

Weitere varianzanalytische Designs

MEHRFAKTORIELLE VA

Das Prinzip der Varianzanalyse kann auch auf mehr als zwei Faktoren angewendet werden. Allerdings steigt die Zahl der zu überprüfenden Varianzen schnell sehr stark an: Bei drei Faktoren sind schon sieben(!) F -Tests durchzuführen (Haupteffekte A , B und C sowie Wechselwirkungen $A \times B$, $A \times C$, $B \times C$ und $A \times B \times C$). Bei vier Faktoren wären es bereits 15 Tests.

Darüber hinaus steigt mit jedem zusätzlichen Faktor die Zahl der Kombinationen der verschiedenen Faktorstufen, was besonders beim Erheben der Daten einen enormen Mehraufwand bedeuten kann.

UNVOLLSTÄNDIGE VARIANZANALYTISCHE DESIGNS

Aus diesen Gründen verzichtet man bei einer größeren Zahl von Faktoren häufig auf das exakte Untersuchen der Wechselwirkungen und greift auf sogenannte *unvollständige varianzanalytische Designs* zurück. Oft ist man aber auch aufgrund äußerer Umstände nicht in der Lage, ein vollständiges Design zu verwenden, da nicht alle Kombinationen aller Faktorstufen realisierbar sind.

Hier seien exemplarisch die *nested factors* und das *lateinische Quadrat* vorgestellt:

nested factors

Bei den nested factors (auch: *geschachteltes Design*, *hierarchisches Design*) werden - meist weil es die Umstände nicht anders zulassen - nur für bestimmte Kombinationen von Faktorstufen Daten ermittelt. Dadurch, dass für einen Faktor jede Stufe nur mit genau einer Stufe des anderen Faktors kombiniert wird, entsteht eine hierarchische Gliederung.

Beispiel:

Zwei (spezielle) Behandlungsmethoden (Faktor A) für eine bestimmte Krankheit sollen verglichen werden. Diese kommen in verschiedenen Spitälern (Faktor B) zur Anwendung, wobei jedes Krankenhaus nur für eine der Methoden ausgerüstet ist. In Spital B_1 und Spital B_2 wird nur Methode A_1 angewandt, in den Spitälern B_3 , B_4 und B_5 nur A_2 :

A_1		A_2		
B_1	B_2	B_3	B_4	B_5
M	e	s	s	-
w	e	r	t	e

Auswertung:

Wechselwirkungen können hier nicht untersucht werden. Außerdem können die Haupteffekte eines Faktors nur auf die Stufen des anderen Faktors aufgeteilt ermittelt werden (im Bsp. muss der Haupteffekt von B geschachtelt in A ausgewertet werden).

Lateinisches Quadrat

Ein lateinisches Quadrat ist ein $n \times n$ -Feld, das mit n verschiedenen Elementen so aufgefüllt wird, dass jedes Element in jeder Zeile und in jeder Spalte genau einmal vorkommt. Es gibt (außer bei $n = 1$) stets mehrere Möglichkeiten, wie die Elemente angeordnet werden.

Beispiel für ein 4×4 - LQ:

♥	♠	♦	♣
♣	♥	♠	♦
♦	♣	♥	♠
♠	♦	♣	♥

Es wird zur Herstellung eines (möglichst brauchbaren) unvollständigen varianzanalytischen Designs verwendet. Zweck ist, den experimentellen Aufwand zu verringern. Anstatt für jede Kombination von Faktorstufen Messwerte zu ermitteln, werden nur jene Kombinationen, die sich durch eines der möglichen lateinischen Quadrate ergeben, verwendet.

Voraussetzungen:

- 1.) Man hat genau 3 Faktoren.
- 2.) Die Faktoren müssen in gleich viele Stufen unterteilt sein.

Ein mögliches 3×3 - LQ (je 3 Faktorstufen):

	B_1	B_2	B_3
A_1	C_1	C_2	C_3
A_2	C_3	C_1	C_2
A_3	C_2	C_3	C_1

Hier ergeben sich also 9 Kombinationen, für die Werte ermittelt werden müssen. Bei 30 Werten pro Stichprobe sind das 270 Messwerte. Zum Vergleich: Beim entsprechenden vollständigen Design gäbe es 27 Kombinationen, man bräuchte also bei gleicher Stichprobengröße 810 Werte.

Wie bei einer vollständigen VA werden die Varianzen geschätzt und per F -Test mit der Schätzung der Fehlervarianz verglichen. Neben den Schätzungen für die drei Haupteffekte und den Messfehler wird auch die *Residualvarianz* geschätzt, die einer teilweisen Schätzung der Wechselwirkungen entspricht.

Sie wird ebenfalls mit der Fehlervarianz verglichen. Ist sie klein genug (\rightarrow F -Test nicht signifikant), dann kann man davon ausgehen, dass die drei Faktoren unabhängig voneinander wirken. Ist sie aber signifikant größer als die Fehlervarianz, dann sind mit großer Wahrscheinlichkeit Wechselwirkungen vorhanden. In diesem Fall ist die Verwendung des LQ problematisch.