Diskrete Zufallsvariable - Varianz (einer Grundgesamtheit):

⇨ **Zusammenhang zwischen Definition und Berechnung**

Beh.: $$V(X) \overset{Def.}{=} \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \overset{Berechnung}{=} \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2$$

Bew. des Übergangs zwischen Definitions- und Berechnungsformel:

$$V(X) \overset{Def.}{=} \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i^2 - 2x_i\overline{x} + \overline{x}^2\right)$$

$$V(X) \overset{Binom}{=} \frac{1}{n}\sum_{i=1}^{n}x_i^2 + \frac{1}{n}\sum_{i=1}^{n}\left(-2x_i\overline{x}\right) + \frac{1}{n}\sum_{i=1}^{n}\overline{x}^2$$

$$V(X) \underset{\substack{\text{Summenzeichen}\\ \text{"ausklammern"}}}{\overset{\text{Konstanten aus}}{=}} \frac{1}{n}\sum_{i=1}^{n}x_i^2 - 2\overline{x}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i + \frac{1}{n}\overline{x}^2\cdot\sum_{i=1}^{n}1$$

$$V(X) \underset{\sum_{i=1}^{n}1 = n}{\overset{\frac{1}{n}\sum_{i=1}^{n}x_i = \overline{x}}{=}} \frac{1}{n}\sum_{i=1}^{n}x_i^2 - 2\overline{x}\cdot\overline{x} + \frac{1}{n}\overline{x}^2\cdot n = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - 2\overline{x}^2 + \overline{x}^2$$

$$V(X) \overset{Berechnung}{=} \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2$$

Begriff der Kovarianz / Co-Varianz (einer Grundgesamtheit)

Ausgang: Varianz $$V(X) \overset{Def.}{=} \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \overset{Berechnung}{=} \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2$$

Kovarianz:

$$Cov(X,Y) \overset{Def.}{=} \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right) \overset{Berechnung}{=} \frac{1}{n}\sum_{i=1}^{n}\left(x_i\cdot y_i\right) - \overline{x}\cdot\overline{y}$$

Herleitung:

$$Cov(X,Y) \overset{Def.}{=} \frac{1}{n}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i)-\frac{1}{n}\sum_{i=1}^{n}\left(x_i\overline{y}\right)-\frac{1}{n}\sum_{i=1}^{n}\left(\overline{x}y_i\right)+\frac{1}{n}\sum_{i=1}^{n}\left(\overline{x}\,\overline{y}\right)$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i)-\overline{y}\cdot\frac{1}{n}\cdot\sum_{i=1}^{n}x_i-\overline{x}\cdot\frac{1}{n}\cdot\sum_{i=1}^{n}y_i+\overline{x}\,\overline{y}\cdot\frac{1}{n}\cdot\sum_{i=1}^{n}1$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i)-\overline{y}\cdot\overline{x}-\overline{x}\cdot\overline{y}+\overline{x}\,\overline{y}\cdot\frac{1}{n}\cdot n\cdot 1 = \frac{1}{n}\sum_{i=1}^{n}(x_i\cdot y_i)-\overline{x}\cdot\overline{y}$$

Abgewandelte Variablen-Schreibweise (Erwartungswert anstelle von Mittelwert):

$$Cov(X,Y) \overset{Def.}{=} \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_x)(y_i-\mu_y)$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i)-\frac{1}{n}\sum_{i=1}^{n}(x_i\mu_y)-\frac{1}{n}\sum_{i=1}^{n}(\mu_x y_i)+\frac{1}{n}\sum_{i=1}^{n}(\mu_x\mu_y)$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i)-\mu_y\cdot\frac{1}{n}\cdot\sum_{i=1}^{n}x_i-\mu_x\cdot\frac{1}{n}\cdot\sum_{i=1}^{n}y_i+\mu_x\mu_y\cdot\frac{1}{n}\cdot\sum_{i=1}^{n}1$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i)-\mu_y\cdot\mu_x-\mu_x\cdot\mu_y+\mu_x\mu_y\cdot\frac{1}{n}\cdot n\cdot 1$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i\cdot y_i)-\mu_x\cdot\mu_y$$

