




## Inferenzstatistik: Bivariate Zusammenhänge

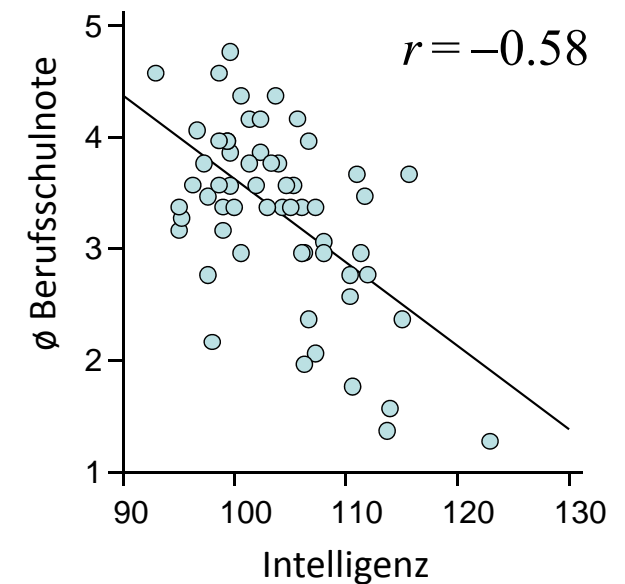
- 1 Bivariate Korrelationen
  - 1.1 Produkt-Moment Korrelation  $r$
  - 1.2 Rangkorrelationen
- 2 Bivariate Regression

## Einführende Literatur

-  Bortz, J. & Schuster, Ch. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Auflage). Berlin: Springer. [Kap. 10.2 & 11.2]
-  Diehl, J. M. & Arbinger, R. (2001). *Einführung in die Inferenzstatistik* (3. Auflage). Eschborn bei Frankfurt: Klotz Verlag. [Kap. 16]
-  Eid, M., Gollwitzer, M. & Schmitt, M. (2015). *Statistik und Forschungsmethoden* (4. Auflage). Weinheim: Beltz. [Kap. 16.4 & 17.9]

# Wiederholung Korrelationen

- Korrelationen sind Maße zur Bestimmung des (bivariaten) Zusammenhangs zwischen zwei Variablen  $X$  und  $Y$ . Zusammenhänge lassen sich in Streudiagrammen visualisieren.
- Korrelationen geben die Stärke und Richtung eines Zusammenhangs an und schwanken (in der Regel) zwischen  $-1$  (=perfekter negativer Zusammenhang) und  $+1$  (=perfekter positiver Zusammenhang), wobei  $0$  einen fehlenden Zusammenhang anzeigt.
- Korrelationen erlauben keinen Rückschluss auf die Kausalrichtung des Einflusses.
- Je nach Skalenniveau der beiden Variablen kommen jeweils unterschiedliche Maße zum Einsatz (vgl. nächste Folie).
- Die **Produkt-Moment Korrelation**  $r$  quantifiziert den **linearen** Zusammenhang zwischen zwei **intervallskalierten** Variablen. Einige Eigenschaften sind:
  - Sie ist invariant gegenüber linearen Transformationen von  $X$  und  $Y$ .
  - Sie steht in folgender Beziehung zum Regressionsgewicht  $r = b \cdot s_X / s_Y$  der linearen Regression.



$$\hat{y}_i = 11.08 - 0.08 \cdot x_i$$

# Wiederholung Korrelationen

Korrelationskoeffizienten in Abhängigkeit vom Skalenniveau der Variablen  $X$  und  $Y$

		Skalenniveau von $X$			
		intervall	ordinal	künstlich dichotom	natürlich dichotom
Skalenniveau von $Y$	intervall	Produkt-Moment Korrelation $r$	u.a. Rangkorrelationen nach Spearman $r_s$ oder Kendall $\tau$	biseriale Korrelation $r_b$	punktbiseriale Korrelation $r_{pb}$
	ordinal			biseriale Rangkorrelation	
	nominal: künstlich dichotom			tetrachorische Korrelation $r_{tet}$	$\phi$ -Koeffizient
	nominal: natürlich dichotom				

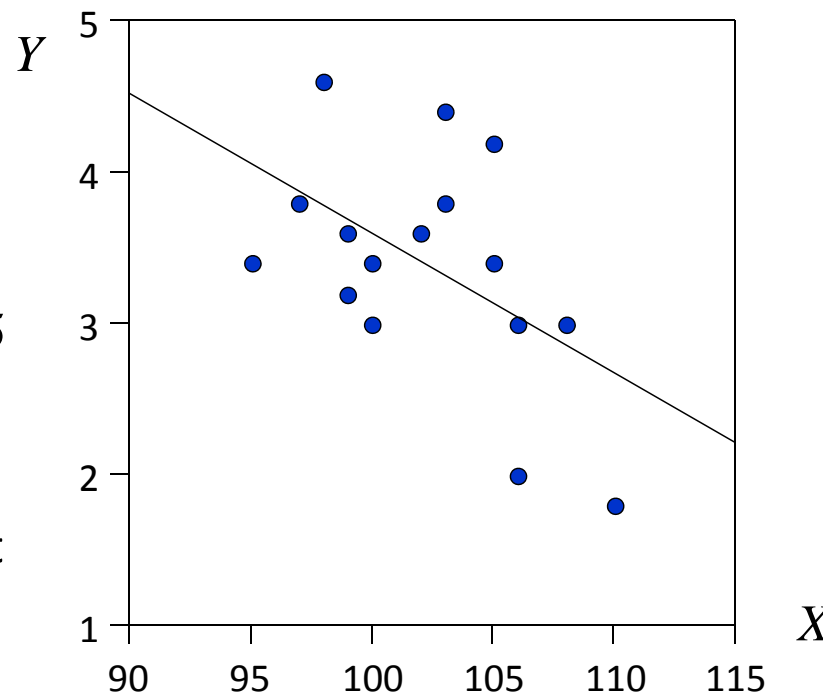
- Für alle hier dargestellten Korrelationen (wie auch die meisten anderen Korrelationen, z.B. die Partialkorrelation) existieren inferenzstatistische Prüfverfahren, die in den Statistik-Büchern dargestellt sind (z.B. Diehl & Arbinger, 2001, S. 388ff; Eid, Gollwitzer & Schmidt, 2015, Kap. 16.4). Im Folgenden werden wir nur die Verfahren für  $r$  und paarbasierte Rangkorrelationen wie  $\tau$  näher behandeln.

# Signifikanztest für Produkt-Moment Korrelation

- **Beispiel:** Es wurde ein neuer Intelligenztest entwickelt und von allen 16 in einem Stahlunternehmen in einem Jahrgang eingestellten Auszubildenden zum Industriemechaniker durchgeführt. Es wird die Hypothese geprüft, dass ein negativer Zusammenhang zwischen den Intelligenzwerten ( $=X$ ) und der Abschlussnote in der Berufsschule besteht ( $=Y$ ).
- Wir gehen (diskussionswürdig) davon aus, dass die Berufsschulnote metrisch ist, betrachten den Zusammenhang zunächst in einem Streudiagramm und berechnen die Produkt-Moment Korrelation (vgl. Statistik I):

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{(n-1) \cdot s_X \cdot s_Y} = -0.515$$

- Frage: Besteht ein statistisch signifikanter Zusammenhang? Ist diese Korrelation statistisch signifikant von 0 verschieden?

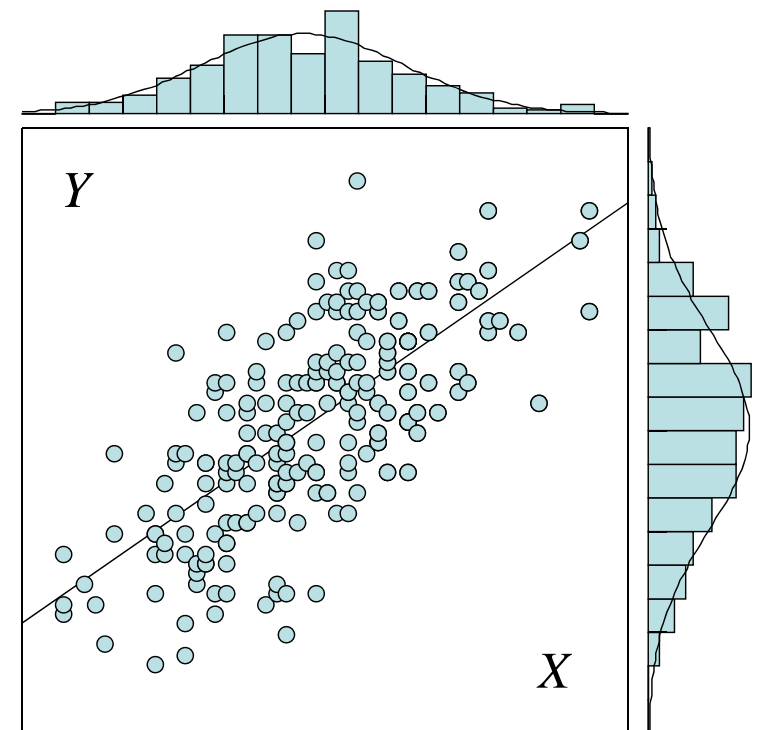


Nr.	X	Y
1	110	1.8
2	99	3.6
3	100	3.4
4	103	4.4
5	100	3.0
6	105	3.4
7	108	3.0
8	95	3.4
9	99	3.2
10	102	3.6
11	98	4.6
12	106	3.0
13	105	4.2
14	103	3.8
15	106	2.0
16	97	3.8

