Il rapporto di correlazione di Pearson

è basato sulla scomposizione della variabilità di una variabile Y quantitativa rispetto alle modalità di una variabile X qualitativa

Devianza e sua scomposizione

$$
\begin{aligned}
Dev(Y) &= \sum_i \sum_j (y_j - \mu_Y)^2 n_{ij} = \\
&= \sum_i \sum_j (y_j - \mu_{Y|x_i} + \mu_{Y|x_i} - \mu_Y)^2 n_{ij} = \\
&= \sum_i \sum_j (y_j - \mu_{Y|x_i})^2 n_{ij} + \sum_i \sum_j (\mu_{Y|x_i} - \mu_Y)^2 n_{ij} + \\
&+ \boxed{2\sum_i \sum_j (y_j - \mu_{Y|x_i})(y_j - \mu_{Y|x_i}) n_{ij}}
\end{aligned}
$$

Sviluppando il quadrato si dimostra che il doppio prodotto è nullo!

Devianza e sua scomposizione

$$
\begin{aligned}
Dev(Y) &= \sum_i \sum_j (y_j - \mu_{Y|x_i})^2 n_{ij} + \sum_i (\mu_{Y|x_i} - \mu_Y)^2 \sum_j n_{ij} \\
&= \sum_i (Dev(Y \mid X = x_i)) + \sum_i (\mu_{Y|x_i} - \mu_Y)^2 n_{i+}
\end{aligned}
$$

$$\boxed{Dev(Y) = Dev(W) + Dev(B)}$$

Devianza interna (Within) — Devianza esterna (between)

Rapporto di correlazione del Pearson

$$\eta^2_{Y/X} = \frac{Dev(B)}{Dev(Y)} = 1 - \frac{Dev(W)}{Dev(Y)} \qquad 0 \le \eta^2_{Y/X} \le 1$$

$$\eta^2_{Y/X} \neq \eta^2_{X/Y} \longrightarrow$$ Il rapporto di correlazione è un indice NON SIMMETRICO

Da Federica, corso prof. Massimo Aria