Korelacja – regresja

l.p.	X	Y	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$y_i-\bar{y}$	$(y_i-\bar{y})^2$	$(x_i-\bar{x})(y_i-\bar{y})$
	$\sum x_i$	$\sum y_i$	Σ = 0	$\sum(x_i-\bar{x})^2$	Σ = 0	$\sum(y_i-\bar{y})^2$	$\sum(x_i-\bar{x})(y_i-\bar{y})$

średnia:

$$\bar{x}=\frac{\Sigma x_i}{n}$$

$$\bar{y}=\frac{\Sigma y_i}{n}$$

wariancja:

$$S_x^2=\frac{\sum(x_i-\bar{x})^2}{n}$$

$$S_y^2=\frac{\sum(y_i-\bar{y})^2}{n}$$

odchylenie standardowe:

$$S_x=\sqrt{S_x^2}$$

$$S_y=\sqrt{S_y^2}$$

współczynnik zmienności:

$$V_{S_x}=\frac{S_x}{\bar{x}}\cdot 100\%$$

$$V_{S_y}=\frac{S_y}{\bar{y}}\cdot 100\%$$

- ***kowariancja:***

$$c(x,y)\overset{\text{(lub)}}{=}\operatorname{cov}(x,y)=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n}=r_{xy}\cdot S_x\cdot S_y=a_x\cdot S_y^2=a_y\cdot S_x^2 \qquad [X\cdot Y]$$

- ***współczynnik korelacji (Pearsona):***

$$r_{xy}=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\cdot\sum(y_i-\bar{y})^2}}=\frac{c(x,y)}{S_x\cdot S_y}=\frac{a_x\cdot S_y}{S_x}=\frac{a_y\cdot S_x}{S_y}=\pm\sqrt{a_x\cdot a_y} \qquad [brak\ \ jednostki]$$

$$\boxed{-1\le r_{xy}\le 1}$$

- ***współczynnik determinacji (R^2) i współczynnik zbieżności (φ^2)***

$$R^2=(r_{xy})^2=a_x\cdot a_y \qquad \boxed{R^2+\varphi^2=1}$$

- ***proste regresji (II-go rodzaju, MNK):***

(Y względem X): $\boxed{\hat{y}=a_y\cdot x+b_y}$

$$a_y=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}=\frac{c(x,y)}{S_x^2}=\frac{r_{xy}\cdot S_y}{S_x}=\frac{a_x\cdot S_y^2}{S_x^2}=\frac{r_{xy}^2}{a_x} \qquad \left[\frac{Y}{X}\right]$$

$$b_y=\bar{y}-a_y\cdot\bar{x} \qquad [Y]$$

(X względem Y): $\boxed{\hat{x}=a_x\cdot y+b_x}$

$$a_x=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(y_i-\bar{y})^2}=\frac{c(x,y)}{S_y^2}=\frac{r_{xy}\cdot S_x}{S_y}=\frac{a_y\cdot S_x^2}{S_y^2}=\frac{r_{xy}^2}{a_y} \qquad \left[\frac{X}{Y}\right]$$

$$b_x=\bar{x}-a_x\cdot\bar{y} \qquad [X]$$