Verschiebungssatz zur Berechnung:

$$Cov(X,Y) \overset{Def.}{=} \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_x)(y_i-\mu_y) = E\left[(X-E(X))\cdot(Y-E(Y))\right]$$

$$Cov(X,Y) = E\left[X\cdot Y-X\cdot E(Y)-E(X)\cdot Y+E(X)\cdot E(Y)\right]$$

$$Cov(X,Y) = E(X\cdot Y)-E(X)\cdot E(Y)-E(X)\cdot E(Y)+E(X)\cdot E(Y)$$

$$Cov(X,Y) = E(X\cdot Y)-E(X)\cdot E(Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i\cdot y_i)-\mu_x\cdot\mu_y$$

Kovarianz und Varianz im Vergleich:

$$Cov(X,Y) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i y_i\right)-\overline{x}\cdot\overline{y}$$

$$Cov(X,Y) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left[x_i-E(X)\right]\cdot\left[y_i-E(Y)\right] = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i y_i\right)-E(X)\cdot E(Y)$$

$$Cov(X,Y) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i-\mu_x\right)\cdot\left(y_i-\mu_y\right) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i y_i\right)-\mu_x\cdot\mu_y$$

$$V(X) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2 = \left[\frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i\right)^2\right]-\overline{x}^2$$

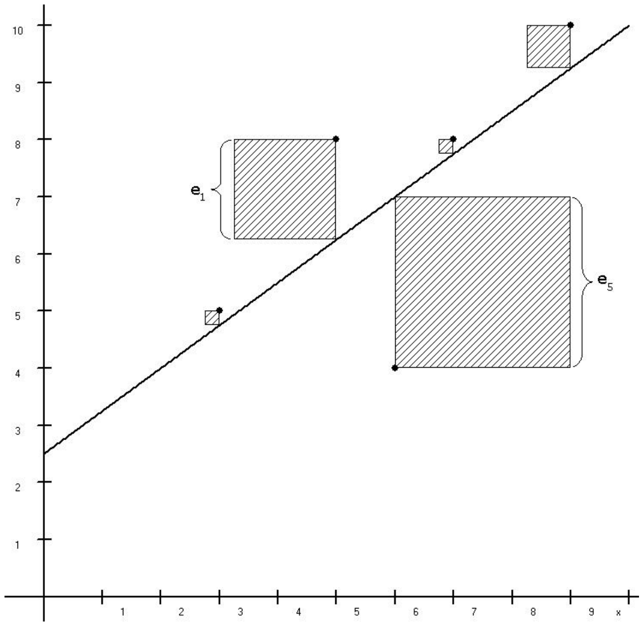
$$V(X) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left[x_i-E(X)\right]^2 = \left[\frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i\right)^2\right]-E(X)^2$$

$$V(X) = \frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i-\mu_x\right)^2 = \left[\frac{1}{n}\cdot\sum_{i=1}^{n}\left(x_i\right)^2\right]-\mu_x^{\,2}$$

Lineare Regression und Korrelation

Anwendung der partiellen Differentiation in der Statistik: Regressionsgerade

Gesucht: ***Eine Gerade, die zu allen Punkten den kleinsten Abstand hat; d.h. die Differenzen e_i müssen insgesamt möglichst klein sein (Optimierungsproblem)***



Ansatz: $$\sum_{i=1}^{n}(e_i)^2 \quad \to \quad \min.$$

$$\textit{Gesucht ist Gerade}: \quad y \;=\; b_0 + b_1 x$$

Es gilt:

$(x_i \mid y_i)$ sind die Punkte des Streudiagramms

ebenso gilt:

$(x_i \mid y_i)$ sind die Punkte auf der Geraden

daraus folgt: $\xrightarrow{y_i \leftrightarrow y_i}$ $y_i \;=\; b_0 + b_1 x_i$

und $\Delta y \;=\; y_i - y_i$ $[=\textit{Seitenlänge des Quadrats}]$

$$\to \quad \sum_{i=1}^{n}(e_i)^2 \;=\; \sum_{i=1}^{n}\left(y_i - y_i\right)^2 \;=\; \sum_{i=1}^{n}\left(y_i - b_0 - b_1 x_i\right)^2$$

