

### Aufgabenblatt 14 (1.7.2010)

1. Es gibt folgende Daten:  $S$  = Geschlecht,  $X$  = Bildungsniveau,  $Y$  = Arbeitslohn.

$S$	$X$	$Y$
0	1	1000
0	2	1500
0	3	1700
0	4	1700
0	5	2000
0	5	2500
0	7	2200
0	7	2600
0	8	2200
0	8	2500
1	1	500
1	2	700
1	3	500
1	3	1000
1	4	1000
1	5	800
1	6	1200
1	7	1300
1	7	1200
1	8	1100
1	8	1200

Eine lineare Regression liefert

$$M(Y|X = x, S = s) = 1088.2 + 180.4x - 613.6s - 82.6xs$$

- Zeichnen Sie die Daten in ein Streuungsdiagramm ein.
- Leiten Sie zwei separate Regressionsfunktionen für  $S = 0$  und  $S = 1$  ab und zeichnen Sie diese Funktionen in das Streuungsdiagramm ein.

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	0
3	3000	0
3	3200	0
2	2500	0
2	2800	0
4	4000	0
4	1000	1
5	2000	1
5	2200	1
4	1500	1
4	1700	1
6	2900	1

Man findet folgende Regressionsfunktionen:

$$M(Y|X = x) = 2635.7 - 64.3x$$

$$M(Y|X = x, Z = z) = 890.0 + 760.0x - 2460.0z - 20.0xz$$

- Zeichnen Sie die Daten in ein Streuungsdiagramm ein.
- Berechnen Sie die gesonderten Regressionsgeraden für  $Z = 0$  und  $Z = 1$ .
- Zeichnen Sie die gemeinsame und die beiden separaten Regressionsgeraden in das Streuungsdiagramm ein.
- Berechnen und interpretieren Sie exemplarisch:  $M(Y|X = 2)$ ,  $M(Y|X = 2, Z = 0)$ ,  $M(Y|X = 2, Z = 1)$  und  $M(Y|X = 3)$ ,  $M(Y|X = 3, Z = 0)$ ,  $M(Y|X = 3, Z = 1)$ .
- Erläutern Sie, inwiefern es sich um ein Beispiel für Simpsons Paradox handelt.