Varianzanalyse* (1)

– Einfaktorielle Varianzanalyse (I) –

Die Varianzanalyse (ANOVA = ***AN****alysis* ***O****f* ***VA****riance*) wird benutzt, um Unterschiede zwischen Mittelwerten von drei oder mehr Stichproben auf Signifikanz zu prüfen. Z.B. lautet die Nullhypothese für fünf Mittelwerte:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

Die Alternativhypothese ist, dass H_0 **als Ganzes** nicht wahr ist, d.h. es bestehen **irgendwelche** Unterschiede zwischen den Mittelwerten.

Eine ANOVA basiert auf der Tatsache, dass anhand der Stichprobendaten zwei unhängige Schätzungen der Populationsvarianz möglich sind:

1. **Geschätzte Varianz innerhalb der Gruppen (*within*-Varianz, Fehlervarianz, "*error*")**: Wie unterscheiden sich die einzelnen ***Werte*** in einer Stichprobe (oder Gruppe) von den übrigen Werten in der gleichen Gruppe?
2. **Geschätzte Varianz zwischen den Gruppen (*between*-Varianz)**: Wie unterscheiden sich die ***Mittelwerte*** verschiedener Stichproben (oder Gruppen) voneinander?

Wenn alle Stichproben der gleichen normalverteilten Population entstammen, sind die Schätzungen der Varianzen zwischen und innerhalb der Gruppen ähnlich (und ähnlich der Populationsvarianz). ***Je größer die Varianz zwischen den Gruppen im Vergleich zur Varianz innerhalb der Gruppen*** ist, desto wahrscheinlicher ist es, dass die Stichproben ***nicht*** aus einer Population mit gleichen Mittelwerten stammen.

*) Angelehnt an: Welkowitz, J., Ewen, R.B. & Cohen, J. (2002). Introductory statistics for the behavioral sciences. New York (5th ed.): Academic Press.

Varianzanalyse (2)

– Einfaktorielle Varianzanalyse (II) –

a) Die Variation **zwischen** den Gruppen und **innerhalb** der Gruppen ist ähnlich groß:

	A	B	C	D	E
	16	16	16	16	14
	16	14	14	14	13
	14	13	12	12	13
	13	13	10	12	10
	12	10	10	12	10
Mittelwerte:	14.2	13.2	12.4	13.2	12.0

Variation innerhalb der Gruppen

Variation zwischen den Gruppen

b) Die Variation **zwischen** den Gruppen ist viel größer als **innerhalb** der Gruppen:

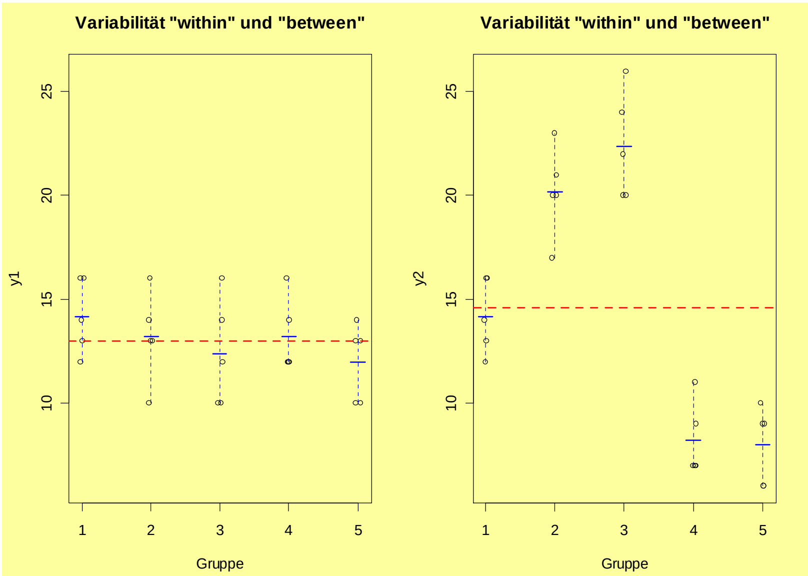
	A	B	C	D	E
	16	23	26	11	10
	16	21	24	9	9
	14	20	22	7	9
	13	20	20	7	6
	12	17	20	7	6
Mittelwerte:	14.2	20.2	22.4	8.2	8.0

