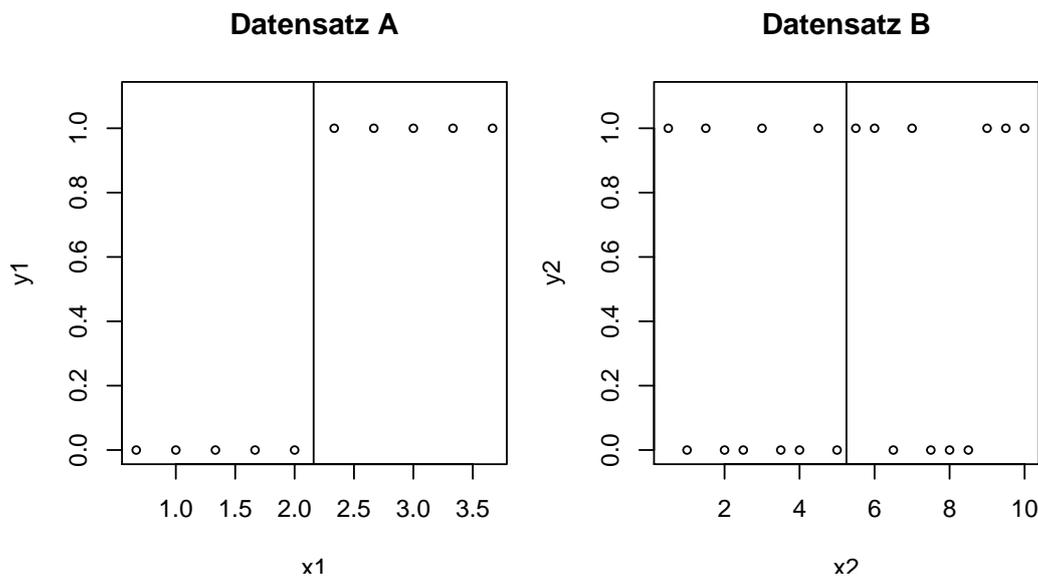


**Aufgabe 29\*** (20 Punkte)

Für sechs Teilnehmer einer Umfrage im Rahmen der Vorlesung zur deskriptiven Statistik sind die Dauer des Weges zur Uni in Minuten  $X$  und die binäre Variable  $Y$ , die beschreibt, ob die Person im Innenraum ( $Y=0$ ) oder im Außenraum ( $Y = 1$ ) des MVVs wohnt, gegeben:

Person $i$	1	2	3	4	5	6
$X$	14	15	40	45	50	70
$Y$	0	0	1	1	0	1

- Zeichnen Sie die ROC-Kurve für  $X$  und  $Y$  und berechnen Sie den AUC-Wert. Beschreiben Sie Ihr Vorgehen!
- Interpretieren Sie das Ergebnis aus Teilaufgabe a).
- Ändert sich der AUC-Wert, wenn die Zeit in Stunden anstatt in Minuten zur Berechnung verwendet wird? Begründen Sie kurz!
- Wie würde die ROC-Kurve aussehen, wenn Person 5 im Außenraum wohnen würde?
- Wie hoch ist jeweils in etwa der AUC-Wert für die Datensätze A und B, die in den nachfolgenden Streudiagrammen abgebildet sind? (Ohne Berechnung, aber mit Begründung!)



**Aufgabe 30**

Gegeben ist die Anzahl  $X$  (in Tsd.) der Geburten eines Landes und die Größe  $Y$  (in Tsd.) der Storchpopulation des Landes über den Zeitraum  $T = 1, \dots, 10$ .

$T$	1	2	3	4	5	6	7	8	9	10
$X$	38	37	37	37	35	34	34	34	33	31
$Y$	49	48	47	46	46	46	44	44	43	40

- Berechnen Sie die Korrelation von  $X$  und  $Y$
- Führen Sie eine Trendbereinigung von  $X$  und  $Y$  wie folgt durch: Rechnen Sie eine Regression von  $X$  bzw.  $Y$  auf  $T$ . Die Residuen der beiden Regressionen entsprechen den trendbereinigten Werten. Berechnen Sie die Korrelation der trendbereinigten Werte und kommentieren sie das Ergebnis.

### Aufgabe 31

Bei einer Erhebung wurden an Personen jeweils zwei Blutdruckmessungen SBD1 und SBD2 mit unterschiedlichen Messgeräten durchgeführt. Es wurde eine Regression berechnet. Das Ergebnis und eine Deskription der Daten sind im angefügten R-Output zu finden.

- Berechnen Sie die Korrelation der Variablen SBD1 und SBD2 nach Pearson und interpretieren Sie diesen Wert.
- Wie lautet die zum angefügten R-Output gehörige Modellgleichung? Geben Sie für dieses Modell die geschätzten Regressionskoeffizienten an und interpretieren Sie den geschätzten Steigungsparameter!
- Geben Sie für dieses Modell  $R^2$  an und interpretieren Sie den Wert!
- Welcher Wert für SBD2 wird auf der Basis dieses Regressionsmodells prognostiziert, wenn der Wert für SBD1 bei 160 liegt?
- Berechnen Sie die Regressionskoeffizienten der orthogonalen Regression!  
Hinweis:  $S_{\text{SBD1}}^2 = 9762.741$ ,  $S_{\text{SBD2}}^2 = 7329.630$ .
- Wie lauten die Modellgleichung und die geschätzten Regressionskoeffizienten für die Umkehrregression?

```
> summary(SBD1)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
100.0  105.5   114.0   122.5  139.0   166.0
> summary(SBD2)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
101.0  108.5   118.0   121.3  128.5   160.0
> ergr<-(lm(SBD2~SBD1))
> summary(ergr)

Call:
lm(formula = SBD2 ~ SBD1)

Residuals:
    Min       1Q   Median       3Q      Max
-28.2997  -5.8952  -0.5398   8.3695  20.9650

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  42.3437    14.3536   2.950  0.00681 **
SBD1         0.6446     0.1158   5.566  8.67e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.44 on 25 degrees of freedom
Multiple R-squared:  0.5535, Adjusted R-squared:  0.5356
F-statistic: 30.99 on 1 and 25 DF,  p-value: 8.674e-06
```