Hypothesen

- Wenn/dann und je/desto Hypothesen
- deterministisch vs. probabilistisch
- Informationsgehalt einer Hypothese
- Im Allgemeinen: Erwartung, wie Untersuchungseinheiten in einem analytischen Raum von Variablen verteilt sind
- Über Drittvariablen bestimmbar: Allgemeinheits- und Komplexitätsgrad einer Hypothese
- Falsifizierbarkeit und Spezifizität einer Hypothese

Wenn – dann – Hypothesen

		wenn: A	wenn: ~A
dann:	B	K I	K II
	~B	F III	K IV

		wenn: A	wenn: ~A
dann:	B	K I	F II
	~B	F III	K IV

K = Konfirmatoren

F = Falsifikatoren

(1) Wenn – dann – Hypothese

(deterministische **Implikation**)

(2) Wenn – und – nur – wenn – dann – Hypothese

(deterministische **Äquivalenz**)

A = hinreichende Bedingung

A = hinreichende und notwendige Bedingung

Quelle: A. Diekmann (1995) Empirische Sozialforschung. Reinbek bei Hamburg: Rowohlt, S. 109

Deterministische und probabilistische Hypothesen

„wenn A, dann B“:

kann in die Erwartung transformiert werden:

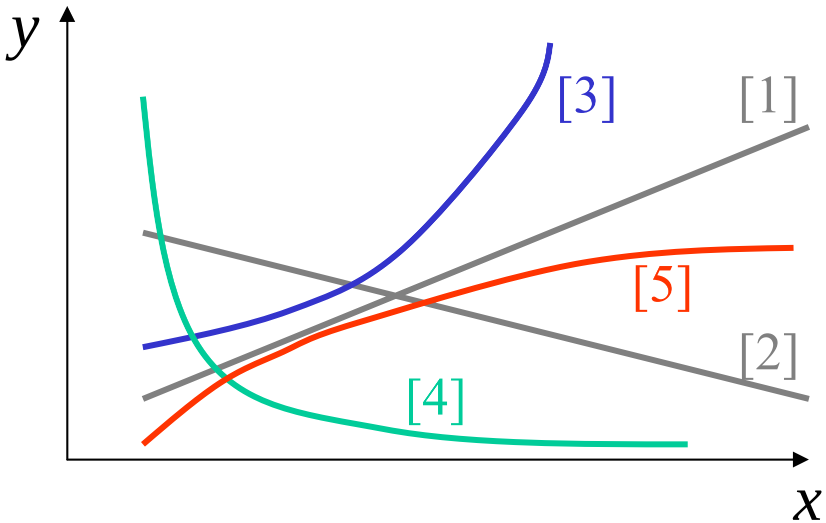
Prob (B) bei Auftreten von A
ist größer als
Prob (B) bei Auftreten von ~A

$$P(B \mid A) > P(B \mid \sim A)$$

Je – desto – Hypothesen

Monoton steigende bzw. monoton fallende Zusammenhänge

Ggf. Spezifikation des Zusammenhangs als spezielle mathematische Funktion möglich (Beispiele):