Variation innerhalb der Gruppen

Variation zwischen den Gruppen

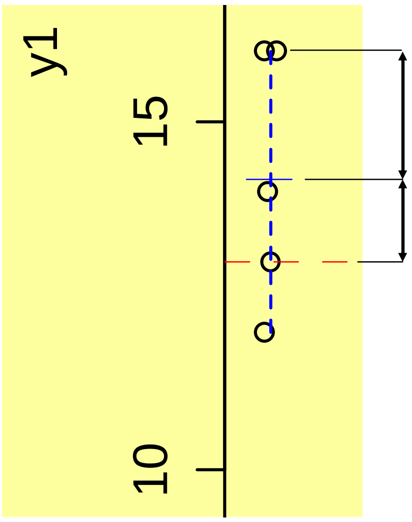
Varianzanalyse (3)

– Einfaktorielle Varianzanalyse (III) –



Varianzanalyse (4)

– Einfaktorielle Varianzanalyse (IV) –



16.0	
Innerhalb-Gruppen Abweichung ("error")	= **1.8**
14.2	**+**
Zwischen-Gruppen Abweichung	= **1.2**
13.0	-------
Totale Abweichung	= **3.0**

Varianzanalyse (5)

– Einfaktorielle Varianzanalyse (V) –

Die ***totale*** Varianz, die ***zwischen-***Gruppen Varianz und ***innerhalb***-Gruppen Varianz drücken die Größe dieser Abweichungen für die Personen z.B. innerhalb eines Experiments aus.

Die Summe der quadrierten Abweichungen der individuellen Werte vom Gesamtmittelwert	SS_T ("**s**ums of **s**quares **t**otal")
	=
Summe der quadrierten Abweichungen der Gruppenmittelwerte vom Gesamtmittelwert ("zwischen Gruppen")	SS_B ("**s**ums of **s**quares **b**etween groups")
	+
Summe der quadrierten Abweichungen der individuellen Werte vom jeweiligen Gruppenmittelwert ("innerhalb Gruppen")	SS_W ("**s**ums of **s**quares **w**ithin groups")

Die Schätzung der ***zwischen-Gruppen Varianz*** schließt ***sowohl*** die ***Effekte des Treatments als auch*** die ***Fehlervarianz*** ein: Die Gruppenmittelwerte sind sowohl von den Gruppenunter-schieden (Treatment) als auch von der Varianz der AV (oder Fehlervarianz der beobachteten Werte) beeinflusst. Die ***innerhalb-Gruppen Varianz*** drückt jedoch ***nur*** die ***Fehlervarianz*** aus.

Varianzanalyse (6)

– Einfaktorielle Varianzanalyse (VI) –

$$F = \frac{Treatment\ Varianz + Fehler\ Varianz}{Fehler\ Varianz}$$

$$= \frac{gesch.\ zwischen\ Gruppen\ Varianz}{gesch.\ innerhalb\ Gruppen\ Varianz}$$

Ist die **H_0 wahr**, gibt es keine Treatment-Varianz. Die zwischen-Gruppen Varianz und die inner-halb-Gruppen Varianz werden ungefähr gleich sein, so dass ***F* ungefähr 1.0** ist. Um so mehr ***F*** größer als **1.0** ist, um so sicherer kann man sein, dass das Treatment einen Effekt auf die AV hat bzw. dass die Gruppen nicht aus der gleichen Population stammen.

Für einen *F*-Wert ***sehr*** viel kleiner 1.0 (z.B. 0.2) gäbe es keine offensichtliche Interpretation außerhalb des Zufalls (oder einer fehlerhaften Annahme der Voraussetzungen des Tests). In diesen Fällen würde H_0 beibehalten und die Möglichkeit, dass sich ein systematischer Faktor (z.B. eine nicht zufällige Stichprobenziehung) eingeschlichen hat, sollte geprüft werden.

