

Studienplanung, -auswertung

Dr. Matthias Kohl
SIRS-Lab GmbH

Ausgangspunkt

Fragestellung(en)/Hypothese(n):

- Hauptfragestellung: Grund für Durchführung der Studie
- Nebenfragestellung(en): Welche Fragestellungen lassen sich zusätzlich beantworten

Empfehlung:

Untersuchung einer präzise formulierten Fragestellung

Studien-Design Erkenntnispyramide

2. Fallstudie
3. Fälle ohne Kontrollen
4. Fälle mit Literaturkontrollen
5. Auswertung von Registerdatenbanken
6. Beobachtungsstudien mit Kontrollgruppe
7. Klinische Studie mit historischen Kontrollen
8. Eine randomisierte kontrollierte klinische Studie
9. Mehrerer randomisierte kontrollierte klinische Studien

Kategorisierung der Studientypen

deskriptiv – analytisch

retrospektiv – prospektiv

transversal – longitudinal

beobachtend – experimentell

Definition der Studienpopulation

- Festlegung der Ein- und Ausschlußkriterien
- Wie werden die Beobachtungseinheiten (i.d.R. Patienten, Probanden) ausgewählt?
- **Randomisierungstechniken:** Wie werden die Beobachtungseinheiten auf die Gruppen aufgeteilt?

Randomisierungstechniken

- **Einfache Randomisation**

(zufällige Zuordnung zu Behandlungsgruppen)

- **Blockrandomisation**

(Behandlungsgruppen gleich groß)

- **Geschichtete Randomisation**

(Schichten von prognostischen Faktoren)

Weitere Schritte

- Festlegung der **Ver- und Entblindung** (einfach, doppelt oder dreifach blind)
- **Fallzahlplanung!**
- **Beschreibung des biometrischen Auswertungsvorgangs!** (inkl. statistischer Schätz- und Testverfahren)

Weitere Angaben im Studienplan (vgl. Schäfer et al. (1999))

- Ermittlung, Dokumentation und Bewertung unerwünschter Ereignisse
- Zeitlicher Studienablauf
- Kriterien für den Abbruch der Studie
- Verfahren zur Kontrolle der Einhaltung des Protokolls und Korrektheit der Dokumentation (Monitoring)
- Ethische und rechtliche Basis
- Quellenangaben

Ver- und Entblindung

Zuordnung der Studienteilnehmer ist

- allen bekannt – **offen**
- dem Teilnehmer nicht bekannt – **einfachblind**
- dem Teilnehmer und Untersucher nicht bekannt – **doppelblind**
- dem Teilnehmer, Untersucher und Statistiker nicht bekannt – **dreifachblind**

Fallzahlplanung

Beispiel: Wirksamkeit zweier Medikamente

Heilungsraten: 25 % (alt) vs. 35 % (neu) (=> **Heilung?**)

Welche Substanz ist besser?

Generell gilt: Je größer die Fallzahl ist, desto kleiner ist die Wahrscheinlichkeit, dass dies ein rein zufälliges Ereignis ist!

Aber auch: Je größer die Fallzahl ist, desto teurer wird die Studie!

Fallzahlplanung – wichtige Begriffe

- **Nullhypothese (H_0):** Hypothese, die man widerlegen will
- **Alternative (H_1):** Hypothese, die man „beweisen“ will
- **Fehler 1. Art (α):** Wahrscheinlichkeit, dass H_0 abgelehnt wird, obwohl zutreffend!
- **Fehler 2. Art (β):** Wahrscheinlichkeit, dass H_0 nicht abgelehnt wird, obwohl falsch!

Ziel der Fallzahlplanung

Fallzahl wird so groß gewählt, dass man mit
möglichst großer Wahrscheinlichkeit einen
signifikanten (klinisch relevanten)
Unterschied beobachtet; d.h. H_0 ablehnt!