$$E(b_1;b_0) \;=\; \sum_{i=1}^{n}\left(y_i^2 - y_i b_0 - y_i b_1 x_i - y_i b_0 + b_0^2 + b_0 b_1 x_i - y_i b_1 x_i + b_0 b_1 x_i + b_1^2 x_i^2\right)$$

$$E(b_1;b_0) \;=\; \sum_{i=1}^{n}\left(y_i^2 - 2y_i b_0 - 2y_i b_1 x_i + b_0^2 + 2b_0 b_1 x_i + b_1^2 x_i^2\right)$$

$$E(b_1;b_0) \;=\; \sum_{i=1}^{n} y_i^2 - 2b_0\sum_{i=1}^{n} y_i - 2b_1\sum_{i=1}^{n} x_i y_i + b_0^2\sum_{i=1}^{n} 1 + 2b_0 b_1\sum_{i=1}^{n} x_i + b_1^2\sum_{i=1}^{n} x_i^2$$

$$E(b_1;b_0) \;=\; \sum_{i=1}^{n} y_i^2 - 2b_0\sum_{i=1}^{n} y_i - 2b_1\sum_{i=1}^{n} x_i y_i + nb_0^2 + 2b_0 b_1\sum_{i=1}^{n} x_i + b_1^2\sum_{i=1}^{n} x_i^2$$

$$\frac{\partial E(b_1;b_0)}{\partial b_1} \;=\; -2\sum_{i=1}^{n} x_i y_i + 2b_0\sum_{i=1}^{n} x_i + 2b_1\sum_{i=1}^{n} x_i^2 \;=\; 0$$

$$\frac{\partial E(b_1;b_0)}{\partial b_0} \;=\; -2\sum_{i=1}^{n} y_i + 2nb_0 + 2b_1\sum_{i=1}^{n} x_i \;=\; 0$$

Lösung der Gleichungen der partiellen Ableitungen:

$$\mathrm{I)} \quad -2\sum_{i=1}^{n} x_i y_i + 2b_0 \sum_{i=1}^{n} x_i + 2b_1 \sum_{i=1}^{n} x_i^2 \quad = \quad 0$$

$$\mathrm{II)} \quad -2\sum_{i=1}^{n} y_i + 2nb_0 + 2b_1 \sum_{i=1}^{n} x_i \quad = \quad 0$$

Es gilt:

$$\sum_{i=1}^{n} x_i \quad = \quad n \cdot \overline{x} \qquad und \qquad \sum_{i=1}^{n} y_i \quad = \quad n \cdot \overline{y} \quad (\text{arithmetisches Mittel})$$

$eingesetzt:$

$$\mathrm{I)} \quad -2\sum_{i=1}^{n} x_i y_i + 2b_0 n \cdot \overline{x} + 2b_1 \sum_{i=1}^{n} x_i^2 \quad = \quad 0$$

$$\mathrm{II)} \quad -2n \cdot \overline{y} + 2nb_0 + 2b_1 n \cdot \overline{x} \quad = \quad 0 \quad \xrightarrow{:(2n)} \quad -\overline{y} + b_0 + b_1 \overline{x} \quad = \quad 0$$

$$\rightarrow \quad b_0 \quad = \quad \overline{y} - b_1 \overline{x}$$

$$\xrightarrow{\mathrm{II)} \quad eingesetzt\ in \quad \mathrm{I)}} \mathrm{I)} \quad -2\sum_{i=1}^{n} x_i y_i + 2\left(\overline{y} - b_1 \overline{x}\right) n \cdot \overline{x} + 2b_1 \sum_{i=1}^{n} x_i^2 \quad = \quad 0$$

$$\xrightarrow{:2} \mathrm{I)} \quad -\sum_{i=1}^{n} x_i y_i + n \cdot \overline{x} \cdot \overline{y} - b_1 n \cdot \left(\overline{x}\right)^2 + b_1 \sum_{i=1}^{n} x_i^2 \quad = \quad 0$$

$$\rightarrow \quad -b_1 n \cdot \left(\overline{x}\right)^2 + b_1 \sum_{i=1}^{n} x_i^2 \quad = \quad \sum_{i=1}^{n} x_i y_i - n \cdot \overline{x} \cdot \overline{y}$$

