Varianzanalyse

Cooler Untertitel, den wir uns noch ausdenken

Henri Neumann & Robert Feldhans

15. Dezember 2016

Experimentelle Psychologie für Nichtpsychologen

Inhalt

1. Einführung

2. Prinzip der Varianzanalyse

3. Interaktion

Einführung

Einführung

Definition

Verfahren, welches die Wirkung einer (oder mehrerer) UV auf eine (oder mehrerer) AV untersucht.

- testet Unterschiede zw. Mittelwerten auf Signifikanz
- Einsatz bei mehr als 2 Stichproben
- Häufig auch als Globaltest bezeichnet

Grundbegriffe

- Zielvariable: abhängige Variable(AV)
- Faktor: unabhängige Variable(UV)
- Faktorstufen: Ausprägungen/Kategorien eines Faktors
- Effekt: Wirkung eines Faktors auf die AV
- Interaktionseffekt: kombinierte Wirkung zweier Faktoren auf die AV

Unterteilung

Abgrenzung anhand von Anzahl abhängige Variablen und Faktoren

Zahl der AVn	Zahl der UVn	Bezeichnung
1	1	Einfaktorielle VA
1	2	Zweifaktorielle VA
1	3	Dreifaktorielle VA
	usw.	
≥ 2	≥ 1	Multivariante VA

Vorraussetzungen

- Fehlerkomponenten sind normalverteilt
- Fehlervarianzen homogen in den Faktorstufen
- Messwerte bzw. Faktorstufen sind unabhängig voneinander

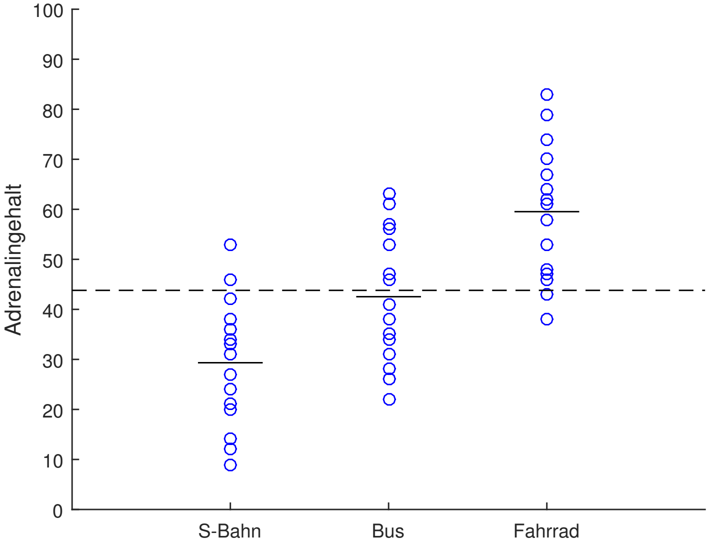
Prinzip der Varianzanalyse

Die gesamte Varianz der AV wird aufgeteilt in:

- Varianz *zwischen* Gruppen:
  Abweichung der Gruppenmittelwerte vom Gesamtmittelwert
  = systematische Varianz
- Varianz *innerhalb* von Gruppen:
  Abweichung einzelner Messwerte vom Gruppenmittelwert
  = unsystematische Varianz, Fehlervarianz

$\Rightarrow$ anschließend Vergleich der Varianzschätzungen

Beispiel



100
90
80
70
60
50
40
30
20
10
0
Adrenalingehalt
S-Bahn
Bus
Fahrrad

Mathematisches Modell

Allgemeines Modell der einfaktoriellen Varianzanalyse

Zielvariable

Effekt der Faktorstufe i
(systematische Varianz)

$$y_k = \mu + \alpha_k + \varepsilon_k$$

Verbleibender Fehler
(Fehlervarianz)

Mittelwert der
Gesamtstichprobe

Anzahl der Faktorstufen des
betrachteten Faktors

$$k = 1, \ldots, K, n = 1, \ldots, N_k$$

Stichprobenumfänge für
die einzelnen Faktorenstufen

Mathematisches Modell

Allgemeines Modell der zweifaktoriellen Varianzanalyse

Effekt der Faktorstufe i
(systematische Varianz A)

Zielvariable

$$y_{jk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{jk}$$

Verbleibender Fehler
(Fehlervarianz)

Mittelwert der
Gesamtstichprobe

Interaktionseffekt zw. der Stufe j
UVA und der Stufe k UVB
(systematische Varianz AxB)

Hypothesen

einfaktoriell

- Nullhypothese:
  Alle Mittelwerte sind gleich oder alle Effekte α_k sind 0.
  Formal: $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ oder $\sum \alpha_k^2 = 0$