Fallzahlplanung – erforderliche Größen

- **Standardtherapie:** Heilungsrate $P_0 = 25 \%$
- **Klinisch relevant:** Jede andere Therapie mit $>30 \%$ Heilungsrate
- **Fehler 1. Art (α):** meist 5% oft auch 1%
- **Fehler 2. Art (β):** meist $10 - 20 \%$
- **Neue Therapie:** Heilungsrate P_1 (z.B. durch Pilotstudie)

Berechnung der Fallzahl

Fallzahlformel für Beispiel:

$$n \geq \left[\frac{u_{1-\alpha/2} + u_{1-\beta}}{P_1 - P_0} \right]^2 [P_1(1 - P_1) + P_0(1 - P_0)]$$

$u_{1-\alpha/2}$: $(1-\alpha/2)$ -Quantil von $N(0, 1)$ (Normalverteilung)

$u_{1-\beta}$: $(1-\beta)$ -Quantil von $N(0, 1)$

D.h.: $u_{0.8} = 0.84$, $u_{0.85} = 1.04$, $u_{0.9} = 1.28$,

$u_{0.975} = 1.96$, $u_{0.995} = 2.56$

Fallzahlen im Beispiel

$$P_0 = 25 \%, \beta = 20 \%$$

	$\alpha = 5 \%$	$\alpha = 1 \%$
$P_1 = 30 \%$	1248	1857
$P_1 = 35 \%$	326	485
$P_1 = 40 \%$	149	222

Fazit: Eine Fallzahl ist i.a. schwierig zu berechnen!

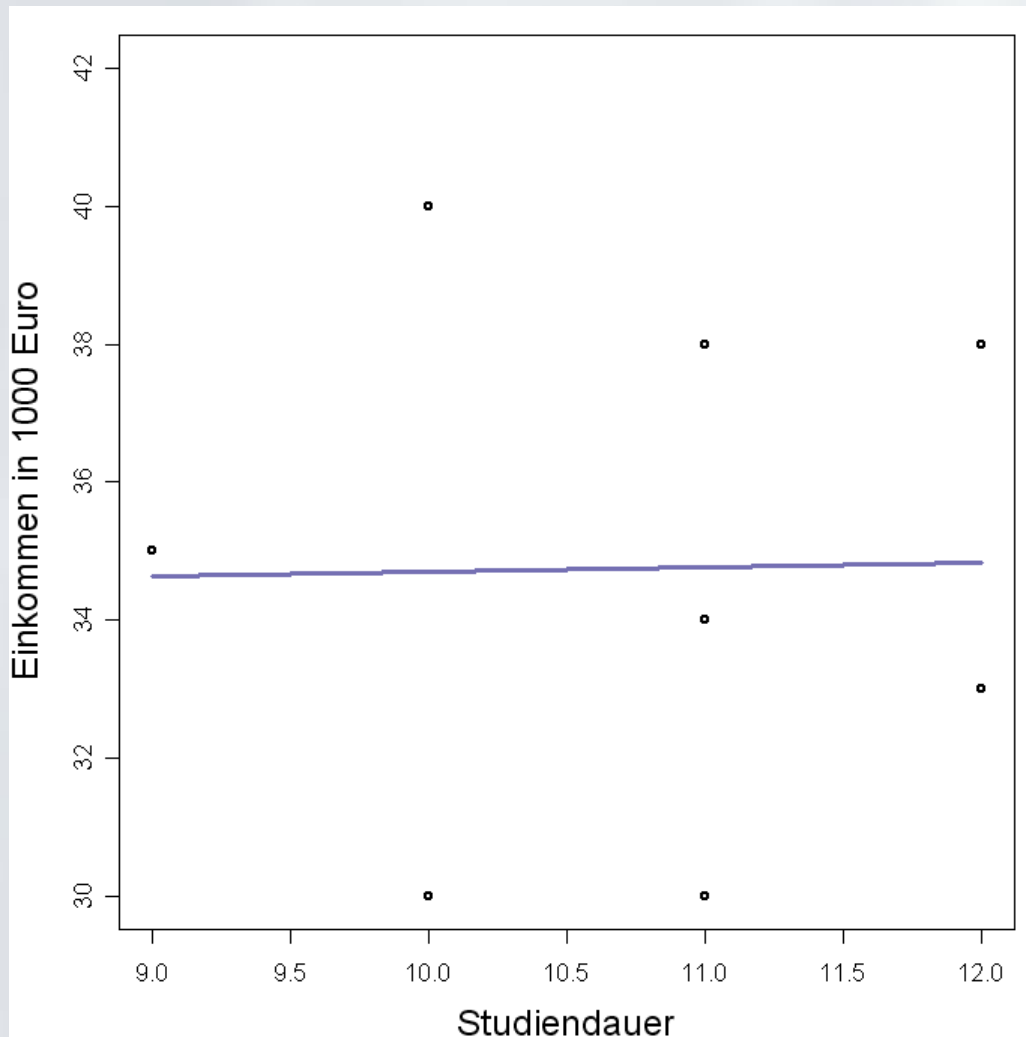
Definition aller Meßgrößen

Man unterscheidet:

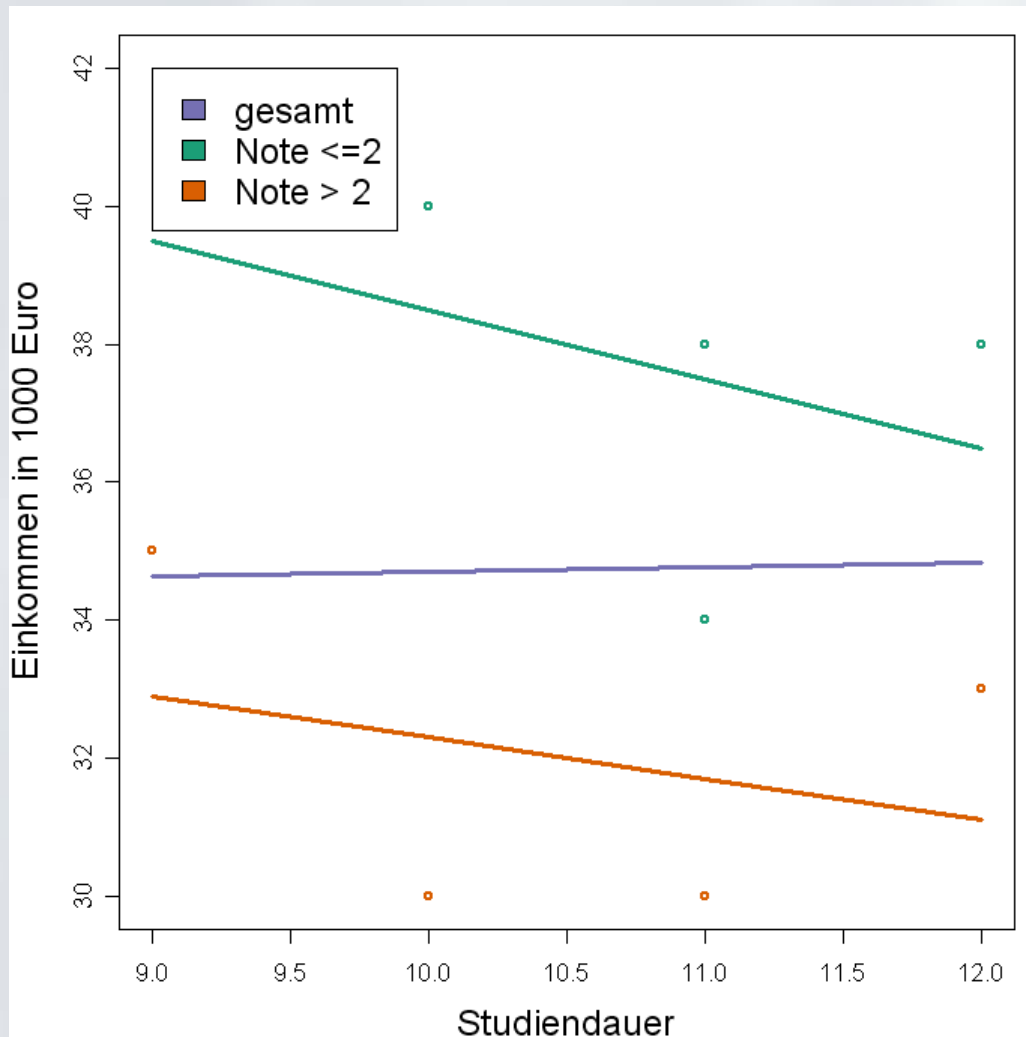
- (Haupt-)Zielgrößen
- Einflussgrößen (Faktoren)
- Störgrößen (verzerrende und nicht-verzerrende)

Ziel: Gewinne mittels der Einflussgrößen Erkenntnisse über die Zielgrößen und kontrolliere dabei die Störgrößen!