$$\rightarrow \quad b_1 \left[\sum_{i=1}^{n} x_i^2 - n \cdot \left(\overline{x}\right)^2 \right] \quad = \quad \sum_{i=1}^{n} x_i y_i - n \cdot \overline{x} \cdot \overline{y}$$

$$\rightarrow \quad b_1 \quad = \quad \frac{\sum_{i=1}^{n} x_i y_i - n \cdot \overline{x} \cdot \overline{y}}{\sum_{i=1}^{n} x_i^2 - n \cdot \left(\overline{x}\right)^2} \quad \underset{\cdot \frac{1}{n}}{\overset{\cdot \frac{1}{n}}{=}} \quad \frac{\frac{1}{n} \cdot \sum_{i=1}^{n} x_i y_i - \overline{x} \cdot \overline{y}}{\frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 - \left(\overline{x}\right)^2} \quad = \quad \frac{Cov(X,Y)}{V(X)}$$

Lineare Regression und Korrelation

Ansatz: $y = b_0 + b_1 x \qquad \mu_x = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \mu_y = \frac{1}{n}\sum_{i=1}^{n} y_i$

$$b_1 = \frac{\sigma_{xy}}{\sigma_{x^2}} = \frac{\sum_{i=1}^{n}(x_i y_i) - n \cdot \mu_x \cdot \mu_y}{\sum_{i=1}^{n}(x_i^2) - n \cdot \mu_x^2} \overset{\cdot \frac{1}{n}}{\underset{\cdot \frac{1}{n}}{=}} \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i y_i) - \mu_x \cdot \mu_y}{\frac{1}{n}\sum_{i=1}^{n}(x_i^2) - \mu_x^2}$$

$$b_1 = \frac{\sigma_{xy}}{\sigma_{x^2}} = \frac{Kov(X,Y)}{V(X)}$$

$$b_0 = \mu_y - b_1 \cdot \mu_x$$

Korrelation nach Brevais-Pearson:

$$r \overset{Definition}{=} \frac{\sum_{i=1}^{n}(x_i - \mu_x)\cdot(y_i - \mu_y)}{\sqrt{\sum_{i=1}^{n}(x_i - \mu_x)^2} \cdot \sqrt{\sum_{i=1}^{n}(y_i - \mu_y)^2}} \overset{Berechnung}{=} \frac{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_i y_i) - \mu_x \cdot \mu_y}{\sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_i^2) - \mu_x^2} \cdot \sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(y_i^2) - \mu_y^2}}$$

$$r = \frac{Kov(X,Y)}{S(X)\cdot S(Y)} = \frac{\sigma(X,Y)}{\sigma(X)\cdot\sigma(Y)}$$

Rangkorrelation nach Spearman $\quad r_S = 1 - \frac{6\sum_{i=1}^{n}(d_i)^2}{n(n^2-1)}$

$r = [-1;1]$ möglicher Wertebereich

$r = [-1;-0,7[$ sehr starke negative Korrelation

$r = [-0,7;-0,3[$ starke negative Korrelation

$r = [-0,3;0,3[$ keine Korrelation (Punktwolke)

$r = [0,3;0,7[$ starke positive Korrelation

$r = [0,7;1]$ sehr starke positive Korrelation

Korrelation nach Brevais-Pearson – Herleitung:

$$r \;=\; \frac{\sum_{i=1}^{n}(x_i-\mu_x)\cdot(y_i-\mu_y)}{\sqrt{\sum_{i=1}^{n}(x_i-\mu_x)^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i-\mu_y)^2}} \;=\; \frac{\sum_{i=1}^{n}(x_iy_i)-n\cdot\mu_x\cdot\mu_y}{\sqrt{\sum_{i=1}^{n}(x_i^2)-n\cdot\mu_x^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i^2)-n\cdot\mu_y^2}}$$

$$r \;\overset{\cdot\frac{1}{n}}{\underset{\cdot\frac{1}{n}}{=}}\; \frac{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_iy_i)-\mu_x\cdot\mu_y}{\frac{1}{n}\cdot\sqrt{\sum_{i=1}^{n}(x_i^2)-n\cdot\mu_x^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i^2)-n\cdot\mu_y^2}} \;=\; \frac{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_iy_i)-\mu_x\cdot\mu_y}{\sqrt{\frac{1}{n}}\cdot\sqrt{\frac{1}{n}}\cdot\sqrt{\sum_{i=1}^{n}(x_i^2)-n\cdot\mu_x^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i^2)-n\cdot\mu_y^2}}$$

$$r \;=\; \frac{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_iy_i)-\mu_x\cdot\mu_y}{\sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_i^2)-\mu_x^2}\cdot\sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(y_i^2)-\mu_y^2}} \;=\; \frac{Kov(X,Y)}{S(X)\cdot S(Y)} \;=\; \frac{\sigma(X,Y)}{\sigma(X)\cdot\sigma(Y)}$$

