Wo liegen die Knackpunkte von stochastischen Begriffen und Modellen?

Manfred Borovcnik, Institut für Statistik, Universität Klagenfurt

1. *xy* Jahre Didaktik der Stochastik – Bilanz 2024

---

2. Evidence-based society
3. Grundsätzliche Missverständnisse
4. Knackpunkte in der Wahrscheinlichkeitsrechnung
5. Knackpunkte in der beurteilenden Statistik
6. Knackpunkte in der beschreibenden Statistik
7. Zum guten Schluss

1. *xy* Jahre Didaktik der Stochastik – Eine „Bilanz“

Unterricht in Stochastik auf allen Ebenen scheint zu verpuffen, so er überhaupt stattfindet. Na ja, seit der Zentralmatura geht das nicht mehr.
Der öffentliche Umgang mit stochastischen Begriffen und Fragestellungen, aber auch mit Daten ist beklagenswert.

Woher die Ablehnung?
Woher dies Unwissen?
Was ist so schwierig an der Stochastik?
Wie überzeugt man Menschen von den Chancen dieser Methoden?

Wo liegen – mit einem Wort – die Knackpunkte?

Nachsatz: Gibt es im Vergleich zur „Rest“mathematik Unterschiede im Wissen, den Fertigkeiten, den Einstellungen, der Ablehnung etc.?

1.1 Ein paar pessimistische Gedanken im Jahr 2024

Brexit-Referendum: Die meisten Institute sagten ein knappes pro-EU Ergebnis voraus, am Ende stand es 51,9% für den Brexit.

COVID 19: Die Vorhersagen für die Verbreitung und die Wirksamkeit der Maßnahmen lagen ziemlich daneben.

Die US-Wahlen 2016: Trump als Außenseiter wurden in den Polls gegen H. Clinton kaum Chancen eingeräumt.

Kann man stochastische Methoden wirklich besser nützen?

Gibt es einen Unterschied im Erfolg zwischen einzelnen Personen und größeren Stakeholdern wie größere Firmen, oder Interessensgruppen oder der Staat selbst?

Mein Vortrag ist als Anregung zur kritischen Diskussion gedacht.

Vielleicht finden wir die Knackpunkte?

1.2 Spezifische Schwierigkeiten der Stochastik

Wie soll ein durchschnittlicher Schüler in die Thematik eintauchen können, wenn alle gescheiten Köpfe ... die Regelmäßigkeiten fallweise nicht einmal geglaubt haben?

(Cardano: In sechs Würfen mit einem Würfel kommen alle sechs Zahlen – dass dies nicht so ist, ist Schicksal).

Gemein ist auch, dass die *Stochastik über den einzelnen, aktuellen Fall nichts sagen kann.* Mein Schwiegerpapa sollte sich einer Routineoperation unterziehen, die seit 18 Jahren nicht misslungen war. Bei ihm ging es schief.

Kinder können mit solchen Unsicherheiten nichts anfangen, bestimmt nicht, wenn sie den momentanen Unterricht genießen – die müssen doch denken, dass das alles nur erfunden ist.

2. Evidence-based society

Erkenntnisse werden aus Daten „geschöpft".

„Signifikante" Ergebnisse werden als Beweise gehandelt, anstatt sie zum Anlass für substanzwissenschaftliche Forschungen zu nehmen.

Daten werden als Fakten missverstanden. Daten sind aber nur

1. im Vergleich interpretierbar;
2. nur aussagekräftig, wenn sie Bedingungen im Entstehen genügen.

Flops in Projekten der Angewandten Statistik sind meist in Versäumnissen in der Systemanalyse (VOR der Datenproduktion) begründet.

- Definition von Merkmalen;
- Abgrenzung von Zielgesamtheit;
- Studium von Confoundern (Störgrößen) etc.

„Trau keiner Statistik, die du nicht selbst gefälscht hast!"

2.1 Umgang mit Daten

In Lehrbüchern: Manipulation die (x+1)te, ohne die Struktur von Manipulationstechniken aufzuarbeiten. Wenig verwunderlich daher:

Noch immer werden aus Zeitreihen „beliebige“ Zeitpunkte als Referenz herangezogen anstatt aktuelle Daten mit langfristiger Variabilität (und Saisonalität) zu evaluieren.

Das führt zu Aktionismus & Heraufbeschwören von „Kausalketten“: die Maßnahmen waren erfolgreich; es müssen Maßnahmen gesetzt ...

Die Definition von Merkmalen wird mehrfach abgeändert. Dennoch werden die Zahlen verglichen und ihr Niveau interpretiert: Arbeitslosigkeit; CT-Werte bei Covid-Tests ein jüngeres Beispiel.

Andere Daten sind nicht verfügbar: So werden Pensionsreformen OHNE Datenbasis diskutiert und wichtige Entscheidungen getroffen.

2.2 Evidence-based society – Ein Beispiel aus Österreich

Man sucht sich immer die günstigere Vergleichsbasis:

In Österreich haben wir 2020 ein „Budgetloch“ von € 30 Mrd. ($30 \cdot 10^9$); es ist 2023 auf € 12·Mrd. geschrumpft? Das entspricht 2,7% des BIP. Allerdings sind die Staatsschulden auf 369 Mrd. angestiegen (Vergleich 2019: 280 Mrd.).

Wenn die Daten zur Gegenwart Entscheidungen begründen sollen, so löst man die Definitionsprobleme nicht:

Wir haben daher keine geeignete Bilanz über die Schulden des Staates, weil wir uns nicht einigen können, was als Schulden zu zählen hat.

Evidence-based society!

3. Grundsätzliche Missverständnisse

Vorausblick

3.1 Daten sind Fakten.

3.2 Rivalisierende Konzepte & Vorstellungen sind irrational.

3.3 Perspektive des Individuums – Perspektive von oben.

3.4 Wahrscheinlichkeit ist eine Eigenschaft von Objekten.

3.5 Kleine Wahrscheinlichkeiten sind „Business as usual“.

3.6 Axiomatik rechtfertigt eine Konzeption von Wahrscheinlichkeit.

3.7 Risiko und Entscheidungen.

3.8 Reichweite des Ansatzes und seiner Methoden.

3.9 Objektiver Stellenwert von Modellen.

3. Grundsätzliche Missverständnisse

3.1 Daten sind Fakten

Von wegen.

Das bedeutet aber auch keineswegs, dass man sie „verhandeln“ kann.

Der Mythos von der Repräsentativität von Daten ist Geschäftsgrundlage der Stochastik.

Tatsächliche Relevanz von Daten jedoch ist schwer zu erreichen und noch schwieriger zu überprüfen.

Es hilft auch nicht, zu sagen, das sind dann eben keine Daten.

Daten existieren und werden – unabhängig von ihrer Qualität – als Argumentationsgrundlage verwendet

3.2 Rivalisierende Konzepte & Vorstellungen sind irrational

Magie des Zufalls hat mehr Faszination als abstrakte Begriffe und wird zu Unrecht abqualifiziert.

Das beginnt mit der Abwertung von Laplacens Dämon, der alles weiß und Wahrscheinlichkeit dem Ignoranten überlässt;

es endet nicht mit der Einschätzung von Altersstarrsinn, mit welchem sich Einstein zum Zufall in der Physik äußert: „Gott würfelt nicht".

Die Leute spielen in Glücksspielen und Wetten aus anderen Gründen und gelegentlich durchaus gegen den Rat von Stochastikern ganz gut.

Das Glück (Schicksal) herauszufordern entspricht archetypischen Entwicklungsmustern und stellt eine wichtige Triebfeder für Fortschritt dar.

Da muss man die Stochastik aber auch richtig positionieren:

Was sie leistet und was sie nicht kann, was andere Ansätze leisten.

3.3 Perspektive des Individuums – Perspektive von oben

Für das System bietet Stochastik viel eher gute Lösungen als für ein Individuum. One-off gegen long run-Perspektive.

Die Frage „Ist das für mich relevant?“ ist berechtigt.

Im klassischen Versicherungsvertrag ist die Frage berechtigt,
„Wie kann etwas gut für die Versicherung sein und dann auch noch gut für mich?“

Ein Arzt ist dem System verpflichtet und auch juristisch haftbar.
Der Patient hat an den Folgen von Entscheidungen direkt zu tragen.

Die Frage bezüglich der Perspektive kann man auf jedes Glücksspiel, auf das Wettgeschäft, aber auch auf den Finanzmarkt beziehen.

Die Lösung der Black-Scholes-Gleichung ermöglicht die Bepreisung von Optionen und anderen Futures auf dem Finanzmarkt;
dieselben Methoden liefern aber keineswegs so zuverlässige Prognosen über den Ablauf einzelner Geschäfte.

3.4 Wahrscheinlichkeit ist eine Eigenschaft von Objekten

Wahrscheinlichkeiten existieren. Man muss sie nur finden & geeignet messen.

Wahrscheinlichkeiten können objektgebunden, subjektgebunden oder Eigenschaften von Szenarien sein.

