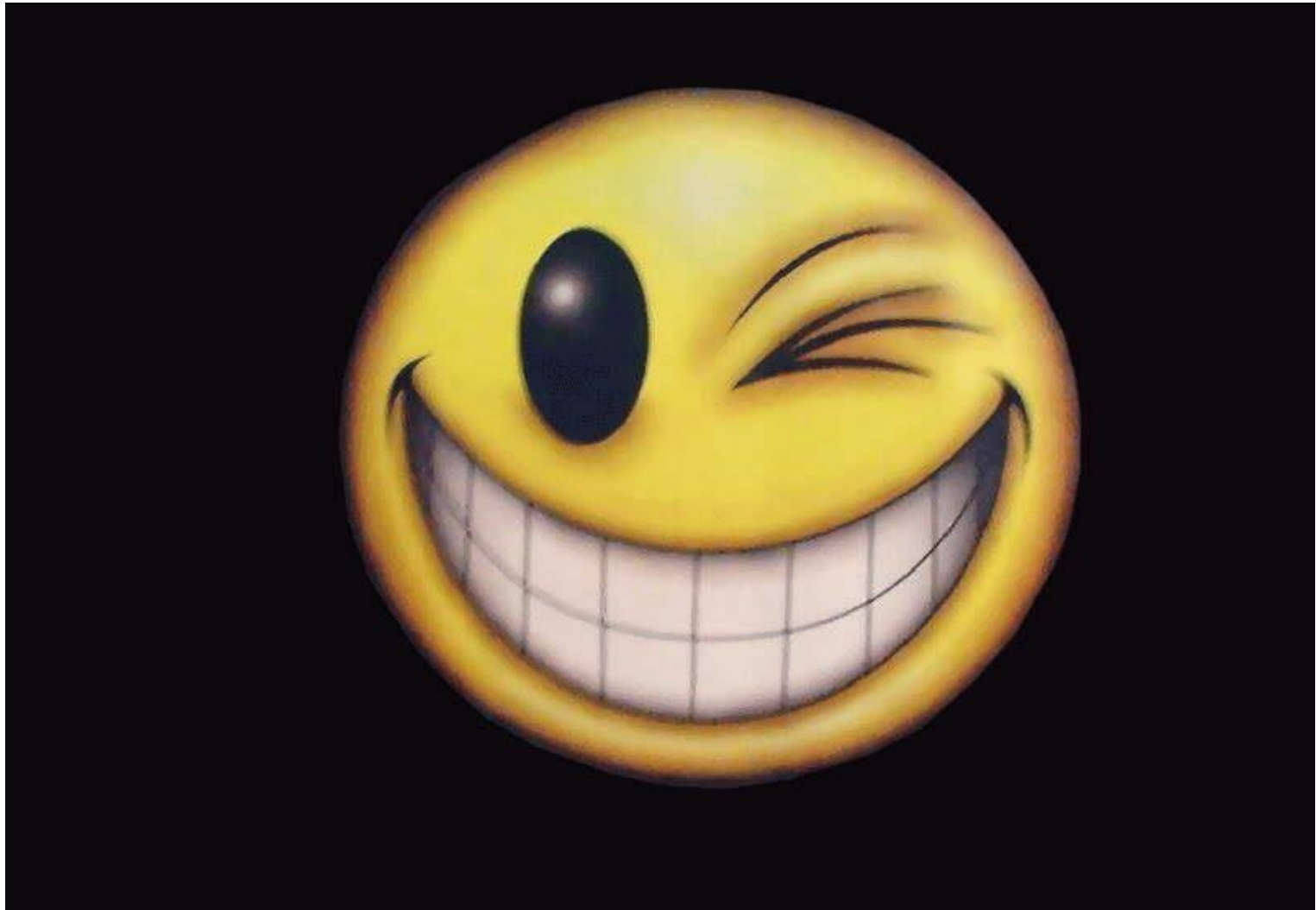


Deskriptive Statistik

Statistik (Biol./Pharm./HST) – FS 2014



Halbzeit !



Lernziele

Gemacht

- Grundlagen:
 - Wa.modell
 - Unabh.
 - Bed. Wa.
 - diskrete ZV
 - Verteilungen
 - Kennzahlen
- Am Bsp vom Binomialtest:
 - Test (Fehler 1. Art, Macht)
 - p-Wert
 - Vertrauensintervall

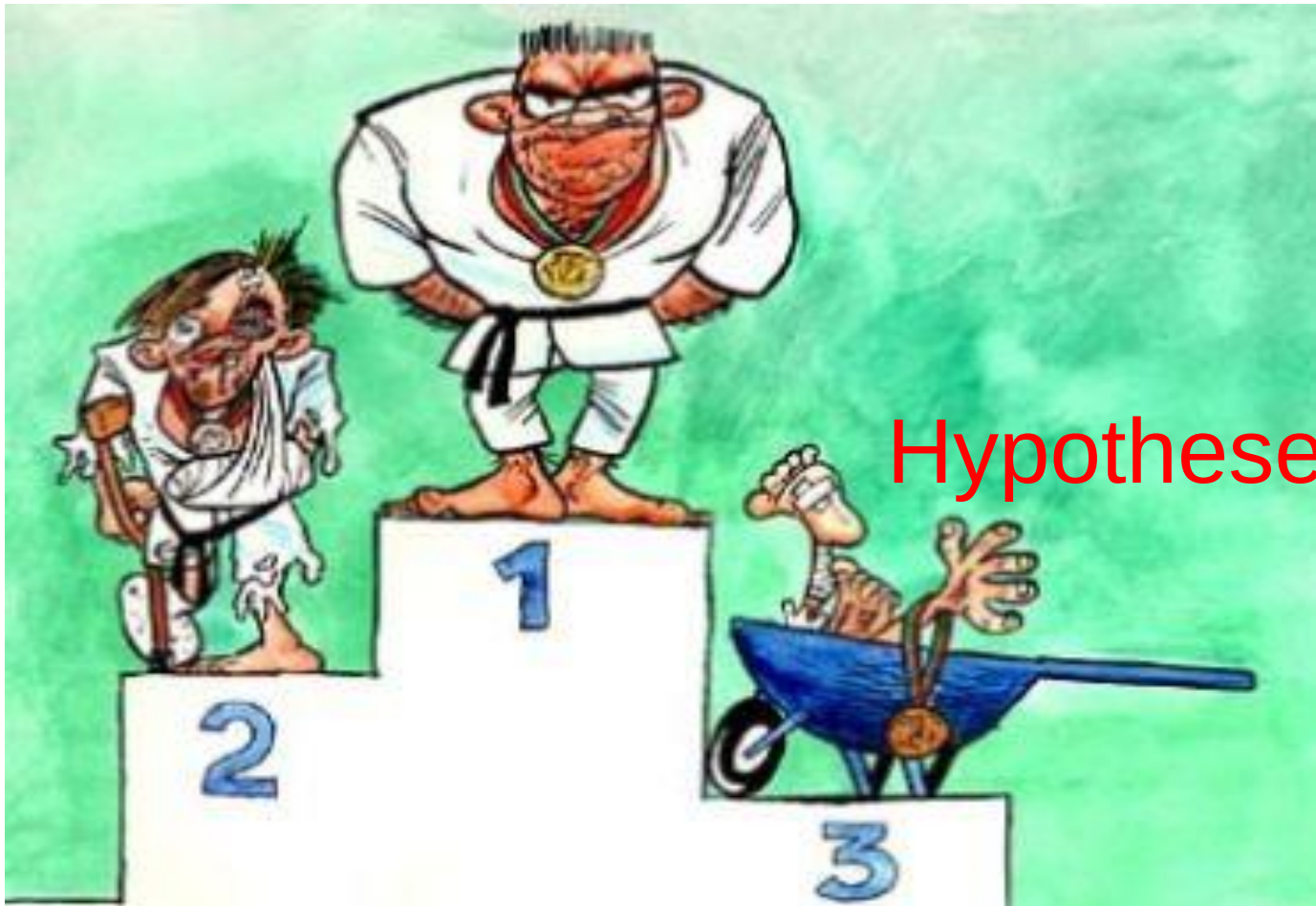
To Do

- Deskriptive Statistik:
 - Zahlen
 - Graphiken
- Am Bsp vom t-Test:
 - Test (Fehler 1. Art, Macht)
 - p-Wert
 - Vertrauensintervall
- Lineare Regression
 - einfach
 - multiple

