorhandene sein?

- Sie wollen ein zweiseitiges 99%-Konfidenzintervall für den Mittelwert bei gegebener Varianz berechnen. Welches Quantil müssen Sie dafür heraussuchen?

- Es soll ein zweiseitiges 90%-Konfidenzintervall für den Mittelwert berechnet werden. Die Standardabweichung bestimmen Sie aus der Stichprobe von 101 Probanden. Welches Quantil verwenden Sie?

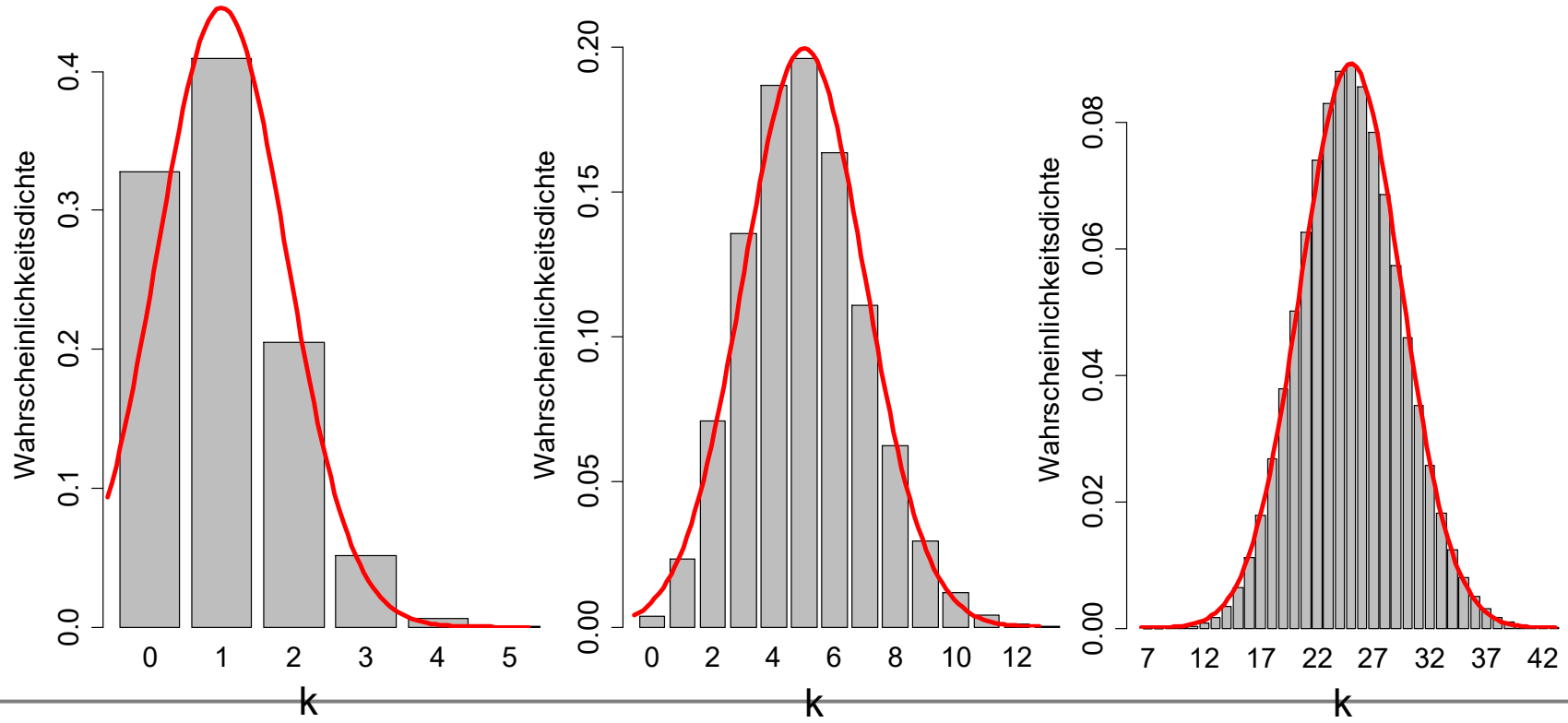
V 6

6.4 Konfidenzintervalle für Anteile

Rückblick: Zentraler Grenzwertsatz

– Für $n \to \infty$ geht die Binomial- in die Normalverteilung über:

$p = 0{,}2;\ n = 5$ $p = 0{,}2;\ n = 25$ $p = 0{,}2;\ n = 125$



Zielstellung

- Szenario: Die Wahrscheinlichkeit p eines Ereignisses in einer Grundgesamtheit soll mit Konfidenzintervall geschätzt werden
- Gegeben
  - Stichprobe mit n Probanden
  - x haben ein Ereignis

Einfachste Lösung: Für große n wird die Binomialverteilung durch eine Normalverteilung approximiert (vgl. VL 5)

- Ist $X \sim B(n, p)$, ist
- $\frac{x}{n} \sim N(\hat{p}, \frac{\hat{p} \cdot (1-\hat{p})}{n})$ approximativ normalverteilt (de Moivre 1730 / Laplace 1812)

Konfidenzintervall für Anteile

- Wir schätzen die Wahrscheinlichkeit *p* mit der relativen Häufigkeit $\hat{p} = \frac{x}{n}$ (Punktschätzer)
  - Anzahl der Erfolge: *x*
  - Umfang der Stichprobe: *n*

$$\hat{p} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

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

Standardabweichung von $\hat{p}$

- Voraussetzung: $n \cdot \hat{p}(1-\hat{p}) \geq 9$
- Ist diese Bedingung erfüllt, sind die Abweichungen von der Normalverteilung klein genug.

Konfidenzintervall für Anteile

Beispiel: neue Therapie für 100 Tumorpatienten, Ansprechen bei 60 Patienten

Gesucht: Punktschätzer für die Wahrscheinlichkeit des Ansprechens mit 95%-C.I.

- Punktschätzer für die Erfolgswahrscheinlichkeit p: $\hat{p} = \frac{x}{n} = \frac{60}{100} = 0{,}6 = 60\%$
  (relative Häufigkeit)
- Prüfung der Bedingung $n \cdot \hat{p}(1-\hat{p}) \geq 9$:
- Konfidenzintervall: $\hat{p} - z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Konfidenzintervall – Diskussion