Szenarien sind anders als Modelle. Modelle sollen möglichst gut passen. Szenarien spielt man durch, um Möglichkeiten abzuschätzen.

Eingangsgrößen in vielen Anwendungen werden keineswegs aus Experimenten geschätzt (Spezifität, Sensitivität bei medizinischen Tests etwa; Prävalenz von BSE oder COVID), sondern in Form von Szenarien *angenommen*.

- Bei subjektivistischen Ws. ist die Tätigkeit der Revision von Urteilen in Form von bedingten Wahrscheinlichkeiten wichtig.
- Bei objektivistischen Ws. die Schätzung durch relative Häufigkeiten.

Der Stellenwert von bedingter Wahrscheinlichkeit ist entsprechend unterschiedlich.

3.5 Kleine Wahrscheinlichkeiten sind „Business as usual“

Kleine Wahrscheinlichkeiten werden so interpretiert, als hätte man zuverlässige Schätzungen darüber.

Etwa aus Analysen von Netzwerken, welche mit Hilfe von *Unabhängigkeitsannahmen* über die Einzelteile berechnet werden.

Aus Daten könnte man sie nie und nimmer schätzen.

Historisch gab es bis in die Mitte des 20. Jh. Vorschläge, Wahrscheinlichkeiten unterhalb einer *moralischen* Grenze überhaupt wegzulassen.

Nehmen wir an, es gehe um eine Ws. von 10^{-4}. Und nehmen wir an, es gäbe eine *Zufalls*stichprobe vom Umfang 10.000. Dann gibt es folgende Wahrscheinlichkeiten für folgende Schätzwerte:

Schätzwert	0	0,0001	0,0002	0,0003	0,0004
Ws. %	36,8	36,8	18,4	6,3	1,8

3.6 Axiomatik rechtfertigt *eine* Konzeption von Ws.

Wahrscheinlichkeit wird reduziert auf relative Häufigkeiten

Hier wird die Möglichkeit einer Interpretation mit der Ausschließlichkeit verwechselt.

Der Fokus auf die relativen Häufigkeiten und deren Entwicklung hat eine negative Auswirkung auf das intuitive Verständnis.

- Statt zu klären, dass Muster der Entwicklung der rel. Häuf. keine Rolle spielen, lässt sie diese Muster geradezu wichtig werden.
- Bedingte Wahrscheinlichkeiten spielen eine Hilfsrolle und können vernachlässig werden.

Es gibt auch ein Axiomensystem für Präferenzen, in welchem bedingte Wahrscheinlichkeiten eine Schlüsselrolle übernehmen.

Die wesentlichen Kenngrößen von statistischen Entscheidungen sind bedingte Wahrscheinlichkeiten.

3.7 Risiko und Entscheidungen

Sprache ist verräterisch:

- Das Risiko gehe ich nicht ein.
  Das maximal Unangenehme, den schlimmsten Fall abwehren. Die Maßnahmen zu BSE in den 2000ern und die jüngeren COVID-Maßnahmen sind nur zwei Beispiele dafür.
- No risk – no fun.
  Die Macht des Schicksals herausfordern. Eine große Triebfeder.
  Schon Kahneman & Tversky zeigen, dass hoher Impakt (Folgekosten, Nutzen) die Perzeption von (kleinen) Wahrscheinlichkeiten stört.

Statt in einem Szenario relative Risiken gegeneinander abzuwägen (weigh the evidence) und Entscheidungskriterien offenlegen, wird irrational …

„Risiko“ ist auch für Wahrscheinlichkeiten von Fehlentscheidungen (bei statistischen Tests) in Gebrauch.

Elementarisieren statistischer Tests durch Weglassen der Zusammenhänge (Typ I und II) sowie ein Transfer auf die Datenebene – durch Simulation – *zerstört* die Begriffe.

3.8 Reichweite des Ansatzes und seiner Methoden

Stochastische Modelle können immer eingesetzt werden

Es gibt einen normalen „Geschäftsverlauf" und Brüche in Entwicklungen. Bei diesen Brüchen versagen stochastische Methoden. Da hilft viel mehr Wissen über den Kontext oder, noch besser, Insiderwissen.

Korrelation oder Assoziation hat (etwas) mit Kausalität zu tun

Damit erspart man sich substanzwissenschaftliche Forschung über Wirkmechanismen. Dabei sind signifikante Korrelationen nur ein Anlass, ernsthaft über sachliche Zusammenhänge zu forschen.

ABC-Ansatz

Die wichtigsten Dinge zuerst erledigen. Statistik ist auf Level *C*.

Experten sind der Wahrheit verpflichtet

Experten haben natürlich Interessen und sind auch ganz anders von Entscheidungen betroffen (etwa sind sie juristisch haftbar).

3.9 Objektiver Stellenwert von Modellen

Es gibt ein – bestes – Modell

Vielmehr werden stochastische Modelle oft parallel durchgerechnet um mehrere plausible Szenarien für die Entwicklung durchzuspielen. Das ist auch der Standardfall in der Beurteilenden Statistik.

Man kann Modelle wissenschaftlich (objektiv) evaluieren

Dazu sind statistische Tests ja da.

Mitnichten: die „Rationalitätskriterien“ versagen für solche Tests, die dadurch zu bloßen Plausibilitätsprüfungen degradiert werden.

Viel *zielführender* sind hier *fundamentale Ideen*,

welche die jeweiligen Modelle beschreiben und die auch zum Schlüssel für die Anwendung oder Verwerfung von bestimmten Modellen für spezielle Situationen werden.

4. Knackpunkte in der Wahrscheinlichkeitsrechnung

Vorausblick

4.1 Stellung der Gleichwahrscheinlichkeit

4.2 Wozu Wahrscheinlichkeiten?

4.3 Kleine Wahrscheinlichkeiten

4.4 Information in Wahrscheinlichkeitsverteilungen

4. Knackpunkte in der Wahrscheinlichkeitsrechnung

4.1 Stellung der Gleichwahrscheinlichkeit

- Ist extrem wichtig.
- Ist nur im Einklang mit der Deutung als relative Häufigkeit und mit Ws. als Chancenverhältnissen (subjektiver Grad des Vertrauens) sinnvoll. Gleichwahrscheinlichkeit kann nur über relative Häufigkeiten erfahren werden – Ws. auf lange Sicht.
- Gleichwahrscheinlichkeit kann zur Kalibrierung des subjektiven Grads des Vertrauens in eine unsichere Sache dienen – Ws. für die Einzelentscheidung.
- Gleichwahrscheinlichkeit kann zur Hypothesengenerierung dienen – Varga (1983) untersucht bereits im Anfangsunterricht – vor der Ws. – ob ein Münzwurfprotokoll echt oder ausgedacht ist (und führt intuitiv zum Signifikanztest).

Es ist notwendig, einen möglichst allgemeinen Ws.begriff zu erarbeiten und den equiprobability bias hinterfragen zu lernen.

4.2 Wozu Wahrscheinlichkeiten?

P(*E*) = ½, P(*E*) = 1/6, was soll das bedeuten?
Wozu kann eine solche Aussage dienen?

Man kann damit Wetteinsätze bemessen: 1:1 bzw. 5:1 (netto).
In Serien kann aber alles passieren, das Feedback ist nur bedingt verwertbar.

Jemand, der diese Aussagen in vielen Situationen anwenden kann (GgZ),
kann damit ein Risiko weitgehend ausschalten,
jemand, der nur eine Einzelentscheidung trifft, trägt das volle Risiko.

Stakeholder der einen oder anderen Seite können, ja müssen ganz andere Entscheidungskriterien verfolgen.

B: Entscheidung über die Auflagenhöhe einer Zeitschrift
B: Abschluss einer Kasko-Versicherung.

Optimieren einer Entscheidungssituation

Die Nachfrage N nach einer Zeitschrift sei – der Einfachheit halber – so modelliert. Der Preis pro Heft beträgt 1,60. Die Kosten abhängig von der Auflage betragen $C(a_j)$.

Man kann 1.000, 2.000, ... , 5.000 auflegen. Wie viele soll man?

N	1.000	2.000	3.000	4.000	5.000
$p(n_i)$	0,40	0,30	0,20	0,06	0,04
Auflage	1.000	2.000	3.000	4.000	5.000
$C(a_j)$	2.000	2.200	2.400	2.600	2.800

Wir vergleichen die Optionen; als Kriterium für die Entscheidung nehmen wir den erwarteten Profit (Nutzen).

Vergleichen der zukünftigen Entwicklung

Profitmatrix U		Entscheidung: Auflagenhöhe a_j					Preis/Heft
		1.000	2.000	3.000	4.000	5.000	1,6
	Kosten C (a_j)	2.000	2.200	2.400	2.600	2.800	
	n_i						Ws (n_i)
Nachfrage	1.000	-400	-600	-800	-1.000	-1.200	0,40
	2.000	-400	1.000	800	600	400	0,30
	3.000	-400	1.000	2.400	2.200	2.000	0,20
	4.000	-400	1.000	2.400	3.800	3.600	0,06
	5.000	-400	1.000	2.400	3.800	5.200	0,04
	E(Profit | a_j)	-400	360	640	600	464	
	Maximaler Verlust (a_j)	-400	-600	-800	-1.000	-1.200	

Entscheidung hängt vom verwendeten Kriterium ab.

