

Statistik



<https://iuk.one/1012-1047>

Clemens H. Cap
ORCID: 0000-0003-3958-6136

Department of Computer Science
University of **Rostock**
Rostock, Germany
clemens.cap@uni-rostock.de

Version 3.1



Ziel: Aus Sicht der Wissenschaftstheorie für Probleme zu sensibilisieren.

Nicht-Ziel: Eine Statistik-Vorlesung zu ersetzen.

Methodik: Anhand einiger Beispiele nachdenklich werden.

1. Simpsonsches Paradoxon
2. Berksonsches Paradoxon
3. Kontextualisierung
4. Spezifität und Sensitivität
5. Law of Small Numbers
6. Schlußfolgerung

1. Simpsonsches Paradoxon

Beachte: Unterteilungen einer Gruppe können Trends erzeugen, die in der Gesamtheit nicht oder anders vorhanden sind.

1. Simpsonsches Paradoxon

2. Berksonsches Paradoxon

3. Kontextualisierung

4. Spezifität und Sensitivität

5. Law of Small Numbers

6. Schlußfolgerung

1. Simpsonsches Paradoxon

Beispiel: Bart und Lisa (1)

Name: Nach Edward Simpson, britischer Statistiker, hier mit den Simpsons geschildert.

Story: Lisa und Bart Simpson arbeiten als Volontäre in einer Redaktion.
In 2 Wochen bearbeiten sie jeweils insgesamt 10 Artikel.
Redaktion analysiert, wie viele Artikel besser wurden.

	In Woche 1		In Woche 2	
	verbessert	bearbeitet	verbessert	bearbeitet
Lisa	0	3	5	7
Bart	1	7	3	3

Tab. 1: Rohdaten aus der Redaktion.

Frage: Wer soll eingestellt werden?

Die Chefs von Lisa und Bart machen ein Spreadsheet.

Mit **demselben Zahlenmaterial** kommen sie zu **verschiedenen** Ergebnissen.

1. Simpsonsches Paradoxon

Beispiel: Bart und Lisa (2)

	In Woche 1		In Woche 2		Insgesamt verbessert
	verbessert	bearbeitet	verbessert	bearbeitet	
Lisa	0	3	5	7	5
Bart	1	7	3	3	4

Tab. 2: Chef von Lisa: **Lisa soll eingestellt werden.** Sie hat insgesamt mehr Artikel verbessert.

	In Woche 1		In Woche 2		Leistung	Leistung
	verbessert	bearbeitet	verbessert	bearbeitet	Woche 1	Woche 2
Lisa	0	3	5	7	0%	71%
Bart	1	7	3	3	14%	100%

Tab. 3: Chefin von Bart: **Bart soll eingestellt werden.** In beiden Wochen war seine Leistung prozentuell höher.

1. Simpsonsches Paradoxon

Beispiel: Bart und Lisa (3)

Beide Argumente sind richtig.

Ohne weitere Selektionspräferenzen kann das Paradoxon nicht aufgelöst werden.

Denkbare Erklärungsmuster wären:

- Barts Arbeitsleistung nimmt von Woche 1 auf 2 stark ab, Lisas Arbeitsleistung nimmt von Woche 1 auf 2 stark zu.
- Bart war bereits zu Beginn prozentuell besser und hat eine bessere Lernkurve.
- Die 100% von Bart in Woche 2 sind zu stark gewichtet, die 0% von Lisa sind zu schwach gewichtet.

Man *kann* sich gleiche Gewichtung aller Arbeiten wünschen.

Nach und **durch** die Entscheidung für Wichtung erhält man ein neues Ergebnis.

In der Praxis ist das Simpsonsche Paradoxon ein Hinweis darauf, daß man noch nicht alle wichtigen Einflußfaktoren oder Gewichtungen kennt.

1. Simpsonsches Paradoxon

Visuelle Erläuterung (1)

x und y sind **negativ** korreliert.

Teilt man die Proben in zwei Gruppen namens **blau** und **rot** dann sind **in jeder Gruppe** x und y **positiv** korreliert.

Die Antwort, die die Statistik gibt, hängt von der Frage ab, die **wir** stellen.

Big Data kann Fragen beantworten, ihnen aber nicht **unseren Sinn** geben.

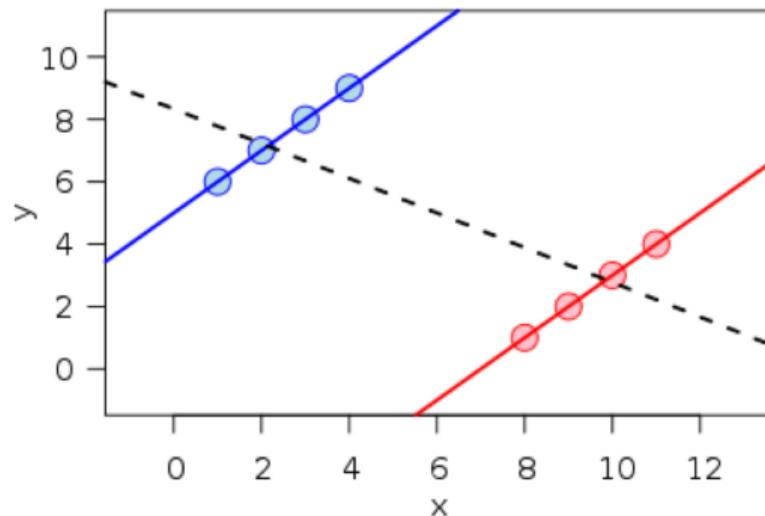


Abb. 1: Illustration des Simpsonschen Paradoxons, wie es bei Einteilung in zwei Gruppen auftreten kann.

© Rechte siehe Anhang.

1. Simpsonsches Paradoxon

Visuelle Erläuterung (2)

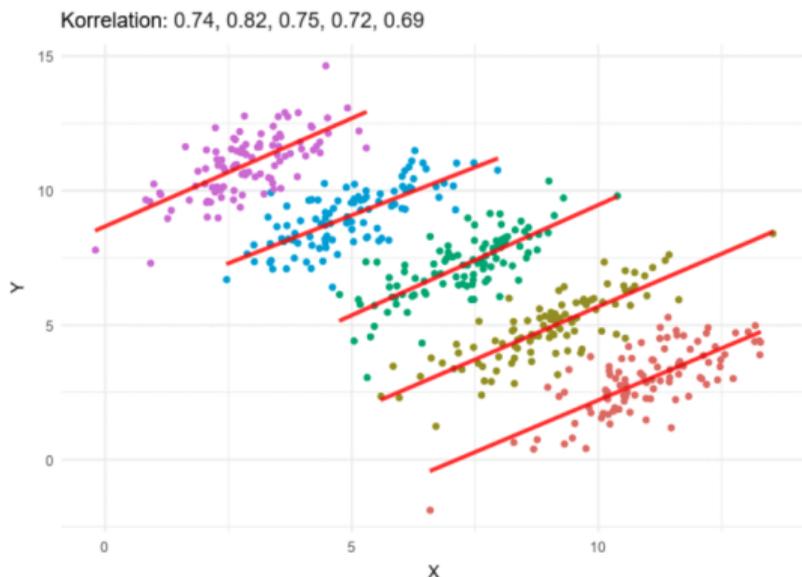


Abb. 2: Korrelation, wie sie sich bei Einteilung in mehrere Gruppen darbieten kann. Korrelationswerte am oberen Bildrand. © Rechte siehe Anhang.

