Bekannte zu prüfende Voraussetzungen beim Berechnen eines Unterschieds

1. Skalenniveau (Aufgabenstellung)
2. Anzahl der Stichproben (Aufgabenstellung)
3. AV oder UV (Aufgabenstellung)
4. Normalverteilung (Schiefe und Exzess)
5. Varianzhomogenität (mittels F-Test)

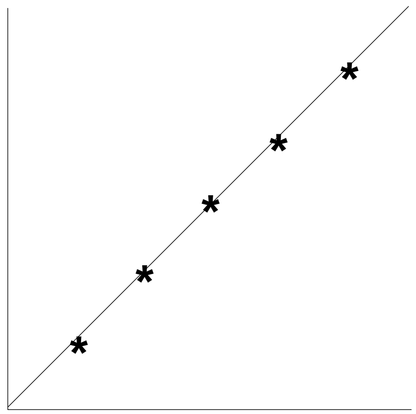
Bekannte Verfahren zur Berechnung eines Unterschieds

	Skalenniveau	Anzahl Stichprobe	UV / AV	Normal-verteilung	Varianz-homogenität
Chi-Quadrat	**Nominal**	**2**	**UV**	**/**	**/**
McNemar	**Nominal**	**2**	**AV**	**/**	**/**
Chi-Quadrat	**Nominal**	**>2**	**UV**	**/**	**/**
Cochran Q	**Nominal**	**>2**	**AV**	**/**	**/**
Mann-Whitney-U	**Ordinal**	**2**	**UV**	**/**	**/**
Wilcoxon	**Ordinal**	**2**	**AV**	**/**	**/**
T-Test AV	**Intervall**	**2**	**AV**	**Ja**	**Ja**
T-Test UV	**Intervall**	**2**	**UV**	**Ja**	**Zwei Tests**
Varianzanalyse	**Intervall**	**>2**	**UV**	**Ja**	**Ja**
Varianzanalyse Messwdh.	**Intervall**	**>2**	**AV**	**Ja**	**Ja**

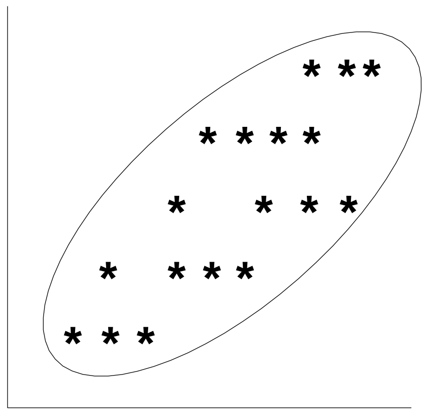
Berechnung von Zusammenhängen: Korrelationen

- Mit der Korrelation wird die Beziehung zwischen zwei Variablen untersucht (bivariate Korrelation).
- Das Ziel der Korrelationsrechnung ist es Maße für den Zusammenhang zweier Messwertreihen zu finden.
- Die Art des Zusammenhangs kann linear sein, oder evtl. die Form einer höheren Funktionsgleichung haben (quadratisch, exponentiell).
- Stärke des Zusammenhangs: Je besser die zugrunde liegende Zusammenhangsfunktion approximiert wird, desto höher ist der Zusammenhang.
- Signifikanz des Zusammenhangs: Es wird danach gefragt, ob die Art und die Stärke bzw. Höhe des Zusammenhangs auch für die Grundgesamtheit anzunehmen ist. Dazu wird der Korrelationskoeffizient aufn seine Signifikanz geprüft.

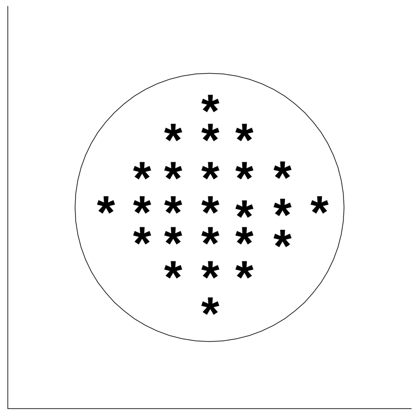
Graphische Darstellung von Zusammenhängen



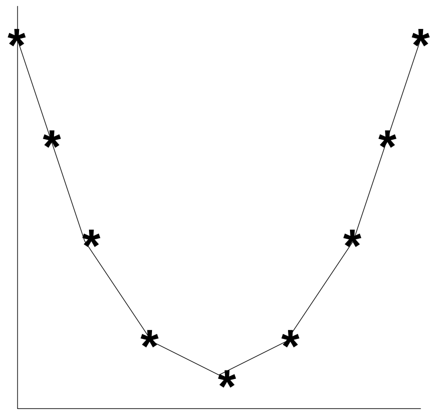
perfekt linear
positiver Zusammenhang
(r=1.00)



hoher positiver
Zusammenhang
(r ca. 0.80)



kein korrelativer
Zusammenhang
(r ca. 0.00)



U-förmige (nicht lineare)
Beziehung, r nicht
berechnen (Anwendungs-
voraussetzung für die
Berechnung von r nicht
erfüllt)

Interpretation von Korrelationsdiagrammen

- Ein Zusammenhang ist umso stärker (enger), je mehr sich die Messwertpaare um eine gedachte Gerade scharen, und umso geringer, je weniger dies der Fall ist.