Varianzanalyse (7)

– Einfaktorielle Varianzanalyse (VII) –

Quadratsummen (sums of squares):

$SS_T = \sum (x - \bar{x})^2$ | $SS_T = (16 - 13.0)^2 + \ldots + (10 - 13.0)^2 = 100$

$SS_B = \sum N_G \cdot (\bar{x}_G - \bar{x})^2$ | $SS_B = 5 \cdot (14.2 - 13.0)^2 + \ldots + 5 \cdot (12.4 - 13.0)^2 = 14.4$

$SS_W = \sum (x_1 - \bar{x}_1)^2 + \ldots + (x_k - \bar{x}_k)^2$ | $SS_W = (16 - 14.2)^2 + \ldots + (10 - 12.0)^2 = 85.6$

$SS_W = SS_T - SS_B$ | $SS_W = 100 - 14.4 = 85.6$

Varianzen (mean squares):

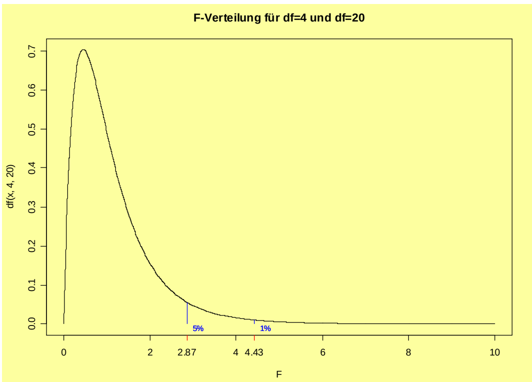
$df_B = k - 1$ $\quad MS_B = \frac{SS_B}{df_B}$ | $df_B = 5 - 1 = 4$ $\quad MS_B = \frac{14.4}{4} = 3.60$

$df_W = N - k$ $\quad MS_W = \frac{SS_W}{df_W}$ | $df_W = 25 - 5 = 20$ $\quad MS_W = \frac{85.6}{20} = 4.28$

***F*-Ratio:** $F = \frac{MS_B}{MS_W}$ | $F = \frac{3.60}{4.28} = 0.84$

Varianzanalyse (8)

– Einfaktorielle Varianzanalyse (VIII) –



Beispiel 1:

Analysis of Variance Table

	SS	df	MS	F	p
Between	14.40	4	3.60	0.841	0.515
Within (error)	85.60	20	4.28		

Beispiel 2:

Analysis of Variance Table

	SS	df	MS	F	p
Between	884.40	4	221.10	51.659	< 0.001
Within (error)	85.60	20	4.28		

Varianzanalyse (9)

– Einfaktorielle Varianzanalyse (IX) –

Wenn der *F*-Wert statistisch signifikant ist, kann mit sogenannten ***post-hoc*** **Tests** (Einzelvergleichen) untersucht werden, ***welche*** der Mittelwerte sich signifikant unterscheiden.

Hierbei dürfen aber einfache *t*-Tests ***nicht*** benutzt werden, da mit wachsender Anzahl von Tests die Wahrscheinlichkeit, einen α-Fehler (Typ I Fehler) zu begehen, steigt!

Daher kommen hierfür nur "geschützte" Tests in Frage. Zu empfehlen ist im Allgemeinen ein Tukey-Test (Tukey HSD = Tukey's honest significant difference) oder Tukey-b.

Mit derartigen Tests können auch Konfidenzintervalle für die jeweiligen Mittelwertsdifferenzen berechnet werden.