Der max. Verlust führt – nicht nur hier – zu unsinnigen Entscheidungen.

Man könnte auch in Nutzen statt Geld rechnen.

Der Versicherungskontrakt

Bei einer Wette auf Ereignis E stehen die Chancen $p:q$. Odds sind das Verhältnis von Wahrscheinlichkeit des Eintretens zum Ausbleiben von E.

a) Zeigen Sie, dass gilt: $P(E)=\dfrac{p}{p+q}$.

b) Sie hätten eine Kaskoversicherung für Ihr neues Auto abzuschließen. Unterscheiden wir der Einfachheit halber nur die Fälle unfallfrei und Totalschaden (mit € 20.000 Schaden).
Schätzen Sie „Ihre“ Wahrscheinlichkeiten für ein Jahr und bestimmen Sie daraus den Wert der Entscheidung keine Versicherung abschließen; vergleichen Sie mit den Kosten der Versicherung (€ 1.000 Prämie). Sollen Sie die Versicherung abschließen?

c) Bestimmen Sie in b) jene Odds für einen Totalschaden, bei denen die Entscheidung zwischen Versicherung ja bzw. nein „kippt“.

Abtausch von Unsicherheit gegen Geld

Kosten [in €]		Entscheidung	
		A_1 = Versicherung ja	A_2 = keine Versicherung
Mögliche Zukunft	T_1 = kein Unfall	1.000	0
	T_2 = Totalschaden	1.000	20.000

Im Nachhinein ist jeder gescheiter: kein Unfall, keine Versicherung!

Minimiert man maximale Kosten, so ist A_1 = Versicherung besser.

Ist man bereit die zukünftigen Ereignisse T_1 und T_2 zu gewichten, etwa durch relative Gewichte von 39 : 1, dann sind die “erwarteten” Kosten für A_2 auf 500 gesunken.

Die Entscheidung hängt von den Gewichten ab.

Welche Gewichte? Es gibt einen “break-even point”: bei 19 zu 1.

4.3 Kleine Wahrscheinlichkeiten

Machen große Probleme, sind aber heute nicht mehr wegzudenken.

Noch in den 1950er Jahren haben namhafte Mathematiker (E. Borel) vorgeschlagen, Ws.en unterhalb eines Schwellenwerts Null zu setzen – historisch lief das unter dem Begriff moralische Wahrscheinlichkeit.

Kleine Ws. können empirisch nicht validiert werden, es gibt einfach keine zuverlässige Information darüber.

Wenn überhaupt, so kann man sie in Form von Szenario-Größen verwenden – auf der Basis, was wäre, wenn …
Überdies können solche Szenarien nur im VERGLEICH interpretiert werden und nicht in ihren absoluten Größen.

B: Prävalenz von BSE oder von Covid. Die Prävalenz ist nicht nur orts-, sondern, im Verlauf des Infektionsgeschehens – auch zeitabhängig.

Nicht ganz so kleine Wsn. sind schwer von den Folge“kosten“ zu unterscheiden und führen sehr häufig zu völligen Fehleinschätzungen.

4.4 Information in Wahrscheinlichkeitsverteilungen 1

Daten zu Wahrscheinlichkeitsverteilungen schwanken viel mehr als man sich vorstellen kann.

Es ist instruktiv, sich bestimmte Verteilungen simulieren zu lassen, um annähernd ein Gefühl für die Schwankungen zu bekommen.
Das macht man am bestem im statistischen Labor,
ansonsten ist man allzu sehr geneigt, kontextgebundene Erklärungen für diese Schwankungen zu suchen.

Weder eine statistische noch eine probabilistische Herangehensweise: Verteilungen nach kleinsten Anomalien & Besonderheiten zu untersuchen.

Wichtig ist der TYP der Verteilung.
Eingipfelig, symmetrisch, schief etc.

Wo liegen „extreme“ und wo „normale“ Werte? Zufallsschwankungen
Wo liegen die Trennpunkte zwischen diesen Werten? Was haben diese Thresholds zu bedeuten? – Modellkritik, Entscheidungen (geht in Richtung Signifikanztest).

Information in Wahrscheinlichkeitsverteilungen 2

TYP der Verteilung.

Wo liegen „extreme“ und wo „normale“ Werte?

Reduktion auf Erwartungswert und Standardabweichung und z-Werte

Für buchhalterische Zwecke (Gesetz der großen Zahlen).

Reduktion besonders einsichtig, wenn die Verteilung ausgeprägt eingipflig & im Wesentlichen symmetrisch ist ⇨ z-Werte spiegeln absolutes Ranking.

Idee hinter der Verteilung, welche zur Modellierung hilfreich ist

Gleichverteilung – Fairness, Symmetrie

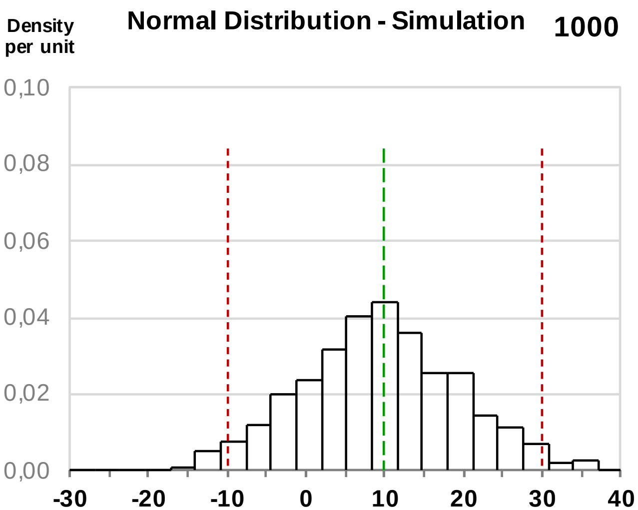
Binomialverteilung – Unabhängige Wiederholung desselben 0-1 Experiments

Poisson- & Exponentialverteilung als Zählen von bzw. Warten auf Ereignisse im völlig zufälligen Entstehen von Ereignissen.

Normalverteilung als Folge der Elementarfehlerhypothese (ZGS)

Diese rechtfertigt die **Normalverteilung als Näherung für fast alle Schätzungen von Parametern aus Stichproben** (= Summen).

Kleine Risken, die Schwellenwerte von 2 s zu überschreiten

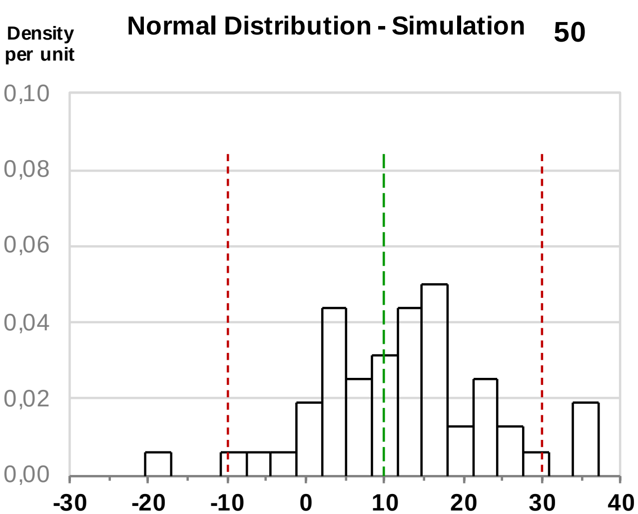


Erstens:

Man beachte die riesige Variabiltiät in der Gestalt

Zweitens

Stichproben-umfang	**% der Daten innerhalb ∝ - 2s und ∝ + 2s:**
1000	**94,9%**
50	**88,0%**



Man sieht:

ca. 95% der Daten liegen innerhalb von 2 Standardabweichungen vom Mittel, egal welche Werte die Parameter haben!

Wenn s klein ist, dann kann schon ein Datum die Lage der Verteilung (Populaton) anzeigen!

Es gibt ein kleines Risiko, dsss die Entfernung größer als zwei Standardabweichungen ist.

Drittens

Für Mittelwerte ist die Normalverteilung approximativ immer gut!

5. Knackpunkte in der beurteilenden Statistik

Vorausblick

5.1 Die Stichprobenverteilung für den Mittelwert

5.2 Konfidenzintervalle

5.3 Hypothesentests – Filtern mit Vorbedingungen

5.4 Hypothesentests ein Problemkind der Wissenschaften

5. Knackpunkte in der beurteilenden Statistik

5.1 Die Stichprobenverteilung – für den Mittelwert

Warum nimmt man Mittelwerte (oder andere Parameter) aus einer Stichprobe als Stellvertreter für die gesamte Verteilung?

