Klausur WS 2018/19

1. Aufgabe

Gegeben sei die Zufallsvariable X mit den Werten A, B und C, die angibt, welche Therapie eine Person erhält und die dichotome Zufallsvariable Y, die angibt, ob die Therapie erfolgreich ($Y = 1$) oder nicht erfolgreich war ($Y = 0$). In der folgenden Tabelle finden Sie die Wahrscheinlichkeiten für alle Kombinationen von Werten, die die beiden Zufallsvariablen annehmen können:

	$X = A$	$X = B$	$X = C$
$Y = 0$	0.1	0.1	0.15
$Y = 1$	0.1	0.3	0.25

Die Regression von Y auf X wird parametrisiert durch

$$E(Y|X) = \alpha_0 + \alpha_1 I_{X=B} + \alpha_2 I_{X=C}$$

Welche der folgenden Aussagen sind korrekt?

- ☐ $E(Y|X = B) = 0.75$
- ☐ $P(Y = 1|X = B \text{ oder } X = C) = 0.55$
- ☐ $\alpha_1 = 0.3$
- ☐ $P(X = B|Y = 1) = 0.3$

2. Aufgabe

Mit der Allgemeinen Linearen Hypothese prüft man Hypothesen der Form $\boldsymbol{A\beta} - \boldsymbol{\delta} = \boldsymbol{0}$. Wir nehmen an, dass das Allgemeine Lineare Modell über Zellenmittelwerte parametrisiert wurde, d.h. $\boldsymbol{\beta} = (\mu_1, \mu_2, \mu_3)'$. Welche der folgenden Aussagen sind korrekt?

- ☐ Mit $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ und $\boldsymbol{\delta} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ prüft man die Hypothese, dass $\mu_1 = \mu_2 = 0$
- ☐ Mit $\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ und $\boldsymbol{\delta} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ prüft man die Hypothese, dass $\mu_1 = \mu_2 = 0$
- ☐ Mit $\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 0 \end{pmatrix}$ und $\boldsymbol{\delta} = \begin{pmatrix} 2 \end{pmatrix}$ prüft man die Hypothese, dass $\mu_1 = 2$ und $\mu_2 = -2$
- ☐ Mit $\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$ und $\boldsymbol{\delta} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ prüft man die Hypothese, dass $\mu_1 = \mu_2 + 1$ und $\mu_2 = \mu_3 + 1$

3. Aufgabe

Gegeben seien die numerischen Zufallsvariablen X, Y und Z. Wir betrachten die Regressionen $E(Y|X)$ und $E(Y|X,Z)$. Es gelte:

- Der Determinationskoeffizient der zweiten Regression sei $R^2_{Y|X,Z} = 0.45$
- $Var(Y) = 20$
- $Var(E(Y|X)) = 8$.

Welche der folgenden Aussagen sind korrekt?

- ◻ Das Modell 2 klärt mehr Varianz auf.
- ◻ Der Determinationskoeffizient der ersten Regression ist $R^2_{Y|X} = 0.4$
- ◻ $Var(E(Y|X,Z)) = 9$
- ◻ $Var(\varepsilon_1) > Var(\varepsilon_2)$, wobei ε_1 das Residuum der ersten Regression, ε_2 das Residdum der zweiten Regression ist.

4. Aufgabe

Laden Sie sich den Output für diese Aufgabe unter folgender Url herunter: `https://eval-serv2.metpsy.uni-jena.de/~andreas/zettel/data/outputLinearesModell.pdf`

In dem Output wird eine Regression der Form

$$E(Y|X_1, X_2) = \alpha_0 + \alpha_1 I_{X_1=Ja} + \alpha_2 I_{X_2=Ja} + \alpha_3 I_{X_1=Ja} I_{X_2=Ja}$$

geschätzt. Die Daten stammen aus einem Versuch, in dem der Abstand Y von Autos, die Fahrradfahrer überholen, in Abhängigkeit verschiedener Kovariaten gemessen wurde. Wir betrachten nur die Kovariate, ob der Fahrradfahrer einen Helm getragen hat ($X_1 = Ja$ oder $X_1 = Nein$) und ob er auf einer Fahrradspur gefahren ist ($X_2 = Ja$ oder $X_2 = Nein$).

Welche der folgenden Aussagen sind korrekt?

- ◻ Das Modell klärt ungefähr 38 Prozent Varianz auf
- ◻ Die Hypothese, dass $E(Y|X_1, X_2) = \alpha_0 + \alpha_1 I_{X_1=Ja} + \alpha_2 I_{X_2=Ja}$ gilt, kann verworfen werden.
- ◻ Die Hypothese, dass $E(Y|X_1, X_2) = \alpha_0 + \alpha_2 I_{X_2=Ja} + \alpha_3 I_{X_1=Ja} I_{X_2=Ja}$ gilt, kann verworfen werden.
- ◻ Der R-Befehl `lm(Überholabstand~Helm+Fahrradspur+Helm:Fahrradspur,data=D)` hätte das gleiche Modell geschätzt.

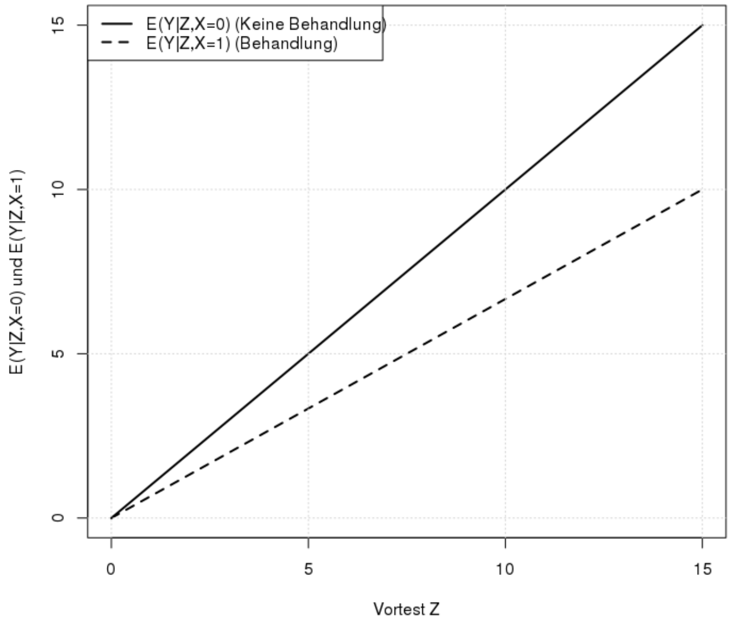
5. Aufgabe

Unter welchen der folgenden Vorraussetzungen gilt, dass $R^2_{Y|X} = 0$?

- ◻ $\varepsilon = 0$
- ◻ $E(\varepsilon|X) = 0$
- ◻ $E(Y|X) = E(Y)$
- ◻ $Var(Y) = Var(E(Y|X))$

6. Aufgabe

Sei Z und Y der Vor- bzw. Nachtest auf einer Depressionsskala. Die Zufallsvariable X mit den Werten $0, 1$ gebe an, ob die Versuchsperson eine Therapie erhalten hat. Höhere Werte für Z und Y bedeuten stärkere Depression. Die folgende Graphik zeigt die Regression $E(Y|X, Z)$



Welche der folgenden Aussagen sind korrekt?

- ◻ Die Regressionsgleichung lautet $E(Y|X, Z) = Z - \frac{1}{3}XZ$
- ◻ Es gibt Interaktion zwischen X und Z
- ◻ Die Graphik zeigt, dass die Therapie jeder einzelnen Person hilft
- ◻ Es handelt sich um eine bedingte lineare Regression