

Kommentare zu den Beispielfiles der ANOVA mit Ansatz nach dem ALM (= Allg. lineares Modell)

A1

- Einen Überblick der aufgeführten Beispiele erhalten Sie anhand der Datenvorbereitungen im File ANOVA.xls.
- Allgemein gilt: Die hier verwendeten Daten stammen aus simulierten Stichproben (approximiert normalverteilt).
So finden Sie im File ANOVA.xls, Tabelle 2, die Stichproben S1-S4 vor. Diese Stichproben sind für die jeweiligen Simulationen (Spalte H und I in Tabelle 2) nach den zu Grunde gelegten Parameterwerten für die Grundgesamtheiten charakterisiert.
Die Stichprobenwerte für Mittelwert, Varianz und Standardabweichung finden Sie für die jeweilige Stichprobe in den Zeilen 44-46 und den Spalten B-E.

DESIGN-1: (Tab. 2, 3, 4 u. 5)

- ANOVA - Ansatz für 3 Stichproben (S1, S2, S3). Die Stichproben sind dummy-codiert in den Variablen X1 u. X2, als abhängige Variable ist Y vorgesehen (siehe Tab. 5)
- Es handelt sich um eine einfaktorische ANOVA.
- Die Daten für die Berechnung mittels Multiple lin. Regression

sind im File ANOVA1.txt und die Rechenergebnisse im File ANOVA1.ERG zu finden.

Graphisch übersichtlich sind die 3 Stichproben in Tab. 4 dargestellt und leicht zu vergleichen.

- Die Nullhypothese (H_0) der Analyse lautet, dass alle 3 Stichproben von ein- und derselben Grundgesamtheit stammen. Die H_0 bleibt, nach Interpretation der Rechenergebnisse, mit einer Irrtumswahrscheinlichkeit von $\approx 0,13126$ für das Globalmodell aufrecht (siehe File ANOVA1.ERG), d.h. es kann kein signifikanter Unterschied zwischen den Stichproben gefunden werden.

DESIGN-2: (Tab. 2, 3, 4 u. 6)

- Es handelt sich wieder um einen Ansatz für eine **emfaktorische ANOVA**, in diesem Fall allerdings für 4 Stichproben, u.z.B. S_1, S_2, S_3 u. S_4 , welche mit 3 Dummy Variablen, nämlich x_1, x_2 u. x_3 codiert sind (siehe Tab. 6).
- Der alleinige visuelle Vergleich in Tab. 4 macht schon sichtbar, dass Stichprobe S_4 von den übrigen Stichproben etwas abgesetzt ist.
- Die Daten für die Regressionsberechnung liegen im File ANOVA2.txt und die Ergebnisse der Berechnung im File ANOVA2.ERG

- Die Interpretation der Rechenergebnisse ergibt (siehe ANOVA2.ERG), dass die Nullhypothese auf Identität der Stichproben abzulehnen ist.
Es besteht mit einer Irrtumswahrscheinlichkeit von $< 0,0001$ für das Globalmodell höchstsignifikanter Unterschied zw. zumindest einer Stichprobe und den übrigen Stichproben hinsichtlich der Mittelwerte in d. Stichproben.

DESIGN-3:

(Tab. 2, 3, 4 u. insbes. 7, 8)

- Es handelt sich um den Ansatz für eine zweifaktorielle ANOVA samt Kreuz-(Interaktions)-termen.
- Es liegen 2 Stichproben, u. zw. S1, S2 u. S3, zugrunde. Die Stichproben sind in den Variablen x_1 und x_2 dummycodiert. Die Variablen x_1 u. x_2 codieren für den 1. Faktor.
- Zusätzlich wird im Ansatz angenommen, dass die Individuen der Stichproben noch nach den Geschlechtern, weiblich u. männlich, zu unterscheiden sind. Die Geschlechtszugehörigkeit soll in der Variable x_3 niedergelegt sein. x_3 codiert für den 2. Faktor.
- Primär kann man nicht ausschließen, dass Interaktionen zwischen Faktor-1 und Faktor-2 existieren. Um diesen Umstand Rechnung zu tragen, müssen die fog. Interaktionskome zwischen den Variablen der Faktoren untersucht werden.

Dafür werden Kreuzterme z.B. den Faktoren mit den Variablen X_4 und X_5 in den Ansatz aufgenommen.

X_4 entspricht d. Multiplikation des Variablenwertes von X_1 mit dem Wert von Variable X_3 für das jeweilige Individuum der Stichprobe. Ebenso wird X_5 im Wert ermittelt: als Multiplikation des Wertes von X_2 mit dem Wert von X_3 für das jeweilige Individuum.

- Die Daten für die Berechnungen befinden sich im File ANOVA3.txt, die Ergebnisse der Berechnungen im File ANOVA3.ERG
- Zur Interpretation d. Ergebnisse:
Für das Globalmodell ergibt sich ein signifikantes Ergebnis.
Die Schätzung der Schätzwerte für die partiellen Regressionskoeffizienten erhält jedoch subtile Einzelheiten der Analyse.
So sind die beiden Kreuzterme (d.h. b_5 und b_6 - sie stehen für X_4 u. X_5) nicht signifikant, d.h. dass sie offenbar keinen direkten Einfluss am Zustandekommen d. signifikanten Ergebnisses im Globalmodell haben.
Hingegen ist der Einfluss d. 2. Faktors (repräsentiert durch b_4 - für X_3 stehend) höchstsignifikant im Einfluss auf die Signifikanz d. Globalmodells.