Warum ist die Normalverteilung so wichtig?

Wichtiger als reale stochastische Phänomene X durch die Normalverteilung zu beschreiben, ist es, Normalverteilung als approximative Verteilung für eine artifizielle Zufallsvariable zu nutzen.

Wir können die Verteilung von X beobachten.

Wir erhalten so die Daten einer Stichprobe: $x_1, x_2, \ldots, x_n$.

Wir berechnen aus der Stichprobe den Mittelwert $\bar{x}$.

Davon haben wir einen *einzigen* Wert.

Da der Mittelwert μ der Verteilung von X – in aller Regel – unbekannt ist, müssen wir uns eine Vergleichsbasis *künstlich* schaffen.

Bestimmung der Stichprobenverteilung

Wir studieren – durch Mathematik oder Simulation im statistischen Labor, wie die Verteilung des realen Phänomens X mit der Verteilung des Mittelwerts $\overline{X}$ einer Stichprobe zusammenhängt.

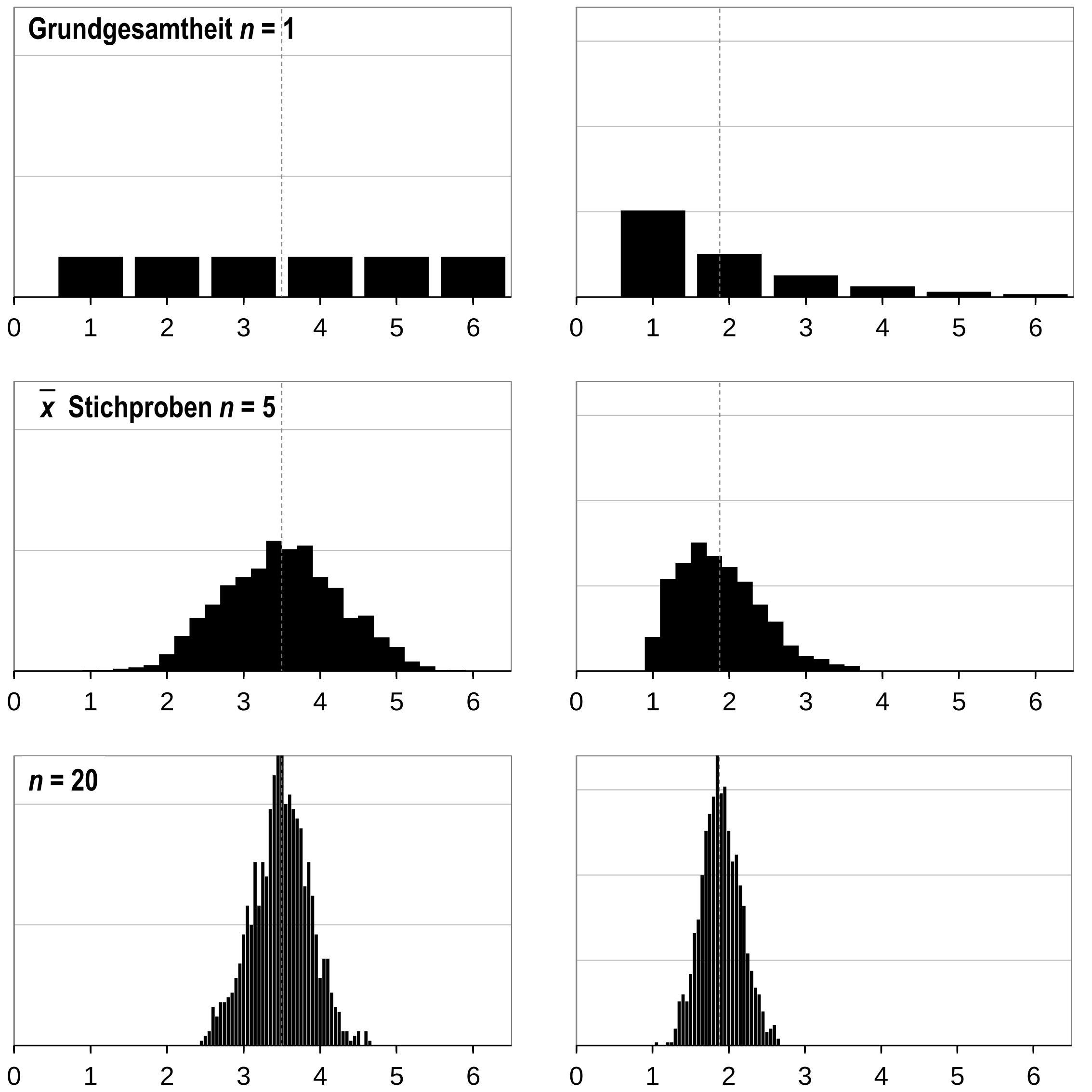
Wir sehen zwei wesentliche Eigenschaften:

1. Die Verteilung der Mittelwerte $\overline{X}$ ist annähernd durch eine Normalverteilung zu beschreiben (ZGS). Es genügt daher, ihren Mittelwert sowie ihre Standardabweichung zu kennen.
2. Die Parameter der Verteilung der Mittelwerte $\overline{X}$ hängen von Mittelwert und Standardabweichung der Ausgangsverteilung X ab:

$$E(\overline{X}) = \mu \quad \text{und} \quad \sigma(\overline{X}) = \sigma/\sqrt{n}$$

Man beachte: Während wir von X immer über die Stichprobe mehrere Werte beobachten, haben wir von $\overline{X}$ immer nur einen einzigen Wert. Damit bekommt man eine artifizielle Verteilung, mit der wir den beobachteten Wert von $\overline{X}$ vergleichen können. Ganz im Sinne der Schwellenwertüberlegung.

Zentraler Grenzverteilungssatz - Stichprobenverteilung



Die Verteilung der Mittelwerte normalisiert sich sehr rasch.

Das macht die Bedeutung der Normalverteilung aus.

Wir benötigen sie insbesondere für die Stichprobenverteilung von Parameterschätzungen.

5.2 Konfidenzintervalle

Wir filtern alle Möglichkeiten für die untersuchte Verteilung (das sind jetzt Hypothesen). Von den vielen Möglichkeiten für μ werden nun – über die Schwellenwertüberlegung – einige wenige aussortiert. Es gibt grundsätzlich zwei Methoden dazu:

Konfidenzintervalle und statistische Tests.

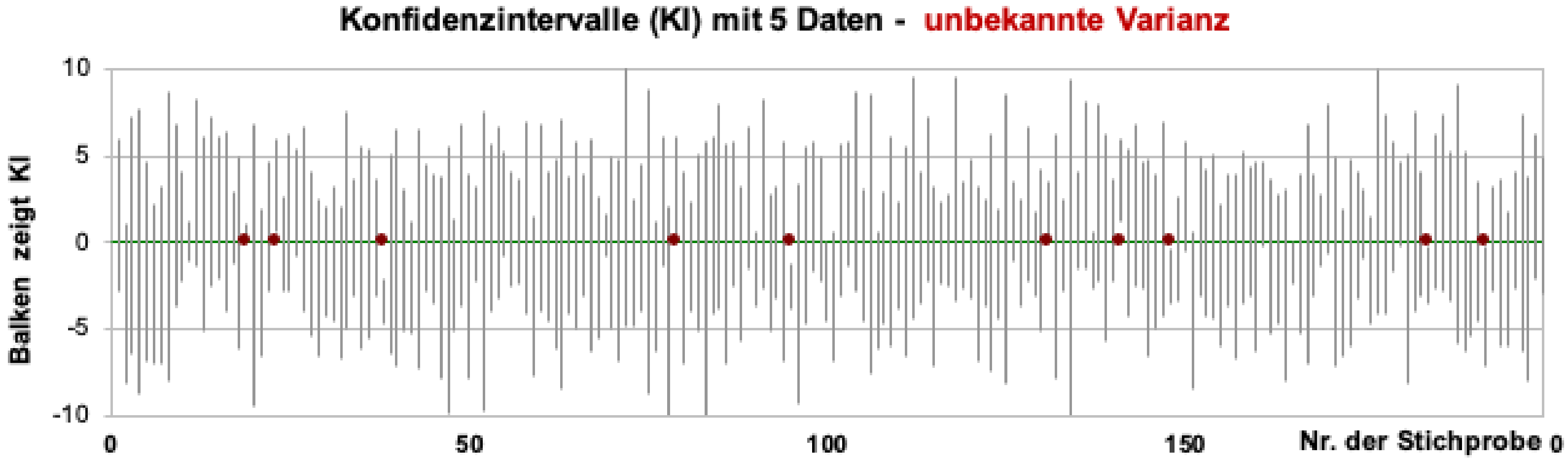
Konfidenzintervall enthält alle jene Werte von μ, welche den beobachteten Wert von $\bar{x}$ innerhalb der Schwellenwerte haben. Man sagt, eine solche Beobachtung ist mit dem Modellwert μ kompatibel.