# Signifikanztest für Produkt-Moment Korrelation

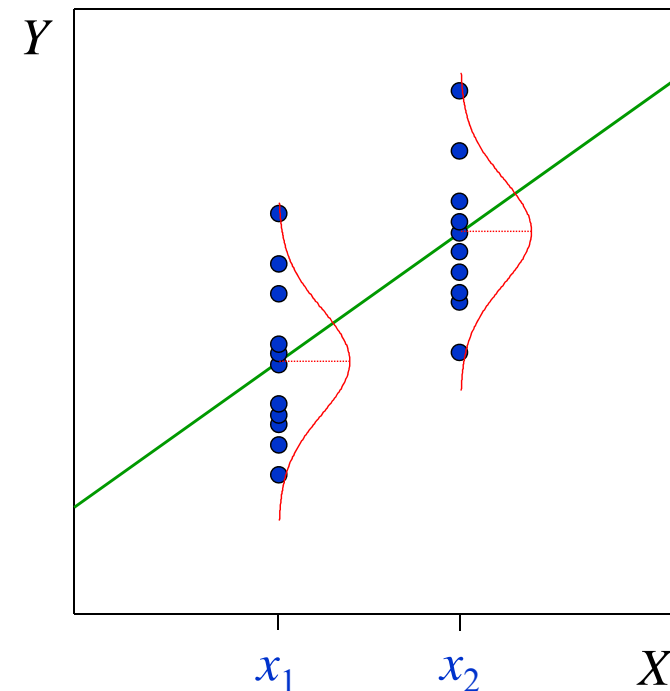
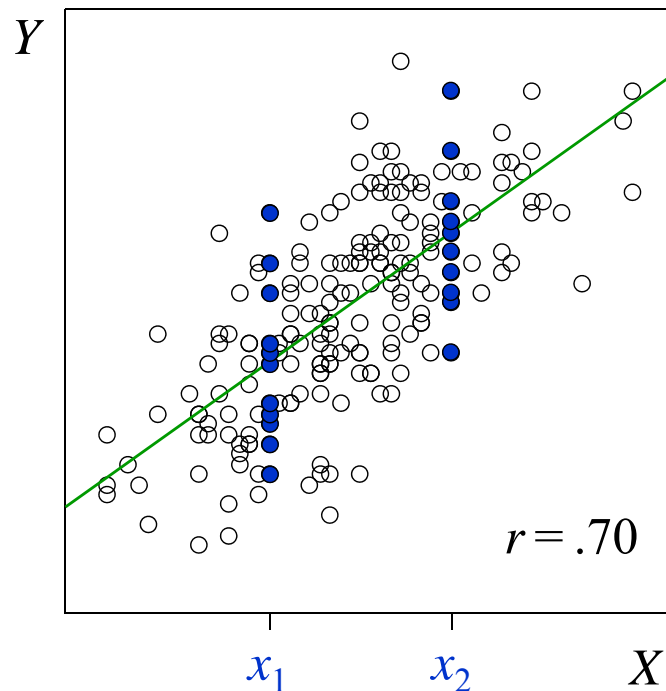
- Bezeichnung: Signifikanztest für Produkt-Moment Korrelation  $r$ .
- Einsatzbereich: Prüfung der Hypothese, dass die Produkt-Moment Korrelation zwischen zwei metrischen Variablen  $X$  und  $Y$  ungleich einem festen Wert  $\rho_0$  ist. Meist wird  $\rho_0 = 0$  betrachtet, d.h. es wird die Hypothese geprüft, dass  $X$  und  $Y$  linear unabhängig sind.
- Hypothesen: allgemein zweiseitig:  $H_0: \rho = \rho_0$  und  $H_1: \rho \neq \rho_0$  oder entsprechend einseitig. Im Folgenden betrachten wir zunächst den Spezialfall:  $H_0: \rho = 0$  und  $H_1: \rho \neq 0$ .
- Voraussetzungen: Die beiden metrischen Variablen  $X$  und  $Y$  müssen sich **bivariat normalverteilen**.

Damit eine bivariate Normalverteilung von  $X$  und  $Y$  vorliegt, müssen nicht nur die (Rand-) Verteilungen der Variablen  $X$  und  $Y$  Normalverteilungen sein. (Dies ist also nur eine notwendige, aber keine hinreichende Bedingung für die bivariate Normalverteilung.)



# Bivariate Normalverteilung

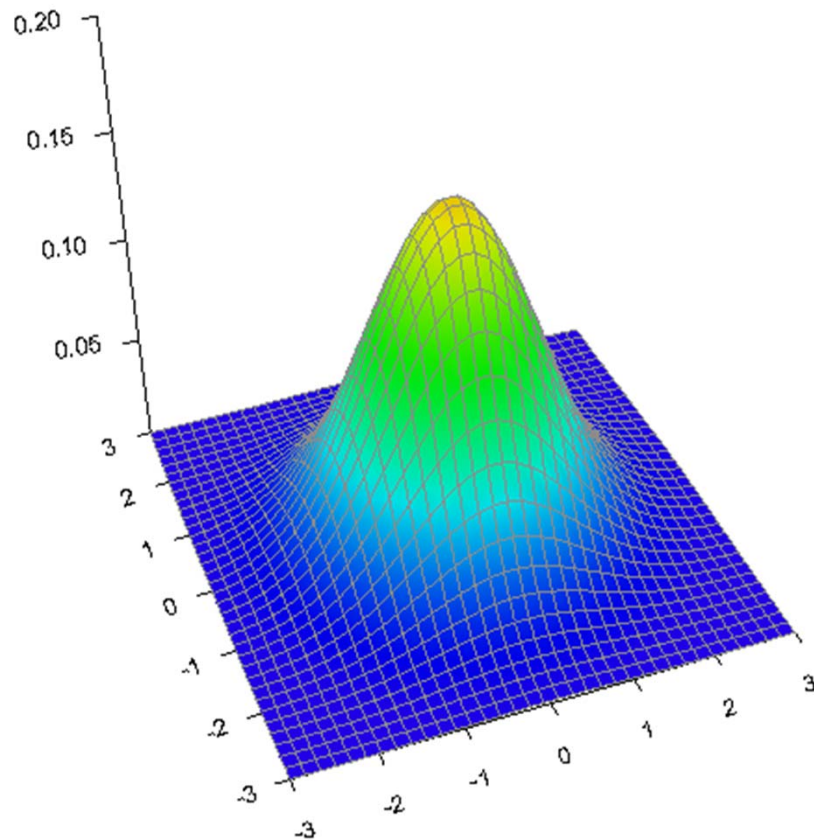
- Voraussetzungen (Fs.): Zudem müssen alle bedingten Verteilungen Normalverteilungen sein, d.h. für jeden beliebigen festen  $x$ -Wert müssen die zugehörigen  $y$ -Werte normalverteilt sein (und umgekehrt). Statt bedingter Verteilung finden sich auch die Bezeichnung konditionale Verteilungen oder Array-Verteilungen.



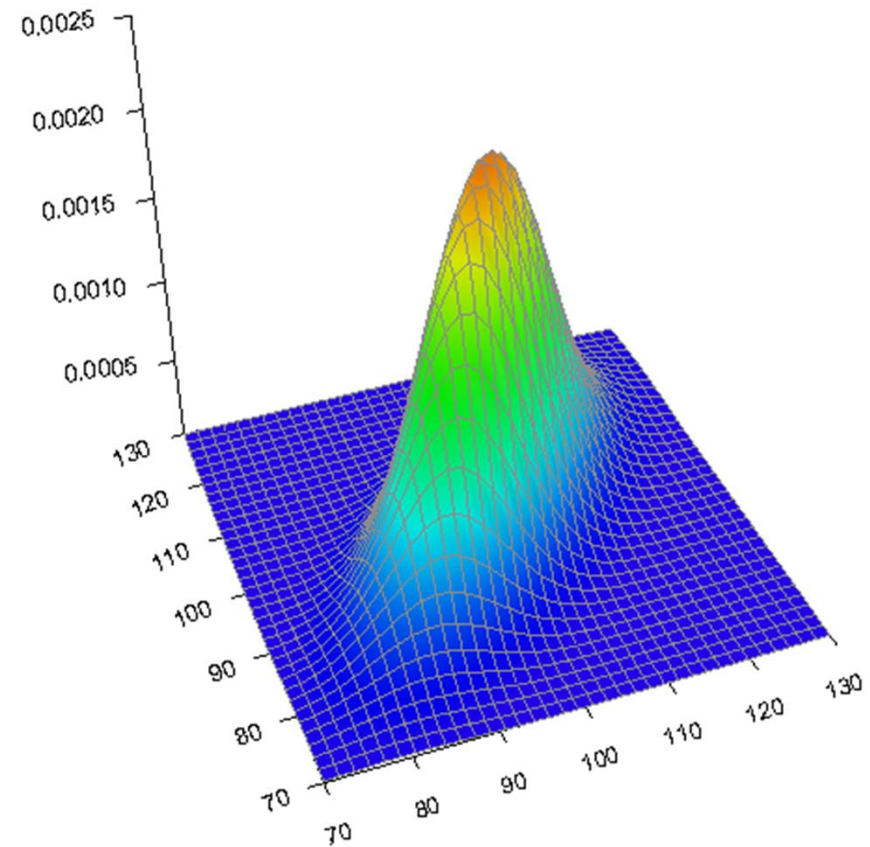


# Bivariate Normalverteilung

Die Dichte der bivariaten Normalverteilung hängt von fünf Parametern ab:  $NV_2(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho)$ .