[1] linear steigend

[2] linear fallend

[3] exponentiell steigend

[4] exponentiell fallend

[5] logarithmisch

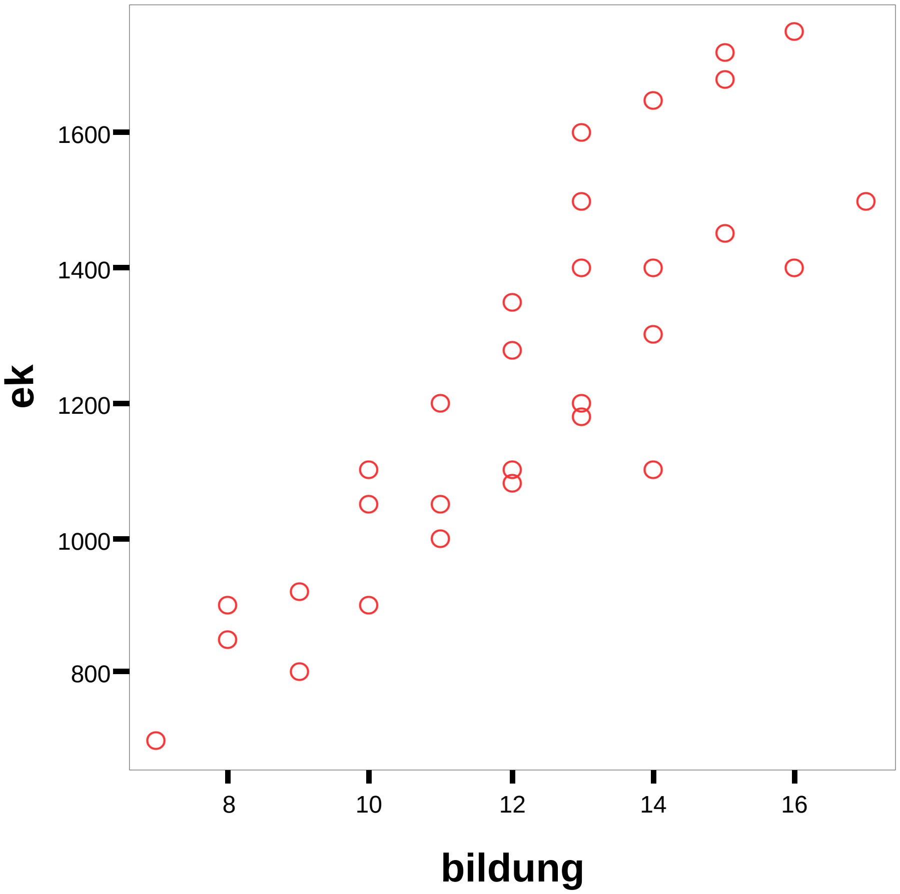
Nach Abb. *IV.3* (S. 115) in Diekmann, Andreas (1995) Empirische Sozialforschung. Reinbek bei Hamburg: Rowohlt

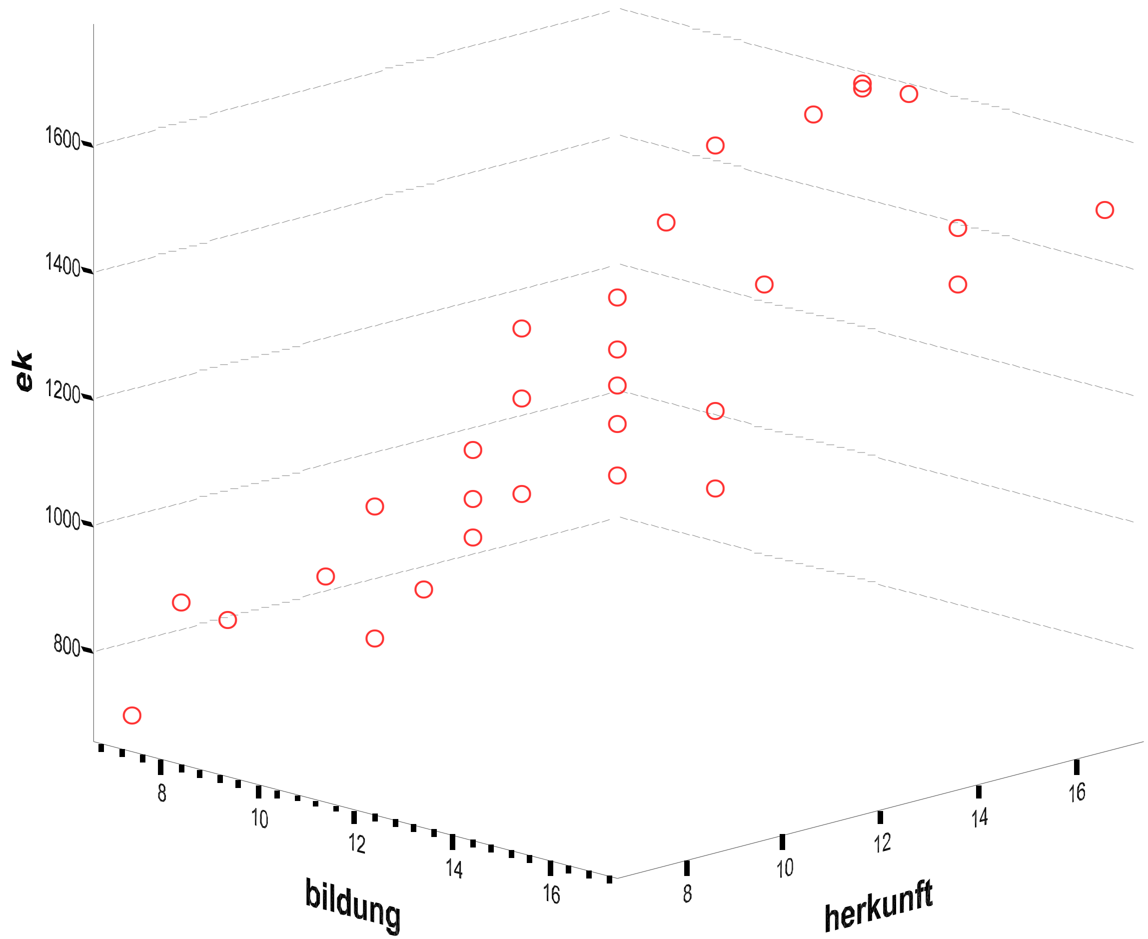
» **Informationsgehalt** (empirischer Gehalt) eines Satzes ist die Menge der von diesem Satz ausgeschlossenen Sätze.