Anhand des *F*-Wertes und der Freiheitsgrade kann auch ein Maß ε für die **Stärke des Zusammenhangs** berechnet werden. ε kann wie ein Korrelationskoeffizient interpretiert werden. Die Berechnung macht allerdings nur Sinn für *F*-Werte > 1.0:

$$\varepsilon = \sqrt{\frac{df_B \cdot (F-1)}{df_B \cdot F + df_W}}$$

Beispiel 2: $$\varepsilon = \sqrt{\frac{4 \cdot (51.659 - 1)}{4 \cdot 51.659 + 20}} = .946$$

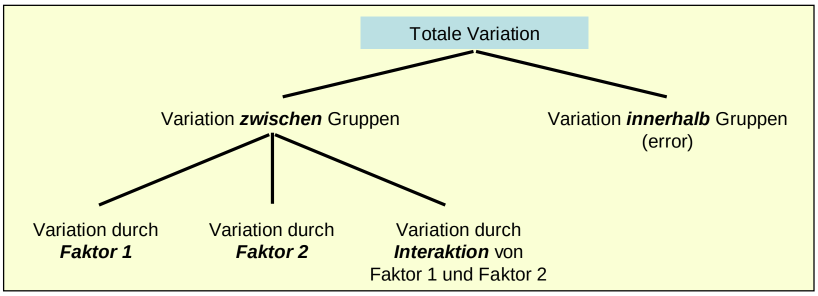
In diesem (fiktiven) Beispiel wäre der Zusammenhang extrem hoch. Dieses Maß würde einer Effektstärke *d* von etwa 5.8 entsprechen (hierbei würde unterstellt, dass Gruppe E (5) die Experimentalgruppe und Gruppe B (2) die Kontrollgruppe darstellten und die übrigen Gruppen ein anderes Treatment bekommen hätten).

Varianzanalyse (10)

– Zweifaktorielle Varianzanalyse (I) –

Mit einer **zweifaktoriellen** Varianzanalyse kann untersucht werden, welchen Einfluss zwei Faktoren (z.B. Faktor 1: Experimental- vs. Kontrollgruppe, Faktor 2: Geschlecht) auf eine abhängige Variable haben.

Hierbei wird die ***Varianz zwischen den Gruppen*** weiter in ***drei Bestandteile*** zerlegt: In Varianz, die auf Faktor 1 zurückzuführen ist (Haupteffekt 1), in Varianz, die auf Faktor 2 zurückzuführen ist (Haupteffekt 2) sowie in Varianz aufgrund der Interaktion von Faktor 1 und 2 (Interaktionseffekt):



Hierbei ist die totale Quadratsumme gleich der Summe der Quadratsummen aller vier Bestandteile: (1) innerhalb Gruppen Quadratsumme, (2) Quadratsumme Faktor 1, (3) Quadratsumme Faktor 2, und (4) Quadratsumme Interaktion Faktor1 und Faktor 2.

Varianzanalyse (11)

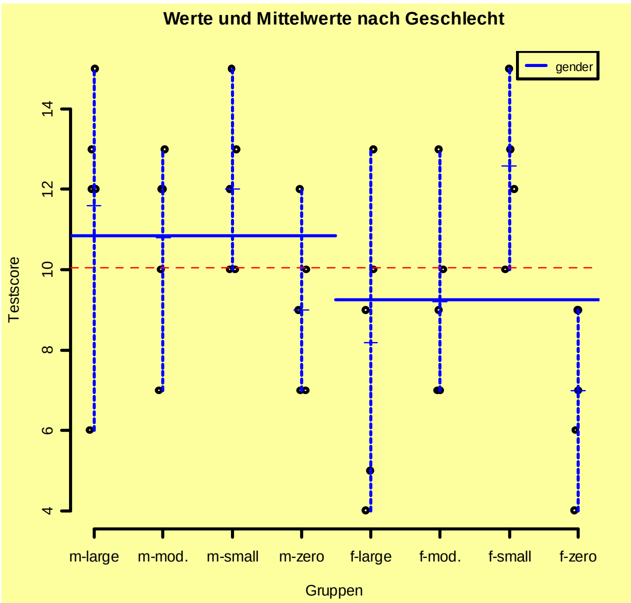
– Zweifaktorielle Varianzanalyse (II) –

Beispiel: In einem fiktiven Experiment wird untersucht, wie der Konsum von Koffein die Leistung in einem Englischtest beeinflusst. Dabei ergeben sich die folgenden fiktiven Messwerte.