- Die Richtung der Geraden lässt Aussagen über die Richtung des Zusammenhangs zu. Bei einer ansteigenden Geraden liegt eine positive, bei einer fallenden Geraden liegt eine negative Korrelation vor.

- Voraussetzung: Linearität

Interpretation der Höhe von Korrelationskoeffizienten

$r = 0$	kein Zusammenhang
$0.00 < \|r\| \leq 0.39$	niedriger Zusammenhang
$0.40 < \|r\| \leq 0.69$	mittlerer Zusammenhang
$0.70 < \|r\| \leq 0.99$	hoher Zusammenhang
$\|r\| \leq 1.00$	perfekter Zusammenhang

Überblick über die wichtigsten Korrelationskoeffizienten in Abhängigkeit vom Skalenniveau

- Intervallskalierung beider Variablen:
  Pearson Produkt-Moment-Korrelationskoeffizient r

- Ordinalskalierung beider Variablen:
  Rangkorrelationskoeffizient rho

- Nominalskalierung beider Variablen:
  Kontingenzkoeffizient C, V, phi

Berechnung des Prosdukt-Moment-Korrelationskoeffizienten

- Kovarianz (exaktes numerisches Maß für den Zusammenhang):

  Zur Berechnung der Kovarianz summiert man die Abweichungsprodukte aller Vpn auf und dividiert durch die um 1 verminderte Anzahl der Vpn:

$$COV(x,y) = \frac{1}{N-1}\sum(x_i - \bar{x})(y_i - \bar{y})$$

  Nachteil der Kovarianz: sie hängt von der Streuung der beiden Messreihen ab.

- Produkt-Moment-Korrelationskoeffizient (normiertes Maß):

  Zur Berechnung der Korrelation dividiert man die Kovarianz durch das Produkt der Standardabweichungen der x- bzw. der y-Werte:

$$r_{xy} = \frac{\frac{1}{N-1}\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{N-1}\sum(x_i-\bar{x})^2}\sqrt{\frac{1}{N-1}\sum(y_i-\bar{y})^2}} = \frac{COV(x,y)}{s_x \times s_y}$$

Berechnung des Produkt-Moment-Korrelationskoeffizienten

- Vereinfachte Berechnungsmöglichkeit:

$$r_{xy} = \frac{N \times \sum x_i y_i - (\sum x_i) \times (\sum y_i)}{\sqrt{N \times \sum x_i^2 - (\sum x_i)^2} \times \sqrt{N \times \sum y_i^2 - (\sum y_i)^2}}$$

- Berechnung mit Hilfe von Standardwerten:

$$r_{xy} = \frac{\sum z_{xi} \times z_{yi}}{N - 1}$$

Berechnung des Rangkorrelationskoeffizienten rho

- Das Verfahren der Rangkorrelation beruht auf einem Vergleich der Rangplätze, die eine Versuchsperson innerhalb der beiden monovariablen Verteilungen einnimmt. Es wird also geprüft, inwieweit Versuchspersonen für zwei unterschiedliche Merkmale in einer Gruppe denselben oder unterschiedliche Rangplätze hinsichtlich ihrer Leistungen einnehmen. Die Prüfgröße lautet:

$$RHO = 1 - \frac{6 \times \sum d_i^2}{N\left(N^2 - 1\right)}$$

Berechung des Produkt-Moment-Korrelationskoeffizienten

1. Problemstellung:
   Es liegen zwei intervallskalierte Messreihen vor.
2. Hypothesen:
   H0: Es gibt keinen signifikanten Zusammenhang zwischen zwei Messreihen.
   H1: Es gibt einen signifikanten Zusammenhang zwischen zwei Messreihen.
3. Voraussetzung:
   Intervallskalenniveau
4. Berechung:
   mit einer der drei obig genannten Formeln
5. Signifikanzprüfung:
   emp Wert > theor. Wert à H1
   emp Wert < theor. Wert à H0
6. Interpretation

Beispiel Produkt-Moment-Korrelationskoeffizient

1. Problemstellung:

Berechnung des Zusammenhangs von Klimmzügen und Liegestützen:

Schüler	Liegestütze (x)	Klimmzüge (y)	Schüler	Liegestütze (x)	Klimmzüge (y)
1	17	12	11	13	4
2	15	10	12	13	4
3	13	9	13	13	4
4	14	8	14	6	3
5	10	7	15	6	3
6	14	7	16	8	3
7	11	6	17	3	2
8	7	5	18	9	2
9	10	5	19	5	1
10	10	5	20	3	0

Beispiel Produkt-Moment-Korrelationskoeffizient

2. Hypothesen:

   H0: Es gibt keinen Zusammenhang zwischen der Leistung bei den Liegestützen und der Leistung bei den Klimmzügen.

   H1. Es gibt einen signifikanten Zusammenhang zwischen der Leistung bei den Liegestützen und der Leistung bei den Klimmzügen.