Intervall kompatibler Werte = Konfidenzintervall.

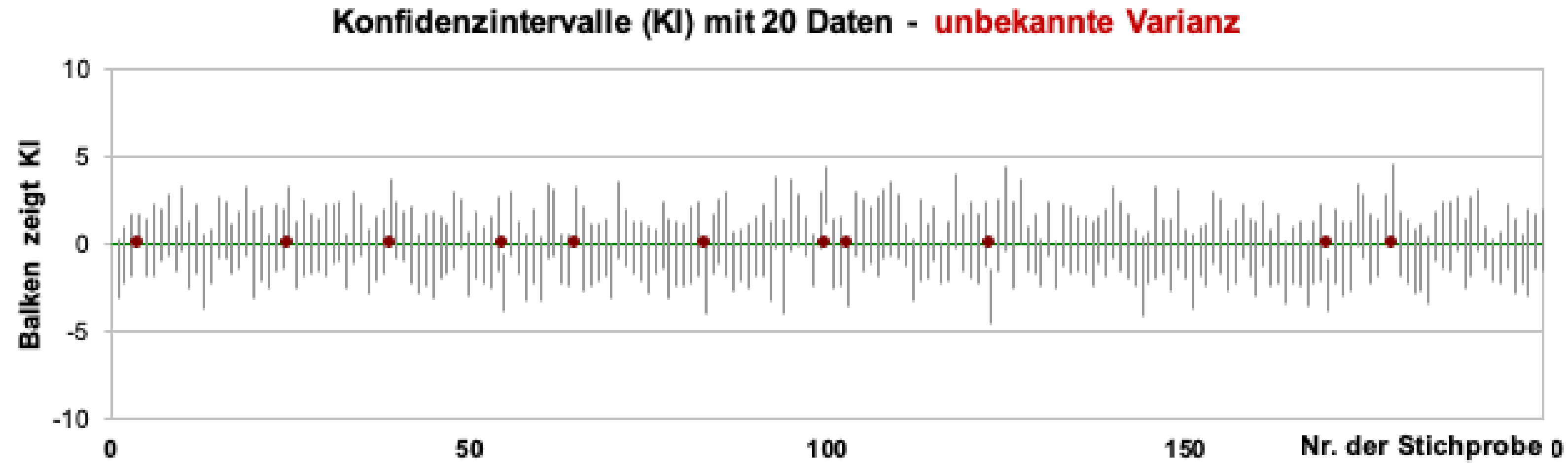
Es wird klar, dass dazu das Risiko des Überschreitens des Schwellenwerts auch noch beachtet werden muss. Die Gegenwahrscheinlichkeit nennt man Überdeckungswahrscheinlichkeit.

Ein grundsätzliches Missverständnis ist hierbei, dass das realisierte Konfidenzintervall den Parameter mit dieser Wahrscheinlichkeit überdeckt.

Überdeckungswahrscheinlichkeit trifft nicht auf das konkrete Intervall zu



	Prozentsatz überdeckender Intervalle			
	ALLE	KURZE	MITTLERE	LANGE
5 Daten	95,0%	85,1%	100,0%	100,0%
20 Daten	94,5%	88,1%	97,0%	98,5%



5.3 Hypothesentests – Filtern mit Vorbedingungen

Anstelle der artifiziellen Stichprobenverteilung der Mittelwerte nehmen wir der Einfachheit halber die Verteilung des Phänomens.

Es gehe um die Trennung von zwei Verteilungen – Gesunde und Kranke – hinsichtlich eines Diagnosemerkmals.

Gesucht ist ein geeigneter Trennpunkt.

Statt mit Wahrscheinlichkeitsverteilungen hantieren wir mit empirischen Verteilungen. Damit kommen wir weg von hypothetischen Verteilungen.

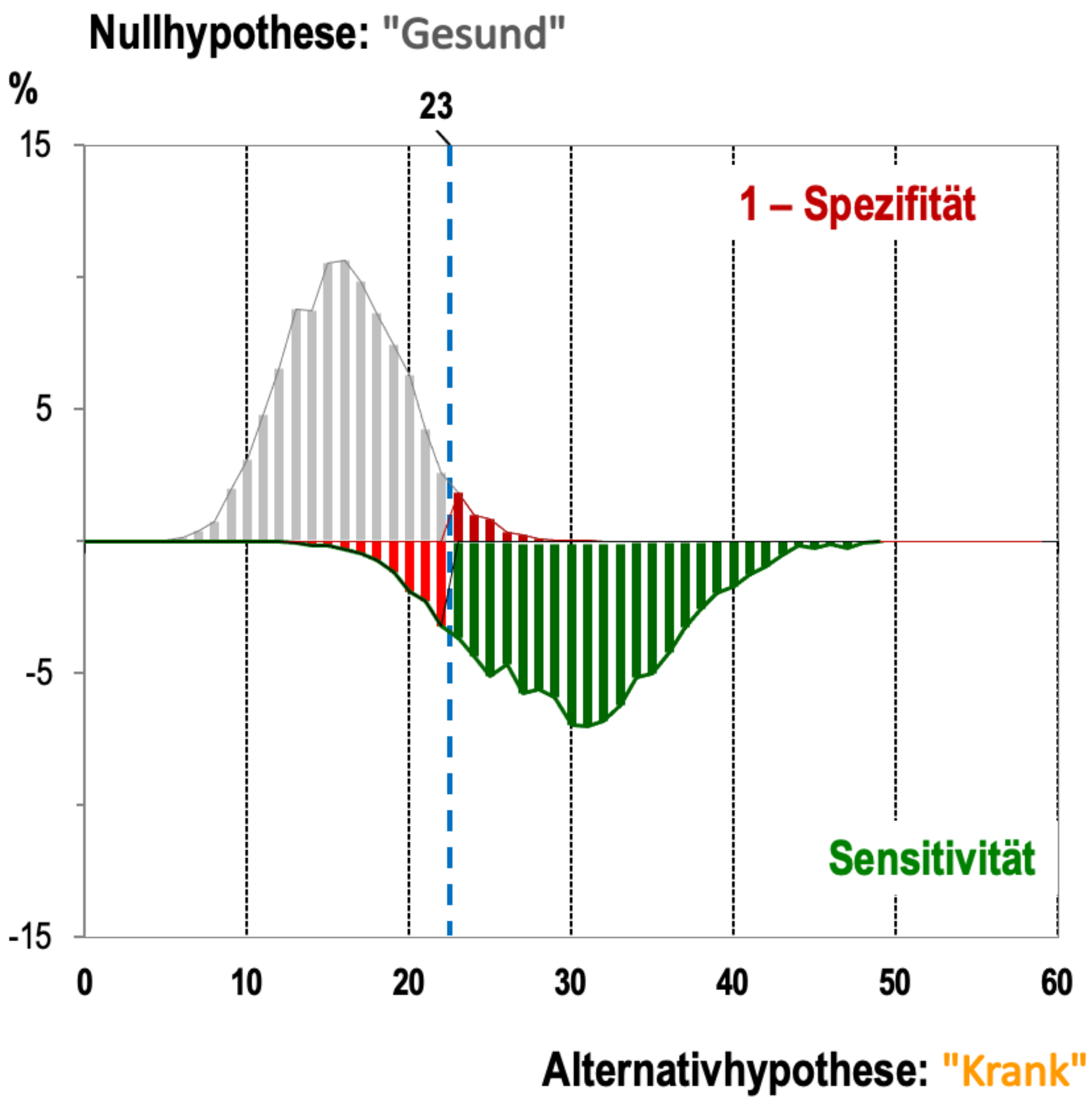
Jeder Trennpunkt hat im Sinne der Diagnose (richtige Klassifikation von Gesunden bzw. Kranken) zwei antagonistische Fehlerraten:

Gesunde werden als Kranke diagnostiziert

Kranke werden als Gesund eingestuft.

Statt den bedingten Wahrscheinlichkeiten für α- und β-Fehler haben wir es mit konkreten relativen Häufigkeit für Fehldiagnosen in zwei unterschiedlichen Szenarios zu tun.

Vorübungen zu Hypothesentests – Trennen zweier Szenarios



Jeder Trennpunkt hat im Sinne der Diagnose zwei antagonistische Fehlerraten.

Die Fehlerraten beziehen sich auf zwei verschiedene Szenarien.

Sie müssen gegeneinander abgetauscht werden.

5.4 Hypothesentests – ein Problemkind der Wissenschaften

Das Festlegen eines Fehlers macht das Problem asymmetrisch und vernachlässigt den anderen Fehler.
Die mathematische Prozedur bei einem Kontinuum von Szenarien: den sogenannten α-Fehler festlegen und den β-Fehler optimieren.
Beim reinen Signifikanztest – Nullhypothese ohne Spezifikation einer Alternativhypothese – fällt das Optimalitätsargument ganz weg.

Leider kann man in den Anwendungen weder auf statistische Tests (zugunsten von Konfidenzintervallen) verzichten, noch kommt man ohne reinen Signifikanztests aus.