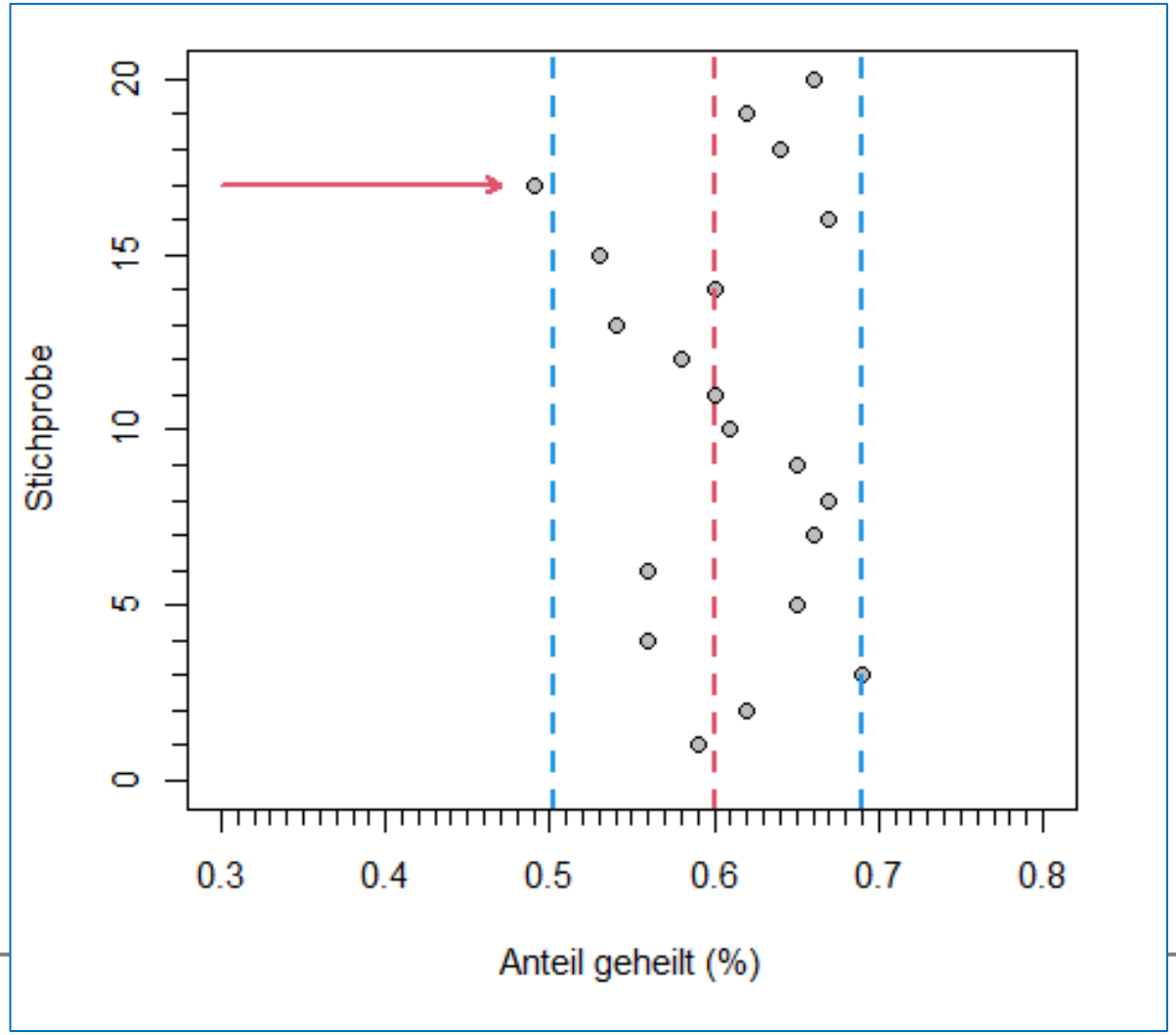
- Die Grenzen des 95%-Konfidenzintervalls sind so gewählt, dass das zufällige Intervall mit ca. 95%-iger Wahrscheinlichkeit den wahren Wert überdeckt.
- Bei vielfacher Wiederholung des Versuches (Long Run) werden 95% der so berechneten Konfidenzintervalle den *wahren Wert* überdecken.
- Es gibt verschieden gute Methoden, das Konfidenzintervall für Wahrscheinlichkeiten zu berechnen (vgl. Newcombe RG, Stat Med 1998;17:857-872)
  - Verschiedene Methoden liefern inkonsistente Ergebnisse, die Approximation durch die NV liefert z. B. gelegentlich Intervallgrenzen über 0 oder 1 hinaus.
  - Diese Methode gilt als antikonservativ (Newcombe 1998).
  - Empfehlenswert ist die Score-Methode nach Wilson (JASA 1927;22:209-212).

Äquivalenz Konfidenzintervall – Test

- 20 Stichproben ($n = 100$)
- geschätzte Wahrscheinlichkeit

Fällt der Schätzwert aus der Stichprobe außerhalb des 95%-Konfidenzintervalls für p, ist dies nicht kompatibel mit der Hypothese $p = 0{,}6$.

Das entspräche einem signifikanten Test zum Signifikanzniveau $\alpha = 0{,}05$.