Rangkorrelation nach Spearman – Herleitung:

Ausgangspunkt: Korrelation nach Brevais-Pearson

Problem: Die Merkmale sind nicht metrisch skaliert sondern ordinal skaliert

Lösungsansatz:

Umordnung der Merkmale in „metrische-skalierte" Ränge als Alternativmerkmale, damit die ordinale Skalierungsidee erhalten bleibt und gleichzeitig eine analytische Betrachtung und Verrechnung im Sinne einer statistischen Merkmalsauswertung (MW- und Varianzberechnung) ermöglicht werden.

$$r \;=\; \frac{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_iy_i)-\bar{x}\cdot\bar{y}}{\sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_i^2)-(\bar{x})^2}\cdot\sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(y_i^2)-(\bar{y})^2}}$$

Bemerkungen: Da hier für x und y Ränge/Rangfolgen von 1, ..., n betrachtet werden, gilt

$$\left.\begin{aligned}
&\sum_{i=1}^{n}x_i = 1+2+3+\dots+n = \frac{n\cdot(n+1)}{2} \xrightarrow{\text{Mittelwert}} \frac{1}{n}\sum_{i=1}^{n}x_i = \frac{1}{n}\cdot(1+2+3+\dots+n) = \frac{1}{n}\cdot\frac{n\cdot(n+1)}{2} = \frac{n+1}{2} = \bar{x}\\
&\sum_{i=1}^{n}y_i = 1+2+3+\dots+n = \frac{n\cdot(n+1)}{2} \xrightarrow{\text{Mittelwert}} \frac{1}{n}\sum_{i=1}^{n}y_i = \frac{1}{n}\cdot(1+2+3+\dots+n) = \frac{1}{n}\cdot\frac{n\cdot(n+1)}{2} = \frac{n+1}{2} = \bar{y}
\end{aligned}\right\}\bar{x}=\bar{y}$$

zudem:

$$\sum_{i=1}^{n}x_i = \sum_{i=1}^{n}y_i = \sum_{i=1}^{n}i = \frac{n\cdot(n+1)}{2} \quad und \quad \sum_{i=1}^{n}x_i^2 = \sum_{i=1}^{n}y_i^2 = \sum_{i=1}^{n}i^2 = \frac{n\cdot(n+1)\cdot(2n+1)}{6}$$

Für Zähler folgt:

$$Z \;=\; \frac{1}{n}\cdot\sum_{i=1}^{n}(x_iy_i)-\bar{x}\cdot\bar{y} \;\underset{\bar{x}=\frac{n+1}{2}}{\overset{\bar{x}=\bar{y}}{=}}\; \frac{1}{n}\cdot\sum_{i=1}^{n}(x_iy_i)-\frac{n+1}{2}\cdot\frac{n+1}{2} \;=\; \frac{1}{n}\cdot\sum_{i=1}^{n}(x_iy_i)-\frac{(n+1)^2}{4}$$

Ergänzung für $\sum_{i=1}^{n}(x_i y_i)$: Binomischer Lehrsatz & Rechenregeln für Summen

$$\sum_{i=1}^{n}(x_i-y_i)^2 = \sum_{i=1}^{n}\left(x_i^2-2x_iy_i+y_i^2\right) = \sum_{i=1}^{n}x_i^2 - 2\sum_{i=1}^{n}(x_iy_i) + \sum_{i=1}^{n}y_i^2 \underset{\sum_{i=1}^{n}x_i^2=\frac{n\cdot(n+1)\cdot(2n+1)}{6}}{\overset{\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}y_i^2}{=}}$$