Statistische Inferenz: Drei Methoden



Statistische Inferenz: Drei Methoden



Hypothesentest

Hypothesentest

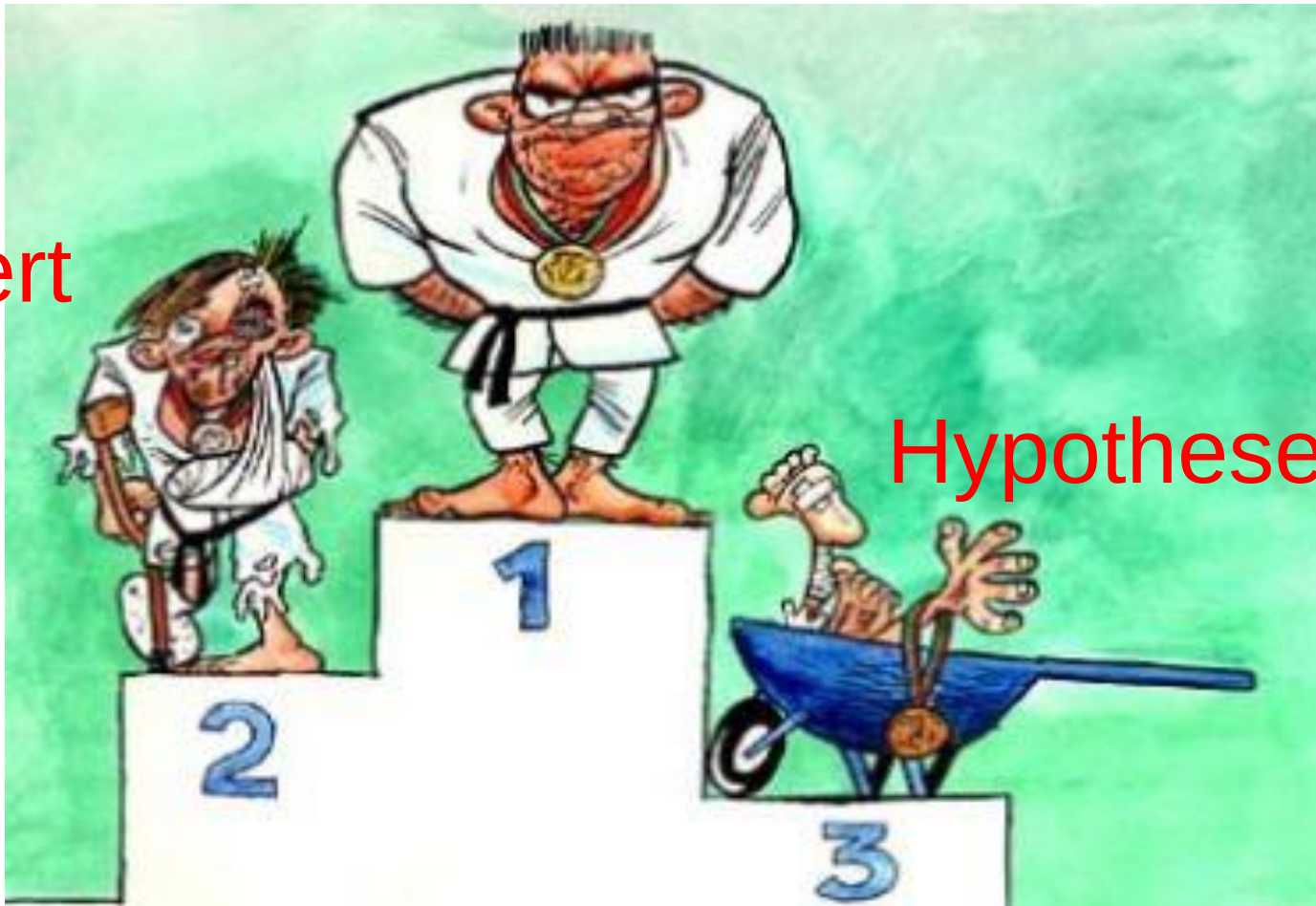
(Input: Sign.niveau, Daten)

1. Modell
2. H_0 , H_A
3. Teststatistik etc.
4. Sign.niveau
5. Verwerfungsbereich etc.
6. Testentscheid

- + Klares Prozedere
- + Klare Aussage über Fehler 1./2. Art
- Wie deutlich wurde verworfen?
- Wie gross ist der wahre Parameter?

Statistische Inferenz: Drei Methoden

p-Wert



Hypothesentest

p-Wert

(Input: Daten)

- Kleinstes Sign.niveau, bei dem Test gerade noch verwirft
- Gegeben H_0 ; wie plausibel ist die Beobachtung oder etwas noch extremeres?

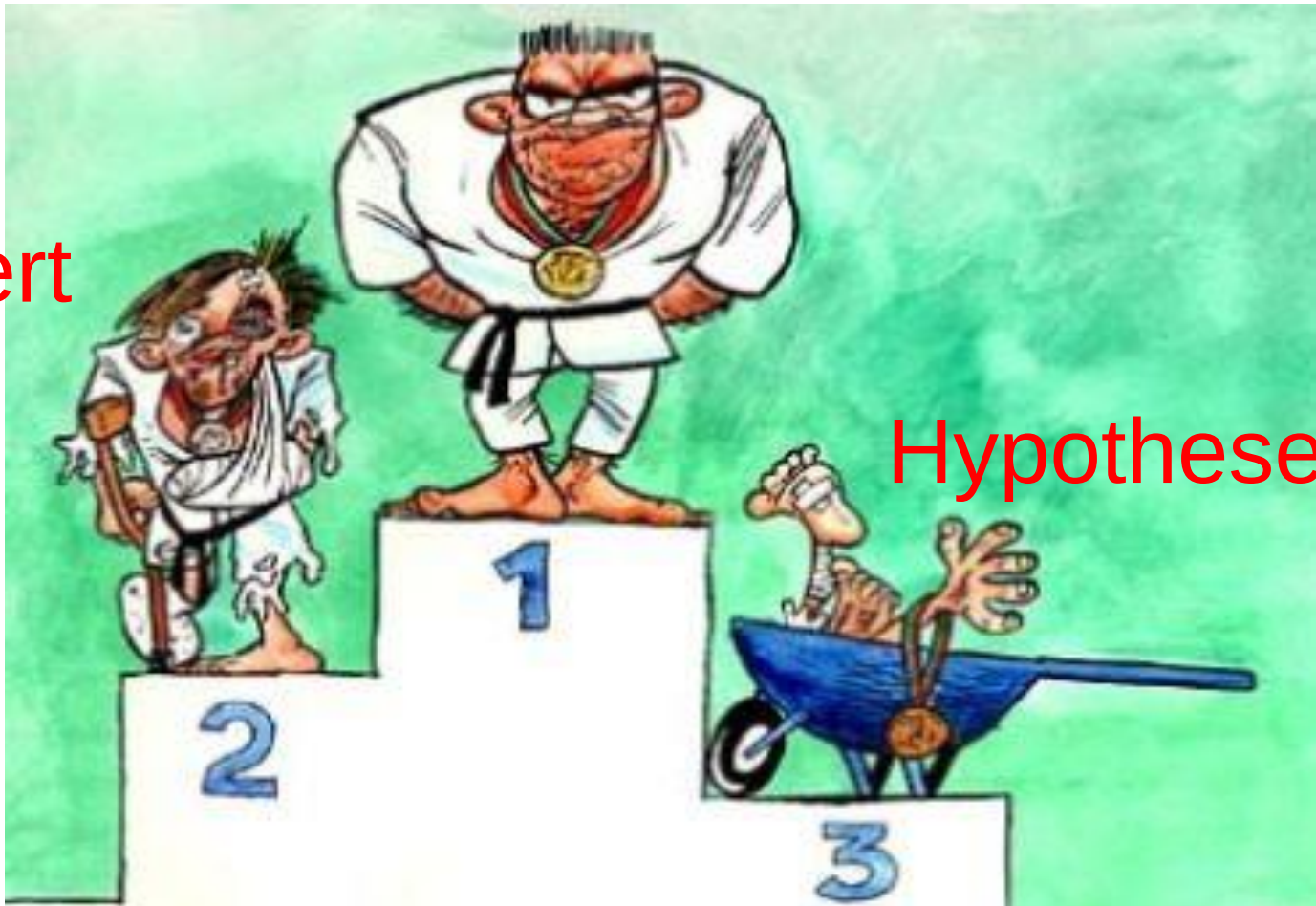
+ Klar, ob und wie deutlich verworfen wird

- Keine klare Aussage über Fehler 1./2. Art
- Wie gross ist der wahre Parameter?