Störgrößen – Simpson Paradoxon



Störgrößen – Simpson Paradoxon



Merkmalsarten

Merkmalsart	Qualitativ (kategorial)		Quantitativ (metrisch)	
Skalenniveau	Nominalskala	Ordinalskala	Intervallskala	Ratioskala
Beispiele	Geschlecht, Blutgruppe, Rhesusfaktor	Schulnoten, med. Score	Temperatur in °C, Intelligenzquotient	Leukozytenanzahl, Körpergröße
Hinweise	niedrigstes Niveau	Rangfolge ist definiert	Skala mit Nullpunkt, Abstand ist definiert	höchstes Niveau, Verhältnis ist definiert
Vergleiche	$A = B, A \neq B$	$A = B, A \neq B, A < B, A > B$	$A = B, A \neq B, A < B, A > B, d = A - B$	$A = B, A \neq B, A < B, A > B, d = A - B, c = A:B$

Univariate Datenbeschreibung

Skalenniveau	Nominalskala	Ordinalskala	Intervallskala	Ratioskala
Lagemaße	Modus	Modus, Median, Quartile, Quantile	Modus, Median, Quartile, Quantile, Mittelwert	
Streuungsmaße	NA	Spannweite, Quartilsabstand	Spannweite, Quartilsabstand, Standardabw.	Spannweite, Quartilsabstand, Standardabw., Variationskoeff.
Formmaße	NA	NA	Schiefe, Wölbung	
Graphische Darstellung	Kreis-, Rechteck-, Balkendiagramme	Kreis-, Rechteck-, Balkendiagramme	diskrete Daten: Balkendiagramm, Stem and Leaf stetige Daten: Histogramm, Häufigkeitspolygon	

Look at the data!

Anscombe's Quartet

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3,000	1,125	2,667	0,026
x1	0,500	0,118	4,241	0,002

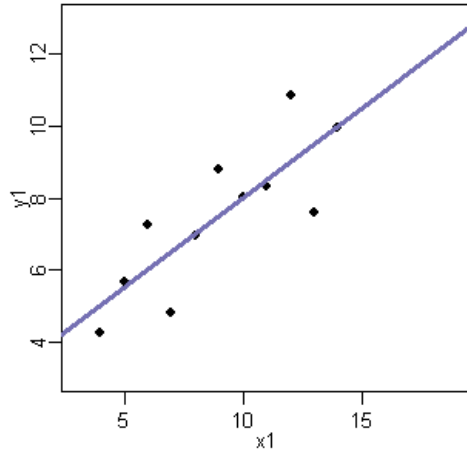
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3,001	1,125	2,667	0,026
x2	0,500	0,118	4,239	0,002

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3,002	1,124	2,670	0,026
x3	0,500	0,118	4,239	0,002

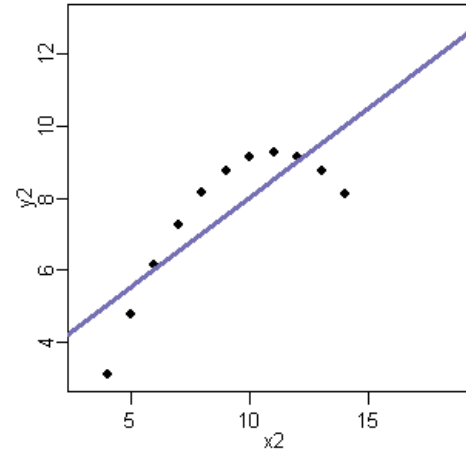
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3,002	1,124	2,671	0,026
x4	0,500	0,118	4,243	0,002

Anscombe's Quartet

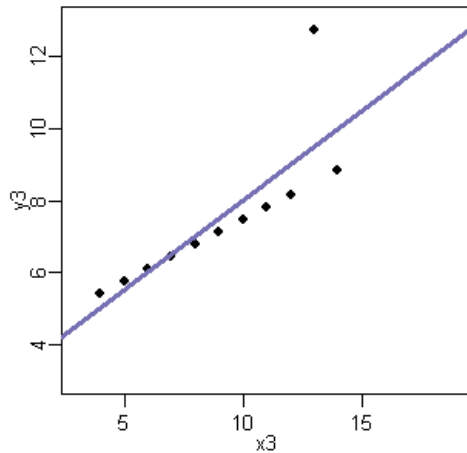
o.k.



