

Streuungsmaße von Stichproben



SPANNWEITE, VARIANZ,
STANDARDABWEICHUNG, QUARTILE,
KOVARIANZ, KORRELATIONSKOEFFIZIENT

Methodenzentrum Sozialwissenschaften - Erziehungswissenschaften
Goethe-Universität, Frankfurt am Main

Streuungsmaße: Streuungsdiagramm

2

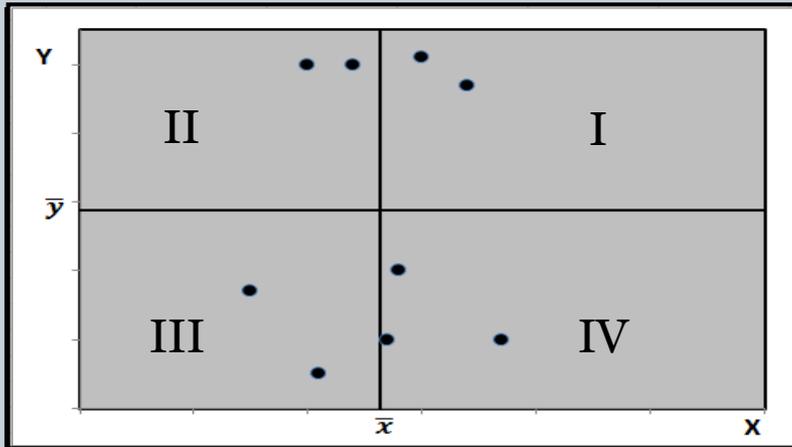
- Streuungsdiagramme werden eingesetzt, um Zusammenhänge zwischen 2 quantitativen Merkmalen X und Y **visuell** zu untersuchen.
- Eine Stichprobe der zwei Merkmale X (das unabhängige Merkmal) und Y (das abhängige Merkmal) wird auf folgende Weise bezeichnet:

$$\{(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)\}$$

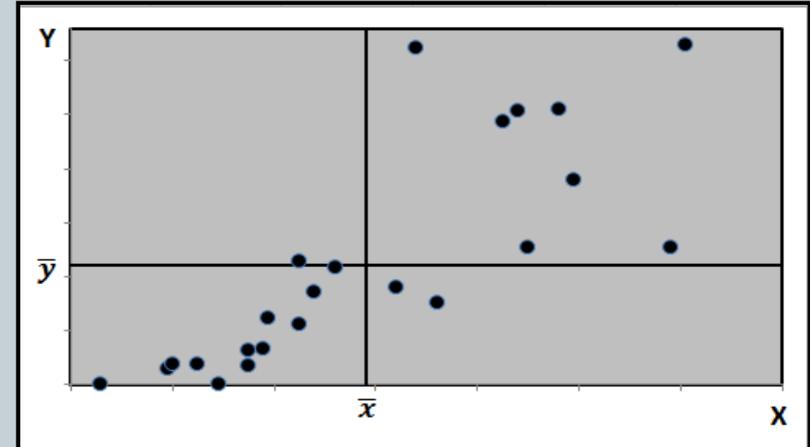
- In diesem Seminar werden wir **lineare Zusammenhänge** behandeln.