Statistische Inferenz: Drei Methoden

Vertrauensintervall

p-Wert

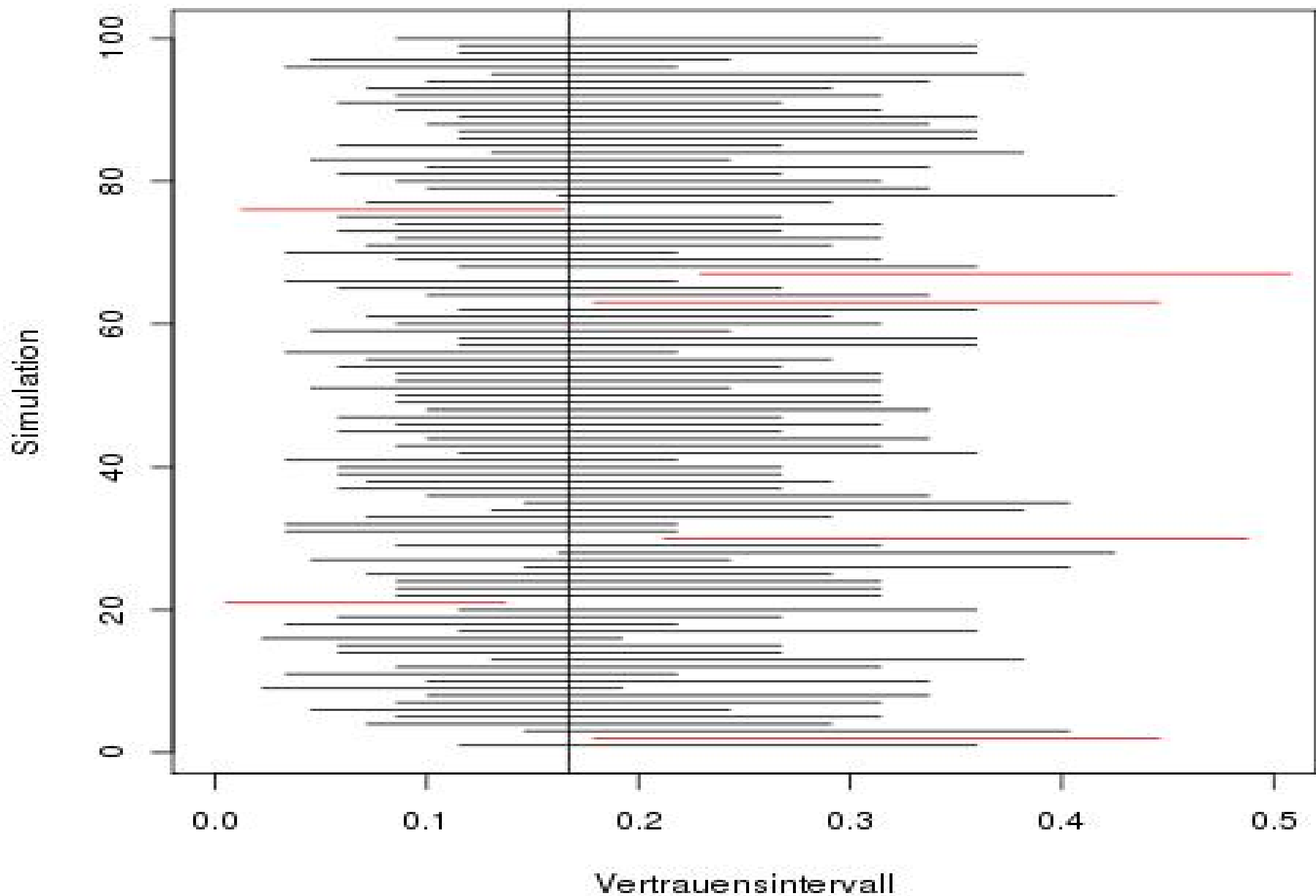


Hypothesentest

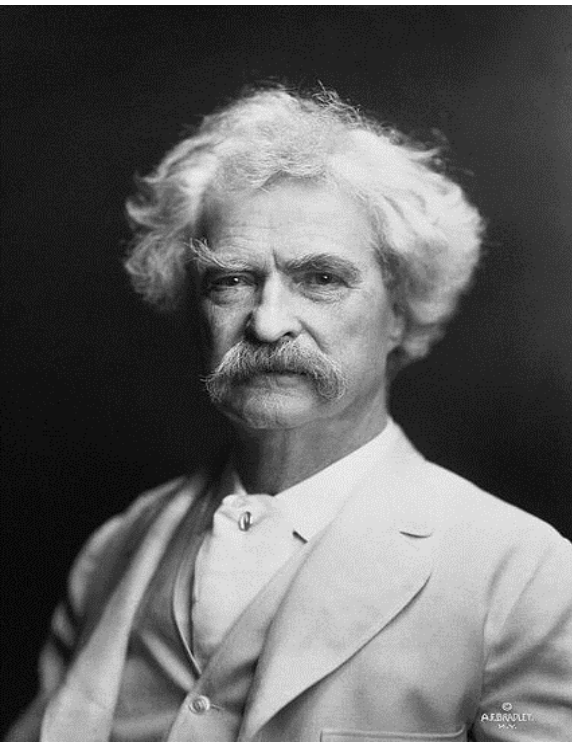
95% - Vertrauensintervall (Input: Sign.niveau, Daten)

- Enthält wahren Parameter mit Wa. 95%
 - Enthält alle Parameter, bei denen ein Test auf dem 5% Niveau NICHT verwirft (= plausible Werte)
- + Klar, ob und wie deutlich verworfen wird
 - + Klar, wie gross der wahre Parameter etwa ist
 - Keine klare Aussage über Fehler 1./2. Art