$NV_2(0,0,1,1,0)$



$NV_2(100,100,10,10,0.70)$

# Bivariate Normalverteilung

- Die Prüfung, ob bei zwei Variablen eine bivariate Normalverteilung (und bei der Verallgemeinerung auf mehr als zwei Variablen eine multivariate Normalverteilung) vorliegt, ist schwierig.
- Eine univariate Testung jeweils der beiden Variablen  $X$  und  $Y$  (z.B. mit dem Lilliefors-Test) ist aus den genannten Gründen nicht hinreichend.
- Häufiger angewendete Tests sind verallgemeinerte Tests für Schiefe und Exzess von Mardia. Neuere Computersimulationen zeigen, dass ein von Henze & Zirkler entwickelter Test bei  $n \geq 75$  gute Eigenschaften aufweist (vgl. Farrell, Salibian-Barrera & Naczk, 2007). Beide sind in SPSS nicht implementiert (aber z.B. im Statistik-Programm SYSTAT in `Basic_Statistics`; die Mardia-Statistiken können auch über einen SPSS-Macro bestimmt werden, vgl. DeCarlo, 1997).
- Da die Testung der Hypothese  $H_0: \rho = 0$  (nicht aber bei  $\rho_0 \neq 0$ ) relativ robust gegenüber der Verletzung der Voraussetzung der bivariaten Normalverteilung ist, wird diese Annahme in der Regel nicht getestet.

# Signifikanztest für Produkt-Moment Korrelation

- Vorgehen: Gilt die  $H_0$ , so resultiert bei  $\rho_0 = 0$  zur Prüfung folgende  $t$ -verteilte Prüfgröße

$$t = \frac{r - \rho_0}{\sigma_r} = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = r \cdot \sqrt{\frac{n-2}{1-r^2}} \quad \text{mit} \quad df = n-2$$

- Entscheidung: Zurückweisung der  $H_0$ , falls  $|t| > t_{crit}$  mit  $t_{crit} = t_{n-2; 1-\alpha/2}$  bei zweiseitiger Prüfung und  $t_{crit} = t_{n-2; 1-\alpha}$  bei einseitiger Prüfung (und korrekter Richtung).
- (Manche Statistikprogramme berichten statt der  $t$ - eine  $F$ -Prüfgröße. Da jede  $t$ -verteilte Prüfgröße mit  $k$  Freiheitsgraden, wenn man sie quadriert, einer  $F$ -verteilten mit  $df_1 = 1$  und  $df_2 = k$  Freiheitsgraden entspricht, resultiert daraus dieselbe Entscheidung.)
- Will man nicht testen, ob eine Korrelation statistisch signifikant von 0 verschieden ist, sondern die Abweichung von einem Wert  $\rho_0 \neq 0$  prüfen, so ist die obige  $t$ -Prüfgröße nicht mehr angemessen.
- Je stärker  $\rho_0$  von 0 abweicht (und je kleiner  $n$  ist), desto schiefer ist die Stichprobenkennwerteverteilung unter der  $H_0$ . Durch eine im Folgenden dargestellte Transformation kann eine Symmetrisierung der Stichprobenkennwerteverteilung erzielt werden.

# Signifikanztest für Produkt-Moment Korrelation

- Man kann dazu die Korrelation mittels der folgenden monotonen **Fisher-Z-Transformation** (nicht zu verwechseln mit der gewöhnlichen z-Transformation) transformieren (für  $|r| < 1$ ):

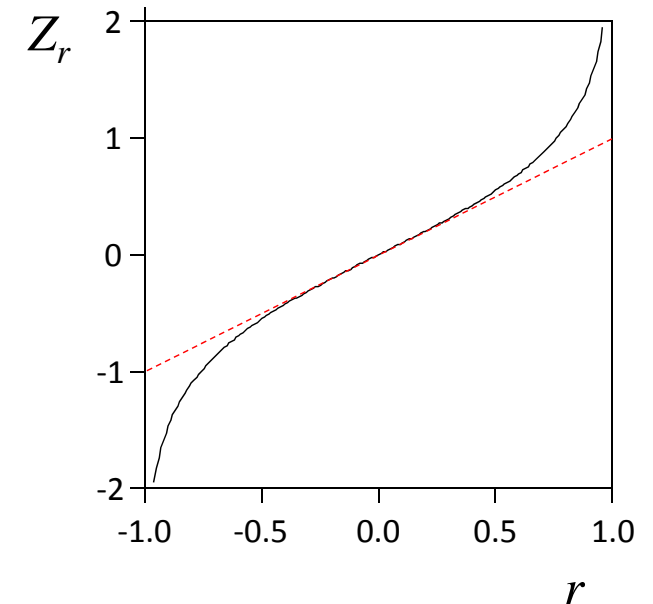
$$Z_r = \frac{1}{2} \cdot \ln \frac{1+r}{1-r}$$

mit  $\ln$  als natürlicher Logarithmus, d.h. als Logarithmus zur Basis  $e \approx 2.718$  und der Eigenschaft  $Z_{-r} = -Z_r$ .

- Die Stichprobenkennwerteverteilung von  $Z_r$  weist (im Gegensatz zu der von  $r$ ) bereits ab  $n \geq 25$  eine gute Annäherung an die Normalverteilung auf; mit dem Standardfehler

$$\sigma_{Z_r} = \frac{1}{\sqrt{n-3}}.$$

- Weder ist  $r$  ein erwartungstreuer Schätzer von  $\rho$  noch  $Z_r$  von  $Z_\rho = 0.5 \cdot \ln(1 + \rho) / (1 - \rho)$ . D.h. der Erwartungswert der Z-transformierten Korrelationen ist nicht gleich der Z-Transformation der Populationskorrelation. Bei größerem  $n$  ist der (unterschätzende) Bias aber klein; daher werden üblicherweise keine Korrekturen vorgenommen (näheres vgl. Schulze, 2004).



# Signifikanztests für Produkt-Moment Korrelation

➤ Die Fisher-Z transformierte Korrelation wird statt  $r$  bei der Beantwortung einer Reihe von statistischen Fragen herangezogen, darunter ...

- Unterscheidet sich die Korrelation statistisch signifikant von einem Wert  $\rho_0 \neq 0$ ?  
Dabei resultiert folgende, approximativ normalverteilte Prüfgröße:

$$z = \frac{Z_r - Z_{\rho_0}}{\sigma_{Z_r}} = \frac{Z_r - Z_{\rho_0}}{\sqrt{\frac{1}{n-3}}}$$

- Unterscheiden sich die Korrelationen zweier Variablen  $X$  und  $Y$  mit den Werten  $r_1$  und  $r_2$  aus zwei unabhängigen Gruppen (mit Stichprobenumfängen  $n_1$  und  $n_2$ ) statistisch signifikant voneinander? Prüfgröße:

$$z = \frac{Z_{r_1} - Z_{r_2}}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}}$$

Es gibt auch Signifikanzprüfungen z.B. für die Fälle, dass die Korrelationen der beiden Variablen  $X$  und  $Y$  aus abhängigen Stichproben stammen (z.B. aufgrund von Messwiederholung) oder die Frage, ob eine Variable  $X$  höher mit  $Y$  als mit einer anderen Variablen  $Z$  korreliert (vgl. Diehl & Arbinger, 2001, S. 380ff).

# Signifikanztests für Produkt-Moment Korrelation

- Die Fisher-Z transformierte Korrelation wird statt  $r$  bei der Beantwortung einer Reihe von statistischen Fragen herangezogen, darunter ... (Fs.)
  - Wie sieht das **Konfidenzintervall** um  $r$  aus? Antwort:

$$Z_r \pm z_{1-\alpha/2} \cdot \sqrt{\frac{1}{n-3}}$$

Da dabei das Konfidenzintervall um  $Z_r$  und nicht um  $r$  gebildet wird, muss man die Grenzen des Konfidenzintervalls anschließend wieder durch folgende Transformation in die  $r$ -Metrik zurückrechnen:

$$r = \frac{e^{2 \cdot Z_r} - 1}{e^{2 \cdot Z_r} + 1} \quad \left( \Leftrightarrow \quad Z_r = \frac{1}{2} \cdot \ln \frac{1+r}{1-r} \right)$$

# Signifikanztests für Produkt-Moment Korrelation

- Die Fisher-Z transformierte Korrelation wird statt  $r$  bei der Beantwortung einer Reihe von statistischen Fragen herangezogen, darunter ... (Fs.)
  - der **Mittlung von  $k$  Korrelationen**, indem man zunächst alle Korrelationen Fisher-Z transformiert und dann, gewichtet mit dem Stichprobenumfang, wie folgt mittelt

$$\bar{Z}_r = \frac{\sum_{j=1}^k (n_j - 3) \cdot Z_{r_j}}{\sum_{j=1}^k (n_j - 3)}$$

Bei gleichen  $n_j$  in allen Gruppen vereinfacht sich die Formel zu:

$$\bar{Z}_r = \frac{\sum_{j=1}^k Z_{r_j}}{k}$$

und dann mittels der Formel auf der vorangegangenen Folie wieder in die  $r$ -Metrik zurückrechnet (vgl. Eid, Gollwitzer & Schmidt, 2010, S. 543ff)

- Insgesamt ist umstritten, ob man beider Mittlung von Korrelationen die Fisher-Z-Transformation (oder eine andere Transformation) anwenden soll (vgl. Schulze, 2004).

# Signifikanztests für Produkt-Moment Korrelation

- **Beispiel:** Es wird die Hypothese geprüft, dass ein negativer Zusammenhang zwischen der Intelligenz ( $=X$ ) und der Berufsschulnote besteht ( $=Y$ ). Es resultiert  $r = -0.515$  ( $n = 16$ ).
- Wir testen zweiseitig bei  $\alpha = 0.05$ , ob die Korrelation statistisch signifikant von 0 verschieden ist und setzen dazu ein

$$t = r \cdot \sqrt{\frac{n-2}{1-r^2}} = -0.515 \cdot \sqrt{\frac{16-2}{1-(-0.515)^2}} = -2.25$$

- Der tabellierten  $t$ -Verteilung entnehmen wir für  $\alpha = 0.05$  und  $df = n - 2 = 14$  sowie zweiseitiger Testung den Wert  $t_{crit} = t_{14;0.975} = 2.145$ .