Die Angabe von p-Werten ist – speziell bei reinen Signifikanztests sehr verwirrend und lädt zu Fehlinterpretationen über die Gültigkeit der Nullhypothese ein.

In wissenschaftlichen Zeitschriften wird – alle 10 Jahre – die Praxis von reinen Signifikanztests und p-Werten verboten.

Statistische Tests durch den Bayes-Ansatz verstehen lernen

Fehlendes Bindeglied: Die Prävalenz und Bayes

Radiologische Klinik

	–	+	
H_0: No Ca	96 Spezifität →	4 Falsch pos. →	100
H_1: Ca	20 Falsch neg. →	80 Sensitivität →	100
	116	84	**200**

Screening

	–	+	
H_0	95 232 NPV ↑	3 968	99 200
H_1	160	640 PPV ↑	800
	95 392	4 608	**100 000**

	Radiologische Klinik		Screening		
Prävalenz		**50 %**	**0.8 %**		
Sensibilität →		80/100 =	80.0%	80.0%	P (+| Ca)
Spezifität →		96/100 =	96.0%	96.0%	P (– | No Ca)
Pos. prädikt. Wert (PPV) ↑	80/84 =	**95.2%**	**13.9%**		P (Ca| +)
Neg. prädikt. Wert (NPV) ↑	96/116 =	82.8%	99.8%		P (No Ca| –)

Statistische Tests können die wirklich wichtigen Fragen nicht beantworten

Statistische Tests durch den Bayes-Ansatz verstehen lernen

Fehlendes Bindeglied: Die Prävalenz und Bayes

Radiologische Klinik

	–	+	
H_0: No Ca	96 Spezifität →	4 Falsch pos. →	100
H_1: Ca	20 Falsch neg. →	80 Sensitivität →	100
	116	84	200

Screening

	–	+	
H_0	95 232 NPV ↑	3 968	[illegible]
H_1	160	[illegible]	
	95 392	4 608	100 000

	Radiologische Klinik		Screening		
Prävalenz		50 %	0.8 %		
Sensibilität →	80/100 =		80.0%	80.0%	P (+| Ca)
Spezifität →	96/100 =		96.0%	96.0%	P (– | No Ca)
Pos. prädikt. Wert (PPV) ↑	80/84 =	95.2%	13.9%		P (Ca| +)
Neg. prädikt. Wert (NPV) ↑	96/116 =	82.8%	99.8%		P (No Ca| –)

Die Qualitätsindices haben in beiden Situationen denselben Wert:

- Sensibilität = 80%
- Spezifität = 96%

Aber die Qualität der Entscheidungen ist sehr verschieden:

- In der Klinik vielleicht eine Unterstützung.
- Im Screening vollkommen obsolet .

Die Diskrepanz ist umso größer, je kleiner die Prävalenz ist.

Statistische Tests durch den Bayes-Ansatz verstehen lernen

Fehlendes Bindeglied: Die Prävalenz und Bayes

Radiologische Klinik

	–	+	
H_0: No Ca	96 Spezifität →	4 Falsch pos. →	100
H_1: Ca	20 Falsch neg. →	80 Sensitivität →	100
	116	84	200

Screening

A posteriori Wahrscheinlichkeiten:

- haben ohne Bezug auf a priori Wahrscheinlichkeiten keinerlei Sinn,
- entbehren über ein sehr eng abgestecktes Szeanrio hinaus vollständig einer Häufigkeitsdeutung.

Überdies ist diese Häufigkeitsdeutung für das Individuum wertlos, sie dient nur als Referenzpunkt für das System (auch als Argumentationshilfe, um die diagnostische Prozedur zu "rechtfertigen").

	–	+	
H_0	95 232 NPV ↑	3 968	[illegible]
H_1	160	[illegible]	[illegible]
	[illegible]	4 608	100 000

Prävalenz		50 %		0.8 %		
Sensibilität →	80/100 =		80.0%		80.0%	$P(+ \mid Ca)$
Spezifität →	96/100 =		96.0%		96.0%	$P(- \mid No\ Ca)$
Pos. prädikt. Wert (PPV) ↑	80/84 =	95.2%		13.9%		$P(Ca \mid +)$
Neg. prädikt. Wert (NPV) ↑	96/116 =	82.8%		99.8%		$P(No\ Ca \mid -)$

Statistische Tests durch den Bayes-Ansatz verstehen lernen

Fehlendes Bindeglied: Die Prävalenz und Bayes

Radiologische Klinik

	–	+	
H_0: No Ca	96 Spezifität →	4 Falsch pos. →	100
H_1: Ca	20 Falsch neg. →	80 Sensitivität →	100
	116	84	200

Screening

	–	+	
H_0	95 232 NPV ↑	3 968	99 200
H_1	[illegible]	[illegible] PPV ↑	800
	95 392	4 608	100 000

Der Kontext der Medizin erleichtert das Verständnis dessen, was bei klassischen statistischen Tests fehlt:

Die a priori Wahrscheinlichkeit der Nullhypothese!

Statistische Inferenz verstehen über die Analogie mit der Medizin

		Radiologische Klinik	Screening		
Prävalenz		50 %	0.8 %		
Sensibilität →	80/100 =		80.0%	80.0%	P (+| Ca)
Spezifität →	96/100 =		96.0%	96.0%	P (– | No Ca)
Pos. prädikt. Wert (PPV) ↑	80/84 =	95.2%	13.9%		P (Ca| +)
Neg. prädikt. Wert (NPV) ↑	96/116 =	82.8%	99.8%		P (No Ca| –)

6. Knackpunkte in der beschreibenden Statistik

Vorausblick

5.1 Die Stichprobenverteilung für den Mittelwert

6.1 Multiple-Choice-Test zur Studieneingangs- & Orientierungsphase

  # 3. Boxplot

  # 4. Mittelwert und Standardabweichung

  # 5. Graphische Darstellungen von Daten

  # 6. Mittelwerte

  # 7. Maße für die Lage einer Verteilung

6.2 Antworten und Begründungen

6. Knackpunkte in der beschreibenden Statistik

6.1 Multiple-Choice-Test zur Studieneingangs- & Orientierungsphase

24 Anfänger in „Technische Mathematik“, 17 beim ersten Termin.

Es geht um eine Knock-out-Prüfung; wer sie nicht schafft, darf das Fach nicht weiterstudieren.

Vorlesung zum Teil Statistik (9 Stunden). Skriptum (30 Seiten). Inhalt:

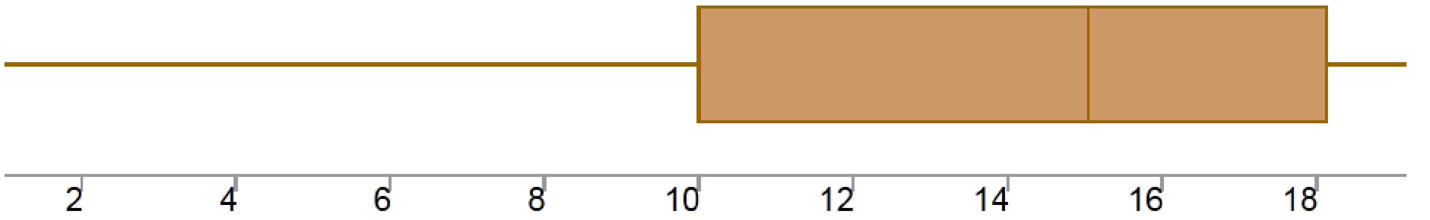
Statistiken
Methoden der Deskription univariater Daten
    Kennziffern, Boxplot, Stabdiagramm, Histogramm, Lorenzkurven
Deskription bivariater Daten
    Vorgehensweise bei Regression und Korrelation

Der Test beinhaltet

Deskription univariater Daten (5 Items).
Lorenzkurven (3 Items).
Regression und Korrelation (2 Items).

#3. Boxplot

Man betrachte den folgenden Boxplot:



Welche der folgenden Aussagen ist wahr?

I. Die Verteilung ist rechtsschief (linkssteil).
II. Der IQR ist ca. 8.
III. Der Median ist ungefähr 10.

- ☐ Lediglich I
- ☐ Nur II
- ☐ Nur III
- ☐ I und III
- ☐ II und III

4. Mittelwert und Standardabweichung

Mittelwert und Standardabweichung von *x*-Daten sind gegeben durch: $\bar{x} = 100013{,}4$, $s = 10{,}23$. Neue *y*-Daten werden durch eine lineare Transformation der *x*-Daten wie folgt festgelegt:

$y = \dfrac{x - 100000}{10}$. Welche der folgenden Aussagen ist wahr?