Anwendungsaufgabe

Beispiel 6.5: Innovatives Diagnoseverfahren

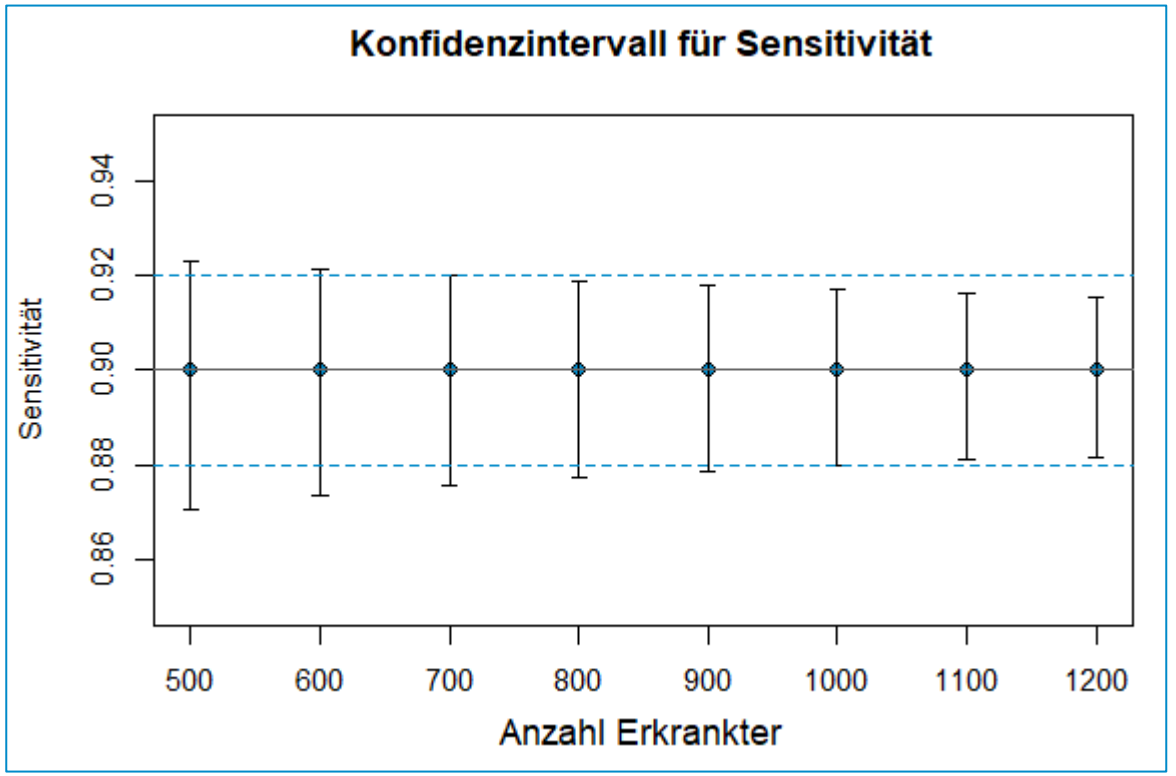


Mit einer diagnostischen Studie am UKL soll die Sensitivität eines neuen Diagnoseverfahrens auf 2% genau bestimmt werden.

- von jährlich 1 000 Patienten werden ca. 600 Erkrankte erwartet
- Vermutung: $Sensitivität \approx 90\%$
- Analysevorschlag: Punktschätzung mit 95%-Konfidenzintervall der Breite 4 %
- Für die vorgesehene Fallzahl errechnete der Biometriker eine Intervallbreite von etwa 4,8%.

Aufgabe: Wie kann die gewünschte Genauigkeit erreicht werden?