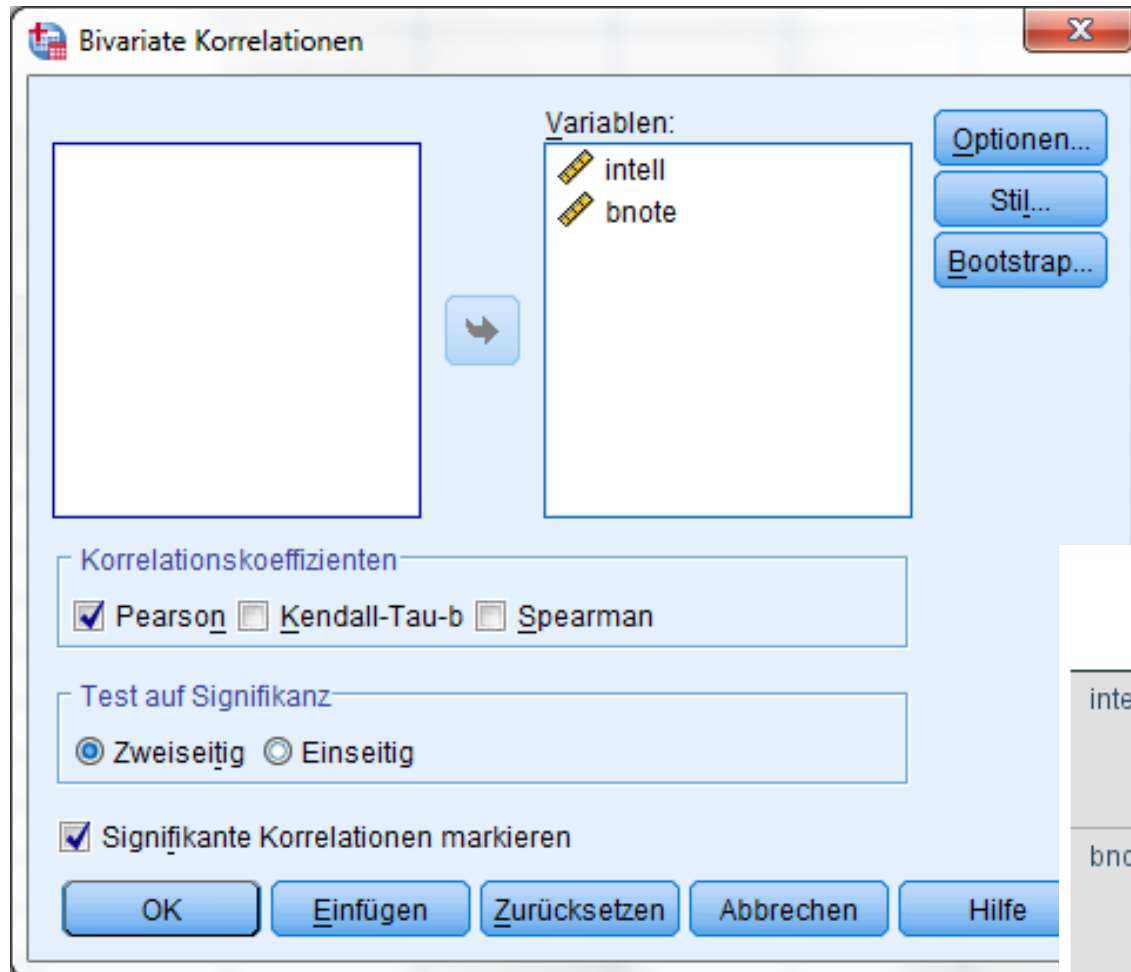
Da  $|t| > t_{crit}$  ist weisen wir die Nullhypothese zurück und schließen, dass ein statistisch signifikanter negativer Zusammenhang zwischen der Intelligenz und der Berufsschulnote besteht.

X	Y
110	1.8
99	3.6
100	3.4
103	4.4
100	3.0
105	3.4
108	3.0
95	3.4
99	3.2
102	3.6
98	4.6
106	3.0
105	4.2
103	3.8
106	2.0
97	3.8



# Signifikanztest für Produkt-Moment Korrelation in SPSS

- Die Produkt-Moment Korrelation  $r$  erhält man unter Analyseieren/Korrelation/Bivariat...



SPSS erlaubt nur die Prüfung der  $H_0: \rho = 0$

Da  $0.041 < \alpha = 0.05$  (zweiseitig), besteht ein statistisch signifikanter negativer Zusammenhang.

		intell	bnote
intell	Korrelation nach Pearson	1	-,515 <sup>*</sup>
	Signifikanz (2-seitig)		,041
	N	16	16
bnote	Korrelation nach Pearson	-,515 <sup>*</sup>	1
	Signifikanz (2-seitig)	,041	
	N	16	16

\*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

# Signifikanztests für Produkt-Moment Korrelation

- **Beispiel:** wie eben. Wir wollen jetzt die Hypothese prüfen, dass unser neu entwickelter Intelligenztest die Berufsschulnoten besser vorhersagt als die bisher dazu eingesetzten Intelligenztests. Einer Metaanalyse entnehmen wir, dass der Zusammenhang zwischen der Intelligenz und Schulnoten bei Auszubildenden typischerweise  $-0.32$  beträgt (Hülshager, Maier, Stumpp & Muck, 2006).
- Wir testen wieder zweiseitig und bei  $\alpha = 0.05$ . Um in die Formel einsetzen zu können konvertieren wir beide Korrelationen mittels der Fisher-Z-Transformation und erhalten:

$$Z_r = \frac{1}{2} \cdot \ln \frac{1+r}{1-r} = \frac{1}{2} \cdot \ln \frac{1+(-0.515)}{1-(-0.515)} = -0.569$$

$$Z_{\rho_0} = \frac{1}{2} \cdot \ln \frac{1+\rho_0}{1-\rho_0} = \frac{1}{2} \cdot \ln \frac{1+(-0.32)}{1-(-0.32)} = -0.332$$

- Einsetzen liefert:

$$z = \frac{Z_r - Z_{\rho_0}}{\sqrt{\frac{1}{n-3}}} = \frac{-0.569 - (-0.332)}{\sqrt{\frac{1}{16-3}}} = -0.86$$

# Signifikanztests für Produkt-Moment Korrelation

- Der Standardnormalverteilung entnehmen wir für  $\alpha = 0.05$  sowie zweiseitiger Testung den Wert  $z_{crit} = z_{1-\alpha/2} = z_{0.975} = 1.96$ .

Da  $|z| = 0.86 < z_{crit}$  ist, behalten wir die Nullhypothese bei und schließen, dass der Zusammenhang zwischen der Berufsschulnote und der mit dem neuen Test erfassten Intelligenz nicht statistisch signifikant höher ausfällt als mit den existierenden Intelligenztests.

# Signifikanztests für Produkt-Moment Korrelation

- **Beispiel:** wie eben. Wie lautet das 95%-Konfidenzintervall für die Korrelation  $r = -0.515$ ?

Wir setzen wieder die Fisher-Z-transformierte Korrelation ein

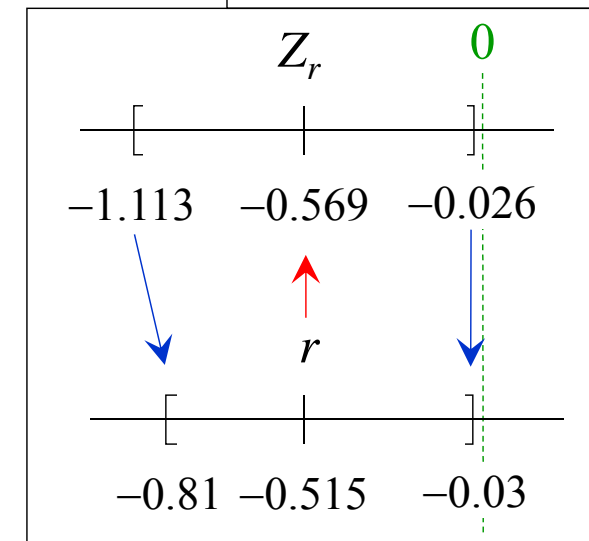
$$Z_r \pm z_{1-\alpha/2} \cdot \sqrt{\frac{1}{n-3}} = -0.569 \pm 1.96 \cdot \sqrt{\frac{1}{16-3}} \Rightarrow [-1.113, -0.026]$$

$$Z_r = \frac{1}{2} \cdot \ln \frac{1+r}{1-r}$$

und rechnen dann die beiden Grenzen in die  $r$ -Metrik zurück:

$$\frac{e^{2 \cdot Z_r} - 1}{e^{2 \cdot Z_r} + 1} = \frac{e^{2 \cdot (-1.113)} - 1}{e^{2 \cdot (-1.113)} + 1} = -0.81 \quad (\text{obere Grenze})$$

$$\frac{e^{2 \cdot Z_r} - 1}{e^{2 \cdot Z_r} + 1} = \frac{e^{2 \cdot (-0.026)} - 1}{e^{2 \cdot (-0.026)} + 1} = -0.03 \quad (\text{untere Grenze})$$



- Das Konfidenzintervall (KI) lautet also  $[-0.81, -0.03]$ . Wir können auch hier wieder an dem KI erkennen, dass der Signifikanztest auf Unterschiedlichkeit von 0 statistisch signifikant ausgefallen ist, da das KI den Wert 0 nicht einschließt. Das KI ist also (wie auch bei den anderen Statistiken) informativer als die Signifikanzprüfung.

$$r = \frac{e^{2 \cdot Z_r} - 1}{e^{2 \cdot Z_r} + 1}$$

- 1 Bivariate Korrelationen
  - 1.1 Produkt-Moment Korrelation  $r$
  - 1.2 Rangkorrelationen
- 2 Bivariate lineare Regression