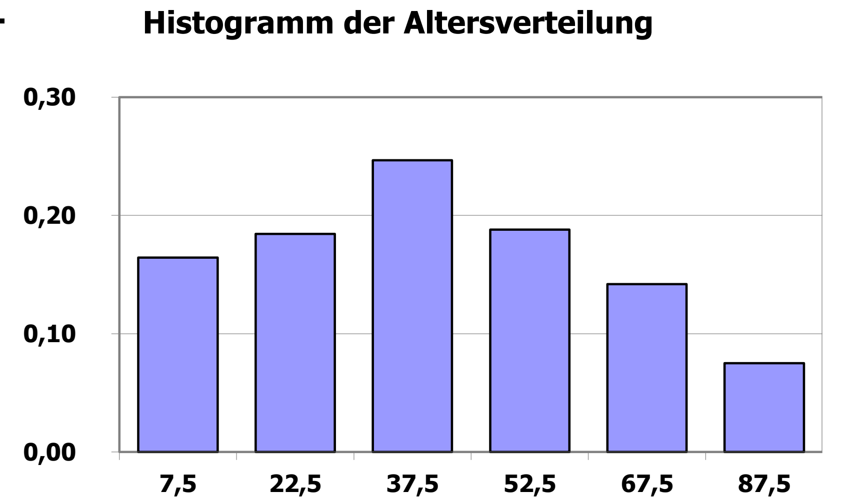
I. Man kann aus den Angaben den Mittelwert der *y*-Daten ausrechnen, nicht aber deren Standardabweichung.
II. Aus den Angaben kann man sowohl Mittelwert als auch Standardabweichung der *y*-Daten ausrechnen.
III. Es gilt: $\bar{y} = 13{,}4$ und $s_Y = 0{,}1023$.

- ☐ Lediglich I.
- ☐ Lediglich II.
- ☐ Lediglich III.
- ☐ II und III.

5. Graphische Darstellungen von Daten

Daten zum Alter aus einer Untersuchung:

Altersbereich von bis			Mitten	Anzahl n_i	Anteil h_i
0	bis inkl.	15	7,5	1.333.505	0,16
15	- " -	30	22,5	1.495.740	0,18
30	- " -	45	37,5	2.002.259	0,25
45	- " -	60	52,5	1.526.110	0,19
60	- " -	75	67,5	1.151.122	0,14
75	- " -	100	87,5	609.018	0,08
				8.117.754	1,000



Ist die Darstellung der Verteilung des Alters korrekt?

- ☐ Ja.
- ☐ Nein.
- ☐ Ja, aber man sollte die Abstände zwischen den Säulen vermeiden.
- ☐ Ja, aber man sollte die Klassenmitten so wählen, dass schönere Zahlen herauskommen, damit die Graphik besser lesbar ist.

6. Mittelwerte

I. Der Mittelwert ist der beste Wert, mit dem man eine Liste von Daten beschreibt.

II. Eine Gruppe von drei Personen hat folgende IQ-Werte (Intelligenquotient): 70, 100, 130. Im Mittel haben sie einen IQ von 100.

III. Der Mittelwert einer Liste von Daten $x_1, x_2, \ldots, x_n$ optimiert das Funktional $\varphi(c) = \sum_{i=1}^{n} |x_i - c|$.

- ☐ Alle drei Aussagen treffen zu.
- ☐ Nur die Aussage III trifft zu.
- ☐ Nur die Aussagen II und III treffen zu.

7. Maße für die Lage einer Verteilung

Aus einer Liste von Daten $x_1, x_2, \ldots, x_n$ hat man folgende Kennziffern berechnet: $\bar{x} = 18$, $x_{med} = 15$, $x_{\mathrm{mod}} = 17$ und $s = 2$.

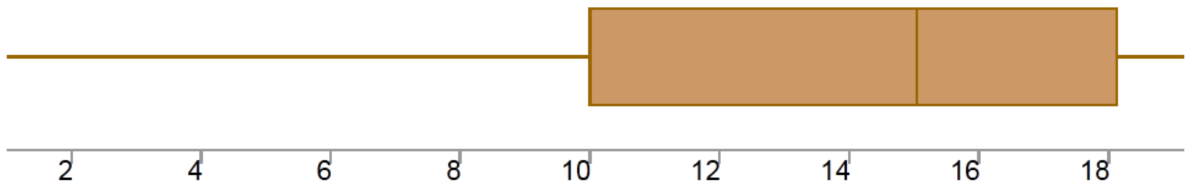
Was sagt das über den Datensatz aus?

- ☐ Die Daten sind linkssteil (rechtschief).
- ☐ Die Daten sind rechtsteil (linksschief).
- ☐ Gar nichts; es muss bei der Berechnung ein Fehler passiert sein.

6.2 Antworten und Begründungen

3. Boxplot

Angegeben wird, wie häufig die Antwortkategorien angekreuzt wurden sowie, ob Begründungen richtig (r) oder falsch (f) waren oder gefehlt haben (√). Ferner eine Auswahl an „Begründungen“.



I. ... ist rechtsschief (linkssteil).
II. Der IQR ist ca. 8.
III. Der Median ist ungefähr 10.

				Begründung	
	4	Lediglich I			
×	**12**	Nur II		r	**8**
	0	Nur III		f	**7**
	0	I und III		√	**2**
	1	II und III			

Median = 10 / 14. Rechtsschief wegen $x_{0,5} < x_{0,75}$. Linksschief → mehr Werte rechts.

 # 4. Mittelwert und Stabw: $\bar{x} = 100013{,}4$, $s = 10{,}23$. $y = \dfrac{x - 100\,000}{10}$.

I. Man kann ... Mittelwert der *y* ausrechnen, nicht aber ... s_Y.

II. ... sowohl Mittelwert als auch Standardabweichung der *y*-Daten ...

III. Es gilt: $\bar{y} = 13{,}4$ und $s_Y = 0{,}1023$.

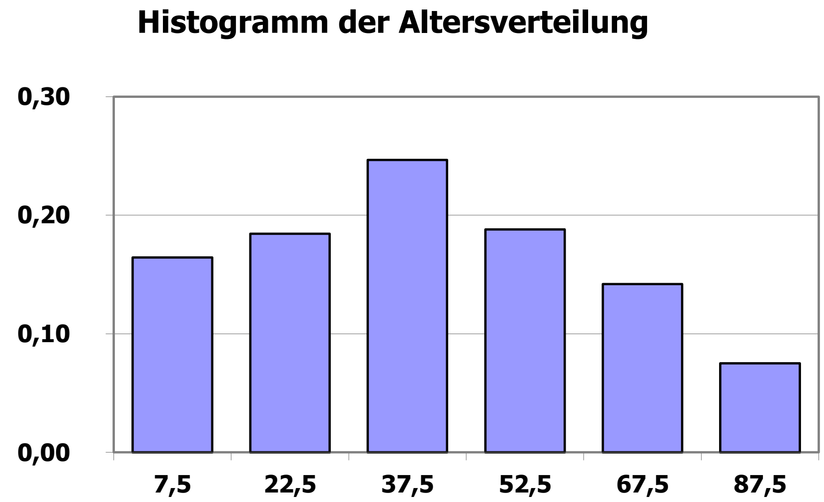
	6	Lediglich I	r	**7**
×	**6**	Lediglich II	f	**5**
	0	Lediglich IIII	√	**5**
	2	II und III		
	1	keine		
	1	alles falsch		
	1	nicht bearbeitet		

s kann berechnet werden, wenn man zuvor alle *y* ermittelt, allein aus dieser Formel jedoch NICHT.

alles falsch, *s* und $\bar{x}$ bleiben gleich. – … [bräuchte man] mehr Angaben.

5. Graphische Darstellungen

Altersbereich von bis			Mitten	Anzahl n_i	Anteil h_i
0	bis inkl.	15	7,5	1.333.505	0,16
15	- " -	30	22,5	1.495.740	0,18
30	- " -	45	37,5	2.002.259	0,25
45	- " -	60	52,5	1.526.110	0,19
60	- " -	75	67,5	1.151.122	0,14
75	- " -	100	87,5	609.018	0,08
				8.117.754	1,000



	4	Ja	r	**5**
×	**8**	Nein	r*	**6**
	5	Ja, aber man sollte die Abstände …	f	**5**
	0	Ja, aber man sollte die Klassenmitten …	√	**1**

Da die Werte metrisch sind, *benötigt* man keine Abstände.

Man könnte das Diagramm falsch verstehen.

Man kann aus dieser Grafik die Werte gut ablesen.

Abstände sind egal.

6. Mittelwerte

I. Der Mittelwert ... beste Wert, ... Liste von Daten beschreibt.

II. ... IQ-Werte: 70, 100, 130. Im Mittel haben sie einen IQ von 100.

III. Der Mittelwert ... optimiert ... $\varphi(c) = \sum_{i=1}^{n} | x_i - c |$.

	1	Alle drei Aussagen treffen zu	r	**3**
	0	Nur Aussage III trifft zu	f	**9**
	14	Nur Aussagen II und III treffen zu	$\sqrt{}$	**5**
×	**1**	keine Aussage ist richtig		
	1	nur II		

Mittelwert dient zur Veranschaulichung.