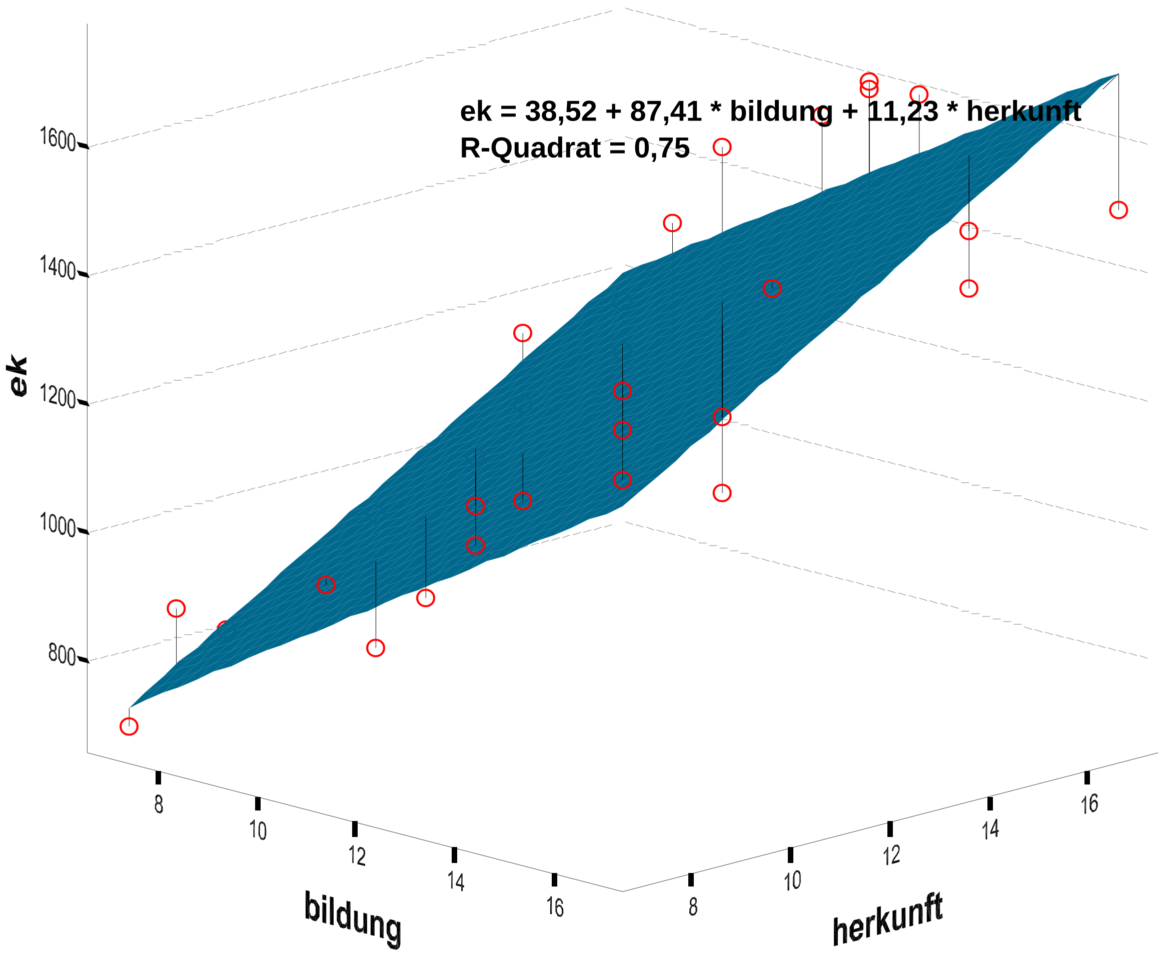
» Der Gehalt einer Wenn-dann-Hypothese wächst **gleich**sinnig mit dem Gehalt der Dann-Komponente und **gegen**sinnig zum Gehalt der Wenn-Komponente

Nach Diekmann (1995) Empirische Sozialforschung.
Reinbek bei Hamburg: Rowohlt, S. 132f.