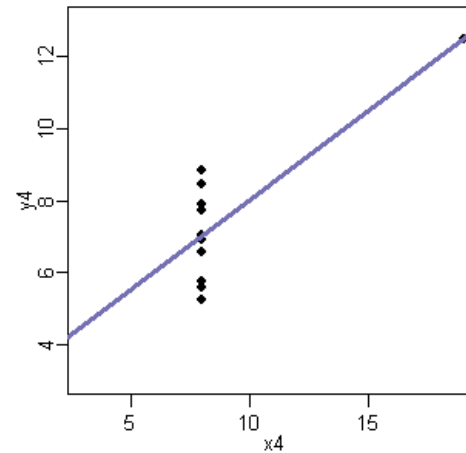
Modell?



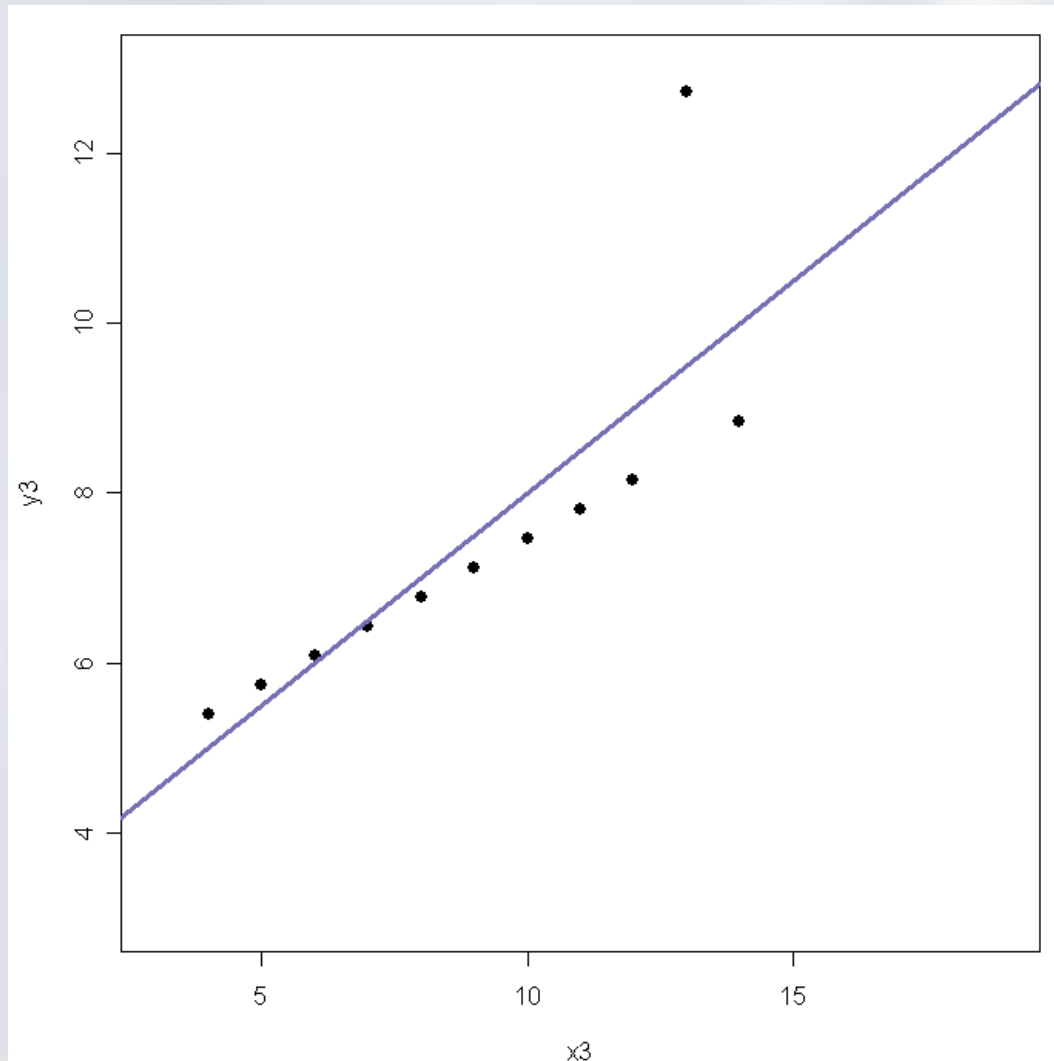
Ausreißer?