Streuungsmaße: Streuungsdiagramm

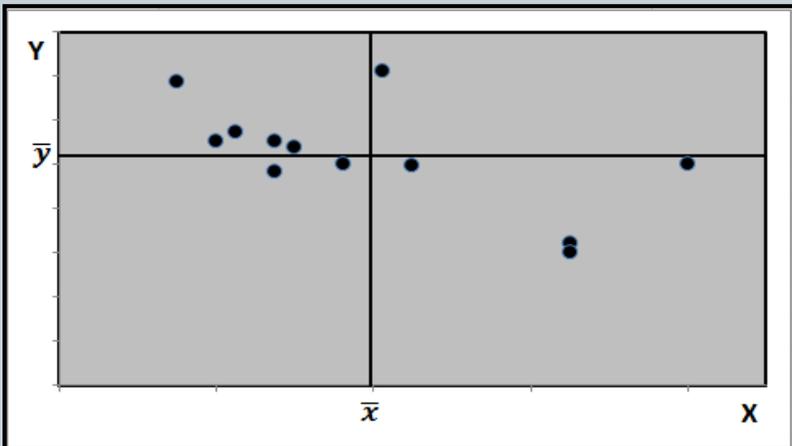
3



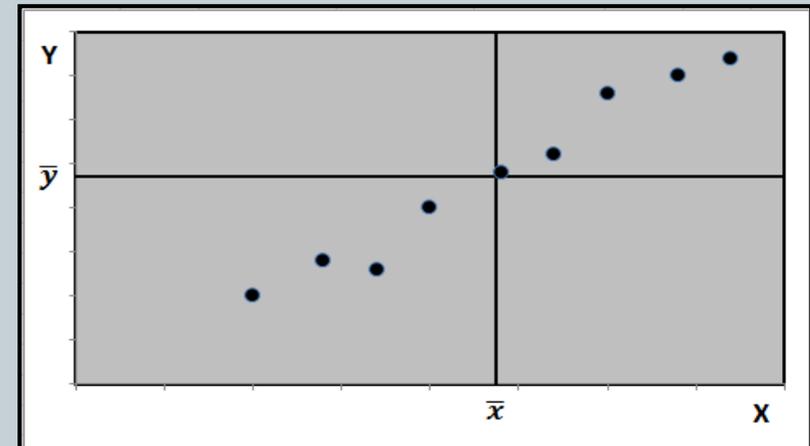
A: Kein Zusammenhang



B: positiver linearer Zusammenhang



C: Negativer (inverser) linearer Zusammenhang



D: positiver linearer Zusammenhang (stärker als in B)

Streuungsmaße: Zusammenhang zwischen 2 quantitative Merkmale

4

- Um den linearen Zusammenhang zwischen 2 quantitativen Merkmalen zu analysieren, gibt es 2 Streuungsmaße, die 2 verschiedene Komponenten des Zusammenhangs beschreiben:
 - Richtung des Zusammenhangs: **Kovarianz**
 - Stärke des Zusammenhangs: **Korrelationskoeffizient**

Streuungsmaße: Kovarianz

5

- Die Kovarianz (bezeichnet als s_{xy}) ist definiert als:

$$s_{xy} = \frac{\sum_{i=1}^n (X_i - \bar{x})(Y_i - \bar{y})}{n-1}$$

- $\sum_{i=1}^n (X_i - \bar{x})(Y_i - \bar{y})$ wird auch SAQ_{xy} genannt. Außerdem:

$$SAQ_{xy} = \sum_{i=1}^n X_i Y_i - n\bar{x}\bar{y}$$

- Die Kovarianz zeigt die **Richtung des Zusammenhangs**:
 - Wenn $s_{xy} > 0$: Es gibt einen positiven Zusammenhang zwischen X und Y.
 - Wenn $s_{xy} < 0$: Es gibt einen negativen Zusammenhang zwischen X und Y.

Streuungsmaße: Korrelationskoeffizient

6

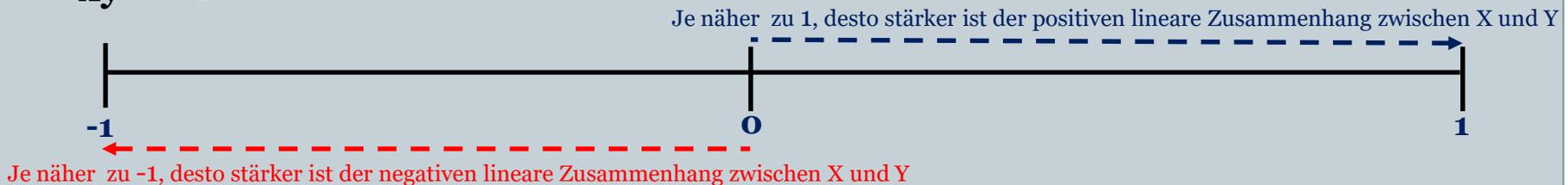
- Aber: Wie stark ist der Zusammenhang? Dafür berechnen wir **Pearsons Korrelationskoeffizient**:
- Korrelationskoeffizient (bezeichnet als r_{xy}) ist definiert als

bzw.

$$r_{xy} = \frac{s_{xy}}{s_x s_y}$$

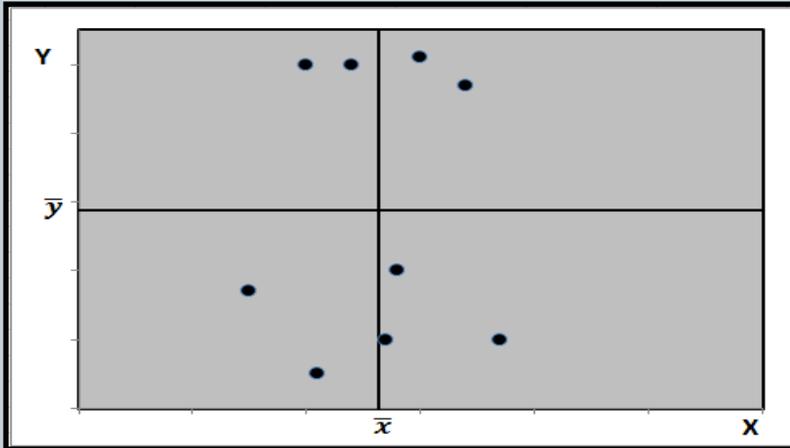
$$r_{xy} = \frac{SAQ_{xy}}{\sqrt{SAQ_x SAQ_y}}$$

