

Statistik 3

Übung 5

Einleitung und Ziele der Übung

Herzlich Willkommen zur fünften Übung Statistik 3. Diese Übung beschäftigt sich mit der Interaktion zwischen nominalen Variablen. Dabei wiederholen wir auch die Dummykodierung.

Agenda

- Regression nominaler Daten mit Interaktion

1. Regression nominaler Daten mit Interaktion - Einführung

Der Befehl `lm()` kann auch für Regressionen mit Interaktionen verwendet werden.

```
1 Beispiel <- data.frame(VP = 1:10, Y = rnorm(10, 5, 1), X1 = rnorm(10, 5, 1),
2                        X2 = rnorm(10, 5, 1), X3 = rnorm(10, 5, 1))
3 #Regression mit 2 UV ohne Interaktionen
4 lm(Y~X1+X2, Beispiel)
5 #Regression mit einer Interaktion zwischen X1 und X2
6 lm(Y~X1+X2+X3+X1:X2, Beispiel)
7 #Kurzschreibweise
8 lm(Y~X1*X2, Beispiel)

1 #Regression mit 3 UV ohne Interaktionen
2 lm(Y~X1+X2+X3, Beispiel)
3 #Regression mit einer Interaktion zwischen X1 und X2
4 lm(Y~X1+X2+X3+X1:X2, Beispiel)
5 #Regression mit allen Interaktionen
6 lm(Y~X1+X2+X3+X1:X2+X1:X3+X2:X3+X1:X2:X3, Beispiel)
7 #Kurzschreibweise
8 lm(Y~X1*X2*X3, Beispiel)
9 #Beide Varianten können auch kombiniert werden
10 lm(Y~X1*X2+X3+X1:X3+X2:X3+X1:X2:X3, Beispiel)
```

1. Regression nominaler Daten mit Interaktion - Übung 1

Gegeben sei der Datensatz Uebung_Statistik3_5A.csv. In diesem wurde die Konzentrationsleistung von Probanden unter Berücksichtigung des Schlafverhaltens (kein Schlaf, Schlaf) und des getrunkenen Kaffees (Ohne Koffein, Mit Koffein) untersucht. Eine Dummykodierung wurde bereits vorgenommen (die Spalten Koffein und Schlaf im Datensatz).

Stellen Sie die Regressionsgleichung zur Beschreibung der Konzentration durch die Dummyvariablen Koffein (X_1) und Schlaf (X_2) einmal mit und einmal ohne Interaktionsterm auf. Denken Sie daran, die Regressionsgleichung ist nicht der R-Befehl. Verwenden Sie bitte die Variante mit den Kleinbuchstaben und dem Index i .

Fahren Sie dann auf der nächsten Seite fort.

1. Regression nominaler Daten mit Interaktion - Übung 2

Die Regressionsgleichungen sollte also folgend lauten:

$$y_i = b_0 + b_1 * x_{1i} + b_2 * x_{2i} + e_i$$

$$y_i = b_0 + b_1 * x_{1i} + b_2 * x_{2i} + b_3 * x_{1i} * x_{2i} + e_i$$

Falls Sie die Populationsvariante mit β statt b verwendet haben, ist das natürlich auch korrekt.

Überlegen Sie nun welche inhaltlichen und welche statistischen Alternativhypothesen für b_0, b_1, b_2 bzw. b_0, b_1, b_2, b_3 jeweils getestet werden.

Fahren Sie dann auf der nächsten Seite fort.

1. Regression nominaler Daten mit Interaktion - Übung 3

Ohne Interaktionsterm Hinweis: c...gleicher Wert

- b_0
 - Die Konzentrationsleistung ohne Schlaf und ohne Koffein ist von 0 verschieden.
 - $\beta_0 \neq 0$ oder $\mu_{x_1=0, x_2=0} \neq 0$
- b_1
 - Die Konzentrationsleistung für Personen mit gleichem Schlafverhalten unterscheidet sich je nachdem, ob sie Kaffee mit oder ohne Koffein getrunken haben.
 - $\beta_1 \neq 0$ oder $\mu_{x_1=0, x_2=c} \neq \mu_{x_1=1, x_2=c}$
- b_2
 - Die Konzentrationsleistung für Personen mit gleichem Kaffeekonsum unterscheidet sich je nachdem, ob sie geschlafen haben oder nicht.
 - $\beta_2 \neq 0$ oder $\mu_{x_1=c, x_2=0} \neq \mu_{x_1=c, x_2=1}$

1. Regression nominaler Daten mit Interaktion - Übung 4

Mit Interaktionsterm Hinweis: c...gleicher Wert

- b_0
 - Die Konzentrationsleistung ohne Schlaf und ohne Koffein ist von 0 verschieden.
 - $\beta_0 \neq 0$ oder $\mu_{x_1=0, x_2=0} \neq 0$
- b_1
 - Die Konzentrationsleistung für Personen ohne Schlaf unterscheidet sich je nachdem, ob sie Kaffee mit oder ohne Koffein getrunken haben.
 - $\beta_1 \neq 0$ oder $\mu_{x_1=0, x_2=0} \neq \mu_{x_1=1, x_2=0}$
- b_2
 - Die Konzentrationsleistung für Personen ohne Koffein im Kaffee unterscheidet sich je nachdem, ob sie geschlafen haben oder nicht.
 - $\beta_2 \neq 0$ oder $\mu_{x_1=0, x_2=0} \neq \mu_{x_1=0, x_2=1}$

...