Coverage = 0.94



Mark Twain zitiert britischen Premierminister Benjamin Disraeli



„There are three kinds of lies:
lies,
damned lies,
and statistics.“

Frederick Mosteller
Gründer Statistik-Department, Harvard

„It's easy to lie with statistics;
it is **easier to lie without them.**“

Bildschirm		D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	V		
Ganzer Bildschirm		V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24	V25	V26	V27	V28	V29	V30	V31	V32	V		
4	5	0.56	0.77	-1.95	-1.35	-1.94	-2.02	0.45	-0.12	-0.91	-1.41	-0.64	0.18	-1.79	-0.18	1.49	-0.43	-0.78	1	-0.01	1	-0.58	0.99	-1.09	-1.58	-2.02	-1.77	-0.72	-0.39	-0.33	V		
		-1.92	-0.46	1.46	0.3	-0.7	-0.29	2.85	-0.62	-0.68	0.67	-0.85	1.77	1.37	-0.76	0.64	0.5	0.19	0.63	0.62	0.3	-0.86	-1.16	0.15	-0.38	0.74	-0.94	0.43	-0.98	-0.42	V		
6	7	0.8	-0.44	0.2	-0.67	1.7	0.18	-0.07	0.81	-0.78	-0.71	1.13	-0.37	-1.78	0.83	-1.83	0.97	-2.14	-1.31	-2.78	1	-1.54	0.01	-0.58	-0.75	-0.19	2.32	0.05	0.42	-0.32	-1.24	-0.22	-0.01
8	9	-1.49	-0.38	-0.87	0.95	1.12	-1.04	0.48	0.51	1.5	-0.3	0.82	-0.32	0.48	0.13	-0.34	-0.31	0.76	-0.44	-0.37	0.72	1.45	0.22	-0.34	0.34	-2.51	-0.13	0.77	1.05	-1.42	-0.36	0.16	0.27
10	11	1.24	-1.06	0.12	0.22	1.47	-1.94	0.49	1.19	0.41	0.02	1.69	0.67	-0.02	0.1	-1.35	0.16	-0.25	-0.5	0.38	0.8	0.1	1.48	-0.91	1.05	0.06	1.46	1.59	0.95	-0.47	-0.11	2.75	0.11
12	13	0.75	-0.04	-0.85	0.52	-0.28	-2.09	-0.52	0.05	0.17	0.45	1.35	-0.32	-0.01	-0.67	-0.28	-1	-0.93	-2.43	-0.17	0.27	-0.21	2.11	1.2	-0.6	-1.46	-0.38	0.77	1.13	-0.93	-1.8	-0.45	-1.32
14	15	0.67	-0.26	-0.93	-0.34	-0.94	-0.02	1.4	0.14	-0.08	0.68	0.1	1.81	-0.41	-1.42	1.87	0.7	0.85	-0.17	0.96	0.12	-0.35	-1.88	-0.37	-0.46	-1.57	2.52	0.17	-0.25	-0.27	0.77	-0.15	0.83
16	17	-2.15	1.29	-1.62	-1.09	-0.31	-2.57	1.85	-0.53	0.79	0.62	-0.69	-0.14	1.03	-0.98	0.25	1.32	-0.45	-0.8	-0.31	1.15	-0.31	-0.05	0.3	1.62	0.27	1.6	-0.97	-1.41	1.19	-1.82	-1.41	-1.18
18	19	0.99	-2.43	2.01	-0.76	-0.29	0.33	-0.34	-0.11	0.49	1.39	-0.77	-0.11	-0.58	1.32	-0.15	0.63	0.37	-0.08	-0.31	0.22	-0.07	0.93	0.51	-2.47	-1.89	-1.92	-0.49	-1.19	0.68	1.37	0.43	0.99
20	21	1.83	-0.52	-0.24	0.34	0.19	-0.57	1.35	0.12	0.5	-0.02	-1.88	1.36	-0.65	0.1	-0.17	-0.06	-1.39	-0.14	-0.28	-1.87	-0.64	-0.02	-1.28	0.5	-0.84	-2.59	0.73	-0.51	0.42	0.1	0	-0.79
22	23	-3.01	1.08	0.21	-0.18	0.12	-1.85	-0.56	0.22	-0.87	-1	-1.93	0.1	-0.86	0.56	-0.31	1.17	0.35	0.73	-0.8	1.45	0	-0.02	1.72	0.66	1.73	0.08	-1.25	1.54	-1.75	-0.76	-0.47	1.35
24	25	-0.23	0.51	-1.34	1.51	0.15	-0.92	1.04	1.19	-0.65	1.79	-0.93	0.1	-0.15	1.71	-0.09	-0.59	0.74	-0.15	-0.38	1.32	-1.12	-0.75	1.56	1.42	0.52	0	2.14	-1.87	-0.18	-0.88	0.21	-1.01
26	27	-0.64	-0.59	0.17	0.92	-1.44	-0.01	-0.23	-1.26	0.34	-0.53	-2.3	1.83	-0.5	0.94	-0.26	0.39	0.3	-0.27	0.1	0.02	0.15	0.54	0.61	-0.49	-0.23	1.03	0.47	1.13	-2.32	1.3	0.58	-1.94
28	29	1.85	-0.96	1.18	0.34	-0.1	-1.23	-0.3	-0.69	0.35	0.17	0.52	0.42	-0.57	-0.89	-0.28	-1.45	-0.25	0.85	0.42	-0.86	-0.86	-1.91	1.68	-1.46	1.36	-1.57	1.3	-0.61	1.9	-0.37	-0.8	-1.8
30	31	-0.36	0.33	-0.33	-0.84	2.34	0.63	-0.04	0.09	0.03	0.03	0.61	-1.43	-0.31	2.19	0.09	-0.64	-1.32	-0.34	1.89	0.07	1.82	-0.3	0.85	0.37	-0.83	0.87	1.92	0.39	-0.68	0.25	0.04	-0.57
32	33	-0.17	0.03	-1.07	-1.81	0.69	0.59	-0.17	0.41	-0.86	0.88	-0.19	1.42	0.95	0.85	1.02	-1.19	-1.05	-0.18	0.29	-0.22	-0.28	0.26	-2.22	0.28	-0.47	2.04	0.05	-0.4	-0.9	-1.85	0.53	-1.47
34	35	1.81	-0.67	0.53	0.23	1.72	-0.38	-0.46	0.65	-0.94	0.1	-0.15	-0.89	0.04	1.63	0.9	1.04	-0.04	0.03	-0.59	-0.38	-0.57	1.5	0.55	0.07	1.24	0.49	0.39	0.84	0.24	0.06	1.57	-0.36
36	37	0.97	0.18	-1.3	-1.63	-0.44	-1.01	-0.57	0.25	-0.57	-0.59	0.3	-0.42	0.91	-1.06	0.31	-0.69	-0.48	0.54	0.23	0.68	0.59	-1.09	0.59	0.62	0.17	-1.14	0.29	-0.45	-1.28	1.16	0.2	-1.05
38	39	0.76	-1.19	-0.38	-0.06	-0.5	-0.78	0.68	0.76	-0.07	0.4	0.39	1.31	1.5	0.03	0.26	0.12	-0.63	0.82	-0.96	1.01	1.69	-0.1	0.95	-0.46	1.12	-0.59	-1.42	-1.35	-1.22	-0.77	1.17	0.48
40	41	0.15	-0.32	-0.14	1.56	-1.6	0.66	-1.04	0.19	-1.46	-0.65	-1.79	0.31	0.76	-1.63	1.47	-0.37	-0.32	-0.04	0.48	-1.41	0.53	1.38	0.54	1.24	-0.79	-1.31	1.07	-0.23	1.23	0.2	-0.27	-1.89
42	43	0.03	1.89	0.37	0.63	-1.58	0.89	0.33	1.97	-0.77	0.66	-0.91	0.1	1.51	-2.24	0.75	-0.7	1.35	-0.64	0.17	1.43	-0.07	-1.03	-0.11	1.02	1.14	-0.29	-0.33	-0.6	1.14	0.57	0.37	0.81
44	45	-1.03	0.37	-0.09	1.28	1.01	-1.57	-2.45	0.41	-1.57	0.11	-0.98	-0.52	0.1	-0.93	-0.78	-0.37	-0.84	0.21	1.14	0.16	-0.81	0.13	0.57	-0.3	-0.39	0.03	0.06	-0.76	-0.12	-0.98	0.52	-2.15
46	47	0.28	1.03	1.35	1.11	2.23	-0.83	1.73	-0.79	-0.35	1.32	-0.52	0.3	0.03	1.01	0.34	0.14	0.57	-1.03	-0.5	0.14	-0.43	-0.2	0.98	0.22	0.57	1.77	0.46	-0.04	-0.77	0.31	0.7	-0.55
48	49	-1.1	1.25	-0.5	0.4	0.27	2.13	-0.46	0.71	-0.39	0.53	-1.34	0.4	0.8	-1.68	-0.06	0.16	-1.41	0.64	-1.38	1.26	-0.46	0.54	0.29	-0.33	-0.46	1.97	-0.31	-1.97	-0.1	-0.34	-0.37	-1.68
50	51	-0.34	1.25	-1.46	-2.1	1.94	1.45	-0.3	-1.82	-1.23	-1.16	0.41	2.1	-1.3	0.3	-0.81	-1.11	0.27	-0.29	-2.35	0.93	-0.75	1.66	-0.1	-0.8	-0.05	-0.86	-1.29	0.93	-0.33	-0.81	-0.83	-0.38
52	53	-0.84	1.4	0.1	-0.52	-1.4	0.46	1.38	-1.96	0.83	0.0	0.3	-1.54	-0.95	-1.4	0.55	-1.49	0.91	-0.61	-0.59	0.68	0.94	1.3	-0.53	0.01	-1.47	3.35	-1.05	0.9	-1.49	0.07	0.95	-1.29
54	55	1.47	-0.38	-0.5	0.46	-1.3	-0.19	0.4	-2.12	1.33	-1.1	-0.3	0.42	-0.49	-0.38	-0.3	-0.55	1.49	-0.33	-0.53	-0.46	0.25	1.79	0.68	0.18	1.6	0.73	-0.01	0.77	1.84	-0.22	1.2	-0.51
56	57	-0.37	1.56	0.45	0.51	-0.28	-0.34	-2.07	-0.65	0.45	-0.91	-0.11	1.49	0.17	0.63	-0.07	-0.05	0.45	0.26	0.39	-0.56	0.15	1	-0.3	-0.82	0.66	0.72	-0.15	0.33	0.8	0.42	-0.06	1.41
58	59	0.8	-1.57	0.02	-0.16	-2	0.34	0.97	-1.12	1.27	0.08	0.39	0.5	0.03	0.03	-0.05	-1.59	0.21	-1.79	0.45	-0.2	-1.18	-1.06	0.89	0.3	-0.56	0.41	-0.16	-0.52	0.43	0.5	0.24	
60	61	-0.85	0.21	-1.51	1.31	2.83	-0.41	0.38	-0.41	1.24	0.63	-0.71	-0.29	-0.41	0.63	0.1	-0.24	-0.89	0.27	-0.33	1.35	-1.13	-0.61	0.32	-1.05	-0.02	-1.29	-2.64	-1.13	-0.18	0.28	1.29	1.82
62	63	1.28	-0.95	-0.48	-0.18	-0.07	0.01	-0.69	0.61	-0.34	0.61	-1.25	0.02	0.3	0.1	0.19	0.51	0.06	-0.91	1.92	0	-0.08	-0.02	0.63	0.89	-2.17	-0.68	-1.48	0.73	-0.13	-0.03	-2.26	1.34
64	65	1.1	-0.24	0.99	0.53	-0.85	-1.81	-0.68	0.4	-0.14	1.06	1.22	-1.3	0.61	0.4	0.7	0.05	-0.92	0.22	-0.67	-1.27	0.12	-1.23	0.95	-1.57	-0.85	-0.84	-0.3	-1.64	0.18	-1.16	0.88	
66	67	-0.96	-0.75	-1.1	-1.61	0.32	-0.92	-0.52	-0.32	-0.76	-0.12	0.7	1.87	0.44	0.0	-0.12	0.09	-0.13	0.0	2.13	0.99	-1.03	-0.72	0.96	0.57	0.32	-0.51	0.48	-2.18	-1.18	2.29	0.86	0.01
68	69	-0.43	-0.58	-0.27	0.8	-0.31	-0.48	-1.6	-1.3	0.56	1.4	-0.8	-2.08	-0.52	0.67	1.01	1.43	0.52	0.19	0.83	-1.55	-0.16	0.05	1.85	-0.65	-1.1	0.33	0.33	-0.46	0.06	-0.59	0.78	0.64
70	71	-0.24	0.55	-0.69	-0.26	1.11	0.04	-0.77	1.42	0.66	-0.19	0.1	-0.77	-1.14	1.36	-0.53	-0.2	-0.58	0.3	0.59	0.51	0.31	-0.49	0.41	0.38	0.42	0.57	0.62	1.17	-1.45	0.51	-0.47	-0.85
72	73	0.8	-0.54	-0.51	0.59	0.27	0.78	0.11	2.71	0.25	-0.13	1.3	-0.01	-0.85	0.29	1.32	-0.22	-0.17	-0.26	0.0	0.74	-0.4	-0.12	-0.97	-1.32	-1.61	-0.25	-0.1	-0.86	-0.77	0.8	-0.02	0.63
74	75	-0.21	0.95	0.99	-0.3	0.03	-1.22	1.36	-2.36	0.81	0.97	0.1	-0.01	0.02	0.37	0.5	1.84	-0.21	-0.34	0.6	0.3	-0.56	-0.07	-0.51	-0.12	1.61	1.05	-0.3	1.3	0.18	-0.93	0.96	-0.64
76	77	0.66	-1.35	0.83	1.92	1	-1.51	-1.83	-1.24	-0.85	-0.1	1.94	-0.42	-0.58	0.29	-0.08	0.9	0.81	1.57	0.0	-0.3	2.16	-0.3	0.21	-0.75	-0.58	-0.15	0.19	-0.12	2.03	0.04	-0.1	-0.96
78	79	-0.49	0.57	0.91	1.51	-0.5	-0.99	-0.49	0.59	2.04	-0.43	1.43	1.5	0.0	0.57	1.19	-0.5	0.33	-1.59	0.24	0.04	-0.25	-0.51	-0.38	0.55	0.54	0.35	-1.15	-0.23	0.09	-0.41	0.7	-0.87
80	81	1.58	0.47	1.06	0.49	-0.51	1.23	-1.74	-0.36	0.87	0.35	0.91	-0.83	0.68	0.3	1.49	0.45	-0.54	-0.61	1.9	0.47	0.65	-0.7	0.81	1	0.							