- Alternativhypothese:
  Nicht alle Mittelwerte sind gleich oder mindestens ein Effekt α_i ist ungleich Null.
  Formal: $H_1 : \sum(\mu_k - \mu)^2 > 0$ oder $\sum \alpha_k^2 > 0$

Hypothesen

zweifaktoriell

Für jeden Faktor wird eine Nullhypothese überprüft

- Faktor A:
  Alle Zeilenmittelwerte sind gleich oder alle Effekte α_j sind 0.
  Formal: $H_0 : \mu_{1\cdot} = \mu_{2\cdot} = \cdots = \mu_{J\cdot}$ oder $\sum \alpha_j^2 = 0$
- Faktor B:
  Alle Spaltenmittelwerte sind gleich oder alle Effekte β_k sind 0.
  Formal: $H_0 : \mu_{\cdot 1} = \mu_{\cdot 2} = \cdots = \mu_{\cdot K}$ oder $\sum \beta_k^2 = 0$
- Interaktion AB:
  Die Wirkung der einzelnen UVn auf die AV ist voneinander abhängig.
  Formal: $H_0 : \bar{y}_{jk} = \mu_{\cdot k} + \mu_{j\cdot} - \mu + \varepsilon$

Zweifaktorielle Varianzanalyse

		UV B				Zeilenmittel
		B_1	B_2	$\ldots$	B_K	HE A
UV A	A_1	μ_{11}	μ_{12}	$\ldots$	μ_{1K}	$\mu_{1\cdot}$ $= \mu + \alpha_1$
	A_2	μ_{21}	$\ldots$	$\ldots$	$\ldots$	$\mu_{2\cdot}$ $= \mu + \alpha_2$
	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\mu_{j\cdot}$ $= \mu + \alpha_j$
	A_J	μ_{J1}	$\ldots$	$\ldots$	μ_{JK}	$\mu_{J\cdot}$ $= \mu + \alpha_J$
Spalten-mittel	HE B	$\mu_{\cdot 1}$ $= \mu + \beta_1$	$\mu_{\cdot 2}$ $= \mu + \beta_2$	$\mu_{\cdot k}$ $= \mu + \beta_k$	$\mu_{\cdot K}$ $= \mu + \beta_K$	μ

Prinzip der Varianzanalyse

Prinzip der Varianzanalyse

Die gesamte Varianz der AV wird aufgeteilt in:

- Varianz *zwischen* Gruppen:
  Abweichung der Gruppenmittelwerte vom Gesamtmittelwert
  = systematische Varianz
- Varianz *innerhalb* von Gruppen:
  Abweichung einzelner Messwerte vom Gruppenmittelwert
  = unsystematische Varianz, Fehlervarianz

$\Rightarrow$ anschließend Vergleich der Varianzschätzungen

Summe der Abweichungsquadrate

Repräsentiert die Unterschiedlichkeit der Werte der AV. Drei relevante Formen:

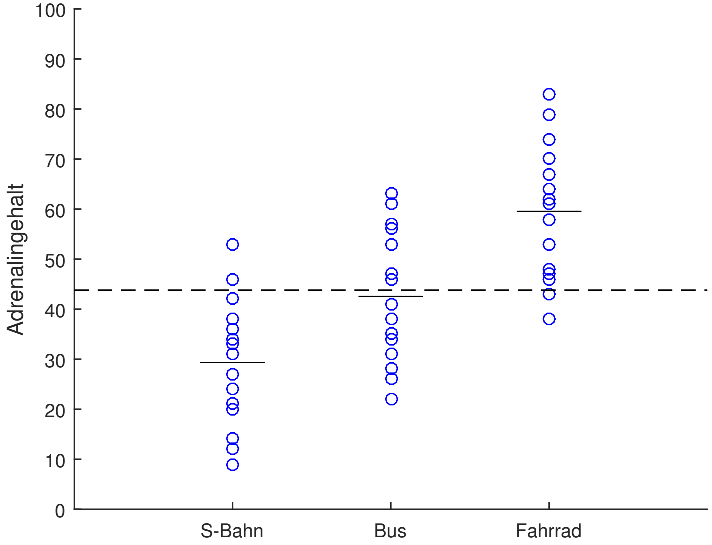
- SAQ_{Gesamt}: Die Gesamtvariabilität.
  Formal: $SAQ_{Gesamt} = \sum(y - \bar{y})^2$