Fahren Sie auf der nächsten Seite fort.

1. Regression nominaler Daten mit Interaktion - Übung 6

- b_3
 - Der Einfluss von Koffein auf die Konzentrationsleistung ist abhängig vom Schlafverhalten ODER Der Einfluss des Schlafverhaltens auf die Konzentrationsleistung ist abhängig vom Koffein.
 - $\beta_3 \neq 0$ oder
 - $(\mu_{x_1=1,x_2=1} - \mu_{x_1=1,x_2=0}) \neq (\mu_{x_1=0,x_2=1} - \mu_{x_1=0,x_2=0})$ ODER
 - $(\mu_{x_1=1,x_2=1} - \mu_{x_1=0,x_2=1}) \neq (\mu_{x_1=1,x_2=0} - \mu_{x_1=0,x_2=0})$

Fahren Sie auf der nächsten Seite fort.

1. Regression nominaler Daten mit Interaktion - Übung 7

Bestimmen Sie die Regressionskoeffizienten für beide Regressionsgleichungen mittels R. Welche Effekte sind jeweils signifikant für $\alpha = .05$? Wie erklären Sie sich den Unterschied?

Ohne Interaktionsterm

- $p_{b_0} <$
- $p_{b_1} =$
- $p_{b_2} <$

Mit Interaktionsterm

- $p_{b_0} <$
- $p_{b_1} =$
- $p_{b_2} <$
- $p_{b_3} =$

Fahren Sie auf der nächsten Seite fort.

1. Regression nominaler Daten mit Interaktion - Übung 7

Bestimmen Sie für beide Regressionen folgende Simple Slopes. Rechnen Sie dafür mit allen Werten auf 2 Nachkommastellen genau. Denken Sie daran, die Simple Slopes basieren auf jenen Teilungen der Gleichung, die mit der betreffenden Variable multipliziert werden:

- Den Einfluss von Koffein auf die Konzentration bei keinem Schlaf $b_{X_1|X_2=0}$
- Den Einfluss von Koffein auf die Konzentration bei Schlaf $b_{X_1|X_2=1}$
- Den Einfluss von Schlaf auf die Konzentration bei Kaffee ohne Koffein $b_{X_2|X_1=0}$
- Den Einfluss von Schlaf auf die Konzentration bei Kaffee mit Koffein $b_{X_2|X_1=1}$

Ihre Lösungen können Sie auf den nächsten Seite überprüfen. Die Herleitung finden Sie immer auf der folgenden Seite

1. Regression nominaler Daten mit Interaktion - Übung 8

Ohne Interaktionsterm

- Den Einfluss von Koffein auf die Konzentration bei keinem Schlaf $b_{X_1|X_2=0} =$
- Den Einfluss von Koffein auf die Konzentration bei Schlaf $b_{X_1|X_2=1} =$
- Den Einfluss von Schlaf auf die Konzentration bei Kaffee ohne Koffein $b_{X_2|X_1=0} =$
- Den Einfluss von Schlaf auf die Konzentration bei Kaffee mit Koffein $b_{X_2|X_1=1} =$

1. Regression nominaler Daten mit Interaktion - Übung 9

Herleitung:

Die Regressionsgleichung lautet:

$$y_i = b_0 + b_1 * x_{1i} + b_2 * x_{2i} + e_i$$

Simple Slope für X_1 : b_1

Simple Slope für X_2 : b_2

Wir sehen der Einfluss von X_1 ist nur von b_1 abhängig und unabhängig von der Ausprägung von X_2 .

Wir sehen ebenfalls der Einfluss von X_2 ist nur von b_2 abhängig und unabhängig von der Ausprägung von X_1 .

1. Regression nominaler Daten mit Interaktion - Übung 10

Mit Interaktionsterm

- Den Einfluss von Koffein auf die Konzentration bei keinem Schlaf $b_{X_1|X_2=0} =$
- Den Einfluss von Koffein auf die Konzentration bei Schlaf $b_{X_1|X_2=1} =$
- Den Einfluss von Schlaf auf die Konzentration bei Kaffee ohne Koffein $b_{X_2|X_1=0} =$
- Den Einfluss von Schlaf auf die Konzentration bei Kaffee mit Koffein $b_{X_2|X_1=1} =$

1. Regression nominaler Daten mit Interaktion - Übung 11

Herleitung:

Die Regressionsgleichung lautet:

$$y_i = b_0 + b_1 * x_{1i} + b_2 * x_{2i} + b_3 * x_{1i} * x_{2i} + e_i$$

Simple Slope für X_1 : $(b_1 + b_3 * x_{2i})$

Simple Slope für X_2 : $(b_2 + b_3 * x_{1i})$

Wir sehen der Einfluss von X_1 ist nicht mehr nur von b_1 abhängig, sondern auch von b_3 und der Ausprägung von X_2 .

Wir sehen ebenfalls der Einfluss von X_2 ist nicht mehr nur von b_2 abhängig, sondern auch von b_3 und der Ausprägung von X_1 .

Daher unterscheiden sich die Simple Slopes hier.