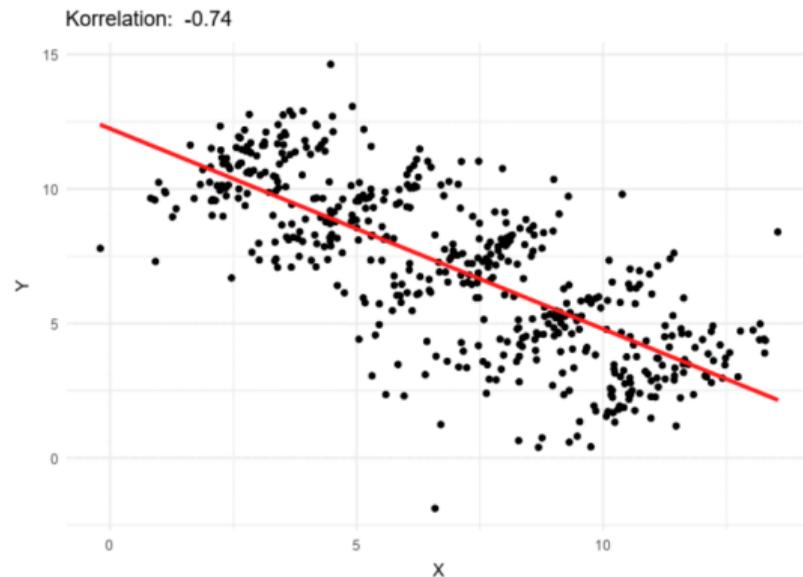


Abb. 3: Korrelation, wie sie beim gleichen Datensatz ohne Einteilung in Gruppen aussieht. Korrelationswerte am oberen Bildrand. © Rechte siehe Anhang.

1. Simpsonsches Paradoxon

Beispiel: Medikamentendosierung

x Dosis eines Medikaments
 y Überlebensjahre des Patienten
Farbe Geschlecht des Patienten

Bei Kenntnis des Geschlechts:
Mehr (Medizin) wirkt mehr.

Ohne Kenntnis des Geschlechts:
Das Zeug verkürzt die Lebensdauer!

Interpretation:

Offenbar ist die sinnvolle Dosis hier geschlechtsabhängig.

Die richtige Auswertung nutzt daher die Kenntnis des Geschlechts.

Vergleiche **Konventionalismus**: Die Fragen bestimmen die Antworten (teilweise) (mit).

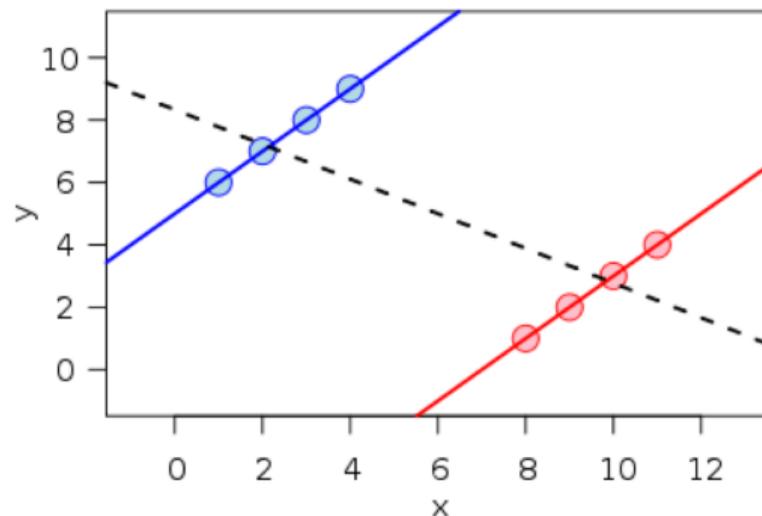


Abb. 4: Das Simpsonsche Paradoxon erzeugt bei spezifischer Anwendungslage möglicherweise schwierige Situationen. © Rechte siehe Anhang.

1. Simpsonsches Paradoxon

Beispiel: Kriminalstatistik in Mittelstadt (1)

Mittelstadt hat 200.000 Einwohner und besteht aus Villenviertel & Vergnügungsviertel.

Tageszeitung 1

140.000 Inländer begingen	59 Verbrechen	⇒ Quote 0,43 Promille
60.000 Ausländer begingen	51 Verbrechen	⇒ Quote 0,85 Promille

Tageszeitung 2

100.000 Vergnügungsviertel	100 Verbrechen, darunter	
50.000 Ausländer mit	50 Verbrechen	⇒ Quote 1 Promille
50.000 Inländer mit	50 Verbrechen	⇒ Quote 1 Promille
100.000 Villenviertel	10 Verbrechen, darunter	
10.000 Ausländer mit	1 Verbrechen	⇒ Quote 0,1 Promille
90.000 Inländer mit	9 Verbrechen	⇒ Quote 0,1 Promille

Beispiel: Kriminalstatistik in Mittelstadt (2)

Dieselben Zahlen – zwei Analysen – stark unterschiedliche Schlußfolgerungen.

Einführen einer weiteren Variablen ändert das in uns ausgelöste interpretative Bild.

Aber Achtung: Tageszeitung 2 hat nicht "mehr" Recht.

Woher wissen wir, daß das es nicht noch andere Einflußaspekte gibt?

Tageszeitung 3 setzt einen Statistiker auf das Problem an und berichtet:

Alle 110 Verbrechen wurden von Angehörigen der Gruppe "Kriminelle" begangen.

1. Simpsonsches Paradoxon

Theoretische Einordnung

Leitsatz 1

Ein Trend in der Gesamtheit kann **durch Unterteilung in Gruppen** verschwinden oder sich umkehren, sogar in allen Teilgruppen.

Leitsatz 2

Ein Trend, der in mehreren oder allen Gruppen besteht kann verschwinden oder sich umkehren, wenn man die **Gruppen vereinigt**.

Leitsatz 3

Die Erklärung besteht in der **unterschiedlichen Gewichtung** durch die jeweilige Größe der Gruppen.

2. Berksonsches Paradoxon

Beachte: Vorauswahlen können Korrelationen erzeugen, die in der Grundgesamtheit nicht bestehen.