Beschreibe mit einem Wort!



Pferde ?

Marc ?

blau ?

Ohne Reduktion ist
kein Überblick möglich

Bei jeder Datenreduktion
kann wichtige Information
verloren gehen

Simpson's Paradox

(Bsp: Aufgenommene Studenten an der UC Berkeley in 1973)

	Bewerber	Aufgenommen
Männer	8442	44%
Frauen	4321	35%

Werden Frauen benachteiligt?

Simpson's Paradox

(Bsp: Aufgenommene Studenten an der UC Berkeley in 1973)

Dept	Männer		Frauen	
	Bew.	Akz.	Bew.	Akz.
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	272	6%	341	7%

Simpson's Paradox

(Bsp: Aufgenommene Studenten an der UC Berkeley in 1973)

Dept	Männer		Frauen	
	Bew.	Akz.	Bew.	Akz.
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	272	6%	341	7%

Simpson's Paradox

(Bsp: Aufgenommene Studenten an der UC Berkeley in 1973)

Dept	Männer		Frauen	
	Bew.	Akz.	Bew.	Akz.
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	272	6%	341	7%

Nein: Frauen bewerben sich mehr bei „schwierigen“ Departments!

Arrow's Paradox: Demokratie

Einzelmeinungen



→ Wahlergebnis

Arrow's Paradox: A, B oder C?

- 17 Personen müssen zw. A, B oder C entscheiden.
- Präferenzen:
 - 5 $A > B > C$
 - 2 $A > C > B$
 - 2 $C > A > B$
 - 4 $C > B > A$
 - 4 $B > C > A$
- Mehrheit (1,0,0):
A gewinnt
- Platz 1 oder 2 (1,1,0):
B gewinnt
- Borda Count (2,1,0):
C gewinnt
- Paarweise:
 $A > B$, $B > C$, $C > A$
Unentschieden

?!

Wahlmethode entscheidend !

Finde gute Wahlmethode: Anforderungen

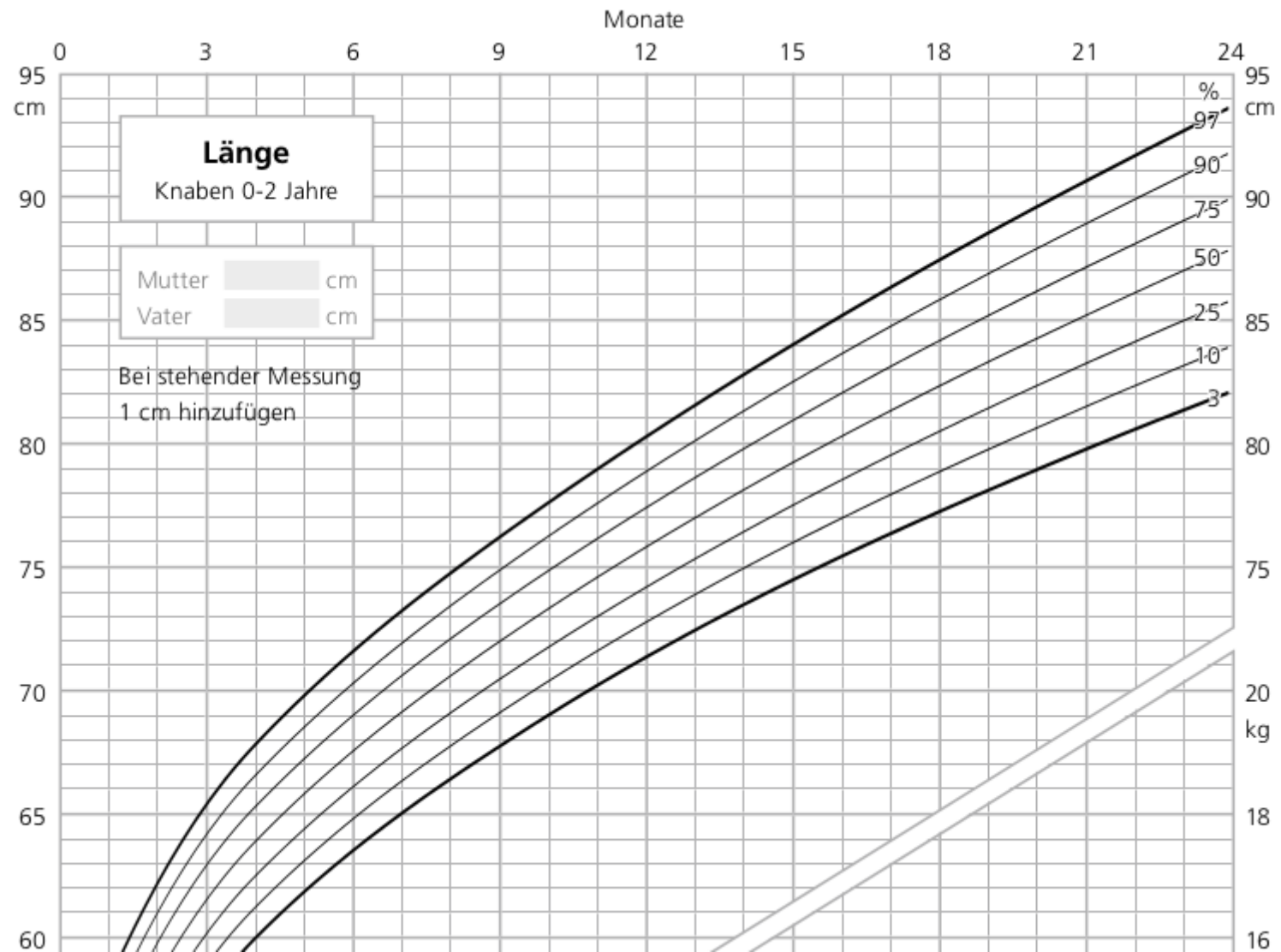
- **Transitiv:**
Wenn $A > B$, $B > C$ dann $A > C$
- **Pareto effizient:**
Wenn $A > B$ für *alle* Wähler, dann muss $A > B$ auch im Wahlerscheid sein
- **Independence of Irrelevant Alternative (IIA):**
Wenn $A > B > C$ und C stirbt, soll $A > B$ bleiben

Arrow's Impossibility Theorem

Falls **3 oder mehr** Optionen:

Es gibt kein Wahlsystem, das alle drei Anforderungen erfüllt.