Eine **Hypothese** ist eine Behauptung darüber, wie ein Set **S** von Untersuchungseinheiten in einem Raum von Variablen $X_1, X_2, \ldots, X_n$ verteilt ist.




ek
1600
1400
1200
1000
800
8
10
12
14
16
bildung
8
10
12
14
16
herkunft


ek = 38,52 + 87,41 * bildung + 11,23 * herkunft
R-Quadrat = 0,75
ek
1600
1400
1200
1000
800
bildung
8
10
12
14
16
herkunft
8
10
12
14
16

Allgemeinheitsgrad …

.. ist bestimmt durch das Set von Bedingungen **B**, unter denen die Hypothese haltbar ist

„[intended] field of tenability“

Forschungsleitende **Frage**

„Gegeben die Hypothese H, zeige mir das Set von Bedingungen **B**, unter denen sie haltbar ist“

Bedingungen B = Ausprägungen von Drittvariablen

Nach: Galtung, Johan (1969) Theory and Methods of Social Research. Oslo: Universitetsforlaget, pp. 309 -339

Komplexitätsgrad …

.. über die **Anzahl** der Variablen, die den untersuchtenVariablenraum aufspannen

H (X,Y) getestet für T=$\mathbf{t_1}$

? ***H***(X,Y) für T=$\mathbf{t_2}$

Mögliche **Ergebnisse**:

H hängt von T ab (gilt z.B. nur für T=$\mathbf{t_1}$) und ist entsprechend zu **spezifizieren**

H hängt **nicht** von T ab und ist entsprechend auf T=$\mathbf{t_2}$ **generalisiert**

Falsifizierbarkeit …

$$F = \frac{n_F}{E}$$

Anteil falsifizierender Ergebnisse an allen möglichen Ergebnissen
- bei a priori Annahme, dass alle E Verteilungen gleich wahrscheinlich sind

Spezifizität

Anzahl möglicher Unterscheidungen, die in den empirischen Ergebnissen getroffen werden können

Nach: Galtung, Johan (1969) Theory and Methods of Social Research. Oslo: Universitetsforlaget, p. 188

Im Fall probabilistischer Hypothesen:

Anzahl der Partitionen von N in R Teile:

$$E = \binom{N+R-1}{R-1}$$

Anzahl der möglichen Verteilungen eines Samples **S** mit **N** Fällen über die **R** [= $r_1 \times r_2 \times \ldots \times r_k$ -] Ausprägungs-kombinationen der *k* Variablen, die den analytischen Raum (Ereignisraum) aufspannen

Nach: Galtung, Johan (1969) Theory and Methods of Social Research. Oslo: Universitetsforlaget, p. 188

Beispiel

Mögliche Verteilungen oder Partitionen von 3 Fällen in 2 Teile (**R** = 2; **N** = 3):

3 + 0 = 3

2 + 1 = 3

1 + 2 = 3

0 + 3 = 3

$$E = \binom{3+2-1}{2-1} = \binom{4}{1} = \frac{4!}{1!(4-1)!} = \frac{4\times 3\times 2\times 1}{1\times 3\times 2\times 1} = \frac{24}{6} = 4$$

$$E = \binom{N+R-1}{R-1}$$

Größe des Ereignisraumes, **R**

Stichproben-Umfang, **N**

	1	2	3	4	5	…10
0	1	1	1	1	1	.
1	1	2	3	4	5	.
2	1	3	6	10	15	.
3	1	4	10	20	35	.
4	1	5	15	35	70	.
5	1	6	21	56	126	.
.						.
.						.
10	.	.	.	.	.	92.378

1. Vereinfachung von **E** durch ..

.. **deterministische** Repräsentation probabilistischer Hypothesen über eine **Korrelations**sprache

---

2. Abbildung des „Ereignissets“ auf ein Set mit den Ausprägungen

wahr – unentschieden – falsch
Korrespondenz – unentschieden – Nicht Korresp.
bestätigt – unentschieden – entkräftet