# Wiederholung Rangkorrelationen

- Weisen zwei Variablen  $X$  und  $Y$  (mindestens) Ordinalskalenniveau auf, so können zur Quantifizierung des monotonen Zusammenhangs Rangkorrelationen eingesetzt werden.
- Als Rangkorrelationen kommen dazu in Frage
  - Spearmans Rangkorrelation  $r_s$ , die der Produkt-Moment Korrelation der jeweils in eine Rangordnung gebrachten Variablen  $X$  und  $Y$  entspricht:  $r_s = r[\text{Rg}(X), \text{Rg}(Y)]$ .
  - Korrelationen, die auf dem Vergleich aller Paare von Messwerten basieren, und die sich in der Behandlung von Ties in den Daten unterscheiden, z.B.  $\tau_a$ ,  $\tau_b$  oder  $\gamma$ .
- Im Folgenden werden wir nur die inferenzstatistischen Prüfungen für die paarbasierten Koeffizienten behandeln. Wir beginnen zunächst mit dem Koeffizienten  $\tau_a$ , bei dem die Differenz aus den konkordanten und diskordanten Paaren  $C - D$  an der Gesamtzahl aller möglichen Paare normiert wird:

$$\tau_a = \frac{C - D}{n \cdot (n - 1) / 2}$$

- $\tau_a$  ist geeignet, wenn keine Ties in beiden Variablen vorliegen; nur in diesem Fall kann der Koeffizient prinzipiell die Werte  $\pm 1$  annehmen.

# Beispiel Rangkorrelation

- **Beispiel** (aus Bortz, Lienert & Boehnke, 2008, S. 425): Es soll überprüft werden, ob zwei Beurteiler (Variablen  $X$  und  $Y$ ) die physische Kondition von 7 Ratten nach dreiwöchiger Mangeldiät übereinstimmend beurteilen. Beide Beurteiler haben die 7 Ratten bzgl. der Kondition in eine Rangreihe gebracht.

- Aufgrund der Erhebungsprozedur können keine Ties auftreten. Wir zählen  $C$  und  $D$  aus und setzen dann ein:

$$\tau_a = \frac{C - D}{n \cdot (n - 1) / 2} = \frac{18 - 3}{7 \cdot (7 - 1) / 2} = 0.714$$

- Fragestellung: Besteht ein statistisch signifikanter Zusammenhang zwischen den Einschätzungen beider Urteiler? Oder anders formuliert: Ist  $\tau_a$  statistisch signifikant von 0 verschieden? Die Fragestellung soll zweiseitig mit  $\alpha = 0.05$  untersucht werden.

Ratte	$X$	$Y$
A	4	4
B	1	2
C	3	1
D	5	6
E	6	5
F	2	3
G	7	7

Z.B. ist das Paar  $(A, B)$  konkordant, weil  $x_1 > x_2$  und  $y_1 > y_2$  und das Paar  $(B, C)$  diskordant, weil  $x_2 < x_3$  und  $y_2 > y_3$ .

# Signifikanztests für Rangkorrelationen

- Bezeichnung: Signifikanztest für Kendalls  $\tau = \tau_a$ .
- Einsatzbereich: Prüfung der Hypothese, dass Kendalls  $\tau_a$  zwischen zwei mindestens ordinalskalierten Variablen  $X$  und  $Y$  ungleich 0 ist.
- Hypothesen: allgemein zweiseitig:  $H_0: T = 0$  und  $H_1: T \neq 0$  oder entsprechend einseitig (mit  $T$  als Korrelation in der Population)
- Voraussetzungen: Die beiden mindestens ordinalskalierten Variablen  $X$  und  $Y$  weisen keine Ties auf.
- Vorgehen: Kendalls  $\tau_a$  und auch die anderen auf paarweisen Vergleichen basierenden Rangkorrelationen (wie  $\tau_b$ ,  $\gamma$  etc.) basieren alle auf der Differenz der konkordanten und der diskordanten Paare  $C - D$ . Diese werden dann jeweils unterschiedlich normiert, z.B. bei  $\tau_a$  an der Gesamtzahl an Paaren:

$$\tau_a = \frac{C - D}{n \cdot (n - 1) / 2}$$



# Signifikanztests für Rangkorrelationen

- Vorgehen (Fs.): Bei hinreichend großer Stichprobe ( $n \geq 10$ ) verteilt sich die Differenz  $C - D$  unter der  $H_0$  der Unabhängigkeit beider Variablen normal (Kendall & Gibbons, 1990). Die allgemeine Formel für den Standardfehler  $\sigma_{C-D}$  (siehe unten) vereinfacht sich für den bei  $\tau_a$  angenommenen Fall fehlender Ties zu

$$\sigma_{C-D} = \sqrt{\frac{n \cdot (n-1) \cdot (2 \cdot n + 5)}{18}} \quad \text{falls keine Ties existieren}$$

Dann erlaubt die folgende Prüfgröße eine Testung der Nullhypothese, dass der Zusammenhang 0 ist:

$$z = \frac{C - D}{\sqrt{\frac{n \cdot (n-1) \cdot (2 \cdot n + 5)}{18}}}$$

(Manchmal wird noch eine Kontinuitätskorrektur vorgenommen, d.h. stattdessen im Zähler  $|C - D| - 1$  eingesetzt.)

- Entscheidung: Zurückweisung der  $H_0$ , falls  $|z| > z_{crit}$  mit  $z_{crit} = z_{1-\alpha/2}$  bei zweiseitiger Prüfung und  $z_{crit} = z_{1-\alpha}$  bei einseitiger Prüfung (und korrekter Richtung).

Alternativ erhält man durch Umformung folgende Prüfgröße:

$$z = \frac{\tau_a}{\sqrt{\frac{2 \cdot (2 \cdot n + 5)}{9 \cdot n \cdot (n-1)}}}$$

# Signifikanztests für Rangkorrelationen

- Beispiel: Für  $\tau_a$  resultiert

$$\tau_a = \frac{C - D}{n \cdot (n - 1) / 2} = \frac{18 - 3}{7 \cdot (7 - 1) / 2} = 0.714$$

- Einsetzen in eine der beiden Formeln für die Prüfgröße liefert:

$$z = \frac{C - D}{\sqrt{\frac{n \cdot (n - 1) \cdot (2 \cdot n + 5)}{18}}} = \frac{18 - 3}{\sqrt{\frac{7 \cdot (7 - 1) \cdot (2 \cdot 7 + 5)}{18}}} = 2.25$$

$$z = \frac{\tau_a}{\sqrt{\frac{2 \cdot (2 \cdot n + 5)}{9 \cdot n \cdot (n - 1)}}} = \frac{0.714}{\sqrt{\frac{2 \cdot (2 \cdot 7 + 5)}{9 \cdot 7 \cdot (7 - 1)}}} = 2.25$$

Ratte	X	Y
A	4	4
B	1	2
C	3	1
D	5	6
E	6	5
F	2	3
G	7	7

C = 18	D = 3
n = 7	

- Bei zweiseitiger Fragestellung und  $\alpha = 0.05$  ist  $z_{crit} = 1.96 < |z| = 2.25$ . Wir weisen also die  $H_0$  zurück und schließen, dass ein statistisch signifikanter Zusammenhang besteht.

# Signifikanztests für Rangkorrelationen

- Bei kleineren Stichproben lässt sich auch ein **exakter Test** durchführen, indem man für alle  $n!$  möglichen Permutationen der Daten jeweils den  $\tau_a$ -Koeffizienten bestimmt und prüft, ob der empirisch gegebene  $\tau$ -Wert zu den extremsten  $100 \cdot \alpha / 2$  % auf beiden Seiten der so erzeugten „Stichprobenkennwerteverteilung“ gehört.
- Im Beispiel ergibt sich also eine Verteilung von  $7! = 5040$  möglichen  $\tau$ -Werten (gegeben die Randverteilungen in  $X$  und  $Y$ ). Bei  $\alpha = 0.05$  und zweiseitiger Fragestellung ist der kritische  $\tau$ -Wert für positives  $\tau$  bei aufsteigend sortierten  $\tau$ -Werten der 4914. Wert ( $=5040 \cdot 0.975$ ); hier  $\tau_{crit} = \tau_{(4914)} = 0.619$ . Da  $\tau_{crit} < \tau$ , ergibt sich auch hier ein statistisch sign. Zusammenhang.

Ratte	X	Y
C	3	1
B	1	2
F	2	3
A	4	4
E	6	5
D	5	6
G	7	7

Daten  
 $\tau = 0.714$

X	Y
3	1
1	2
2	3
4	4
6	5
5	6
7	7

Perm. 1  
 $\tau = 0.714$

X	Y
3	2
1	1
2	3
4	4
6	5
5	6
7	7

Perm. 2  
 $\tau = 0.810$

...

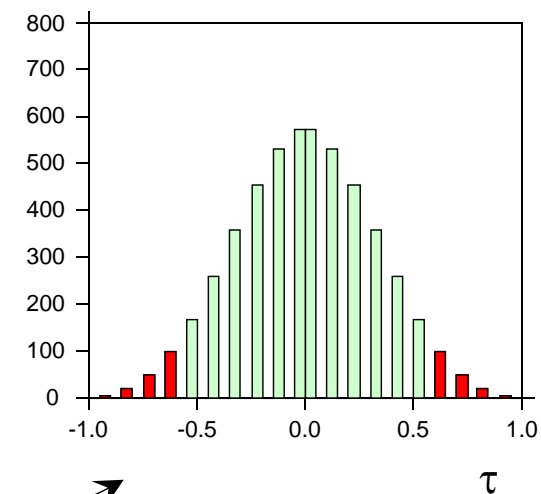
X	Y
3	5
1	3
2	2
4	4
6	1
5	6
7	7

Perm. 199  
 $\tau = 0.333$

...