$$\sum_{i=1}^{n}(x_i-y_i)^2 = \frac{n\cdot(n+1)\cdot(2n+1)}{6} - 2\sum_{i=1}^{n}(x_iy_i) + \frac{n\cdot(n+1)\cdot(2n+1)}{6} = \frac{n\cdot(n+1)\cdot(2n+1)}{3} - 2\sum_{i=1}^{n}(x_iy_i)$$

$$\xrightarrow[\sum_{i=1}^{n}(x_iy_i)]{\textit{Auflösen nach}} \sum_{i=1}^{n}(x_iy_i) = \frac{n\cdot(n+1)\cdot(2n+1)}{6} - \frac{1}{2}\cdot\sum_{i=1}^{n}(x_i-y_i)^2$$

Für Nenner folgt:

$$N = \sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_i^2)-(\bar{x})^2}\cdot\sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(y_i^2)-(\bar{y})^2} \underset{\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}y_i^2}{\overset{\bar{x}=\bar{y}}{=}} \sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_i^2)-(\bar{x})^2}\cdot\sqrt{\frac{1}{n}\cdot\sum_{i=1}^{n}x_i^2-(\bar{x})^2}$$

$$N \underset{\textit{zusamemnfassen}}{\overset{\textit{Wurzel}}{=}} \frac{1}{n}\cdot\sum_{i=1}^{n}(x_i^2)-(\bar{x})^2 \underset{\sum_{i=1}^{n}(x_i^2)=\frac{n\cdot(n+1)\cdot(2n+1)}{6}}{\overset{\bar{x}=\frac{n+1}{2}}{=}} \frac{1}{n}\cdot\frac{n\cdot(n+1)\cdot(2n+1)}{6}-\frac{1}{4}(n+1)^2 = \frac{(n+1)\cdot(2n+1)}{6}-\frac{(n+1)^2}{4}$$

$$N = \frac{2n^2+3n+1}{6}-\frac{n^2+2n+1}{4} = \frac{1}{3}n^2+\frac{1}{2}n+\frac{1}{6}-\frac{1}{4}n^2-\frac{1}{2}n-\frac{1}{4} = \frac{1}{12}n^2-\frac{1}{12} = \frac{1}{12}\left(n^2-1\right)$$

Daraus folgt:

$$r_S = \frac{\frac{1}{n}\cdot\sum_{i=1}^{n}(x_iy_i)-\frac{(n+1)^2}{4}}{\frac{1}{12}\left(n^2-1\right)} = \frac{\frac{1}{n}\cdot\left[\frac{n\cdot(n+1)\cdot(2n+1)}{6}-\frac{1}{2}\cdot\sum_{i=1}^{n}(x_i-y_i)^2\right]-\frac{(n+1)^2}{4}}{\frac{1}{12}\left(n^2-1\right)}$$

$$r_S = \frac{\frac{(n+1)\cdot(2n+1)}{6}-\frac{1}{2n}\cdot\sum_{i=1}^{n}(x_i-y_i)^2-\frac{(n+1)^2}{4}}{\frac{1}{12}\left(n^2-1\right)} = \frac{\frac{\left(2n^2+3n+1\right)}{6}-\frac{1}{2n}\cdot\sum_{i=1}^{n}(x_i-y_i)^2-\frac{\left(n^2+2n+1\right)}{4}}{\frac{1}{12}\left(n^2-1\right)}$$

$$r_S = \frac{\frac{1}{12}\cdot\left[4n^2+6n+2-\frac{6}{n}\cdot\sum_{i=1}^{n}(x_i-y_i)^2-3n^2-6n-3\right]}{\frac{1}{12}\cdot\left(n^2-1\right)} = \frac{n^2-1-\frac{6}{n}\cdot\sum_{i=1}^{n}(x_i-y_i)^2}{\left(n^2-1\right)}$$

$$r_S = \frac{n^2-1}{\left(n^2-1\right)}-\frac{\frac{6}{n}\cdot\sum_{i=1}^{n}(x_i-y_i)^2}{\left(n^2-1\right)} = 1-\frac{6\cdot\sum_{i=1}^{n}(x_i-y_i)^2}{n\cdot\left(n^2-1\right)} \overset{x_i-y_i=d_i}{=} 1-\frac{6\cdot\sum_{i=1}^{n}d_i^2}{n\left(n^2-1\right)}$$