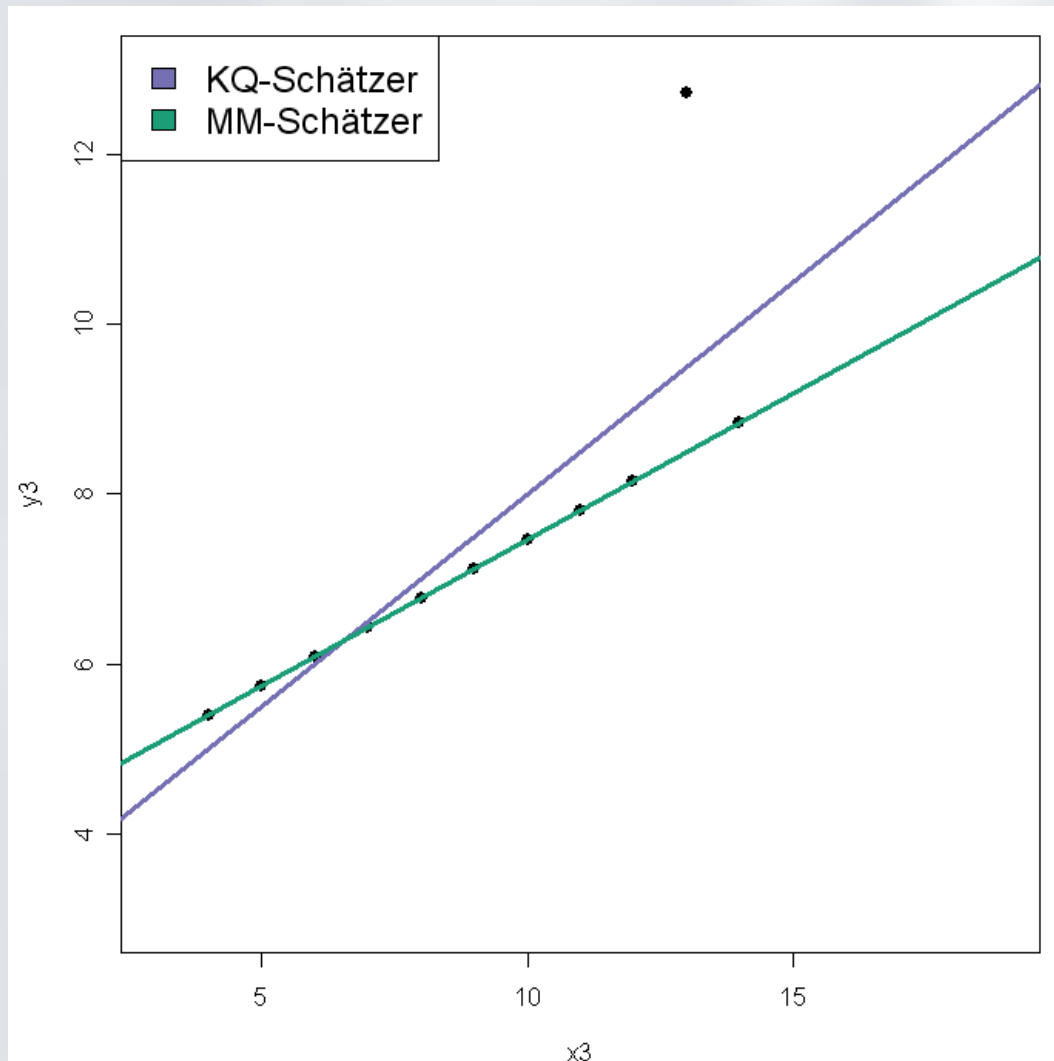
Design?