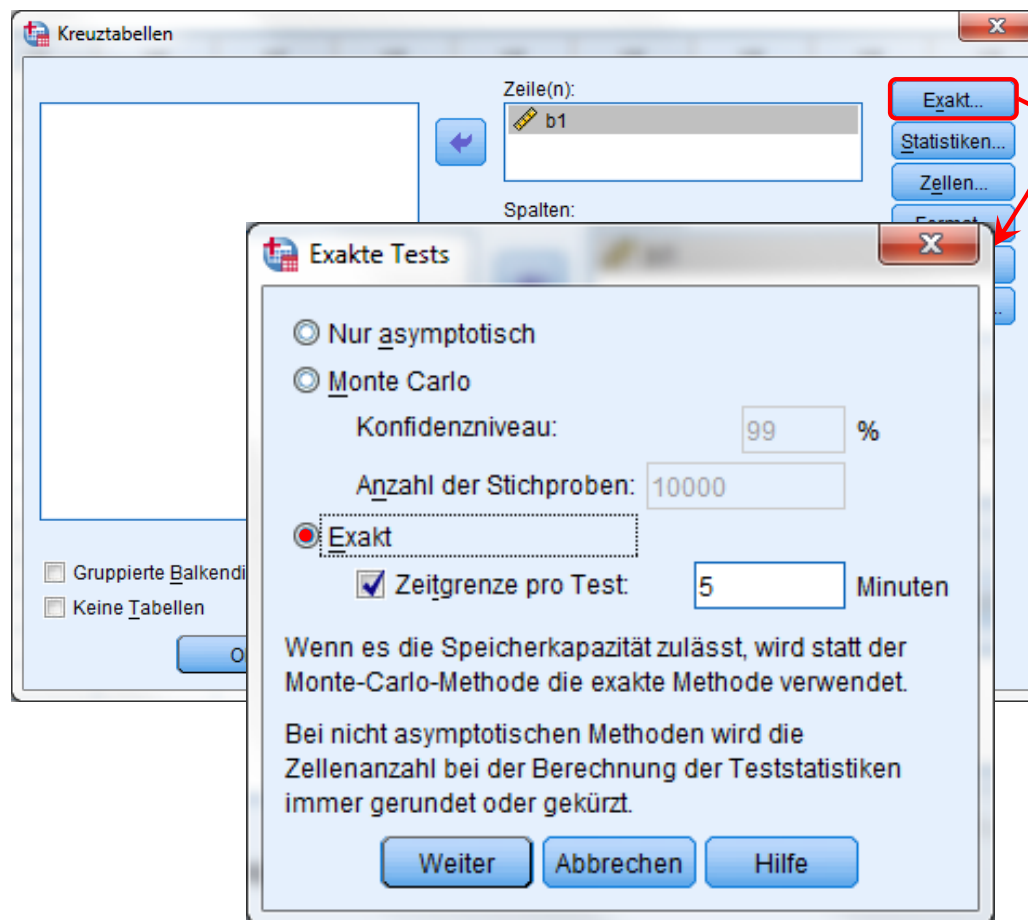
X	Y
3	7
1	6
2	5
4	4
6	3
5	2
7	1

Perm. 5040  
 $\tau = -0.714$



# Paarbasierte Rangkorrelationen in SPSS

- In SPSS lassen sich unter Analyse/Korrelation/Bivariate... der Koeffizient  $\tau_b$  und unter Analyse/Deskriptive\_Statistiken/Kreuztabellen... (Prozedur Crosstabs) die Koeffizienten  $\tau_b$  und  $\gamma$  anfordern. (Nur) Wenn keine Ties in den Daten auftreten, gilt  $\tau_a = \tau_b = \gamma$  und man kann irgendeinen der Koeffizienten zur Bestimmung von  $\tau_a$  wählen.



- Das SPSS-Zusatzmodul „Exakte Tests“ erlaubt in der Prozedur Crosstabs die Berechnung exakter Tests.

Zu beachten ist aber, dass bei größer werdendem Stichprobenumfang  $n!$  schnell sehr groß werden kann (z.B.  $20! = 2.432.902.008.176.640.000$ ) und die exakten Berechnung zu rechenzeit- bzw. speicherintensiv werden.

Unter (Exakt...) ist dann die Option „Exakt“ zu aktivieren und ggf. die Zeitgrenze hochzusetzen.

# Paarbasierte Rangkorrelationen in SPSS

- Für die unter Analyse/Deskriptive/Statistiken/Kreuztabellen... angeforderten Koeffizienten resultieren bei fehlenden Ties wie in den Daten entsprechend gleiche Werte:

**Symmetrische Maße**

		Wert	Asymptotischer Standardfehler <sup>a</sup>	Näherungsweise t <sup>b</sup>	Näherungsweise Signifikanz	Exakte Signifikanz
Ordinal- bzgl. Ordinalmaß	Kendall-Tau-b	,714	,161	4,437	,000	,030
	Gamma	,714	,161	4,437	,000	,030
Anzahl der gültigen Fälle		7				

a. Die Null-Hyphothese wird nicht angenommen.

b. Unter Annahme der Null-Hyphothese wird der asymptotische Standardfehler verwendet.

$$\tau_a = \tau_b = \gamma = 0.714$$

Hier wird der zweiseitige  $p$ -Werte des exakten Tests ausgegeben, der statistisch signifikante Korrelationen anzeigt,  $p < .05$  (und für alle paarbasierten Rangkorrelationen gleich ist).

# Signifikanztests für Rangkorrelationen

- Die Korrelation  $\tau_a$  ist ein Spezialfall einer Familie von paarbasierten Rangkorrelationen, die sich darin unterscheiden, wie sie auf Ties in  $X$  und  $Y$  reagieren. Wenn Ties in einer Variable durch die Prozedur der Datenerhebung erzwungen werden (z.B. bei einer 3-stufigen Antwortskala), sollte man einen Koeffizienten wählen, bei dem sich dies in der betreffenden Variable nicht mindernd auswirkt.
- Bei allen unten aufgeführten Koeffizienten bildet die Differenz  $C - D$  den Zähler; der Nenner unterscheidet sich wie in der Tabelle aufgeführt.

Koeffizient	Nenner	Minderung durch ...		
		Ties in $X$	Ties in $Y$	Ties in $X$ und $Y$
Kendall $\tau = \tau_a$	$C + D + T_X + T_Y + T_{XY} = n \cdot (n - 1) / 2$	ja	ja	ja
Kendall $\tau_b$	$\sqrt{(C + D + T_X) \cdot (C + D + T_Y)}$	ja	ja	nein
Wilson $e$	$C + D + T_X + T_Y$	ja	ja	nein
Somers $d_{XY}$	$C + D + T_X$	ja	nein	nein
Somers $d_{YX}$	$C + D + T_Y$	nein	ja	nein
Goodman & Kruskals $\gamma$	$C + D$	nein	nein	nein

$T_X$  = Zahl an Ties nur in  $X$ ;  $T_Y$  = Zahl an Ties nur in  $Y$ ;  $T_{XY}$  = Zahl an Ties gleichzeitig in  $X$  und  $Y$

# Signifikanztests für Rangkorrelationen

- **Beispiel:** In einem Wettkampf bewerten zwei Punktrichter ( $X$  und  $Y$ ) die Ausführung der Übung (B-Note) von  $n = 8$  Bodenturnern auf einer Skala von 1 bis 10 (=Bestnote).
- Es soll geprüft werden, ob ein statistisch signifikanter Zusammenhang zwischen den beiden ordinalskalierten Variablen besteht. In diesem Fall kommt  $\tau_a$  nicht infrage, da Ties in beiden Variablen existieren (und bei mehr als 10 Turnern zwangsläufig durch die Prozedur „erzwungen“ werden).
- Ein sinnvoller Koeffizient wäre also  $\gamma$ .
- Die oben für  $\tau_a$  als Spezialfall dargestellte Testprozedur ist für alle auf der Differenz  $C - D$  im Zähler basierenden Koeffizienten gleich. Im Gegensatz zu dem Fall, bei dem keine Ties vorliegen, ist der verallgemeinerte Standardfehler (bzw. dessen Quadrat, der Varianzfehler), dessen Formel beim Vorliegen von Ties verwenden werden muss, anhand einer umfangreicheren Formel zu berechnen, die auf der folgenden Seite dargestellt wird.

Turner	X	Y
A	6	4
B	7	8
C	6	7
D	9	6
E	7	8
F	6	5
G	8	9
H	5	3

# Signifikanztests für Rangkorrelationen

- Der **Varianzfehler** für die Differenz  $C - D$  ist:

$$\sigma_{C-D}^2 = \frac{n \cdot (n-1) \cdot (2 \cdot n + 5) - \sum_i T_i \cdot (T_i - 1) \cdot (2 \cdot T_i + 5) - \sum_j T_j \cdot (T_j - 1) \cdot (2 \cdot T_j + 5)}{18} + \frac{\left[ \sum_i T_i \cdot (T_i - 1) \cdot (T_i - 2) \right] \cdot \left[ \sum_j T_j \cdot (T_j - 1) \cdot (T_j - 2) \right]}{9 \cdot n \cdot (n-1) \cdot (n-2)} + \frac{\left[ \sum_i T_i \cdot (T_i - 1) \right] \cdot \left[ \sum_j T_j \cdot (T_j - 1) \right]}{2 \cdot n \cdot (n-1)}$$

X	Y
6	4
7	8
6	7
9	6
7	8
6	5
8	9
5	3

$C = 20$	$D = 4$	$n = 8$
----------	---------	---------

- Dabei bezeichnet  $T_i$  die Zahl aller Messwerte in  $X$ , die sich im Tie-Block  $i$  befinden (und  $T_j$  analog die Zahl aller Messwerte in  $Y$ , die sich im Tie-Block  $j$  befinden).
- Im Beispiel ergeben sich zwei Tie-Blocks in  $X$  (mit 3 bzw. 2 Messwerten) und ein Tie-Block in  $Y$  (mit 2 Messwerten).