- SAQ_{Effekt}: auch $SAQ_{zwischen}$; Variabilität zwischen Bedingungen.
  Formal: $SAQ_{Effekt} = n_k \sum_{k=1}^{K}(\bar{y}_k - \bar{y})^2$

- SAQ_{Fehler}: auch $SAQ_{innerhalb}$; Variabilität innerhalb einer Bedingung.
  Formal: $SAQ_{Fehler} = \sum(y - \bar{y}_k)^2$

Es gilt $\mathbf{SAQ_{Gesamt} = SAQ_{Effekt} + SAQ_{Fehler}}$

Beispiel



100
90
80
70
60
50
40
30
20
10
0
Adrenalingehalt
S-Bahn
Bus
Fahrrad

Freiheitsgrade (FG)

Anzahl der frei variierbaren Werte oder auch
Anzahl der in die SAQ eingehenden Werte

- SAQ_{Gesamt}: $N - 1$
- SAQ_{Effekt}: $K - 1$ K: Anzahl Faktorstufen
- SAQ_{Effekt}: $K \cdot (n - 1)$

Es gilt $\mathbf{FG_{Gesamt} = FG_{Effekt} + FG_{Fehler}}$

Mittlere Quadratsumme (MQ)

- Die mittlere Quadratsumme entspricht der Varianz
- MQ_{Fehler}: Schätzung der Populationsvarianz
- MQ_{Effekt}: Schätzung der Populationsvarianz wenn H_0 gilt
- Mittlere Quadratsummen sind **nicht** additiv

$$MQ = \frac{SAQ}{FG}$$

Bedeutung der MQ

- $MQ_{Effekt} = MQ_{Fehler}$: H_0 ist gültig
- $MQ_{Effekt} \gg MQ_{Fehler}$: H_0 ist ungültig, MQ_{Effekt} enthält systematische Varianz

Aber: Wann ist MQ_{Effekt} überzufällig größer als MQ_{Fehler}?

Bedeutung der MQ

- $MQ_{Effekt} = MQ_{Fehler}$: H_0 ist gültig
- $MQ_{Effekt} \gg MQ_{Fehler}$: H_0 ist ungültig, MQ_{Effekt} enthält systematische Varianz

Aber: Wann ist MQ_{Effekt} überzufällig größer als MQ_{Fehler}?

$\Rightarrow$ Prüfen mit F-Verteilung

$$F = \frac{MQ_{Effekt}}{MQ_{Fehler}}, FG = K - 1, K(n-1)$$

F-Verteilung

Wenn $F_{empirisch} > F_{kritisch} \Rightarrow$ Ablehnung von H_0

n \ m	1	2	3	4	5	6	7	8	9	10	12	15	20	30	40	50
1	161	200	216	225	230	234	237	239	241	242	244	246	248	250	251	252
2	18,5	19,0	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,62	8,59	8,58
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,75	5,72	5,70
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,50	4,46	4,44
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,81	3,77	3,75
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,58	3,51	3,44	3,38	3,34	3,32
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,08	3,04	3,02
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,86	2,83	2,80
10	4,96	4,10	3,71	3,48	3,33	3,22	3,13	3,07	3,02	2,98	2,91	2,85	2,77	2,70	2,66	2,64
11	4,84	3,98	3,59	3,36	3,20	3,10	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,57	2,53	2,51
12	4,75	3,88	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,47	2,43	2,40
13	4,67	3,81	3,41	3,18	3,02	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,38	2,34	2,31
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,31	2,27	2,24
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,25	2,20	2,18
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,19	2,15	2,12
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,15	2,10	2,08
18	4,41	3,56	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,11	2,06	2,04
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,15	2,07	2,03	2,00
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,04	1,99	1,97
21	4,33	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,01	1,97	1,94
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	1,98	1,94	1,91
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,38	2,32	2,27	2,20	2,13	2,05	1,96	1,91	1,88
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,35	2,30	2,25	2,18	2,11	2,03	1,94	1,89	1,86
25	4,24	3,38	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,92	1,87	1,84

Interaktion

Was ist Interaktion

- blablalba

Verschiedene Arten der Interaktion

- Nullinteraktion
- ordinale Interaktion
- disordinale Interaktion
- semidisordinale Interaktion

Nullinteraktion I

- keine Interaktion
- Auswirkungen einer UV sind auf allen Stufen der anderen UV gleich
- Beispiel
- UVs wirken unabhängig voneinander auf die AV
- Die Kenntniss der wirkung beider UVs reicht aus, um den Mittelwert jeder Zelle voraussagen zu können

ordinale Interaktion I

- blablalba

disordinale Interaktion I

- blablalba

semidisordinale Interaktion I

- blablalba