- r_{xy} liegt zwischen -1 und 1

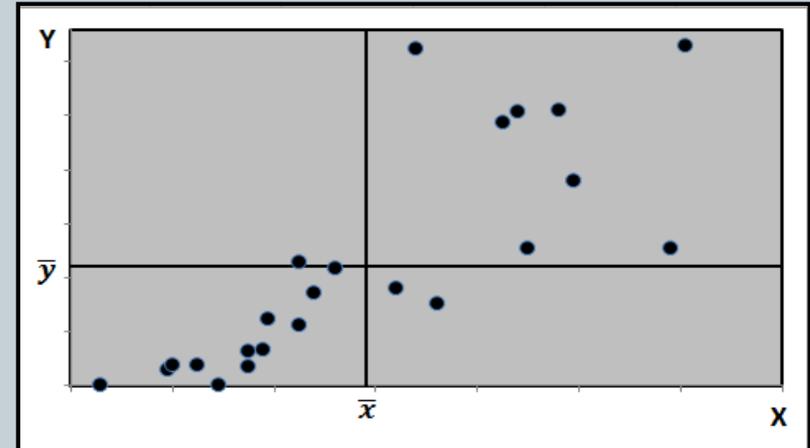


Streuungsmaße: Korrelationskoeffizient

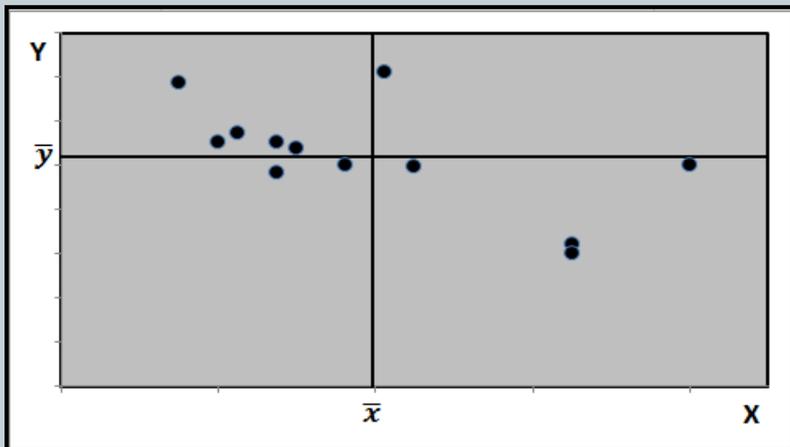
7



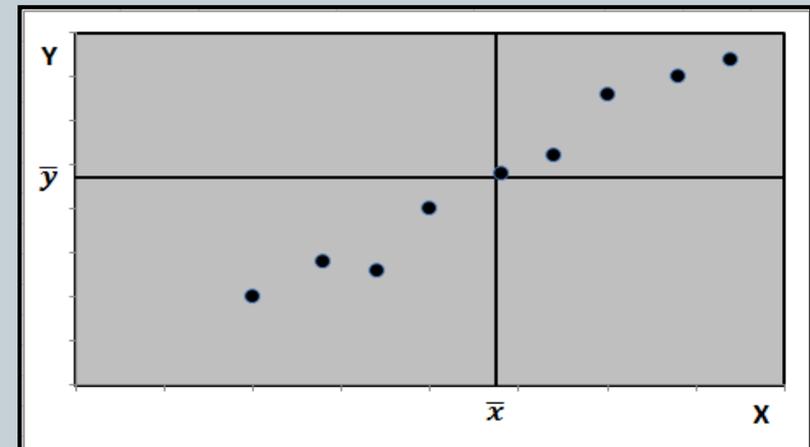
A: $r_{xy} = 0.063$



B: $r_{xy} = 0.806$



C: $r_{xy} = -0.622$



D: $r_{xy} = 0.986$

Streuungsmaße: Korrelationskoeffizient

8

- Die Interpretation von r_{xy} variiert je nach wissenschaftlicher Fachdisziplin. Für die Sozialwissenschaften kann man sich zum Beispiel an folgende Interpretationen orientieren:

nach Brosius:	
Betrag des Koeffizienten	Interpretation
0	keine Korrelation
$0 < r < 0,2$	sehr schwach
$0,2 < r < 0,4$	schwach
$0,4 < r < 0,6$	mittel
$0,6 < r < 0,8$	Stark
$0,8 < r < 1$	sehr stark
1	perfekt

nach Bühl / Zöfel:	
Betrag des Koeffizienten	Interpretation
0	keine Korrelation
$0 < r < 0,2$	sehr gering
$0,2 < r < 0,5$	gering
$0,5 < r < 0,7$	mittel
$0,7 < r < 0,9$	hoch
$0,9 < r < 1$	sehr hoch
1	perfekt