1. Simpsonsches Paradoxon
2. **Berksonsches Paradoxon**
3. Kontextualisierung
4. Spezifität und Sensitivität
5. Law of Small Numbers
6. Schlußfolgerung

2. Berksonsches Paradoxon

Beispiel: Auswahl im College (2)

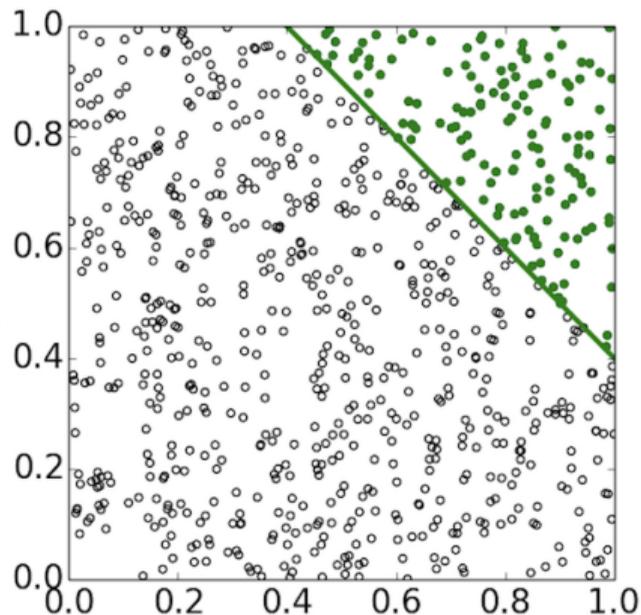


Abb. 5: In der Grundgesamtheit aller Punkte besteht keine Korrelation zwischen Sport und Noten. In der Gesamtheit mit Selektion, also unter den zugelassenen Studenten (grün), besteht negative Korrelation zwischen Sport und Noten. © Rechte siehe Anhang.

3. Kontextualisierung

Beachte: Kontexte können dramatische Effekte in der Interpretation bewirken. Wir betrachten hier Beispiele in der Visualisierung von Daten.

1. Simpsonsches Paradoxon
2. Berksonsches Paradoxon
- 3. Kontextualisierung**
4. Spezifität und Sensitivität
5. Law of Small Numbers
6. Schlußfolgerung

3. Kontextualisierung

Umsatzzahlen ohne Nachbearbeitung

Ziel: Gehaltserhöhung

Ein Marketing Mitarbeiter will den Chef von einer Gehaltserhöhung überzeugen.

Er versucht, seine Umsatzzahlen durch **stets korrekte, aber geeignet kontextualisierte Visualisierung** in ein dafür geeignetes Licht zu rücken.

1. Visualisierung

Die ohne Nachbearbeitung dargestellten Umsatzzahlen geben wenig her.

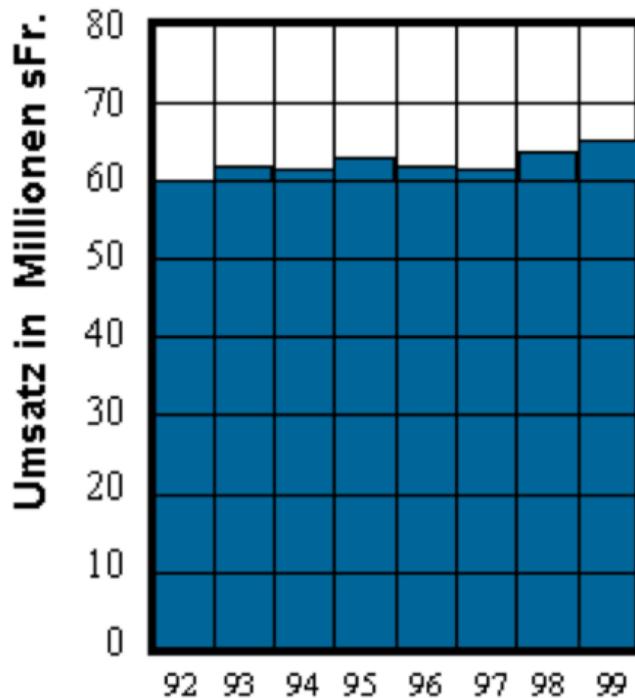


Abb. 6: Die Rohdaten © Rechte siehe Anhang.

2. Visualisierung

Abschneiden der vertikalen Achse erlaubt eine Skalierung und damit Vergrößerung der Unterschiede.

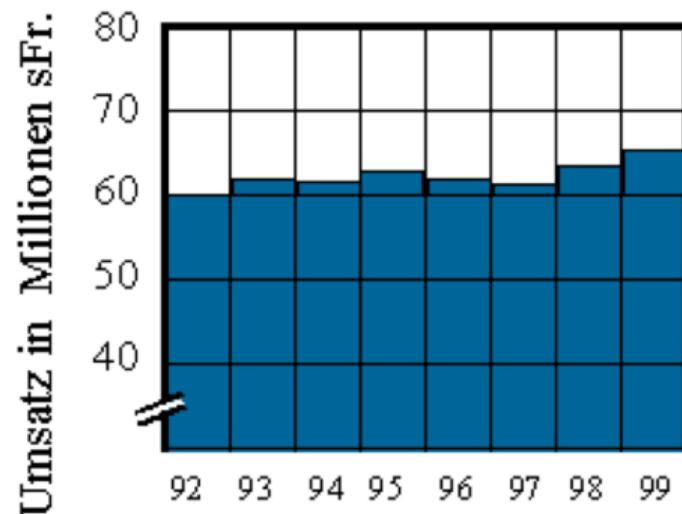


Abb. 7: Die erste Bearbeitung © Rechte siehe Anhang.

3. Visualisierung

Weiteres Ausblenden nicht benötigter Abstände erlaubt ein besseres Herausarbeiten der wesentlichen Informationen.

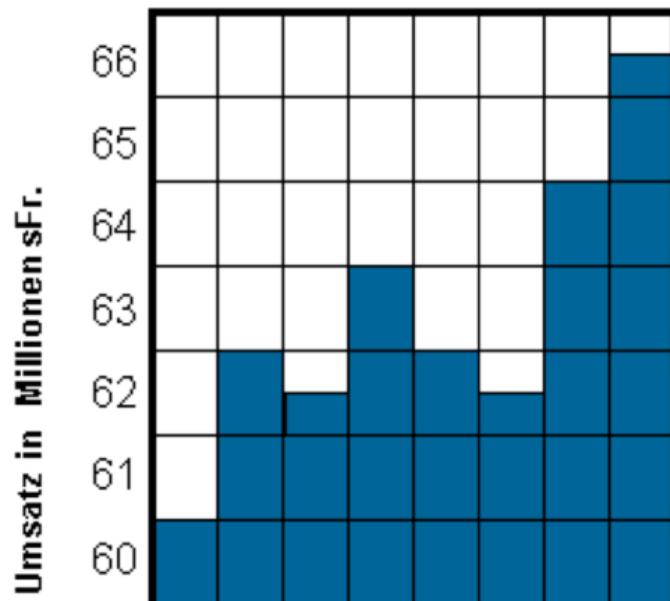


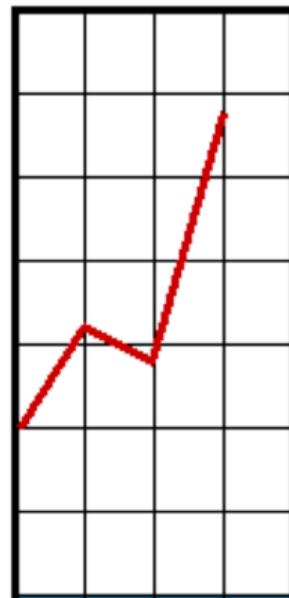
Abb. 8: Die zweite Bearbeitung © Rechte siehe Anhang.

4. Visualisierung

Darstellung als Kurvendiagramm. Zusammenfassung von jeweils zwei Jahren und Angabe des Durchschnitts.

Entfernung der Achsenbeschriftung macht eine eigenständige analytische Bewertung der Zahlen unmöglich.

Vergleiche nun mit dem 1. Bild!