		Koffein-Dosis (Faktor 1)				
		groß	*mäßig*	*klein*	*Null*	*Zeilen-mittel*
Geschlecht (Faktor 2)	*männlich*	6	12	10	9	
		15	10	13	10	
		12	12	15	7	10.85
		12	13	12	12	
		13	7	10	7	
	weiblich	10	9	12	4	
		13	7	13	7	
		4	10	15	6	9.25
		9	7	10	9	
		5	13	13	9	
Spalten-mittel		9.9	10.0	12.3	8.0	10.50

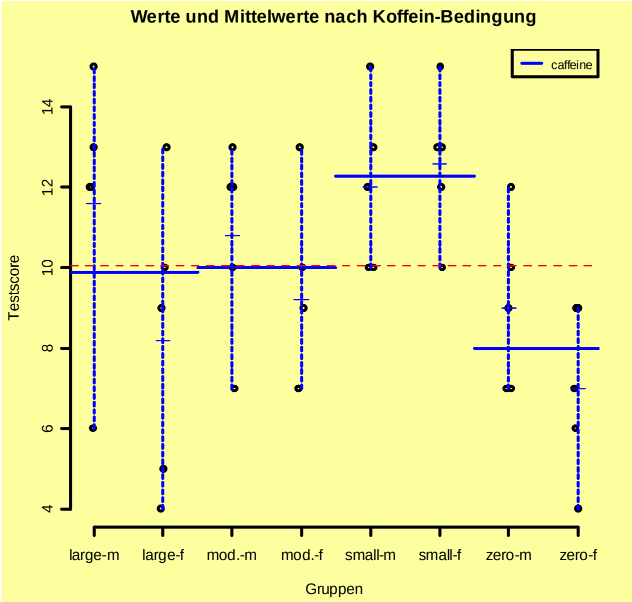
Varianzanalyse (12)

– Zweifaktorielle Varianzanalyse (III) –



Varianzanalyse (13)

– Zweifaktorielle Varianzanalyse (IV) –



Varianzanalyse (14)

– Zweifaktorielle Varianzanalyse (V) –

Für die Berechnung benötigt man außer dem Gesamtmittelwert sowie den Zeilen- und Spaltenmittelwerten noch die Mittelwerte der Zellen:

		Koffein-Dosis (Faktor 1)			
		groß	mäßig	klein	Null
Geschlecht (Faktor 2)	männlich	11.6	10.8	12.0	9.0
	weiblich	8.2	9.2	12.6	7.0

Quadratsummen (sums of squares):

$SS_T = \sum (x - \bar{x})^2$ | $SS_T = (6 - 10.05)^2 + ... + (9 - 10.05)^2 = 353.9$

$SS_B = \sum N_G \cdot (\bar{x}_G - \bar{x})^2$ | $SS_B = 5 \cdot (11.6 - 10.05)^2 + ... + 5 \cdot (12.6 - 10.05)^2 = 139.1$

$SS_W = SS_T - SS_B$ | $SS_W = 353.9 - 139.1 = 214.8$

$SS_{F1} = \sum N_{F1} \cdot (\bar{x}_{F1} - \bar{x})^2$ | $SS_{F1} = 10 \cdot (9.9 - 10.05)^2 + ... + 10 \cdot (8.0 - 10.05)^2 = 92.9$

$SS_{F2} = \sum N_{F2} \cdot (\bar{x}_{F2} - \bar{x})^2$ | $SS_{F2} = 20 \cdot (10.85 - 10.05)^2 + 20 \cdot (9.25 - 10.05)^2 = 25.6$

$SS_{F1 \cdot F2} = SS_B - SS_{F1} - SS_{F2}$ | $SS_{F1 \cdot F2} = 139.1 - 92.9 - 25.6 = 20.6$

Varianzanalyse (15)

