

**Aufgabenblatt 4 (24.5.2011)**

1. Wenn man für das Beispiel in Tabelle 4 des QCA-Aufsatzes ein Logitmodell

$$\Pr(S = 1|R, L, E) = \frac{\exp(\alpha + R\beta_R + L\beta_L + E\beta_E)}{1 + \exp(\alpha + R\beta_R + L\beta_L + E\beta_E)}$$

berechnet, findet man folgende Parameter:  $\alpha = 15.6554$ ,  $\beta_R = -17.9722$ ,  $\beta_L = 1.3398$ ,  $\beta_E = -16.6323$ . Berechnen Sie für die in der Tabelle des Aufsatzes angegebenen Konstellationen die durch das Modell geschätzten Wahrscheinlichkeiten.

2. Es gibt folgende Daten für drei statistische Variablen:  $X$  (Bildungsniveau),  $Y$  (Höhe des Arbeitseinkommens) und  $Z$  (Indikator für Gruppe).

$X$	$Y$	$Z$
2	2000	1
3	3000	1
3	3200	1
2	2500	1
2	2800	1
4	4000	1
4	1000	2
5	2000	2
5	2200	2
4	1500	2
4	1700	2
6	2900	2

- Zeichnen Sie die Daten in ein Streuungsdiagramm ein.
- Berechnen sie (mit OLS) lineare Regressionsgraden: erstens für alle Daten, dann gesondert für die Gruppen ( $Z = 1$  und  $Z = 2$ ).
- Zeichnen Sie die Regressionsgraden in das Streuungsdiagramm ein.
- Berechnen und interpretieren Sie exemplarisch:  $M(Y|X = 2)$ ,  $M(Y|X = 2, Z = 1)$ ,  $M(Y|X = 2, Z = 2)$  und  $M(Y|X = 3)$ ,  $M(Y|X = 3, Z = 1)$ ,  $M(Y|X = 3, Z = 2)$ .
- Berechnen Sie für die unter (d) angegebenen Größen Schätzwerte aus den in (b) berechneten Modellen.
- Erläutern Sie, inwiefern es sich um ein Beispiel für Simpsons Paradox handelt.