Tie-Block in $X$	Wert	$T_i$
$i = 1$	$x = 6$	3
$i = 2$	$x = 7$	2

Tie-Block in $Y$	Wert	$T_j$
$j = 1$	$y = 8$	2



# Signifikanztests für Rangkorrelationen

➤ Für die Summen der Tie-Blöcke ergibt sich:

$$\sum_i T_i \cdot (T_i - 1) \cdot (2 \cdot T_i + 5) = 3 \cdot (3 - 1) \cdot (2 \cdot 3 + 5) + 2 \cdot (2 - 1) \cdot (2 \cdot 2 + 5) = 66 + 18 = 84$$

$$\sum_j T_j \cdot (T_j - 1) \cdot (2 \cdot T_j + 5) = 2 \cdot (2 - 1) \cdot (2 \cdot 2 + 5) = 18$$

$$\sum_i T_i \cdot (T_i - 1) \cdot (T_i - 2) = 3 \cdot (3 - 1) \cdot (3 - 2) + 2 \cdot (2 - 1) \cdot (2 - 2) = 6 + 0 = 6$$

$$\sum_j T_j \cdot (T_j - 1) \cdot (T_j - 2) = 2 \cdot (2 - 1) \cdot (2 - 2) = 0$$

$$\sum_i T_i \cdot (T_i - 1) = 3 \cdot (3 - 1) + 2 \cdot (2 - 1) = 6 + 2 = 8$$

$$\sum_j T_j \cdot (T_j - 1) = 2 \cdot (2 - 1) = 2$$

Tie-Block in X	Wert	$T_i$
$i = 1$	$x = 6$	3
$i = 2$	$x = 7$	2
Tie-Block in Y	Wert	$T_j$
$j = 1$	$y = 8$	2

und dann bei Einsetzen in die Formel für den Standardfehler

$$\sigma_{C-D}^2 = \frac{8 \cdot (8 - 1) \cdot (2 \cdot 8 + 5) - 84 - 18}{18} + \frac{6 \cdot 0}{9 \cdot 8 \cdot (8 - 1) \cdot (8 - 2)} + \frac{8 \cdot 2}{2 \cdot 8 \cdot (8 - 1)} = 59.810$$

# Signifikanztests für Rangkorrelationen

- Insgesamt ergibt sich damit als verallgemeinerte Prüfgröße

$$z = \frac{C - D}{\sqrt{\sigma_{C-D}^2}} = \frac{20 - 4}{\sqrt{59.810}} = 2.07$$

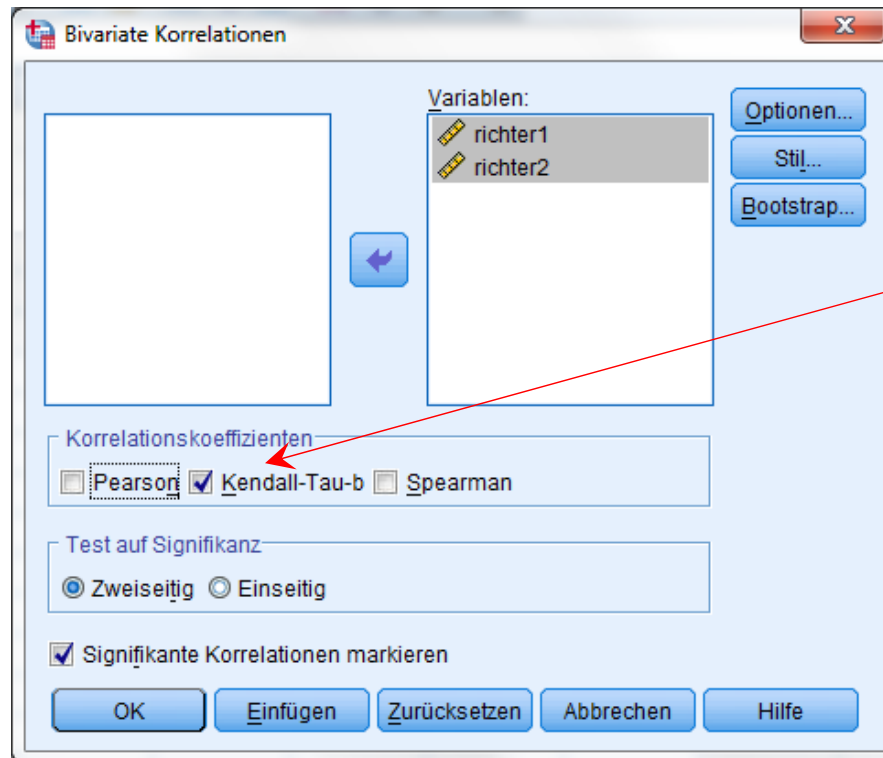
- Bei zweiseitiger Fragestellung und  $\alpha = 0.05$  ist  $z_{crit} = z_{1-\alpha/2} = 1.96 < |z| = 2.07$ . Wir weisen also die  $H_0$  zurück und schließen, dass ein statistisch signifikanter Zusammenhang besteht.

Oder anders ausgedrückt:  $\gamma$  und alle übrigen in der vorangegangenen Tabelle gelisteten paarweisen Koeffizienten sind statistisch signifikant von 0 verschieden.

- **Sonderfall:** Soll nicht die Verschiedenheit der Korrelation von 0 getestet werden (sondern von einem anderen Wert) oder ein Konfidenzintervall bestimmt werden, so ergeben sich nochmals komplexere Formeln (vgl. Woods, 2007).

# Paarbasierte Rangkorrelationen in SPSS

- Unter Analysieren/Korrelation/Korrelation/Bivariate... kann man sich in SPSS  $\tau_b$  ausgeben lassen:



Dazu sind die beiden Variablen anzugeben und unter „Korrelationskoeffizienten“ die Option „Kendall-Tau-b“ zu aktivieren.

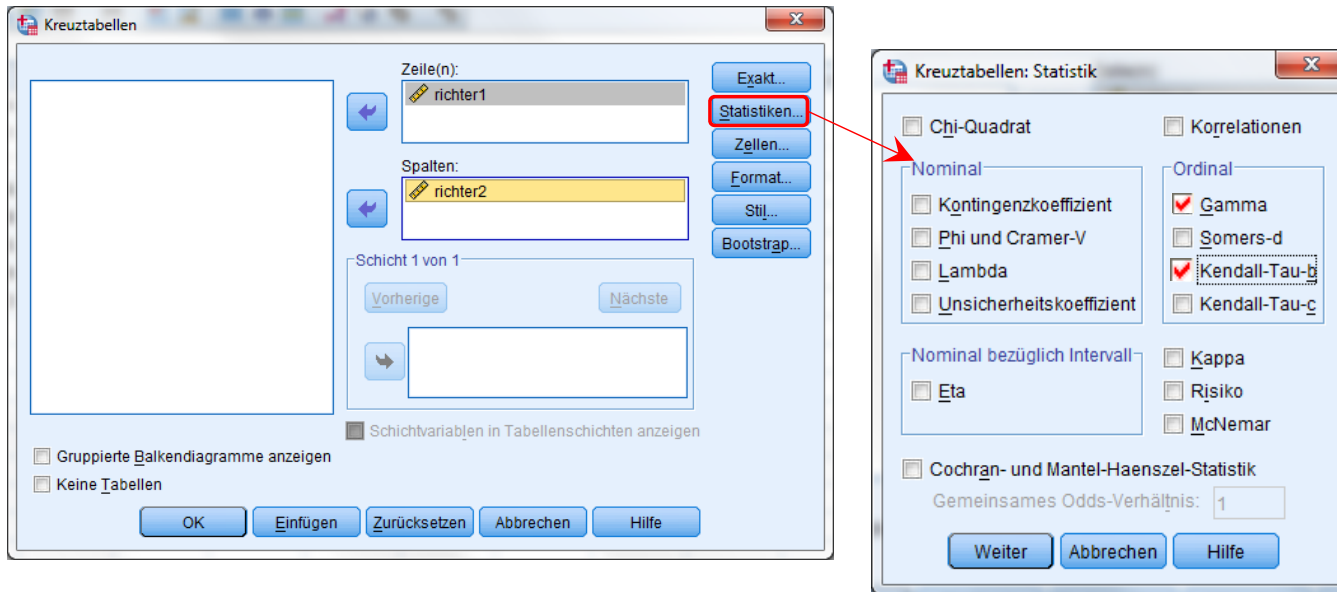
Dabei wird der oben dargestellte Signifikanztest durchgeführt:  $\tau_b = 0.63$ ,  $p < .05$ . Er gilt auch für alle anderen paarweisen Rangkorrelationen.

		richter1	richter2
Kendall-Tau-b	richter1	Korrelationskoeffizient	1,000
		Sig. (2-seitig)	,039
		N	8
richter2		Korrelationskoeffizient	,629*
		Sig. (2-seitig)	,039
		N	8

\*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

# Paarbasierte Rangkorrelationen in SPSS

- $\tau_b$  bzw.  $\gamma$  kann man unter Analysieren/Deskriptive Statistiken/Kreuztabellen... anfordern:



Unter (Statistiken) sind die Koeffizienten und ggf. unter (Exakt) der exakte Test anzufordern.

Hier wird der zweiseitige  $p$ -Wert des exakten Tests ausgegeben, der signifikante Korrelationen anzeigt.

Es resultiert  $\gamma = 0.667$  und  $\tau_b = 0.629$ . (Die Testung in der Mitte der Tabelle erfolgt nach einem anderen, hier nicht behandelten, Verfahren.)