– Zweifaktorielle Varianzanalyse (VI) –

Für die Berechnung der Varianzen (*mean squares*) werden die Freiheitsgrade benötigt, die analog der Quadratsummen zerlegt werden:

- total $df = N - 1 = \mathbf{39}$
  - df zwischen Gruppen = $k - 1 = 7$
    - df Koffein = $4 - 1 = \mathbf{3}$
    - df Geschlecht = $2 - 1 = \mathbf{1}$
    - df Interaktion = $3 \cdot 1 = \mathbf{3}$
  - df innerhalb Gruppen = $N - k = \mathbf{32}$

Varianzen (mean squares):

$$MS_W = \frac{SS_W}{df_W} \qquad MS_W = \frac{214.8}{32} = 6.71 \qquad MS_{F1} = \frac{SS_{F1}}{df_{F1}} \qquad MS_{F1} = \frac{92.2}{3} = 30.97$$

$$MS_{F2} = \frac{SS_{F2}}{df_{F2}} \qquad MS_{F2} = \frac{25.6}{1} = 25.6 \qquad MS_{F1 \cdot F2} = \frac{SS_{F1 \cdot F2}}{df_{F1 \cdot F2}} \qquad MS_{F1 \cdot F2} = \frac{20.6}{3} = 6.87$$

Varianzanalyse (16)

– Zweifaktorielle Varianzanalyse (VII) –

***F*-Ratios:**

Koffein $F = \frac{MS_{F1}}{MS_W}$ $\qquad F = \frac{30.97}{6.71} = 4.62$

Geschlecht $F = \frac{MS_{F2}}{MS_W}$ $\qquad F = \frac{25.60}{6.71} = 3.82$

Interaktion $F = \frac{MS_{F1 \cdot F2}}{MS_W}$ $\qquad F = \frac{6.87}{6.71} = 1.02$

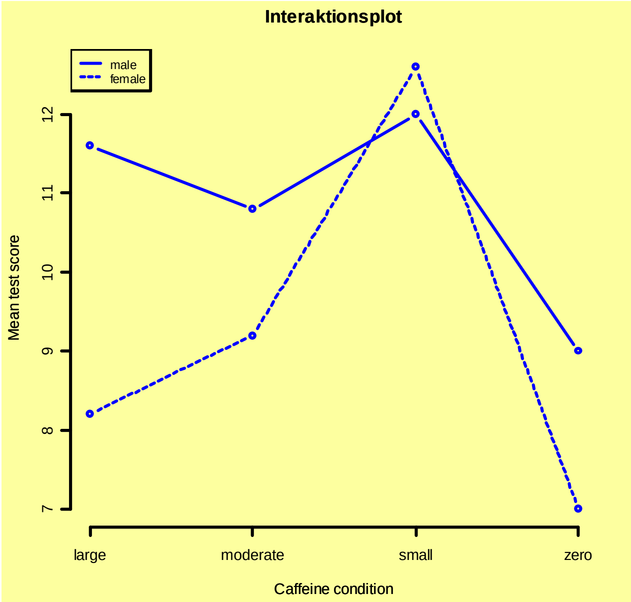
ANOVA-Tabelle:

Analysis of Variance Table

	SS	df	MS	F	p
Koffein	92.9	3	30.97	4.61	.009
Geschlecht	25.6	1	25.60	3.81	.059
Interaktion	20.6	3	6.87	1.02	.395
Within (error)	214.8	32	6.71		