Mittelwert (70+100+130)/3 = 100.

Der Mittelwert passt nicht, wenn Ausreißer …

7. Maße für die Lage einer Verteilung

... berechnet: $\bar{x} = 18$, $x_{med} = 15$, $x_{\mathrm{mod}} = 17$ und $s = 2$.

	2	Daten sind linkssteil	r	**9**
	5	Daten sind rechtssteil	f	**6**
×	**10**	Gar nichts, es muss ein Fehler passiert sein	√	**2**

Es befinden sich ~~mehr~~ höhere Daten im 2. Quartil → Steigung auf der rechten Seite.

Einige führen die Bedingung *mean* > *med* bzw. *mean* > *mod* an, was am Kern des Problems vorbeigeht.

Richtige Begründungen treffen wesentliche Teile und können auch mit falschen Antworten verbunden sein.

Vorschläge für passende Begründungen

3 Der linke Teil der Box ist breiter als der rechte, daher ist die Verteilung linksschief; der Strich für den Median liegt bei 15, die Box hat eine Breite von 18–10 = 8.

4 Zwischen den Kennziffern der Daten und der affin-linear transformierten Daten besteht ein direkter Zusammenhang; die Rechnung in III ist aber falsch.

5 Merkmal ist metrisch, daher keine Abstände, Achse ist falsch unterteilt, die Höhen (anstelle der Flächen) stellen die r. H. dar.

6 I hängt vom verwendeten Kriterium ab; II Daten sind *nicht* metrisch; III Mittelwert minimiert *quadrierte* (!) Abweichungen.

7 Die angegebenen Kennziffern verletzen die Bedingung *mean* < *med* < *mode*.

7. Zum guten Schluss

7.1 Öffentlicher Umgang mit (stochastischen) Aussagen

Ein Promi sagt, tut, ...

(Angelina Jolie und die 80%).

Der Experte rät ...

(gehen Sie mindestens ab dem Alter von 30 jährlich zur Mammographie, besonders, wenn Sie spezifische Risiken in der Familie haben).

Am Ende des Rankings: Belege

(Wie bestimmte Genkonstellationen den körpereigenen Stoffwechsel verändern, sodass, wann und warum präcancerogene Zelldysplasien in invasives Tumorwachstum ausarten).

Wozu auch? Wir leben in einer evidence-based society.

7.2 Vorhersage wirtschaftlicher Entwicklungen

Mein alter Mentor in Graz, der leider schon lange verstorbene Sepp Gölles hat immer die ABC – Regel gepredigt und hat hierbei die Statistik auf dem Niveau C eingestuft:

Zuerst die groben Ungereimtheiten mit der allergrößten Wirkung beseitigen, dann … . Bei der Statistik hat man dann eigentlich nur mehr Feinabstufungen und vergleichsweise wenig Wirkung.

Der Konflikt zwischen den Lehman Brothers und Goldman Sachs und die unterschiedlichen Auswirkungen auf USA und Europa, das ist A-Wissen. Ja, Insiderwissen ist viel wichtiger als die Statistik, die zudem noch schwierig und unzuverlässig ist.

Nicht gerade eine Werbung, den langen Weg dorthin zu gehen. Außerdem kommt der Austausch von Insiderwissen und das Lancieren von Intrigen den menschlichen Bedürfnissen viel mehr entgegen als die zumalen trockene Statistik.

7.3 Vorhersage von Erdbeben

Die Geologen arbeiten doch schon längst auch mit Stochastikern zusammen. Ich kenne ernstzunehmende Arbeiten schon aus den 70ern. Die hätten doch schon längst Daten genug, wenn der Ansatz wirklich erfolgreich wäre.

Dass immer wieder so vage Vorhersagen von großen Beben gemacht werden (Kassandra hat am Ende immer Recht), steigert den Ruf der Profession nicht wirklich.

Dass auf der anderen Seite so riesige Beben wir das vor Sumatra oder Fukushima nicht vorhergesagt werden (nicht einmal im Ansatz hat man an so starke Beben dort gedacht), ist auch kein Ruhmesblatt.

Die Aktion der italienischen Gerichte mit der Anklage gegen Erdbebenforscher (Emiglia Romagna) zeigt die Frustration der Öffentlichkeit über die hoch mit Forschungsgeldern ausgestatteten Experten.

7.4 Zu den Wahlprognosen

Ich meine nicht, die Feinheiten (oder Unsinnigkeiten) des US-amerikanischen Wahlsystems ausnutzend eine ordentliche Prognose zu erstellen. Der Prozentsatz global war als Gewinn von H. Clinton vorhergesagt.

Entweder ist den Wahlforschern das öffentliche Ausspielen ihrer Vorzüge unter-sagt oder sie haben einfach versagt.

Es gibt immer wieder – auch bei uns – Flops von Wahlprognosen.

7.5 Public Health

…

Auf viele Daten in diesem Bereich trifft der alte Spruch zu

„Garbage in – garbage out“.

7.6 Gedanken, die Unterricht und Überzeugen berühren

Mit der Stochastik scheint die Perspektive von oben verbunden zu sein:

Während es für die Organisation auf größerer Ebene, für Entscheidungen und Bemessungen relevante Teile zusammenträgt,

bleibt der Einzelne draußen.

a. Lotterie:

Wie muss ich sie gestalten, damit sie interessant und ertragreich wird?

Für den einzelnen: ich spiele nicht, weil der Betreiber eh gewinnt.

Aber dann spielen einzelne (gar nicht so wenige) doch.

Was sind ihre Beweggründe und warum kommen die in der Stochastik nicht oder nicht geeignet vor?

b. Versicherung:

Der Preis einer Polizze basiert auf statistischer Information und Daumen mal pi (zur Absicherung).

Für die Versicherungsgesellschaft klappt das nach dem Gesetz der großen Zahlen hervorragend.

Für den Einzelnen wiederum gilt: seine Kriterien werden üblicherweise nicht abgebildet. Er bleibt dem Risiko voll ausgesetzt.

Warum lässt sich der Einzelne versichern?

c. Marktforschung:

Die Klicks (insbesondere die Folge-Klicks) bei Amazon (wer bestellt was gemeinsam) und ein geeignetes Programm zur Aufzeichnung und Analyse lässt Amazon situationsbedingt Angebote erstellen (viele Firmen machen das im Internet).

Der einzelne kann sich das weder leisten noch würde es ihm viel bringen (oder?) Kleinere Anbieter haben den Nachteil, dass ihre Daten noch wenig Struktur enthalten und daher wenig zuverlässig sind.

d. PISA:

Die Betreiber haben eine Folgeindustrie errichtet, beginnend von getesteten Aufgabenpaketen (die dann auch für zentrale Reifeprüfungen gebraucht werden können) einschließlich Lösungen.

Dass dabei so viele Fehler (s. Erich Neuwirth) passieren, kann man erst beurteilen, wenn man sich jahrelang mit den Modellen beschäftigt.

Die einzelnen Staaten können es sich nicht leisten, vom System abzukop-peln und reformieren andauernd ihre Schulsysteme (aufgrund von kaum interpretierbaren Rangunterschieden), bis sie …

e. Wirtschafts- und Sozialstatistiken

Diese werden andauernd geändert, wenn sie nicht den „Erfordernissen“ der „Betreiber“ entsprechen. Dennoch werden sie als Basis tagespolitischer Entscheidungen herangezogen.

Es gibt unendlich viele Tricks, die Zahlen so zu biegen bis sie passen.

Für viele organisatorische Planungen hingegen fehlen – aus Mutwillen – die nötigen Statistiken.

Etwa hat man vor 10 Jahren die Lehrer mit Überangeboten in die Pension gelockt, …. Heute sucht man händeringend Lehrer und nimmt viele auch ohne Prüfung.

Dabei ist der Bestand der Lehrkräfte nach Alter …, die Ausfälle durch … gut schätzbar, die Daten der Schüler gibt es auch … früher, sie gehen schon in die Krabbelstube). Und noch immer schafft man es nicht, die Flüsse zu schätzen und den Bedarf zu prognostizieren.

Meine Knackpunkt-Fragen

Als Verkaufsstory – wozu das Ganze wirklich gut ist – fällt mir derzeit nichts ein, außer dass man beim „Lügen“ nicht mehr mit von der Partie ist, wenn man über zu wenig Wissen aus der Statistik verfügt.

Es gilt, Knackpunkte zu suchen und finden, welche die Art des Zugangs in der Stochastik wenigstens als eine (von mehreren) als probat erscheinen lässt.

Ich glaube auch, dass wir die Unterweisung noch lange nicht auf einen solchen Punkt gebracht haben.

Worin besteht der eigene Ansatz? (Voraussetzungen; Ziel-Kriterien)
Was bieten alternative Ansätze?

Was kann man mit besser als ohne stochastische Methoden?

Rückmeldungen willkommen: manfred.borovcnik@aau.at
Mehr: https://www.researchgate.net/profile/Manfred-Borovcnik