Entwicklung 1992-1999

Abb. 9: Die dritte Bearbeitung © Rechte siehe Anhang.

3. Kontextualisierung

Nochmal im direkten Vergleich

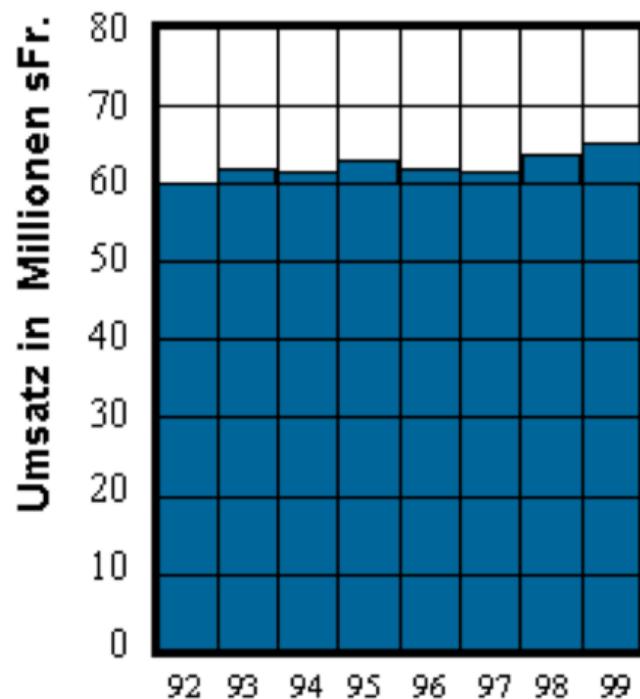
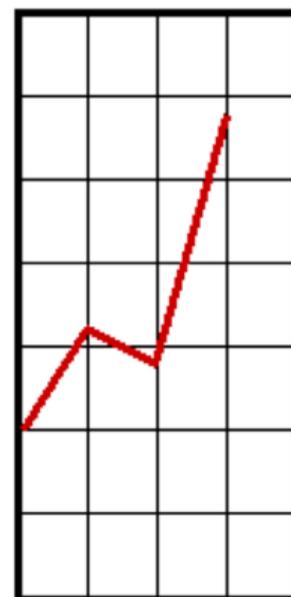


Abb. 10: Die Rohdaten © Rechte siehe Anhang.



Entwicklung 1992-1999

Abb. 11: Die bearbeiteten Daten © Rechte siehe Anhang.

4. Spezifität und Sensitivität

Beachte: Ohne Angaben von Spezifität und Sensitivität können Tests nicht sinnvoll bewertet werden.

1. Simpsonsches Paradoxon
2. Berksonsches Paradoxon
3. Kontextualisierung
4. Spezifität und Sensitivität
5. Law of Small Numbers
6. Schlußfolgerung

4. Spezifität und Sensitivität

Beispiel: Krebstest (1)

Eine Nobelpreis-würdige Erfindung!

Eine Frau kommt zu uns und will einen Test auf Brustkrebs machen.

Wir nehmen Blut ab und gehen kurz ins Labor.

Wir kommen zurück und sagen ihr (wie jeder Frau): "Sie haben Brustkrebs".

Wir erkennen mit diesem Test **jeden** Fall von Brustkrebs.

Unser Test hat eine **Treffer-Rate von 100%**.

Wir sind Genies und erwarten den Nobelpreis für Medizin.

Analyse

Problem 1: Eine Treffer-Rate von 100% klingt unglaubwürdig.

Problem 2: Jede Frau, die sich testen läßt, bekommt die Diagnose Brustkrebs.

Aber: Wir verdienen an den weiteren Behandlungen ganz gut...

4. Spezifität und Sensitivität

Beispiel: Krebstest (2)

100% Treffer-Rate sind verdächtig – wir müssen den Test "verbessern"

Wir werfen zwei faire Münzen (mit Ergebnissen "Kopf" und "Zahl").

Wenn mindestens einmal "Kopf" dabei ist, sagen wir der Frau: "Sie haben Brustkrebs".

Unser Test hat jetzt eine Treffer-Rate von 75%. Das klingt realistischer!

Analyse

Problem: Wir haben immer noch eine Falschalarm-Rate von 75%.

" Lösung": Wir informieren darüber nicht.

Nochmalige kritisch Nachdenken: Stimmt unsere Analyse?

Vermutlich ist unsere Falschalarm-Rate doch ein *klein wenig* kleiner.

Denn: Frauen die zum Arzt gehen haben ev. einen Grund, der sie beunruhigt (Knötchen ?)

Die Vorauswahl des Samples in unserer Analyse ist nicht randomisiert.

Erst eine randomisierte Studie bringt Aufschluß!

Gleichwohl: Bei einer *flächendeckenden* Vorsorgeuntersuchung stimmt die Analyse!

Erkenntnis

Zur Bewertung eines Testverfahrens müssen wir **Treffer-Rate** und **Falschalarm-Rate** kennen.

Und wir müssen wissen, wie diese genau definiert sind.

Jede andere Darstellung ist nicht unbedingt gefälscht oder manipulativ,
– wir sollten nicht immer böse Absicht unterstellen –
sie ist aber mindestens wissenschaftlich grob fahrlässig.

Spezifität und Sensitivität

Kontingenztafel

	Merkmal vorhanden	Merkmal nicht vorhanden
Merkmal behauptet	TP = True Positive	FP = False Positive
Merkmal nicht behauptet	FN = False Negative	TN = True Negative

Sensitivität: $TP / \text{Zahl der Merkmalsträger} = TP / (TP + FN)$

Wie empfindlich (**sensitiv**) reagiert der Test auf das Merkmal Wahrscheinlichkeit, daß ein vorhandenes Merkmal erkannt wird, Treffer-Rate

Spezifität: $TN / \text{Zahl der Nicht-Merkmalsträger} = TN / (TN + FP)$

Wie eingrenzend (**spezifisch**) reagiert der Test nur auf das vorhandene Merkmal und nicht auch sonst noch. Wahrscheinlichkeit, daß ein nicht-vorhandenes Merkmal korrekt als das erkannt wird. 1 - Falschalarmrate

In der Biometrie definiert man:

False Match Rate: $FP / (FP + TN)$

False Non Match Rate: $FN / (FN + TP)$

4. Spezifität und Sensitivität

Maßzahlen für die Güte von Tests

		True condition			
		Condition positive	Condition negative		
Predicted condition	Total population			Prevalence = $\frac{\sum \text{Condition positive}}{\sum \text{Total population}}$	Accuracy (ACC) = $\frac{\sum \text{True positive} + \sum \text{True negative}}{\sum \text{Total population}}$
	Predicted condition positive	True positive	False positive, Type I error	Positive predictive value (PPV), Precision = $\frac{\sum \text{True positive}}{\sum \text{Predicted condition positive}}$	False discovery rate (FDR) = $\frac{\sum \text{False positive}}{\sum \text{Predicted condition positive}}$
	Predicted condition negative	False negative, Type II error	True negative	False omission rate (FOR) = $\frac{\sum \text{False negative}}{\sum \text{Predicted condition negative}}$	Negative predictive value (NPV) = $\frac{\sum \text{True negative}}{\sum \text{Predicted condition negative}}$
		True positive rate (TPR), Recall, Sensitivity, probability of detection, Power = $\frac{\sum \text{True positive}}{\sum \text{Condition positive}}$	False positive rate (FPR), Fall-out, probability of false alarm = $\frac{\sum \text{False positive}}{\sum \text{Condition negative}}$	Positive likelihood ratio (LR+) = $\frac{\text{TPR}}{\text{FPR}}$	Diagnostic odds ratio (DOR) = $\frac{\text{LR+}}{\text{LR-}}$
		False negative rate (FNR), Miss rate = $\frac{\sum \text{False negative}}{\sum \text{Condition positive}}$	Specificity (SPC), Selectivity, True negative rate (TNR) = $\frac{\sum \text{True negative}}{\sum \text{Condition negative}}$	Negative likelihood ratio (LR-) = $\frac{\text{FNR}}{\text{TNR}}$	
				F ₁ score = $2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$	