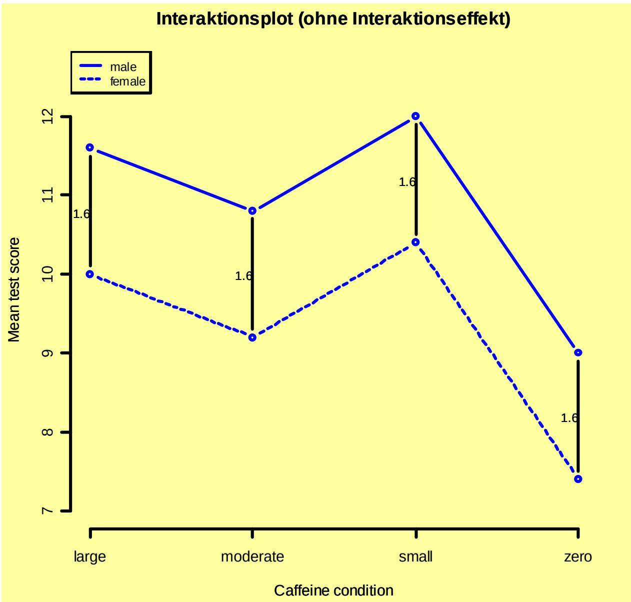
Varianzanalyse (17)

– Zweifaktorielle Varianzanalyse (VIII) –



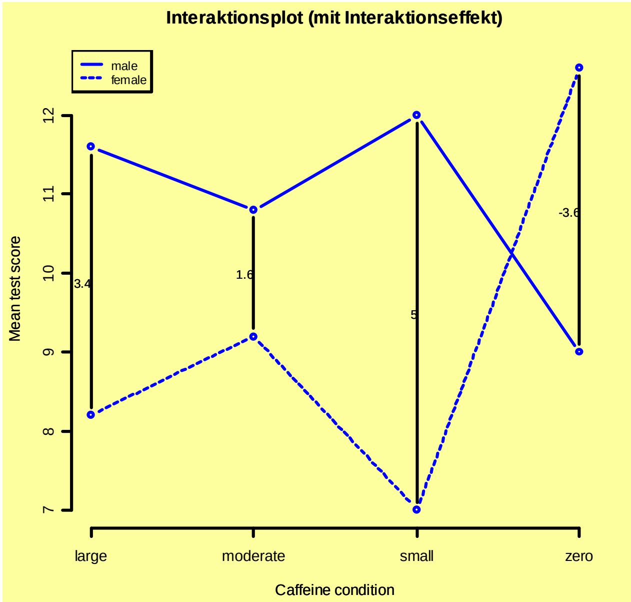
Varianzanalyse (18)

– Zweifaktorielle Varianzanalyse (IX) –



Varianzanalyse (19)

– Zweifaktorielle Varianzanalyse (X) –

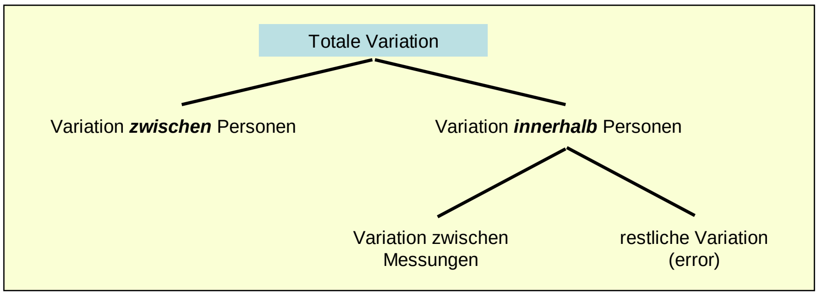


Varianzanalyse (20)

– Einfaktorielle Varianzanalyse mit Messwiederholung –

Bei den bisherigen Varianzanalysen handelte es sich um ***unabhängige Messungen***. Wird eine Variable zwei- oder mehrmals gemessen, gibt es also **abhängige Messungen**, können die ***Veränderungen*** im Rahmen einer **Varianzanalyse mit einem Messwiederholungsfaktor** analysiert werden.

Bei einer einfaktoriellen Varianzanalyse mit Messwiederholung werden die Quadratsummen in Quadratsummen zwischen und innerhalb ***Personen*** zerlegt. Letzere wird in die Quadratsumme zwischen den Messungen und einen Rest (Residuum oder Fehlervarianz) zerlegt:



Der *F*-Wert zur Prüfung signifikanter Veränderungen wird aus dem Verhältnis der Varianz (*mean squares*) der Messungen und der Fehlervarianz gebildet, mit *df* Messungen = $p - 1$ und *df* Fehlervarianz = $(N - 1) \cdot (p - 1)$.