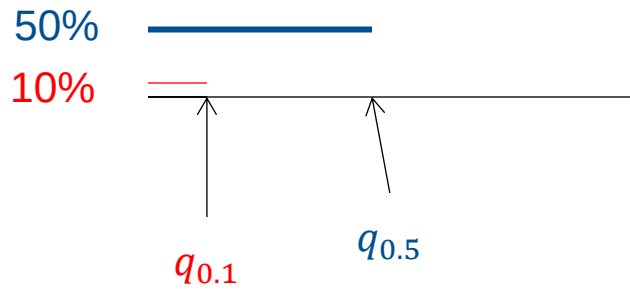
Quantil - Bsp



Quelle: Kinderspital Zurich

Quantil - Intuition

1, 1, 3, 5, 7, 23, 1089



Quantil - Definition

- x_i ist der i-kleinste Datenpunkt ($x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$)

- Falls $\alpha * n$ eine ganze Zahl ist:

$$q_\alpha = \frac{1}{2} (x_{(\alpha n)} + x_{(\alpha n + 1)})$$

- Sonst

$$q_\alpha = x_{(k)} \text{ mit } k = \alpha n + \frac{1}{2}$$

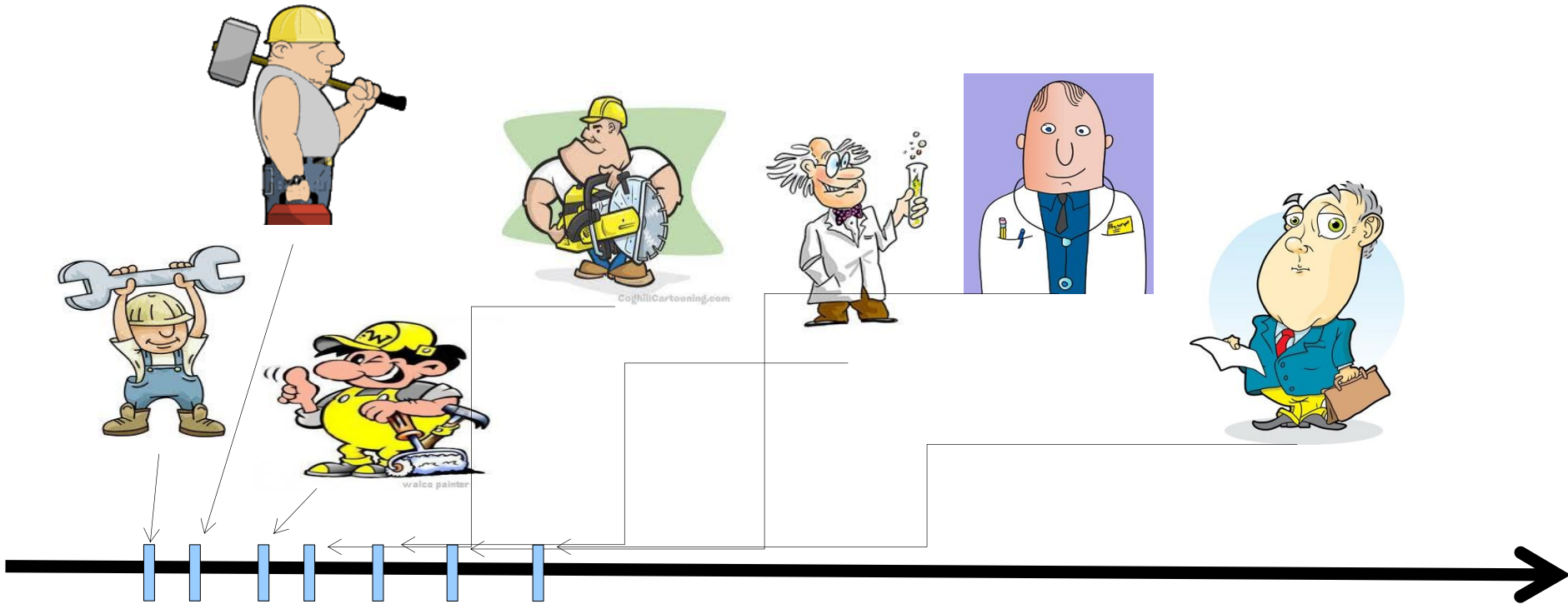
- Bsp: 1,1,3,5,7,23,1089; $n = 7$, $\alpha = 0.5$