3. Voraussetzung:

   Intervallskala liegt vor.

4. Berechung:

   - In der Vorlesung
   - Determinationskoeffizient r^2:

   Er erklärt die Größe des Anteils wechselseitig erklärter Varianz der beiden Merkmale, d.h. …% Varianzanteil der einen Variablen ist durch die andere Variable determiniert.

Tabelle zur Berechnung

Vp	x-xquer	y-yquer	(x-xq)*(y-yq)	Vp	x-xquer	y-yquer	(x-xq)*(y-yq)
1	7	7	49	11	3	-1	-3
2	5	5	25	12	3	-1	-3
3	3	4	12	13	3	-1	-3
4	4	3	12	14	-4	-2	8
5	0	2	0	15	-4	-2	8
6	4	2	8	16	-2	-2	4
7	1	1	1	17	-7	-3	21
8	-3	0	0	18	-1	-3	3
9	0	0	0	19	-5	-4	20
10	0	0	0	20	-7	-5	35
∑				∑	0	0	197

Tabelle zur Berechnung

Vp	$(x-xq)^2$	$(y-yq)^2$	x^2	y^2	Vp	$(x-xq)^2$	$(y-yq)^2$	x^2	y^2
1	49	49	289	144	11	9	1	169	16
2	25	25	225	100	12	9	1	169	16
3	9	16	169	81	13	9	1	169	16
4	16	9	196	64	14	16	4	36	9
5	0	4	100	49	15	16	4	36	9
6	16	4	196	49	16	4	4	64	9
7	1	1	121	36	17	49	9	9	4
8	9	0	49	25	18	1	9	81	4
9	0	0	100	25	19	25	16	25	1
10	0	0	100	25	20	49	25	9	0
$\sum$					$\sum$	312	182	2312	682

Tabelle zur Berechnung

Vp	xy	z_x	z_y	z_x*z_y	Vp	xy	z_x	z_y	z_x*z_y
1	204	1.73	2.26	3.91	11	52	0.74	-0.32	-0.24
2	150	1.23	1.62	1.99	12	52	0.74	-0.32	-0.24
3	117	0.74	1.29	0.95	13	52	0.74	-0.32	-0.24
4	112	0.99	0.97	0.96	14	18	-0.99	-0.65	0.65
5	70	0.00	0.65	0.00	15	18	-0.99	-0.65	0.65
6	98	0.99	0.65	0.64	16	24	-0.49	-0.65	0.32
7	66	0.25	0.32	0.08	17	6	-1.73	-0.97	1.68
8	35	-0.74	0.00	0.00	18	18	-0.25	-0.97	0.24
9	50	0.00	0.00	0.00	19	5	-1.23	-1.29	1.59
10	50	0.00	0.00	0.00	20	0	-1.73	-1.62	2.80
$\sum$					$\sum$	1197	0.00	0.00	15.74

Beispiel Produkt-Moment-Korrelationskoeffizient

5. Signifikanzprüfung:

   n = 20; p=.05 à r = 0.433 à H1

6. Interpretation:

   Es gibt einen signifikanten Zusammenhang zwischen Klimmzügen und Liegestützen.

Berechung des Rangkorrelationskoeffizienten

1. Problemstellung:
   Es liegen zwei ordinalskalierte Messreihen vor.
2. Hypothesen:
   H0: Es gibt keinen signifikanten Zusammenhang zwischen zwei Messreihen.
   H1: Es gibt einen signifikanten Zusammenhang zwischen zwei Messreihen.
3. Voraussetzung:
   Ordinalskalenniveau
4. Berechung:
   mit der obig genannten Formel
5. Signifikanzprüfung:
   emp Wert > theor. Wert à H1
   emp Wert < theor. Wert à H0
6. Interpretation

Beispiel Rangkorrelationskoeffizient

1. Problemstellung:

   Gleiche Daten wie aus dem Beispiel für den Produkt-Moment-Korrelationskoeffizienten.

2. Hypothesen:

   H0: Es gibt keinen Zusammenhang zwischen der Leistung bei den Liegestützen und der Leistung bei den Klimmzügen.

   H1. Es gibt einen signifikanten Zusammenhang zwischen der Leistung bei den Liegestützen und der Leistung bei den Klimmzügen.

3. Voraussetzung:

   Ordinalskala liegt vor.

4. Berechung:

   - In der Vorlesung

Beispiel Rangkorrelationskoeffizient

5. Signifikanzprüfung:

   n = 20; p=.05 à r = 0.377 à H1

6. Interpretation:

   Es gibt einen signifikanten Zusammenhang zwischen Klimmzügen und Liegestützen.

Regressionsanalyse

- Untersucht die Abhängigkeit einer oder mehrerer (abhängiger) Kriteriumsvariablen von (unabhängigen) Prädiktorvariablen.

- Sämtliche Variablen einer Regressionsanalyse müssen intervallskaliert sein.

- Durch die Regressionsgleichung können die unabhängigen Variablen zur Vorhersage, bzw. zur Schätzung der abhängigen Variablen eingesetzt werden:

$$y = ax1 + bx2 + cx3 + d$$