

Hausaufgaben für LL-08, Statistik zum 16.4.2009

Aufgabe 1

Lese Abschnitt 2.7 *Korrelationsrechnung*.

Formel (2.62): Da sich Faktoren wie $1/(n-1)$ rauskürzen, gilt

$$r_{xy} = \frac{s_{xy}}{s_x s_y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Hinweis: Die Definition des Bestimmtheitsmaßes (2.63) ist sehr unglücklich gewählt (Leider genau so im Buch von Duller, Kapitel 9). Man kann zeigen, dass diese Definition mit der unten angegebenen Formel übereinstimmt, wenn man ein lineares Modell betrachtet. Bei anderen Modellen stimmen die Formeln nicht überein. Der Begriff des Bestimmtheitsmaßes wird verwendet um allgemein die Güte eines Regressionsmodells zu testen. Dazu lautet die Definition

$$B = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

Im Zähler steht die „Varianz“ der Residuen. Das Bestimmtheitsmaß beschreibt grob den Anteil der Varianz, der durch das Modell erklärt wird. Bei Interesse: Mehr zum Bestimmtheitsmaß (=Coefficient of Determination) findet man im Buch von Montgomery und Runger, 11-4.2 und 11-7.2.

Bearbeite Aufgabe 2.18. (Aufgabe 2.17 wurde bereits in der Übung gerechnet). Man berechne für Aufgabe 2.17 das Bestimmtheitsmaß nach der Formel im Buch (Ergebnis 0) und nach der obigen Formel für die Modellgleichung $\hat{y}_i = x_i^2$ (Ergebnis 1). Die Diskrepanz ist offensichtlich.

Aufgabe 2

Lese Abschnitt 3.1 *Kombinatorische Grundlagen*.

Bearbeite Aufgaben 3.1 und 3.2.

Aufgabe 3

Lese im Buch von Duller Kapitel 8 *Multivariate Deskriptive Statistik*. Überfliege dabei die Abschnitte, die zusätzlichen Stoff zur Vorlesung bieten (z.B. Assoziationsmaß χ^2 , Cramersches Assoziationsmaß, Spearmanscher Rangkorrelationskoeffizient. Es genügt, einen Eindruck zu gewinnen, wie viel mehr Konzepte es gibt. Es gibt auch noch viel mehr als hier dargestellt.). Beachte, dass der Korrelationskoeffizient von Bravais-Pearson der ist, den wir in der Vorlesung eingeführt haben.