PAARVERGLEICHE:

A5

(Tab. 9, 10 und 11)

Sehr oft ist es für den Einsatz in der Praxis der pharmazeutischen Forschung von entscheidender Bedeutung, herauszufinden, ob
① ein neu einzuführendes Medikament signifikante Wirkung erzielt oder ob ② ein Wirkstoff nur als Placebo zu bewerten ist. Für die Beantwortung d. Fragestellungen sind Paarvergleiche das vorrangige Mittel d. Wahl.

- Die allgemein gültigen Grundvoraussetzungen für den Ansatz eines Paarvergleiches bestehen darin, eine Stichprobe von Probanden zu finden, welche sich zweimal untersuchen lassen - u. z. B. einmal ohne Medikament - (Placebo-)Gabe und einmal mit Medikament - (Placebo-)Gabe.

WICHTIG IST, dass die Probanden NIE Bescheid wissen (um persönlichen Einfluss auszuschließen!), wann sie die Medikament- (Placebo-) Gabe erhalten und wann nicht. Es muss für jeden Probanden der Zufallszahl überlassen bleiben, ob ein Proband für den ersten Paaransatz schon eine Medikament- (Placebo-) Gabe erhält oder erst im zweiten Paaransatz.

Zu den Rechenbeispielen:

AG

DESIGNVORBEREITUNG: (Tab. 9)

- Es wurde zunächst eine Stichprobe $S5$ von Probanden erzeugt, die für den Fall stehen, dass keine Medikament - (Placebo-) Gabe erfolgt ist (siehe Tab. 9).
Diese Anfangsstichprobe geht mit der nunmehrigen Bezeichnung $S5a$ als Stichprobe d. unbehandelten Probanden in die Analysen ein.
- Demnach wurden für die identischen Stichprobenindividuen (Probanden) eine Stichprobe mit Placebogabe, d. i. $S5b$, und eine mit Medikament-Gabe, d. i. $S5c$, erzeugt.
- Bildlich sind die 3 Stichproben in Tab. 11 dargestellt.
Der visuelle Vergleich ergibt bereits, dass $S5a$ und $S5b$ (Farbe blau und rot) zeitgehend zusammenfallen, hingegen $S5c$ von den übrigen abhebt.

DESIGN - PLACEBO TEST: (Tab. 12)

- Um den Placeboeffekt abschätzen zu können, wurde folgender Ansatz nach dem ALM gewählt.
Die abhängige Variable Y steht für die Messwerte als Ergebnis d. Einfluss der Situationen, einmal mit und einmal ohne Placebogabe.

Die beiden Situationen sind durch Dummycodierung in der Variable x_1 niedergelegt.

Um den INDIVIDUALEFFEKT abschätzen zu können wird eine Variable x_2 als Kovariante in den Ansatz integriert.

WICHTIGE ANMERKUNG: Der Individualeffekt wird durch die individuelle Reaktion auf die beiden Situationen begründet. Das bedeutet, dass z.B. versucht werden kann, einen individuellen Effekt damit zu quantifizieren, indem die mittlere Reaktion eines Individuums auf beide Situationen berechnet wird.

(Siehe Tab. 12, Spalte H zur Erklärung)

- Die Daten für die Berechnung sind in File PAVPLA.TXT abgelegt. Die Ergebnisse d. Berechnungen finden sich in File PAVPLA.ERG
- Für die Interpretation der Ergebnisse sind die partiellen Regressionskoeffizienten heranzuziehen. Der partielle Regress.koeff. für den Individualeffekt (d.i. b_3) fällt höchstfüg. aus, was zu erwarten ist. Für die Frage, ob die Placebogabe einen sign. Einfluss hat, ist hingegen der part. Regr.koeffizient b_2 von entscheidender Bedeutung. Dieser Koeffizient stellt sich als nicht signifikant heraus, worauf die Hauptfrage d. Analyse beantwortet ist: Jene Substanz, von der man schon vorher vermutet hat, dass es sich um ein Placebo handelt, stellt sich als falsch heraus.

- In technisch ähnlicher Form wie beim Placebo-Test ist der nachstehende Ansatz strukturiert.
Die abhängige Variable Y repräsentiert die Maßwerte für die beiden Situationen d. Paralleltests - einmal mit und einmal ohne Medikamentgabe.
Die Codierung d. beiden Situationen ist in der Dummyvariable X_1 zugrundegelegt, die Kovariate für den Individual Effekt wird von X_2 repräsentiert.
- Die Daten für die Berechnung sind im File PAVMED.TXT zu finden. Die Ergebnisse d. Berechnungen finden sich im File PAVMED.ERG
- Für die Interpretation d. Ergebnisse sind wieder die partiellen Regressionskoeffizienten heranzuziehen.
Daß der Individual Effekt (d.i. b_3) höchstsign. ausfällt, ist gleichesum selbstredend. Hingegen zeigt auch der Koeffizient für die Medikamentgabe (d.i. b_2) höchst signif. Zusammenhang.
Daraus ergibt sich die Beantwortung d. Grundlegenden Frage:
Das verabreichte Medikament zeigt nachweislich Wirkung.

MACHTRAG:

A9

In Tab. 14-17 befindet sich (für Interessierte) noch die Darstellung eines alternativen Zuganges zu den Paarvergleichen durch schrittweises AUSPARTIALISIEREN des Individual effektes mittels einfach lin. Regression.