$$\rightarrow \alpha n = 3.5, \quad k = 3.5 + \frac{1}{2} = 4$$

$$\rightarrow q_{0.5} = x_{(4)} = 5$$

Median und arithmetisches Mittel

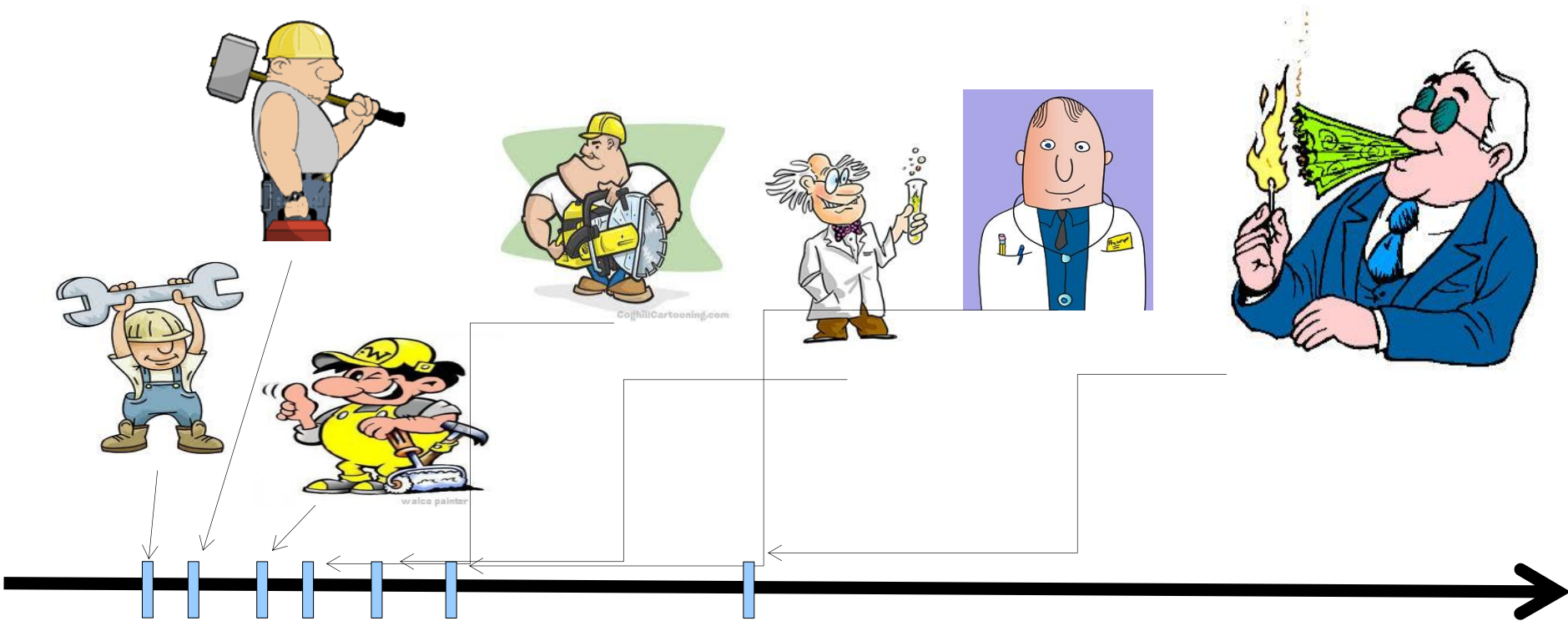
Einkommen von 7 Personen



Median

arithm. Mittel

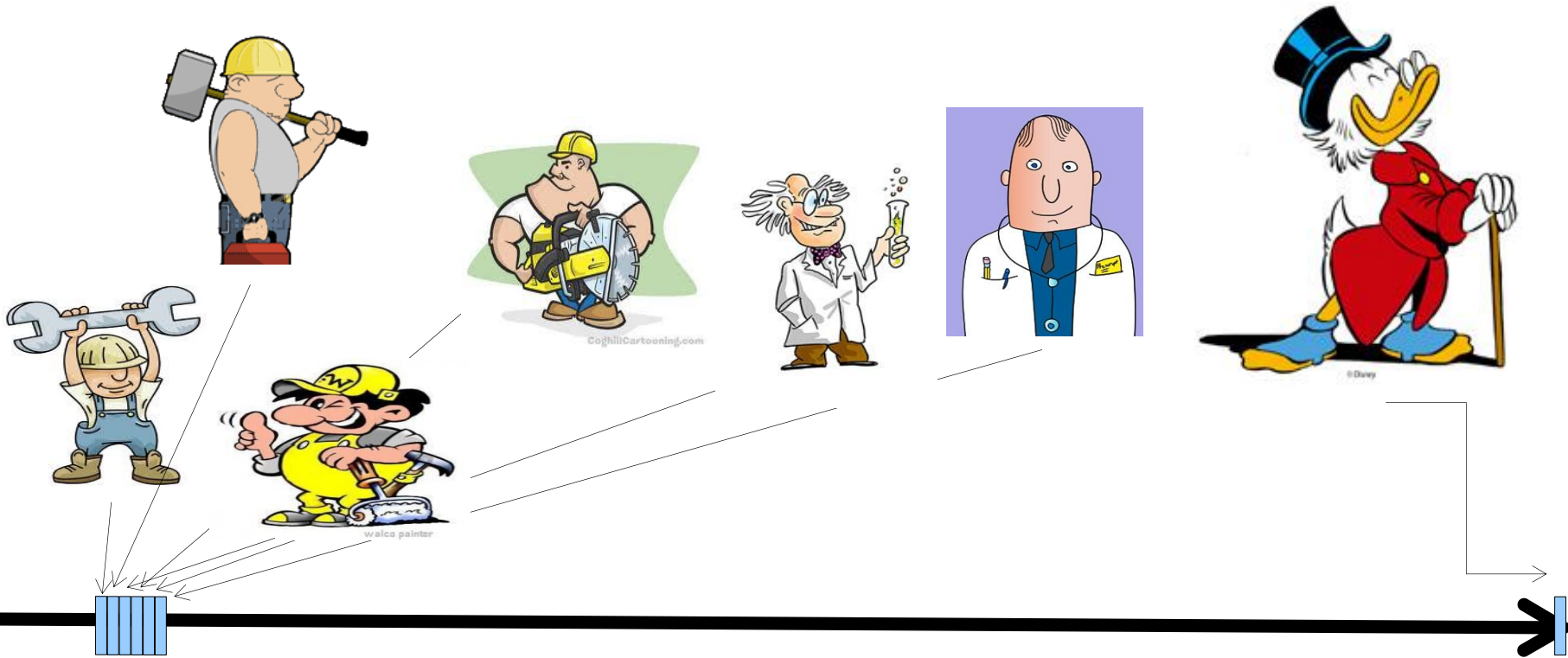
Einkommen von 7 Personen



Median

arithm. Mittel

Einkommen von 7 Personen

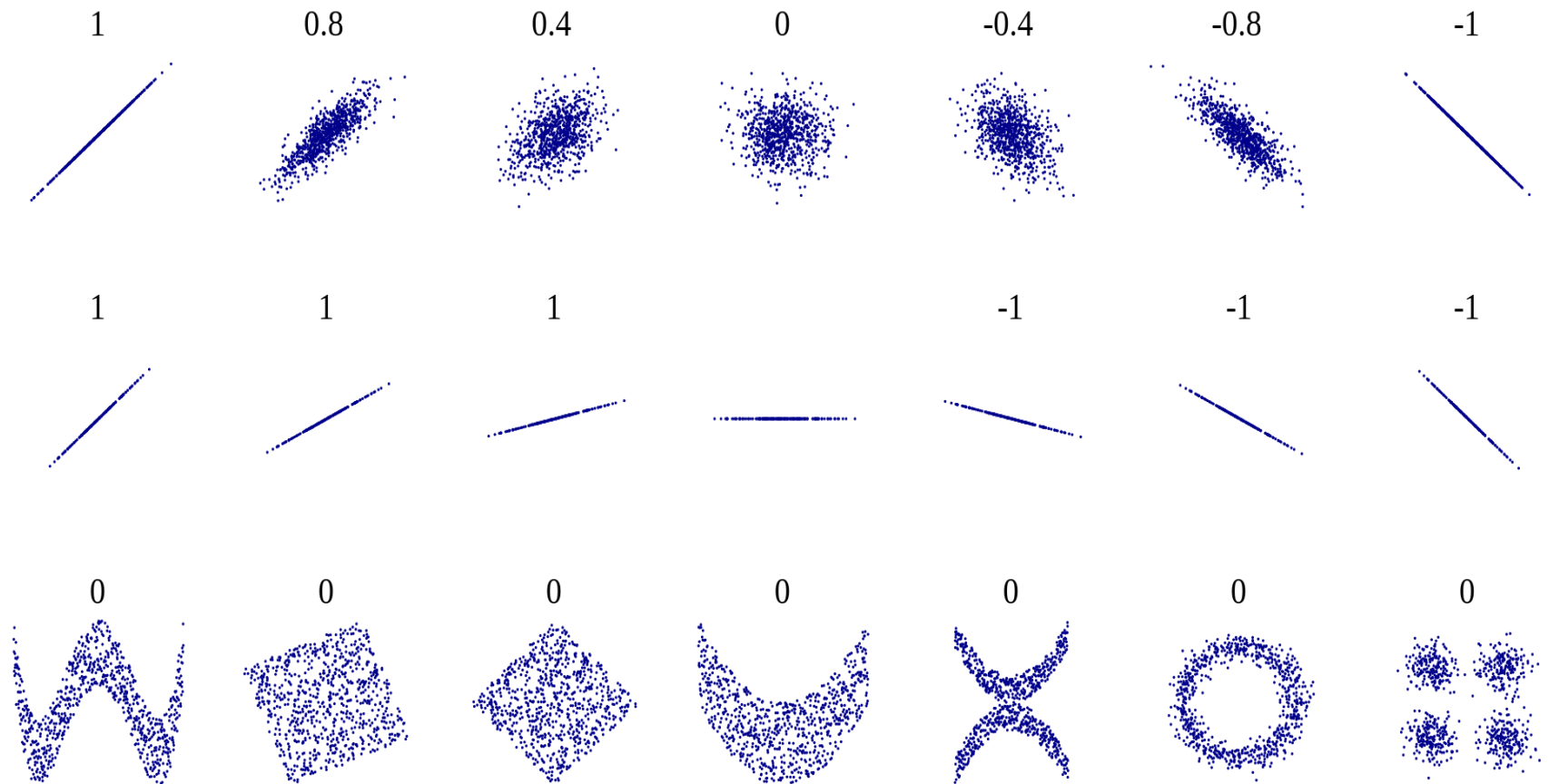


| Median

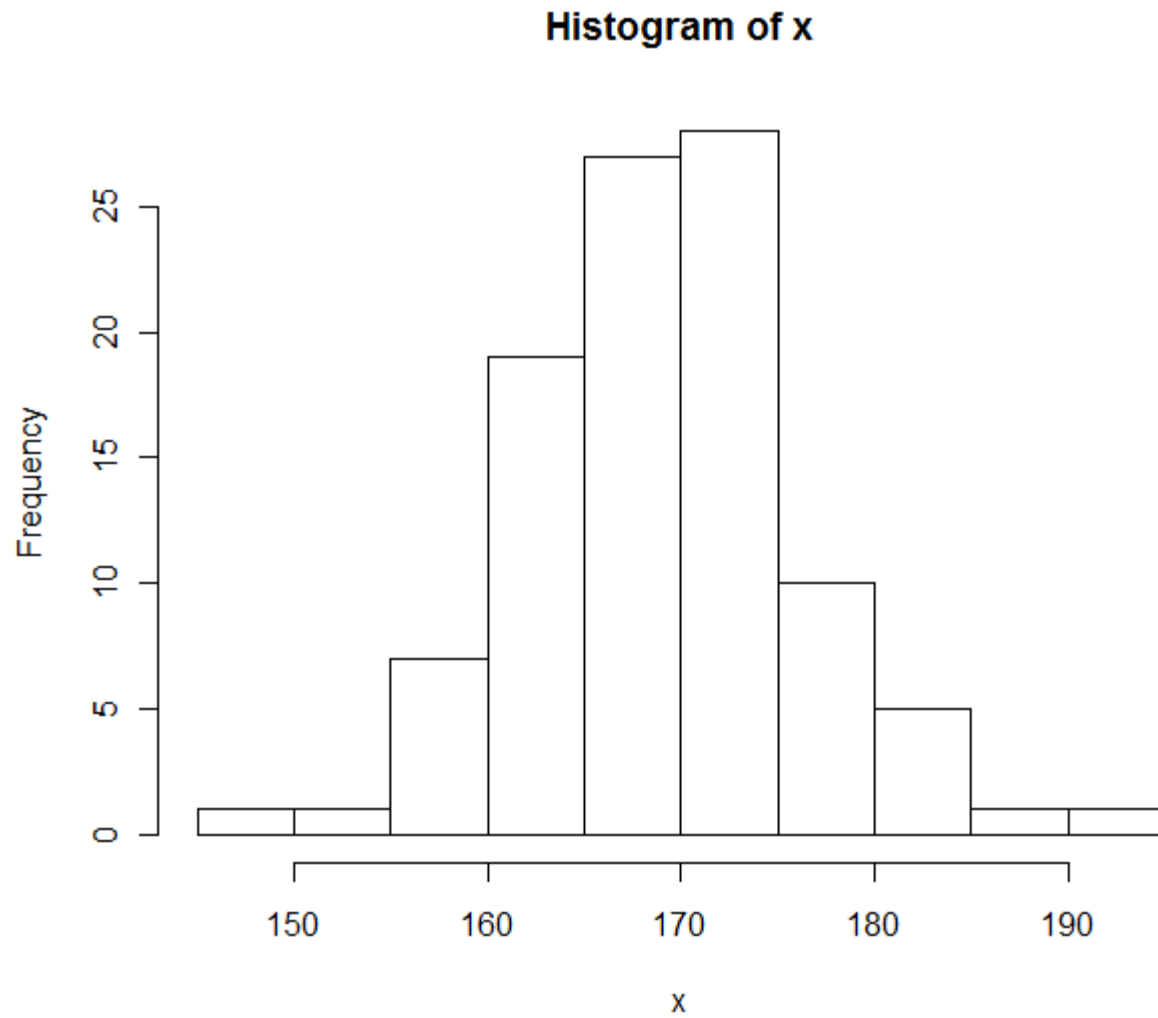