Abb. 12: 14 verschiedene Maßzahlen für die Güte von Tests. Die Abbildung verdeutlicht die Notwendigkeit, in wissenschaftlichen Arbeiten die benutzten Maßzahlen und ihre Definitionen **genau** anzugeben, damit eine korrekte Beurteilung eines Verfahrens möglich wird. Unabhängig von den spezifisch genutzten Zahlen werden immer mindestens zwei Bewertungsrichtungen erforderlich werden. © Rechte siehe Anhang.

Spezifität und Sensitivität als Tradeoff

Man will eine **hohe Sensitivität**: Ein Test soll alle Merkmalsträger herausfinden.

Man will eine **hohe Spezifität**: Der Test soll nur auf das Merkmal reagieren.

Dilemma: Die Ziele sind einander gegenläufig.

Ein-Parameter Tests

Bei sogenannten **Ein-Parameter Tests** ermittelt man einen Untersuchungswert und muß dann eine Grenze festlegen, ab der das Merkmal als zutreffend angesehen wird.

Beispiel: Fingerabdrücke werden mit einem Template verglichen. Dabei wird ein Distanzmaß ausgegeben, das im Intervall $[a, b]$ liegt.

- **"Gleich" nahe bei a :** Kaum 2 Abdrücke gelten als gleich. Test ist sehr spezifisch aber wenig sensitiv. Kaum ein richtiger Finger wird als richtig erkannt.
- **"Gleich" nahe bei b :** Sehr viele Abdrücke werden als gleich erkannt. Test ist sehr sensitiv aber wenig spezifisch. Kaum ein falscher Finger wird zurückgewiesen.
- **"Gleich" in der Mitte:** Kompromiß zwischen beiden Aspekten

Biometrie: Nutzt in den meisten Anwendungen den *point of equal error rate*.
False Match Rate = False Non Match Rate

4. Spezifität und Sensitivität

Receiver Operating Curve ROC

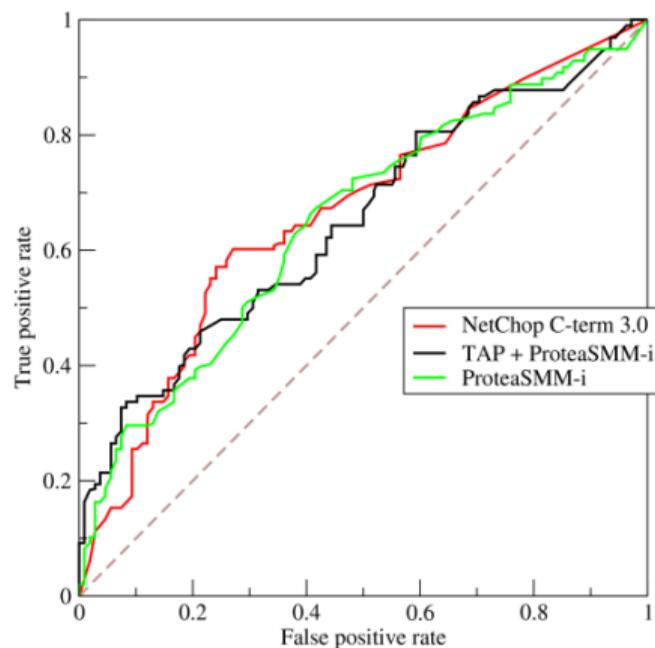


Abb. 13: Die Receiver Operating Curve (ROC) eines Ein-Parameter-Tests gibt für verschiedene Schwellwerte des Parameters die false positive und die true positive rate an. Hier: Vergleich dreier medizinischer Ein-Parameter-Tests.

© Rechte siehe Anhang.

4. Spezifität und Sensitivität

Area under the Curve

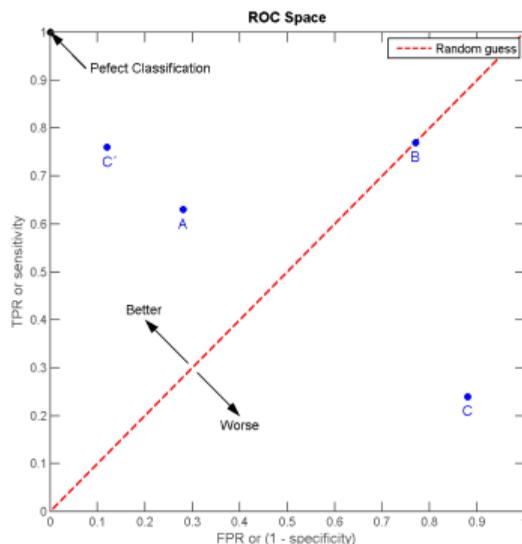


Abb. 14: Kann man eine vollständige ROC nicht angeben, ist **mindestens** der Vergleich mit **(1) reinem Raten** und **(2) perfekter Klassifikation** erforderlich. Als globales Kriterium kann die **Fläche unter der ROC** herangezogen werden. © Rechte siehe Anhang.

4. Spezifität und Sensitivität

Beispiel: Blamage heutiger Forschung und Medien (1)



Abb. 15: Sensationsbericht der BILD-Zeitung, Screenshot vom 4. 10. 2019., Nutzung nach Zitatrecht zum Beweis der behaupteten medialen Problematik. © Rechte siehe Anhang.

4. Spezifität und Sensitivität

Beispiel: Blamage heutiger Forschung und Medien (2)



Abb. 16: Sensationsbericht der BILD-Zeitung, Screenshot vom 4. 10. 2019., Nutzung nach Zitatrecht zum Beweise der behaupteten medialen Problematik. © Rechte siehe Anhang.

4. Spezifität und Sensitivität

Beispiel: Blamage heutiger Forschung und Medien (3)

Heidelberg, 21 Februar 2019



FORSCHER DES UNIVERSITÄTSKLINIKUMS HEIDELBERG
ENTWICKELN ERSTEN MARKTFÄHIGEN BLUTTEST FÜR
BRUSTKREBS

Über HeiScreen GmbH

Die HeiScreen GmbH ist eine Ausgründung des Universitätsklinikums Heidelberg. Mit einem Team aus internationalen Wissenschaftlerinnen und Wissenschaftlern, unter anderem Prof. Dr. Annette Grüters-Kieslich, zielt sie auf klinische Anwendung zu bringen.

Abb. 17: Nach dem Skandal: Aus den Pressemitteilungen der Universität Heidelberg, Screenshots vom 4.10.2019.