Klassische Regressionsanalyse



Robuste Regressionsanalyse



Korrelationsanalyse

Skalenniveau	Art des Zusammenhangs	Nachweis des Zusammenhangs	graphische Darstellung
metr. - metr.	linear	Korrelationskoeffizient nach Pearson	Punktwolke mit Regressionsgerade
metr. - metr., metr. - ordinal	monoton	Korrelationskoeffizient nach Spearman	Punktwolke
metr. - ordinal, ordinal - ordinal	monoton	Korrelationskoeffizient nach Kendall	Punktwolke
nominal - nominal, nominal - ordinal	kontingent	Kontingenzkoeffizient	Balkendiagramm
nominal - nominal	kontingent	Yule'scher Assoziationskoeffizient	Balkendiagramm
nominal - metrisch, nominal - ordinal	kontingent	statistische Tests	Balkendiagramm

Vergleich mehrerer Gruppen

- X wahrscheinlich größer als Y; d.h.

$$WS(X > Y) > 0.5$$

- Y wahrscheinlich größer als Z; d.h.

$$WS(Y > Z) > 0.5$$

⇒ X wahrscheinlich größer als Z; d.h.

$$WS(X > Z) > 0.5$$

NEIN!

Gegenbeispiel

- X nimmt die Werte 9, 5, 1 mit WS $1/3$ an
 - Y nimmt die Werte 8, 4, 3 mit WS $1/3$ an
 - Z nimmt die Werte 7, 6, 2 mit WS $1/3$ an
- X, Y, Z stochastisch unabhängig

Dann:

$$WS(X > Y) = WS(Y > Z) = \text{WS } (X < Z) = 5/9$$

Hypothesentests

- **Nullhypothese (H_0):** Hypothese, die man widerlegen will
- **Alternative (H_1):** Hypothese, die man „beweisen“ will
- **Fehler 1. Art (α):** Wahrscheinlichkeit, dass H_0 abgelehnt wird, obwohl zutreffend!
- **Fehler 2. Art (β):** Wahrscheinlichkeit, dass H_0 nicht abgelehnt wird, obwohl falsch!

Hypothesentests

	H0 ist richtig	H1 ist richtig
Entscheidung für H_0	Richtige Entscheidung $1 - \alpha$ (Sensitivität)	Fehler 2. Art β
Entscheidung für H_1	Fehler 1. Art α (Signifikanzniveau)	Richtige Entscheidung $1 - \beta$ (power, Spezifität)

Tests für eine Stichprobe

Name	Art	Voraussetzungen
t-Test	Lagetest	(approx.) Normalverteilung
Wilcoxon-Test	Lagetest	symm. Verteilung
Binomialtest	Wahrscheinlichkeitstest	Alternativmerkmal
Exakter Test von Fisher	Unabhängigkeitstest	2 Alternativmerkmale
χ^2 -Test	Unabhängigkeitstest	2 Alternativmerkmale

Tests für zwei verbundene Stichproben

Name	Art	Voraussetzungen
t-Test	Lagetest	(approx.) Normalverteilung der Differenzen
Wilcoxon-Test	Lagetest	symm. Verteilung der Differenzen
Vorzeichentest	Lagetest	NA
McNemar-Test	Homogenitätstest	1 Alternativmerkmal

Tests für zwei unverbundene Stichproben

Name	Art	Voraussetzungen
t-Test	Lagetest	(approx.) Normalverteilung mit gleicher Varianz
Welch-t-Test	Lagetest	(approx.) Normalverteilung
Mann-Whitney U-Test	Lagetest	gleiche Verteilungsform
Median-Test	Lagetest	ordinal skalierte Merkmale
Vierfelder-Test	Homogenitätstest	1 Alternativmerkmal
Logrank-Test	Überlebenszeiten	2 Überlebenskurven

Literatur

- **J. A. Knottnerus (Hrsg.) (2002): The Evidence Base of Clinical Diagnosis.** BMJ Books. London.
- **H. Schäfer et al. (1999): Empfehlungen für die Erstellung von Studienprotokollen (Studienplänen) für klinische Studien.** Informatik, Biometrie und Epidemiologie in Medizin und Biologie 30 (3); 141-154.
- **H. J. Trampisch und J. Windeler (Hrsg.) (2000): Medizinische Statistik.** Springer Verlag. 2. Auflage.
- **C. Weiß (2002): Basiswissen Medizinische Statistik.** Springer Verlag. 2. Auflage.