Robust gegen Ausreisser

| arithm. Mittel

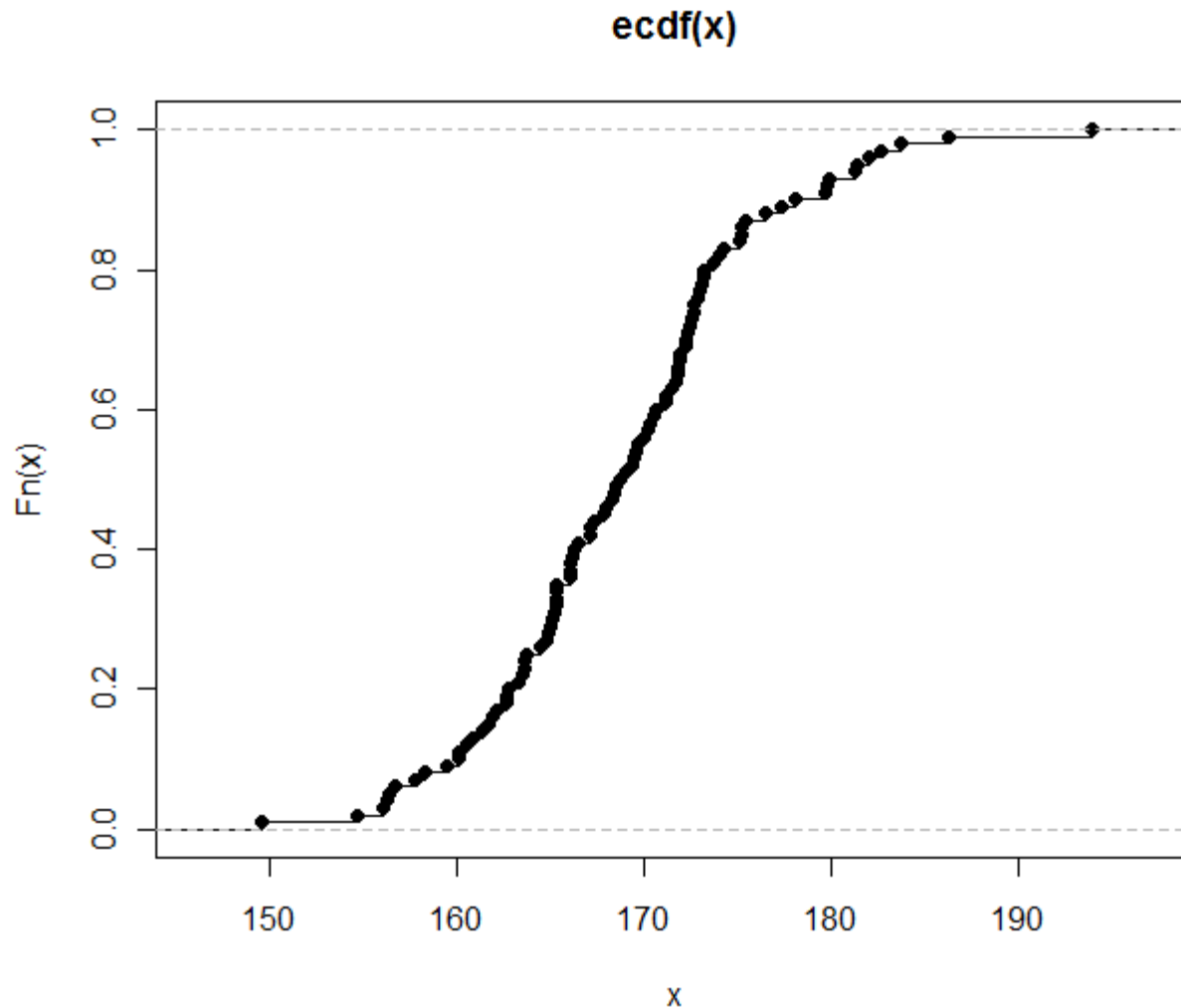
Korrelation



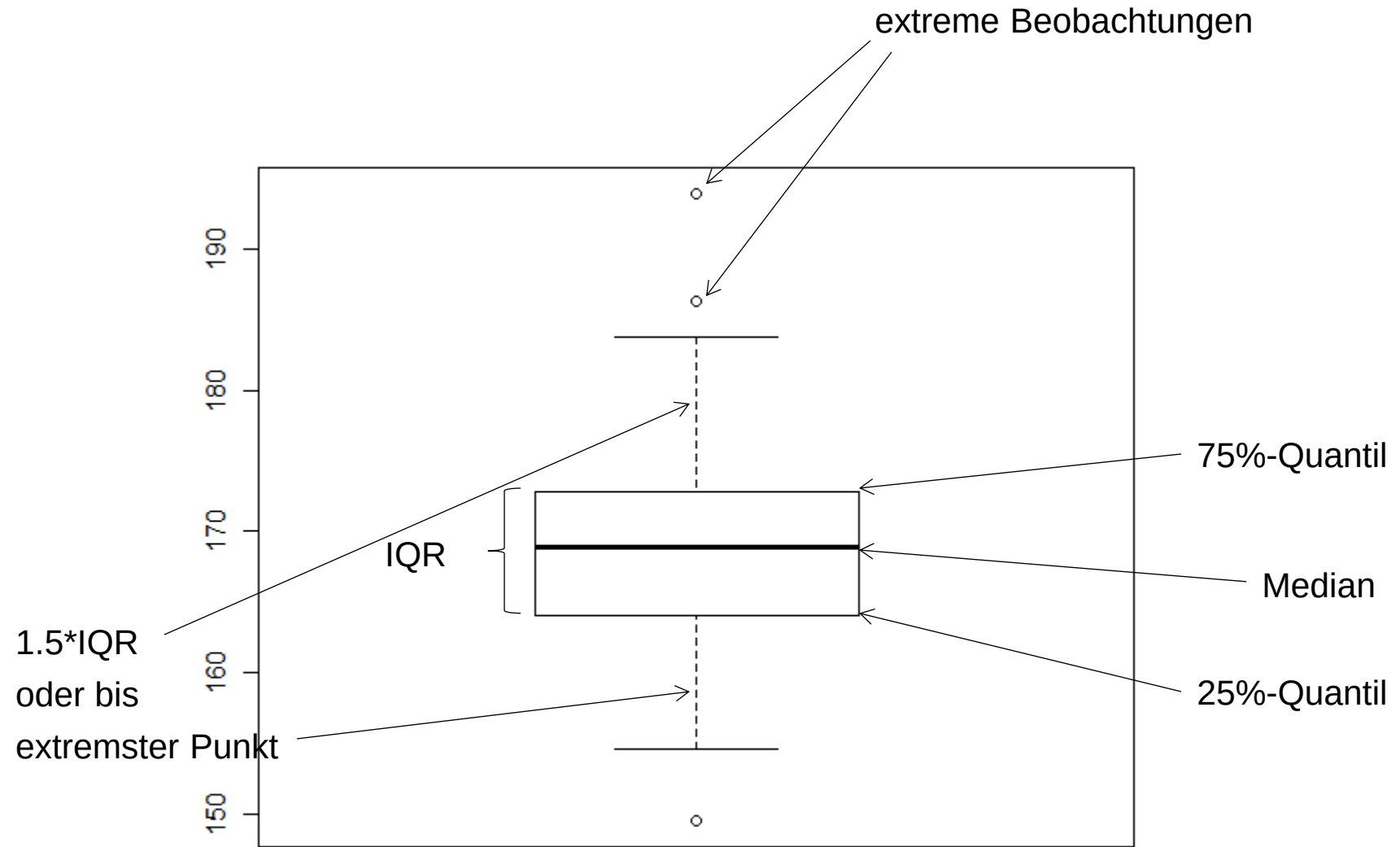
Histogramm



Empirische Verteilungsfunktion (Empirical Cumulative Distribution Function – ECDF)



Boxplot



Prüfung Wi10, A 4.2: Richtige Zuordnung?

- a) A3, B2, C1, D5, E4
- c) A3, B4, C2, D5, E1
- e) A1, B5, C3, D4, E2

- b) A3, B4, C1, D5, E2
- d) A1, B4, C3, D5, E2
- f) A2, B4, C1, D5, E3