© Rechte siehe Anhang.

PERSONELLE ÄNDERUNGEN IM VORSTAND

UNIVERSITÄTSKLINIKUM HEIDELBERG SCHAFFT GRUNDLAGE
FÜR NEUANFANG

30 Juli 2019

Pressemitteilung des Aufsichtsrates: Die Vorstandsvorsitzende des
Universitätsklinikums Heidelberg Prof. Dr. Annette Grüters-Kieslich
und die Kaufmännische Direktorin Dipl.-Volkswirtin Irmtraut ...

[Weiterlesen >](#)

DEKAN DER MEDIZINISCHEN FAKULTÄT
HEIDELBERG IST ZURÜCKGETRETEN

25 Juli 2019

Die Medizinische Fakultät Heidelberg der Universität Heidelberg gibt
bekannt, dass Prof. Dr. Andreas Draguhn am 24. Juli 2019 mit
sofortiger Wirkung vom Amt des Dekans zurückgetreten ist. Er hat
diesen Schritt dem Rektor der Universität ...

[Weiterlesen >](#)

Abb. 18: Nach dem Skandal: Pressemitteilungen der Universität Heidelberg, Screenshot vom 4. 10. 2019.

© Rechte siehe Anhang.

Was ist passiert?

- Forscher geben im Februar 2019 eine Pressemeldung raus.
- Boulevard-Zeitungen berichten darüber.
- Dann kommt raus: Falsch-Alarm-Rate ist 46%.
- Fast die Hälfte aller gesunden Frauen würden falschen Krebsbefund erhalten.
- Aufarbeitung Februar und Mai 2019, siehe etwa:
Gerd Gigerenzer, Unstatistik 91 und Gerd Gigerenzer, Unstatistik 88

4. Spezifität und Sensitivität

Beispiel: Blamage heutiger Forschung und Medien (5)

Krebs-Früherkennung

Bluttest erkennt 10 Tumorarten im
Frühstadium



Bluttest in einem Transfusionszentrum.

dpa/Rainer Jensen/Symbol

Montag, 03.09.2018, 09:53

Abb. 19: Was lernt "die Presse" daraus? Hier: Focus vom 3. 9. 2019, Screenshot vom 4. 10. 2019., Nutzung nach Zitatrecht zum Beweis der behaupteten medialen Problematik. © Rechte siehe Anhang.

Der „Heilige Gral der Krebsforschung“ ist oft kaum besser als ein Münzwurf

Unstatistik vom 04.10.2019

Die Unstatistik des Monats September ist die angebliche Treffsicherheit des Bluttests für zehn verschiedene Krebsarten im Frühstadium, den Forscher der Cleveland Clinic entwickelt haben. Über ihn berichtete unter anderem die Online-Ausgabe des „Focus“ am 3. September unter dem Titel „[Bluttest erkennt 10 Tumorarten im Frühstadium](#)“ und spricht von einer Treffsicherheit bis 90 Prozent. Der Studienleiter nennt den Test im Artikel den „Heiligen Gral der Krebsforschung“. Dabei wird jedoch wie schon beim Heidelberger Brustkrebs-Bluttest (siehe Unstatistik Februar und Mai 2019) lediglich die Trefferrate genannt, nicht jedoch die Falsch-Alarm-Rate. Es bleibt daher völlig unklar, wie häufig der Test bei gesunden Personen fälschlicherweise Krebs diagnostiziert.

Abb. 20: Aufarbeitung durch [Gerd Gigerenzer](#), [Unstatistik 95](#), Screenshot vom 4. 10. 2019. © Rechte siehe Anhang.

Disclaimer und Empfehlung

Disclaimer

Die bewußt ausgewählten negativen Lehr-Beispielen bilden eine didaktisch motivierte, (hoffentlich) nicht-repräsentative Auswahl wissenschaftlicher und medialer Arbeitsweisen.

Empfehlung (von mir als Person! – ich bin kein Arzt)

Vor einer invasiven, möglicherweise risikoreichen Vorsorgeuntersuchung:

- Erkundigen Sie sich, ob Ihr Arzt die hier beschriebene Problematik auch **kennt**.
- Erkundigen Sie sich nach **Sensitivität** und **Spezifität** des Tests.
- Erkundigen Sie sich nach dem **Risiko der Vorsorgeuntersuchung**.
- Erkundigen Sie sich nach besonderen **persönlichen Risikofaktoren**.

^a**Disclaimer:** Es handelte sich hier um eine **didaktisch dramatisierte Veranschaulichung** einer statistischen Problematik, die eine Beratung durch einen Arzt mit Kenntnis Ihrer persönlichen Krankengeschichte nicht ersetzen kann!

Persönliches Fazit

- Mein Vertrauen in das **Sozialsystem** der Wissenschaftswelt ist gesunken.
- Ich erkenne nicht nur im Boulevard sondern auch in den Wissenschaften Mechanismen des **click bait** im Kampf um Reputation, Geld und Macht.
- Ich finde zu Sätzen, welche die Worte **fake news**, **alternative facts** oder **Lügenpresse** enthalten, sehr viele verschiedene Interpretationsansätze.
- Ich bin mit **Begriffen**, **Tatsachen**, **Fakten**, **Beweisen** vorsichtiger geworden, sehe, wie diese Begriffe aus Unkenntnis oder für Manipulation mißbraucht werden.
- Mein Bedürfnis über die Dinge **selber nachzudenken** ist gestiegen, wie auch meine Erkenntnis, das **oft nicht zu können**¹.
- Ich bin nicht mehr bereit, fremde Erkenntnisse als "Wahrheiten" zu akzeptieren.
- Interesse an allem, was eigene Sichtweise in Frage stellen könnte, ist am Steigen.

¹Drei Kern-Ursachen: (1) Eigenes Vorwissen stets stark begrenzt, (2) Zugang zu empirischen Daten fast immer nur mittelbar über Dritte – benötigt Vertrauen in deren Integrität. (3) Zugang zu empirischen Daten oft vorgefiltert und vorinterpretiert.

5. Law of Small Numbers

Beachte: Bei kleinen Samples gibt es eher Ausreißer als bei großen Samples

1. Simpsonsches Paradoxon
2. Berksonsches Paradoxon
3. Kontextualisierung
4. Spezifität und Sensitivität
5. Law of Small Numbers
6. Schlußfolgerung

Beispiel: Krebsrate (1)

Untersuchung der Krebsraten in den 3141 Counties der USA.

Ergebnis: Die Counties mit **minimaler** Krebsrate sind **ländliche** Counties.

Wie erklären wir uns das? Welche Vorurteile rufen wir kognitiv ab?

- Klare Luft
- Besseres Wasser
- Weniger Stress im Beruf
- Gesundere Ernährung
- Mehr Möglichkeiten zu Bewegung in frischer Luft
- ...

Beispiel: Krebsrate (2)

Untersuchung der Krebsraten in den 3141 Counties der USA.

Ergebnis: Die Counties mit **maximaler** Krebsrate sind **ländliche** Counties.

Wie erklären wir uns das? Welche Vorurteile rufen wir kognitiv ab?

