Prof. Dr. Peter von der Lippe

Einige Rekursionsformeln zur multiplen Regression

1. Partielle Regressions- und Korrelationskoeffizienten

Am Beispiel der Zweifachregression

Notation: $\hat{y} = \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 = b_{y.12} + b_{y1.2} x_1 + by_{y2.1} x_2$. Es gelten dann *für die(nicht-standardisierten)* ***partiellen Regressionskoeffizienten*** $b_{y1.2}$ und $b_{y2.1}$ die Rekursionsformeln

$$b_{y1.2} = \frac{b_{y1} - b_{y2} b_{21}}{1 - b_{12} b_{21}} \text{ und } b_{y2.1} = \frac{b_{y2} - b_{y1} b_{12}}{1 - b_{12} b_{21}}$$

(man beachte $b_{21} \neq b_{12}$ und $b_{21}\ b_{12} = r_{12}^2$). *Standardisierte* Regressionskoeffizienten (auch β–Koeffizienten genannt, nicht zu verwechseln mit β in der Notation von v. Auer) sind dagegen

$\beta_{y1.2} = b_{y1.2} \frac{s_1}{s_y}$ und $\beta_{y2.1} = b_{y1.2} \frac{s_2}{s_y}$, mit s_1, s_2 und s_y die Standardabweichungen von x_1, x_2 und y.

Rekursionsformel ***für die partiellen Korrelationskoeffizienten***

$$(1) \quad r_{y1.2} = \frac{r_{y1} - r_{y2} r_{12}}{\sqrt{(1 - r_{y2}^2)(1 - r_{12}^2)}} \quad \text{und} \quad (2) \quad r_{y2.1} = \frac{r_{y2} - r_{y1} r_{12}}{\sqrt{(1 - r_{y1}^2)(1 - r_{12}^2)}}$$

beachte $r_{12} = r_{21}$.

2. Multipler Korrelationskoeffizienten R und multiple Bestimmtheit R^2 (coefficient of determination)

$R_{y.12}$ ist der einfache Korrelationskoeffizient zwischen y und $\hat{y} = f(x_1, x_2)$, entsprechend ist $R_{y.123}$ die Korrelation zwischen y und $\hat{y} = f(x_1, x_2, x_3)$ im Falle einer Dreifachregression. Für die multiple Bestimmtheit gilt die Rekursionsformel

$$(3) \quad R_{y.12}^2 = \frac{r_{y1}^2 + r_{y2}^2 - 2 r_{y1} r_{y2} r_{12}}{1 - r_{12}^2}.$$

Man beachte, dass bei $r_{12} = 0$ (unabhängige Regressoren) gilt $R_{y.12}^2 = r_{y1}^2 + r_{y2}^2$ (sonst [bei korrelierten Regressoren] ist stets $R_{y.12}^2 < r_{y1}^2 + r_{y2}^2$)

$$(4) \quad R_{y.12}^2 = r_{y1}^2 + r_{y2.1}^2 (1 - r_{y1}^2) \quad \text{und} \quad (4a) \quad R_{y.12}^2 = r_{y2}^2 + r_{y2.2}^2 (1 - r_{y2}^2)$$

Nach Gl. 2 ist das $R_{y.12}^2 = r_{y1}^2 + \frac{(r_{y2} - r_{y1} r_{12})^2}{1 - r_{12}^2}$ und ausmultipliziert ergibt das Gl. 3. Ferner gilt

(5) $R_{y.123}^2 = R_{y.12}^2 + r_{y3.12}^2 (1 - R_{y.12}^2)$, woraus folgt dass $R_{y.123}^2 \geq R_{y.12}^2$, entsprechend ist $R_{y.1234}^2 \geq R_{y.123}^2$ usw. Man braucht deshalb ein ***korrigiertes Bestimmtheitsmaß*** *(adjusted R square)*

$$(6) \quad \overline{R}^2 = 1 - \frac{S_{\hat{u}\hat{u}} / (T - K - 1)}{S_{yy} / (T - 1)} = 1 - (1 - R^2) \frac{T - 1}{T - K - 1} = R^2 + \frac{(R^2 - 1) K}{T - K - 1}$$

denn $1 - R^2 = S_{\hat{u}\hat{u}} / S_{yy}$. Offenbar ist auch wegen $R^2 - 1 < 0$ auch $\overline{R}^2 \leq R^2$.