Varianzanalyse (21)

– Zweifaktorielle Varianzanalyse mit Messwiederholung (I) –

Bei Varianzanalysen mit Messwiederholung sind weitere Faktoren denkbar. Ein typisches Beispiel ist ein **Experiment** mit einer Experimental- und Kontrollgruppe (Faktor 1) und einer oder mehreren **Vor- und Nachmessung(en)** (Faktor 2). Hier lässt sich Wirkung des Treatments mittels einer **zweifaktoriellen Varianzanalyse** untersuchen, wobei ein Faktor den Messwiederholungsfaktor darstellt.

Im Gegensatz zu Varianzanalysen, bei denen nur eine Messung pro Person vorliegt, genügt es hierbei nicht, dass sich ein signifikanter Unterschied zwischen Experimental- und Kontrollgruppe (Haupteffekt) zeigt. Falls das Treatment ***wirksam*** ist, müssen sich die **Veränderungen** der Personen in Experimental- und Kontrollgruppe (in bestimmter Weise) unterscheiden. Das wird in einem ***signifikaten Interaktionseffekt*** zwischen dem Messwiederholungs- und dem Treatment-faktor sichtbar.

Handelt es sich dabei nur um **eine Vor- und eine Nachmessung** (hat der Messwiederholungs-faktor also nur zwei Stufen), ist die statistische Prüfung eines signifikanten Interaktionseffekts im Rahmen einer zweifaktoriellen Varianzanalyse mit Messwiederholung identisch mit einem signifikanten Haupteffekt (Treatmentfaktor) einer **einfaktoriellen Varianzanalyse**, wobei die abhängige Variable die **Differenzwerte** zwischen zweiter und erster Messung darstellen.

Die Zerlegung der Quadratsummen einer zweifaktoriellen Varianzanalyse mit Messwiederholung ist komplexer und wird in Bortz (1993, S. 310 ff.) dargestellt.

Varianzanalyse (22)

– Zweifaktorielle Varianzanalyse mit Messwiederholung (II) –

Beispiel: In einem Abstand von einem Jahr werden Aggressionswerte in einer Kontrollgruppe ohne und einer Experimentalgruppe mit einem Anti-Gewalt-Training gemessen (fiktive Werte).

Messwerte:

		Messwiederholung (Faktor 1)	
		Vormessung (t1)	*Nachmessung (t2)*
Treatment (Faktor 2)	*Kontrollgruppe*	52	57
		54	57
		51	55
		51	58
		53	56
	Experimental-gruppe	50	51
		59	53
		48	50
		52	52
		51	51

Mittelwerte:

		Messwiederholung (Faktor 1)	
		Vormessung (t1)	*Nachmessung (t2)*
Treatment (Faktor 2)	*Kontrollgruppe*	52.2	56.6
	Experimentalgruppe	52.0	51.4

Varianzanalyse (23)

– Zweifaktorielle Varianzanalyse mit Messwiederholung (III) –

ANOVA-Tabelle:

(zweifaktorielle Varianzanalyse mit Messwiederholung)

```
Analysis of Variance Table

between subjects:
                  SS    df  MS     F     p
Treatment       36.45   1  36.45  4.70  .062
Residuals       62.00   8   7.75

within subjects:
                  SS    df  MS     F     p
time            18.05   1  18.05  5.73  .044
Treatment*time  31.25   1  31.25  9.92  .014
Residuals       25.20   8   3.15
```

ANOVA-Tabelle:

(einfaktorielle Varianzanalyse der Differenzwerte)

```
Analysis of Variance Table

                  SS    df  MS     F     p
Treatment       62.50   1  62.50  9.92  .014
Residuals       50.40   8   6.30
```

Varianzanalyse (24)

– Zweifaktorielle Varianzanalyse mit Messwiederholung (IV) –