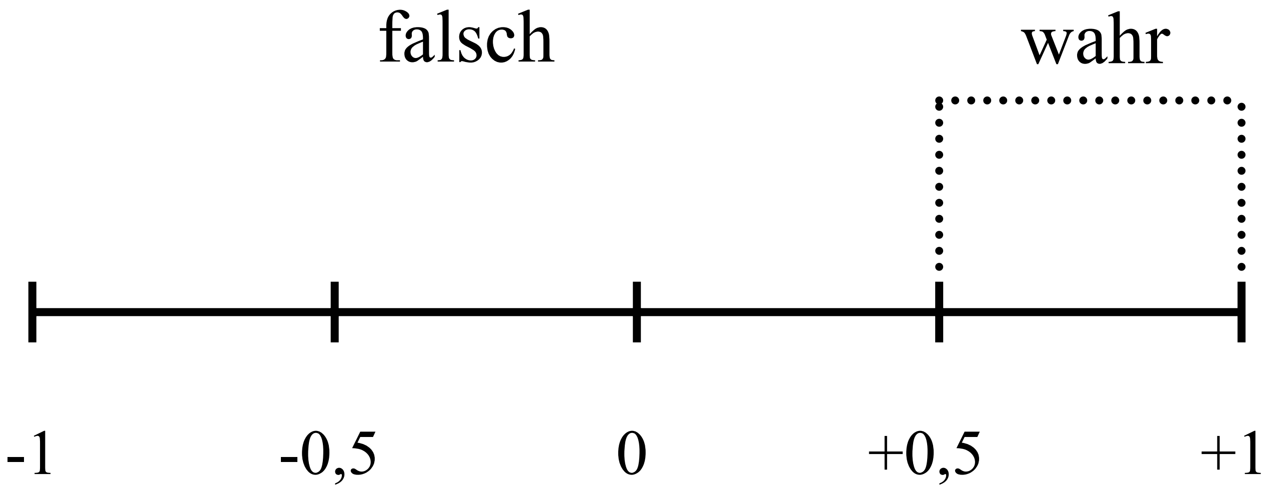
Nach: Galtung, Johan (1969) Theory and Methods of Social Research. Oslo: Universitetsforlaget, pp. 309 - 315

3. Vergleich der **erwarteten** Korrelationen mit den **beobachteten** Korrelationen

$C\ (X,Y)$	$C^{+}\ (X,Y)$
	$C^{-}\ (X,Y)$
	$C^{0}\ (X,Y)$
	$C^{\geq 0,5}(X,Y)$

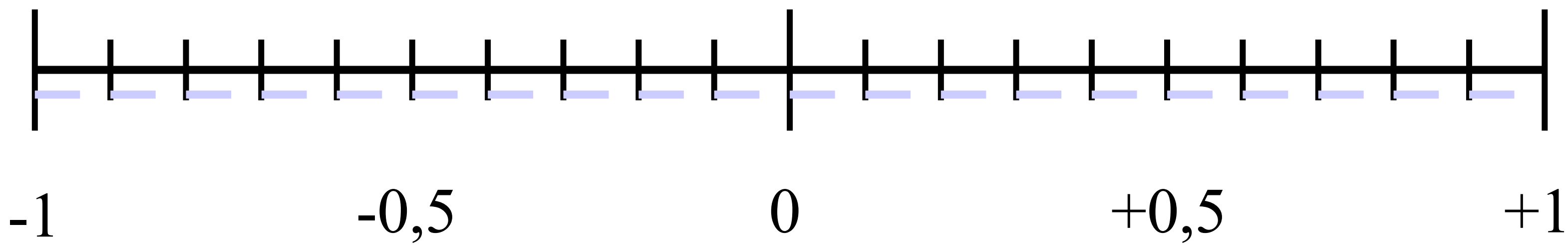
Nach: Galtung, Johan (1969) Theory and Methods of Social Research.
Oslo: Universitetsforlaget, pp. 309 - 315

Beispiel:
Hypothese, derzufolge r ≥ 0,5 erwartet wird



Nach: Galtung, Johan (1969) Theory and Methods of Social Research.
Oslo: Universitetsforlaget, p. 314

Korrelation



$$E=20$$

$\boxed{-1{,}0 \le r < -0{,}9}$ $\boxed{-0{,}9 \le r < -0{,}8}$ $\boxed{-0{,}8 \le r < -0{,}7}$ […

…] $\boxed{0{,}7 \le r < 0{,}8}$ $\boxed{0{,}8 \le r < 0{,}9}$ $\boxed{0{,}9 \le r < 1{,}0}$