(iii) Wenn für zwei Konstanten $a$ und $b$ mit $b < 0$ gilt:

$$y_i = a + bx_i, \quad i = 1, \ldots, n,$$

dann ist $r = -1$.

*Beweis:* Übungsaufgabe 4.10

Interpretation von $r$:

- Ein Wert von $|r| = 1$ bedeutet, dass es einen vollkommen linearen Zusammenhang gibt.
- Ein Wert von $|r| \approx 0.8$ deutet auf einen relativ starken linearen Zusammenhang hin.
- Ein Wert von $|r|$ um 0.3 besagt, dass der lineare Zusammenhang relativ schwach ist.

In unserem Beispiel erhalten wir:

$$\begin{aligned}
\bar{x}_{math} &= \frac{20}{8} = \frac{5}{2} = 2.5 \\
s_{math} &= \sqrt{\frac{1}{7}\sum_{i=1}^{8}(x_i - \bar{x})^2} = \sqrt{\frac{1}{7}(4 \cdot 1.5^2 + 3 \cdot 0.5^2 + 2.5^2)} = \sqrt{\frac{16}{7}} = \frac{4\sqrt{7}}{7} \approx 1.51 \\
\bar{x}_{phy} &= \frac{21}{8} = 2.625 \\
s_{phy} &= \sqrt{\frac{1}{7}(3 \cdot 1.625^2 + 0.625^2 + 2 \cdot 0.375^2 + 2 \cdot 2.375^2)} \approx 1.69
\end{aligned}$$

$$\begin{aligned}
\Rightarrow r = & \\
= & \ \frac{2(-1.5)(-1.625) + (-0.5)(-1.625) + (-1.5)(-0.625) + 0.5 \cdot 0.375 + 1.5 \cdot 0.375 + 2.5 \cdot 2.375 + 0.5 \cdot 2.375}{7 \cdot 1.51 \cdot 1.69} \\
\approx & \ .813
\end{aligned}$$

Bei metrisch- und auch bei nur ordinal-skalierten Werten kann man „Spearmans Korrelationskoeffinzient“ berechnen, das ist der oben definierte Korrelationskoeffizient für die (Durchschnitts-)Ränge. Man setzt in obigen Formel also nicht die Werte selbst, sondern deren (Durchschnitts-)Ränge ein.

Man kann zeigen, dass Spearmans Korrelationskoeffinzient nicht die Stärke des linearen Zusammenhangs misst, sondern misst, inwieweit große bzw. kleine Werte des einen Merkmals mit großen bzw. kleinen Werten des anderen Merkmals korrespondieren (vgl. Fahrmeir et al. 2016, S. 133f). Die Eigenschaften groß und klein sind dabei relativ auf die anderen Werte des Merkmals zu verstehen.

In unserem Beispiel ergibt sich mit den Stichprobenstandardabweichungen der Ränge $s_{Rg(Mathe)} = 2.375$ und $s_{Rg(Physik)} = 2.36$:[1]

$$\begin{aligned}
r_{Sp} &= \frac{2 \cdot (2-4.5)^2 + (3-4.5) \cdot (2-4.5) + (2-4.5) \cdot (3-4.5) + (5.5-4.5)^2 + (7-4.5) \cdot (5.5-4.5) + \ldots}{7 \cdot 2.375 \cdot 2.36} \\
&\approx .815
\end{aligned}$$

**Bemerkung:** Die Korrelation misst den Zusammenhang, **NICHT** die Kausalität!

Siehe R-Beispiel I.10. ABC 8; Ü 4; Moodle 1.8-1.10

[1]Das arithmetische Mittel der Ränge für $n$ Beobachtungswerte ergibt sich unabhängig von der Werten wegen der Formel zur arithmetischen Summe zu: $\overline{rg} = \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$.