- Schlechtere medizinische Betreuung
- Größere Armut der Landbevölkerung
- Schlechtere Lebensverhältnisse
- Mehr Genuß von fettem Fleisch aus der lokalen Landwirtschaft
- Geringerer Bildungsstandard am Land
- ...

Beispiel: Krebsrate (3)

Untersuchung der Krebsraten in den 3141 Counties der USA

Ergebnis: Die Counties mit **maximaler** Krebsrate sind **ländliche** Counties.

Ergebnis: Die Counties mit **minimaler** Krebsrate sind **ländliche** Counties.

Was denn nun??? Ist das Land nun gut oder schlecht???

Es gilt beides.

Ländliche Counties haben weniger Einwohner.

Law of Small Numbers:

Bei kleineren Samples gibt es eher Ausreißer als bei großen Samples.

Zusätzliches Problem hier: Methodische Verwechslung von Korrelation und Kausation.

6. Schlußfolgerung

Leitsatz: Man sollte jeder Statistik mit konstruktiver Skepsis begegnen.

1. Simpsonsches Paradoxon
2. Berksonsches Paradoxon
3. Kontextualisierung
4. Spezifität und Sensitivität
5. Law of Small Numbers
6. Schlußfolgerung

Mahnung zur Vorsicht

Wesentliche Dinge geschehen oft **weit vor** der statistischen Auswertung.

Beispiel: Wer darf ab welchem Alter in einer Demokratie wählen gehen?
Sollen später eintreffende Briefwahlstimmen auch zählen?
Wichtiges Stichwort für Interessierte: Gerrymandering

Über die Fragen der Gruppeneinteilung (Simpson) oder Vorselektion (Berkson) kann und wird man emotionalisiert streiten.

Das Endergebnis gilt dann als "wissenschaftlich gesichert", wurde aber durch Vorfeldentscheidungen herbeigeführt.

Vorsicht!

Ein "folge den Aussagen der Wissenschaft" wird brandgefährlich, wenn emotionalisiert entschiedene ideologische Fragen die Antworten angeblich objektiver Auswertungsmethoden bestimmen und gesellschaftliche Entscheidungen damit den den falschen Status wissenschaftlicher Wahrheiten erhalten und deshalb als alternativlos angepriesen werden.

6. Schlußfolgerung

Schlußfolgerung

Falsche Schlußfolgerung: Destruktive Skepsis

"Im übrigen glaube ich nur an die Statistik, die ich selbst gefälscht habe."

Bischof Otto Dibelius. In: Anekdoten um Bischof Dibelius: Geist und Witz eines großen Kirchenmannes, Bechtle Verlag, 1967, S. 79. Hier zitiert nach [Blogspot: Falschzitate](#).

Ist eine **fahrlässige & irreführende Diskreditierung** eines wissenschaftlichen Instruments.

Richtige Schlußfolgerung: Konstruktive Skepsis

Statistik ist ein wichtiges mathematisches **Auswertungsinstrument**.

Nur Kenntnis der Mechanismen & sog. "Paradoxa" erlaubt die richtigen Schlüsse.

Jede statistische Auswertung muß hinterfragt & diskutiert werden. Das erfordert:

- Vollständige **Dokumentation** der Datenerfassung
- Öffentliche **Verfügbarkeit** aller erfassten Primärdaten
- Vollständige **Dokumentation** der statistischen Auswertung
- Statistische **Sachkenntnis**

Anhang

Übersicht

Verzeichnis aller Abbildungen

Abb

Verzeichnis aller Tabellen

Tab

Rechtsnachweise

©

Rechtliche Hinweise

§

Zitierweise dieses Dokuments

→

Index

Index

Verzeichnis aller Folien



Verzeichnis aller Abbildungen (1/2)

1	Simpsonsches Paradoxon bei zwei Gruppen	8
2	Simpsonsches Paradoxon bei 5 Gruppen (1)	9
3	Simpsonsches Paradoxon bei 5 Gruppen (2)	9
4	Medikamentendosierung	10
5	Berksonsches Paradoxon	16
6	Die Rohdaten	18
7	Die erste Bearbeitung	19
8	Die zweite Bearbeitung	20
9	Die dritte Bearbeitung	21

10	Die Rohdaten	22
11	Die bearbeiteten Daten	22
12	14 Maßzahlen für die Güte von Tests	28
13	Receiver Operating Curve eines Ein-Parameter Tests	31
14	Area under the Curve	32
15	Sensationsbericht (1)	33
16	Sensationsbericht (2)	34
17	Pressemitteilung (1)	35
18	Pressemitteilung (2)	35
19	Sensationsbericht (3)	37
20	Unstatistik 95	38

1 Rohdaten für Simpsonsches Paradoxon mit Bart und Lisa.....	5
2 Bewertung von Lisa	6
3 Bewertung von Bart.....	6

Abb. 1 Quelle: https://commons.wikimedia.org/wiki/File:Simpson%27s_paradox_continuous.svg, Schutz, gemeinfrei.

Abb. 2 Quelle: https://commons.wikimedia.org/wiki/File:Simpsons_paradox_-_animation.gif Pace Swwiki, Nutzung nach CC SA 4.0

Abb. 3 Quelle: https://commons.wikimedia.org/wiki/File:Simpsons_paradox_-_animation.gif Pace Swwiki, Nutzung nach CC SA 4.0

Abb. 4 Quelle: https://commons.wikimedia.org/wiki/File:Simpson%27s_paradox_continuous.svg, Schutz, gemeinfrei.

Abb. 5 Quelle: <https://towardsdatascience.com/statistic-has-made-my-day-or-has-it-982c7d527410>, Alamsyah Hanza, 2018.

Abb. 6 Quelle: http://www.mesosworld.ch/lerninhalte/Univ_GrafischDarst/de/html/unit_ManipulDiagr.html

Abb. 7 Quelle: http://www.mesosworld.ch/lerninhalte/Univ_GrafischDarst/de/html/unit_ManipulDiagr.html

Abb. 8 Quelle: http://www.mesosworld.ch/lerninhalte/Univ_GrafischDarst/de/html/unit_ManipulDiagr.html

Abb. 9 Quelle: http://www.mesosworld.ch/lerninhalte/Univ_GrafischDarst/de/html/unit_ManipulDiagr.html

Abb. 10 Quelle: mesosworld.ch

Abb. 11 Quelle: mesosworld.ch

Abb. 12 Quelle: https://en.wikipedia.org/wiki/Confusion_matrix, Nutzung nach CC SA 3.0, https://en.wikipedia.org/wiki/Wikipedia:Text_of_Creative_Commons_Attribution-ShareAlike_3.0_Unported_License

Abb. 13 Quelle: <https://commons.wikimedia.org/w/index.php?curid=10714489>, BOR, Nutzung nach CC BY-SA 3.0

Abb. 14 Quelle: <https://commons.wikimedia.org/w/index.php?curid=8326140> Abbildung von Kai walz, Nutzung nach CC BY-SA 3.0

Abb. 15 Quelle: <https://www.bild.de/ratgeber/2019/gesundheit/bluttest-erkennt-brustkrebs-warum-dieser-test-eine-weltsensation-ist-60269760.bild.html>

Abb. 16 Quelle: <https://www.bild.de/bild-plus/ratgeber/2019/gesundheit/erster-blut-test-erkennt-zuverlaessig-brustkrebs-60239512>