		Symmetrische Maße				
		Wert	Asymptotischer Standardfehler <sup>a</sup>	Näherungsweise $t^b$	Näherungsweise Signifikanz	Exakte Signifikanz
Ordinal- bzgl. Ordinalmaß	Kendall-Tau-b	,629	,249	2,596	,009	,037
	Gamma	,667	,274	2,596	,009	,037
Anzahl der gültigen Fälle		8				

a. Die Null-Hypothese wird nicht angenommen.

b. Unter Annahme der Null-Hypothese wird der asymptotische Standardfehler verwendet.

## 1 Bivariate Korrelationen

1.1 Produkt-Moment Korrelation  $r$

1.2 Rangkorrelationen

## 2 Bivariate Regression

# Signifikanztests für bivariate lineare Regression

- Im Kontext der bivariaten linearen Regression  $y_i = a + b \cdot x_i + e_i$  lassen sich ebenfalls verschiedene inferenzstatistische Fragen beantworten.
- So kann man prüfen, ob die Regressionskoeffizienten  $a$  und  $b$  statistisch signifikant von 0 verschieden sind.  $a$  ist aber in der Regel nicht von Interesse.  $b$  ist genau dann statistisch signifikant von 0 verschieden, wenn  $r$  es ist; daher werden wir auf diesen Test im Folgenden auch nicht weiter eingehen (vgl. z.B. Eid, Gollwitzer & Schmidt, 2010, S. 578ff).
- Als einzige Fragestellung wollen wir im Folgenden das Konfidenzintervall um einen vorhergesagten Wert betrachten.
- **Voraussetzungen:** Um inferenzstatistische Aussagen treffen zu können, muss beim Vorliegen zweier Zufallsvariablen  $X$  und  $Y$  wieder von der bivariaten Normalverteilung beider Variablen ausgegangen werden.

Ist die Variable  $X$  hingegen keine (stochastische) Zufallsvariable, sondern eine durch die Versuchsplanung festgelegte (deterministische) Variable mit festen Werten, so muss stattdessen gelten, dass die Arrayverteilungen von  $X$  varianzhomogen und normalverteilt um die  $\hat{Y}$ -Werte sind. **Beispiel:** In einem Experiment wird untersucht, wie sich bestimmte verabreichte Mengen von Alkohol ( $=X$ ) auf die Fahrleistung auswirken ( $=Y$ ).

# Signifikanztests für bivariate lineare Regression

- Die beste Prognose für die Ausprägung in  $Y$  auf Basis der Kenntnis der Ausprägung  $x_p$  einer Person  $p$  in der Variablen  $X$  erhalten wir im Modell der linearen Regression durch Einsetzen in die Regressionsgleichung  $\hat{y}_p = a + b \cdot x_p$ . Um diesen Punktschätzer lässt sich unter Rückgriff auf den Standardschätzfehler (vgl. Statistik I)

$$s_E = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}} = s_Y \cdot \sqrt{(1-r^2) \cdot \frac{n-1}{n-2}}$$

folgendes Konfidenzintervall – in diesem Kontext meist als **Vorhersageintervall** (prediction interval) bezeichnet – für gegebenes  $\alpha$  bestimmen:

$$\hat{y}_p \pm t_{df; 1-\alpha/2} \cdot s_E \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{(n-1) \cdot s_X^2}} \quad \text{mit} \quad df = n - 2$$

- Man erkennt (an dem grün hinterlegten Zähler), dass das Konfidenzintervall immer breiter wird, je stärker  $x_p$  vom Mittelwert abweicht.

# Signifikanztests für bivariate lineare Regression

- **Beispiel:** Der bereits mehrfach verwendete (hier auf  $n = 59$  erweiterte) Datensatz soll zur Vorhersage des Ausbildungserfolgs (operationalisiert als Berufsschulnote  $Y$ ) aufgrund der Intelligenz ( $X$ ) der sich bewerbenden Auszubildenden zum Industriemechaniker verwendet werden. Es soll die vorhergesagte Berufsschulnote mit dem Vorhersageintervall für eine neue Person mit einem Intelligenzwert von **105** bestimmt werden.

- Dazu muss zunächst die Regressionsgleichung bestimmt werden. Sie lautet in diesem Fall (vgl. Statistik I):

$$\hat{y}_i = a + b \cdot x_i = 11.1471 - 0.0753 \cdot x_i$$

- Einsetzen ergibt für die Person  $p$  die vorhergesagte Note

$$\hat{y}_p = a + b \cdot x_p = 11.1471 - 0.0753 \cdot 105 = 3.24$$

- Um das Vorhersageintervall zu bestimmen benötigen wir den Standardschätzfehler:

$$s_E = s_Y \cdot \sqrt{(1 - r^2) \cdot \frac{n-1}{n-2}} = 0.789 \cdot \sqrt{[1 - (-0.589)^2] \cdot \frac{59-1}{59-2}} = 0.644$$

Statistiken der Daten	
$n = 59$	$r = -0.589$
$\bar{x} = 103.58$	$\bar{y} = 3.35$
$s_X = 6.171$	$s_Y = 0.789$
$s_X^2 = 38.076$	$s_Y^2 = 0.623$



# Signifikanztests für bivariate lineare Regression

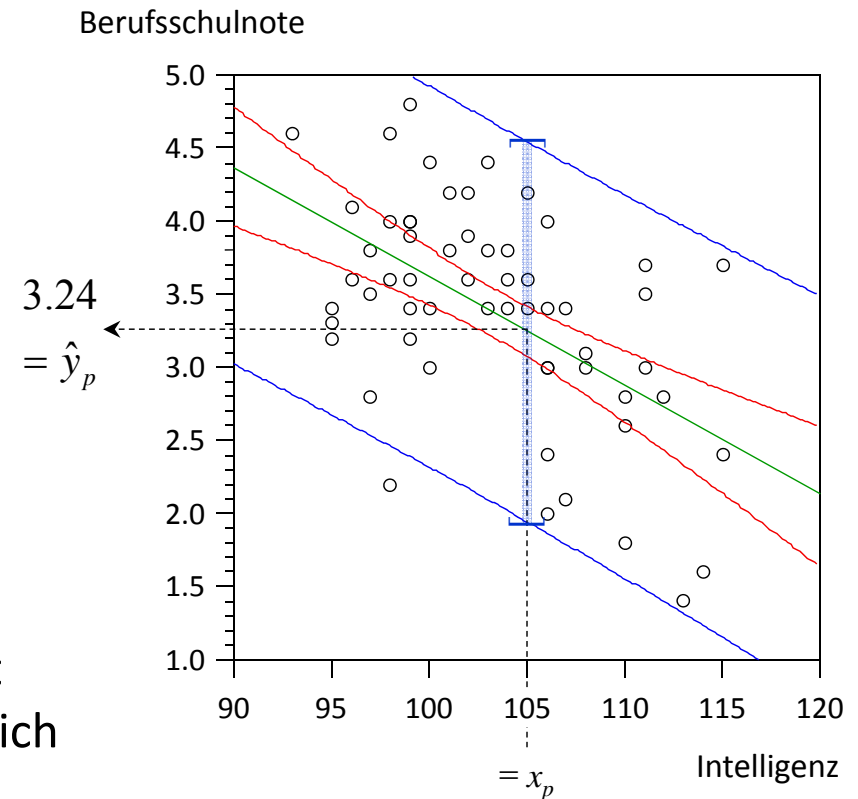
- Einsetzen ergibt dann (für  $\alpha = 0.05$  und  $t_{59-2;1-\alpha/2} = 2.00$ )

$$\hat{y}_p \pm t_{n-2;1-\alpha/2} \cdot s_E \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{(n-1) \cdot s_X^2}}$$

$$3.24 \pm 2.00 \cdot 0.644 \cdot \sqrt{1 + \frac{1}{59} + \frac{(105 - 103.58)^2}{(59-1) \cdot 38.076}}$$

$$\Rightarrow [1.94, 4.54]$$






- Mit einer 95%igen Sicherheit liegt der Vorhersagewert der Berufsschulnote für einen IQ von 105 also im Bereich von 1.94 und 4.54; das Intervall ist also sehr weit.



- **Hinweis:** Das Vorhersageintervall für einen vorhergesagten Wert einer Person ist nicht mit dem schmaleren – oben rechts rot eingezeichneten – Konfidenzintervall um den Mittelwert (Erwartungswert) in  $Y$ , gegeben ein bestimmtes  $x_p$ , zu verwechseln, der sich ganz ähnlich berechnet nach:

$$\hat{y}_p \pm t_{n-2;1-\alpha/2} \cdot s_E \cdot \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{(n-1) \cdot s_X^2}} \Rightarrow [3.07, 3.41]$$

## Zitierte Quellen:

-  DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, 2, 292-307.
-  Farrell, P. J., Salibian-Barrera, M., & Naczk, K. (2007). On tests for multivariate normality and associated simulation studies. *Journal of Statistical Computation & Simulation*, 77, 1065-1080.
-  Hülshager, U. R., Maier, G. W., Stumpp, Th. & Muck, P. M. (2006). Vergleich kriteriumsbezogener Validitäten verschiedener Intelligenztests zur Vorhersage von Ausbildungserfolg in Deutschland. Ergebnisse einer Metaanalyse. *Zeitschrift für Personalpsychologie*, 5, 145-162.
-  Schulze, R. (2004). *Meta-analysis: a comparison of approaches*. Cambridge, MA: Hogrefe & Huber.
-  Woods, C. M. (2007). Confidence intervals for Gamma-family measures of ordinal association. *Psychological Methods*, 12, 185-204.

## Änderungen am 12.04.2017

- Folie 36: SPSS-Tabelle unten rechts korrigiert (falsche Referenz auf exakten Test)
- Folie 41: unten rechts Intervall ergänzt