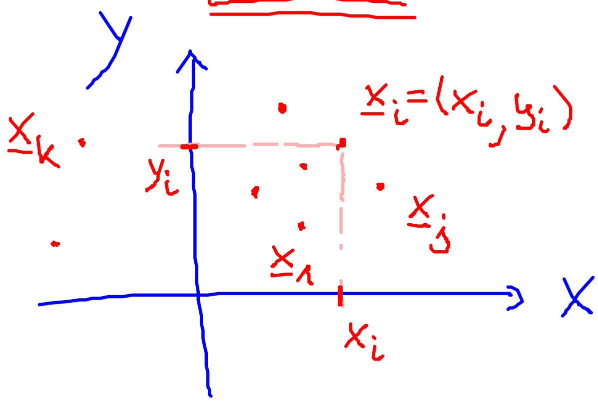
Übung vom 07.11.2012:

Gegeben ein zweidimensionales Merkmal $\underline{X} = (X, Y)$ mit Merkmalsausprägungen $\underline{X}(\omega_i) = (X(\omega_i), Y(\omega_i)) = (x_i, y_i) = \underline{x}_i = \vec{x}_i$

Datensätze = Punkte in der Ebene



Frage: Besteht ein „statistischer" Zusammenhang zwischen X und Y?

Achtung: Statistischer Zusammenhang bedeutet nicht automatisch kausaler Zusammenhang

$\Longrightarrow$ Dazu H16!

Aufgabe Ü15

i	1	2	3	4	5	6
x_i	-1.188	-1.354	-1.854	0.146	-0.354	-0.521
	$\shortparallel$ x_1		$\shortparallel$ x_3			

a) Betrachte dazu Y mit linearer Transformation

$$\boxed{y_i = \alpha + \beta x_i} \qquad \alpha = 0.85433 \;,\; \beta = 1$$

$y_i = x_i + 0.85433$

Standardisierte Daten

Empirische Korrelation: $r(\underline{x},\underline{y}) = \mathrm{cov}(\underline{x}^*,\underline{y}^*) = \dfrac{\mathrm{cov}(\underline{x},\underline{y})}{\sqrt{\mathrm{var}(\underline{x})}\cdot\sqrt{\mathrm{var}(\underline{y})}}$ $\stackrel{!}{=} \dfrac{s_{xy}}{s_x \cdot s_y}$

Standardabweichungen !

Für $\mathrm{var}(\underline{y})$ gilt:

$$\mathrm{var}(\underline{y}) = \mathrm{var}(\alpha\cdot\underline{1} + \beta\underline{x}) = \beta^2\cdot\mathrm{var}(\underline{x})$$

$$\underline{1} = \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$$

$$\underset{\beta=1}{=} \mathrm{var}(\underline{x})$$

Dann:

Nenner zu $r(\underline{x},\underline{y})$: $\sqrt{\mathrm{var}(\underline{x})}\cdot\sqrt{\mathrm{var}(\underline{y})} = \sqrt{\mathrm{var}(\underline{x})}\cdot\sqrt{\mathrm{var}(\underline{x})} = \sqrt{\mathrm{var}(\underline{x})}^2 = \mathrm{var}(\underline{x})$

Zähler: $\mathrm{cov}(\underline{x},\underline{y}) = \frac{1}{n}\cdot\sum_{i=1}^{n}(x_i-\bar{x})\cdot(y_i-\bar{y})$ $\quad y_i = x_i + \alpha \;\Rightarrow\; \bar{y} = \bar{x} + \alpha$

$$= \frac{1}{n}\cdot\sum_{i=1}^{n}(x_i-\bar{x})\cdot\big(\underbrace{(x_i+\cancel{\alpha}) - (\bar{x}+\cancel{\alpha})}_{=x_i-\bar{x}}\big)$$

$$= \frac{1}{n}\cdot\sum_{i=1}^{n}(x_i-\bar{x})(x_i-\bar{x})$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \mathrm{var}(\underline{x})$$

Konkret mit Zahlen:

$$\bar{x} = \frac{1}{6}\cdot\sum_{i=1}^{6}x_i$$

$$= \frac{1}{6}(-1.188 - 1.354 \pm \ldots - 0.521)$$

$$= \frac{-5.125}{6} \approx -0.85417$$

5 Nachkommastellen à la α!!

Also hier: $\boxed{\mathrm{cov}(\underline{x},\underline{y}) = \mathrm{var}(\underline{x})}$

Zusammengenommen: $\boxed{r(\underline{x},\underline{y}) = \frac{cov(\underline{x},\underline{y})^{=var(\underline{x})}}{\underbrace{\sqrt{var(\underline{x})}\cdot\sqrt{var(\underline{y})}}_{=var(\underline{x})}} = \frac{var(\underline{x})}{var(\underline{x})} = 1}$

Teil (b): Analog für

$$u_i = \frac{1}{x_i}, \quad v_i = \frac{1}{y_i} = \frac{1}{x_i + 0.85433}$$

// Achtung u_i, v_i sind nicht mehr über eine lineare Transformation miteinander verbunden!!

Hier muss man tatsächlich

$$\bar{u} = \frac{1}{6}\cdot\sum_{i=1}^{6}\frac{1}{x_i}, \quad \bar{v} = \frac{1}{6}\cdot\sum_{i=1}^{6}\frac{1}{y_i} \quad \text{sowie}$$

$$\boxed{var(\underline{u}) = \frac{1}{6}\cdot\sum_{i=1}^{6}(u_i-\bar{u})^2 = \left(\frac{1}{6}\cdot\sum_{i=1}^{6}u_i^2\right)-\bar{u}^2}, \quad \boxed{var(\underline{v}) = \frac{1}{6}\sum_{i=1}^{6}(v_i-\bar{v})^2 = \left(\frac{1}{6}\sum_{i=1}^{6}v_i^2\right)-\bar{v}^2}$$

Bis auf Rundungsungenauigkeiten erhält man:

$$\boxed{r(\underline{u},\underline{v}) \approx 0} \Rightarrow u_i, v_i \text{ sind unkorreliert!!}$$

Morgen in der Vorlesung mehr zur Kovarianz und empirischen Korrelation.

ENDE der Übung!!