Abb. 17 Quelle: <https://www.klinikum.uni-heidelberg.de/newsroom/forscher-des-universitatsklinikums-heidelberg-entwickeln-ersten-marktfahigen-bluttest-fur-brustkrebs/>

Abb. 18 Quelle: <https://www.klinikum.uni-heidelberg.de/newsroom/alle-nachrichten-zum-thema-heiscreeen/>

Abb. 19 Quelle: https://www.focus.de/gesundheit/ratgeber/krebs/diagnose/bluttest-erkennt-10-krebsarten-forscher-begeistert-wir-haben-hier-vielleicht-den-heiligen-gral-der-krebsdiagnose_id_9022932.html, Nutzung nach Zitatrecht zum Nachweis der Situation

Abb. 20 Quelle: <http://www.rwi-essen.de/unstatistik/95/>

Die hier angebotenen Inhalte unterliegen deutschem Urheberrecht. Inhalte Dritter werden unter Nennung der Rechtsgrundlage ihrer Nutzung und der geltenden Lizenzbestimmungen hier angeführt. Auf das Literaturverzeichnis wird verwiesen. Das **Zitatrecht** in dem für wissenschaftliche Werke üblichen Ausmaß wird beansprucht. Wenn Sie eine Urheberrechtsverletzung erkennen, so bitten wir um Hinweis an den auf der Titelseite genannten Autor und werden entsprechende Inhalte sofort entfernen oder fehlende Rechtsnennungen nachholen. Bei Produkt- und Firmennamen können Markenrechte Dritter bestehen. Verweise und Verlinkungen wurden zum Zeitpunkt des Setzens der Verweise überprüft; sie dienen der Information des Lesers. Der Autor macht sich die Inhalte, auch in der Form, wie sie zum Zeitpunkt des Setzens des Verweises vorlagen, nicht zu eigen und kann diese nicht laufend auf Veränderungen überprüfen.

Alle sonstigen, hier nicht angeführten Inhalte unterliegen dem Copyright des Autors, Prof. Dr. Clemens Cap, ©2020. Wenn Sie diese Inhalte nützlich finden, können Sie darauf verlinken oder sie zitieren. Jede weitere Verbreitung, Speicherung, Vervielfältigung oder sonstige Verwertung außerhalb der Grenzen des Urheberrechts bedarf der schriftlichen Zustimmung des Rechteinhabers. Dieses dient der Sicherung der Aktualität der Inhalte und soll dem Autor auch die Einhaltung urheberrechtlicher Einschränkungen wie beispielsweise **Par 60a UrhG** ermöglichen.

Die Bereitstellung der Inhalte erfolgt hier zur persönlichen Information des Lesers. Eine Haftung für mittelbare oder unmittelbare Schäden wird im maximal rechtlich zulässigen Ausmaß ausgeschlossen, mit Ausnahme von Vorsatz und grober Fahrlässigkeit. Eine Garantie für den Fortbestand dieses Informationsangebots wird nicht gegeben.

Die Anfertigung einer persönlichen Sicherungskopie für die private, nicht gewerbliche und nicht öffentliche Nutzung ist zulässig, sofern sie nicht von einer offensichtlich rechtswidrig hergestellten oder zugänglich gemachten Vorlage stammt.

Wenn Sie Inhalte aus diesem Werk nutzen oder darauf verweisen wollen, zitieren Sie es bitte wie folgt:

Clemens H. Cap: Statistik. Electronic document. <https://iuk.one/1012-1047> 8. 3. 2021.

Bibtex Information: <https://iuk.one/1012-1047.bib>

```
@misc{doc:1012-1047,  
  author      = {Clemens H. Cap},  
  title       = {Statistik},  
  year        = {2021},  
  month       = {3},  
  howpublished = {Electronic document},  
  url         = {https://iuk.one/1012-1047}  
}
```

Typographic Information:

Typeset on March 8, 2021

This is pdfTeX, Version 3.14159265-2.6-1.40.21 (TeX Live 2020) kpathsea version 6.3.2

This is pgf in version 3.1.5b

This is preamble-slides.tex myFormat©C.H.Cap

area under the curve, 32

Auswahl im College, 15

Bart und Lisa, 5

Berksonsches Paradoxon, 15

Brustkrebs, 24

Ein-Parameter-Tests, 30

Equal error rate, 30

Falschalarm-Rate, 26

false match rate, 27

false negative, 27

false non match rate, 27

false positive, 27

Kontextualisierung, 18

Kontingenztafel, 27

Konventionalismus, 10

Krebsrate in den USA, 42

Kriminalstatistik, 11

law of small numbers, 44

Medikamentendosierung, 10

receiver operating curve, 31

Sensitivität, 27

Simpsonsches Paradoxon, 5, 13

Spezifität, 27

Treffer-Rate, 26

true negative, 27

true positive, 27

Umsatzzahlen, 18

Visualisierung, 18

- 1 Titelseite
- 2 Ziel
- 3 Inhaltsverzeichnis

1. Simpsonsches Paradoxon

- 5 Beispiel: Bart und Lisa (1)
- 6 Beispiel: Bart und Lisa (2)
- 7 Beispiel: Bart und Lisa (3)
- 8 Visuelle Erläuterung (1)
- 9 Visuelle Erläuterung (2)
- 10 Beispiel: Medikamentendosierung
- 11 Beispiel: Kriminalstatistik in Mittelstadt (1)
- 12 Beispiel: Kriminalstatistik in Mittelstadt (2)
- 13 Theoretische Einordnung

2. Berksonsches Paradoxon

- 15 Beispiel: Auswahl im College (1)
- 16 Beispiel: Auswahl im College (2)

3. Kontextualisierung

- 18 Umsatzzahlen ohne Nachbearbeitung
- 19 Abschneiden vertikaler Achsen
- 20 Ausblenden nicht benötigter Abstände
- 21 Zusammenfassung zu günstigen Gruppen
- 22 Nochmal im direkten Vergleich

4. Spezifität und Sensitivität

- 24 Beispiel: Krebstest (1)
- 25 Beispiel: Krebstest (2)
- 26 Treffer-Rate und Falschalarm-Rate
- 27 Spezifität und Sensitivität
- 28 Maßzahlen für die Güte von Tests
- 29 Spezifität und Sensitivität als Tradeoff
- 30 Ein-Parameter Tests
- 31 Receiver Operating Curve ROC
- 32 Area under the Curve
- 33 Beispiel: Blamage heutiger Forschung und Medien (1)
- 34 Beispiel: Blamage heutiger Forschung und Medien (2)
- 35 Beispiel: Blamage heutiger Forschung und Medien (3)
- 36 Beispiel: Blamage heutiger Forschung und Medien (4)
- 37 Beispiel: Blamage heutiger Forschung und Medien (5)
- 38 Beispiel: Blamage heutiger Forschung und Medien (6)
- 39 Disclaimer und Empfehlung
- 40 Persönliches Fazit

5. Law of Small Numbers

- 42 Beispiel: Krebsrate (1)
- 43 Beispiel: Krebsrate (2)
- 44 Beispiel: Krebsrate (3)

6. Schlußfolgerung

46 Mahnung zur Vorsicht

47 Schlußfolgerung

Legende:

 Fortsetzungsseite

 Seite ohne Überschrift

 Bildseite