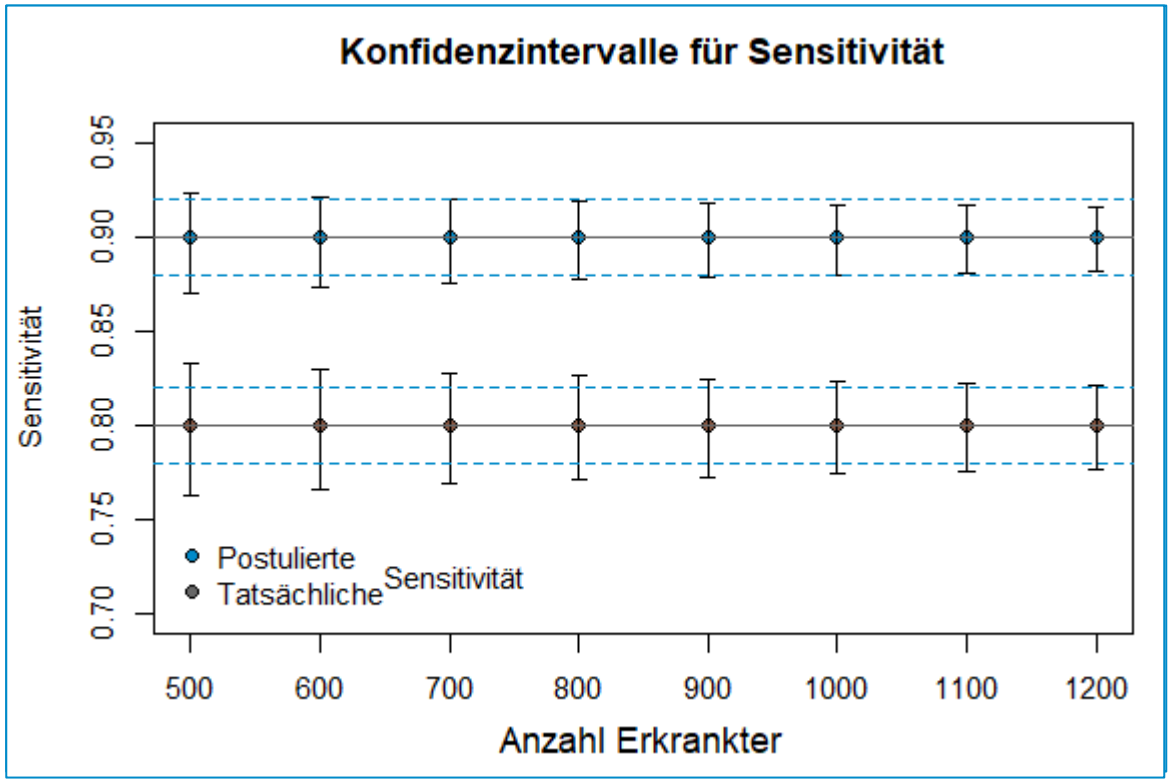
Lösung 1



Hinweise:

1. Da es bei der Sensitivität um Erkrankte geht, kommt eine weitere Planungs-Unsicherheit hinzu: Es ist unklar, wie viele im Zeitraum überhaupt als Erkrankte in die Tabelle eingehen!
2. Die Breite 4% des Konfidenzintervalls wird bereits für 1 000 Erkrankte erreicht. Die untere Grenze des Intervalls läge aber mit 87,98% unter den angestrebten 88%.

Lösung 2



Hinweise:

3. Die Planung ist dennoch unsicher, da offen bleibt, ob der Zielwert 90% erreicht wird. Wäre die tatsächliche Sensitivität z. B. nur 80%, würden die Konfidenzintervalle deutlich breiter!
4. Zielt man auf eine Mindest-Sensitivität von 90%, könnte man auch einseitige Konfidenzintervalle berechnen

Einseitiges Konfidenzintervall



Im Beispiel B 6.5 wäre eine Mindest-Sensitivität sinnvoll. Dann ist ein einseitiges Konfidenz-intervall adäquat.

Aufgabe: Wie groß ist das einseitige 95%-Konfidenzintervall für die Sensitivität, wenn von $n = 1\,100$ Erkrankten $x = 1\,008$ einen positiven Test hatten?

Ergebnis:

- Punktschätzer: $\widehat{Sens} = \frac{1008}{1100} = 0{,}9013 \ldots \approx 90\%$
- Die Methode nach Clopper & Pearson liefert das einseitige 95%-Intervall

  [90,1%; 100%]

Konfidenzintervall – Zusammenfassung

Die Breite eines Konfidenzintervalls ist abhängig von:

- der Anzahl der Beobachtungen in der Stichprobe,
- der Varianz des betrachteten Merkmals
- dem Konfidenzniveau α

Ergänzen Sie diese drei Sätze:

- Erhöht man den Stichprobenumfang, wird das Konfidenzintervall …
- Ist die Varianz größer (als erwartet), wird das Konfidenzintervall …
- Erhöht man das Konfidenzniveau, wird das Konfidenzintervall …



UNIVERSITÄT LEIPZIG

Medizinische Fakultät

VIELEN DANK!

Dr. rer. nat. Meinhard Mende

Senior-Wissenschaftler, Institut für Medizinische Informatik, Statistik und Epidemiologie und Zentrum für Klinische Studien

Härtelstraße 16-18, 04107 Leipzig

Meinhard.mende@zks.uni-leipzig.de