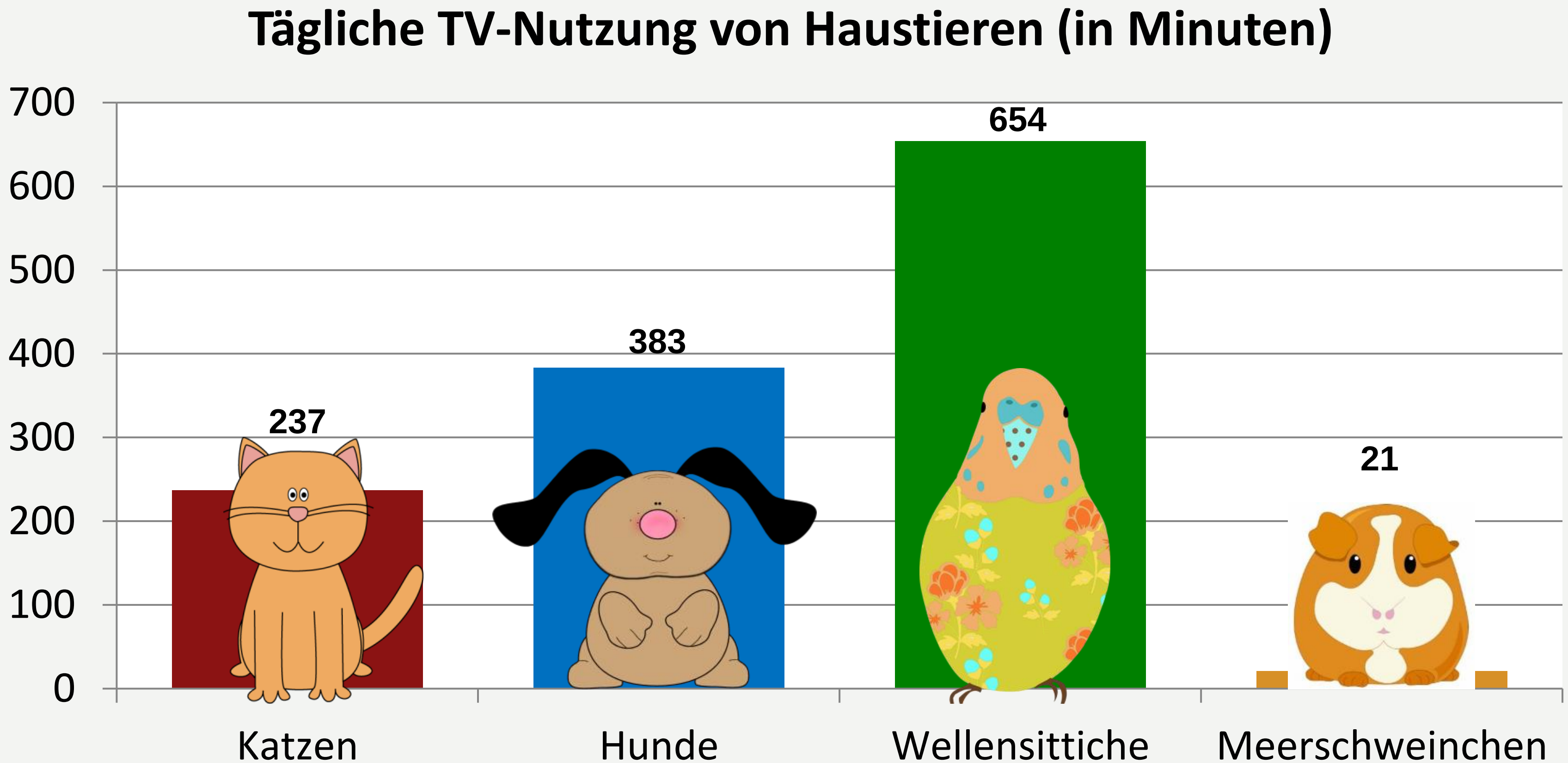
BA KW | Vorlesung

Einführung in die Statistik

Varianzanalyse

Prof. Thomas Hanitzsch

Mittelwertvergleich: grafische Darstellung



Varianzanalyse: Grundlagen

- Die Varianzanalyse untersucht den Einfluss einer oder mehrerer gestufter, unabhängiger Variablen auf eine oder mehrere abhängige Variablen
- Terminologie:
  - *Faktor* = unabhängige Variable (manchmal auch: „Treatment“)
  - *Faktorstufen* = Ausprägungen der unabhängigen Variable
- Arten von Varianzanalysen:
  - *univariat*: 1 abhängige Variable (ANOVA)
    - einfaktoriell: 1 unabhängige Variable
    - mehrfaktoriell: >1 unabhängige Variable
  - *multivariat*: >1 abhängige Variable (MANOVA)

Varianzanalyse vs. t-Tests

Vorteile gegenüber multiplen t-Tests

- Varianzanalyse kann mehrere Mittelwerte simultan miteinander vergleichen
  - für die Betrachtung beliebig vieler Mittelwerte ist also nur noch ein Test nötig
  - es tritt keine α-Fehler-Kumulierung auf (die durch Sequenzen von t-Tests entsteht)
- In den Test gehen gleichzeitig die Werte aller Versuchspersonen mit ein
  - die Teststärke ist sehr viel höher als die einzelner t-Tests

Einfaktorielle Varianzanalyse

- **Funktion:**
  - Prüft, ob die jeweiligen Stichproben aus Grundgesamtheiten stammen, deren Parameter $\mu_1, \mu_2, \ldots, \mu_\mathrm{k}$ identisch sind
  - Mit anderen Worten: die einfaktorielle Varianzanalyse geht der Frage nach, ob sich die Mittelwerte aus den jeweiligen Gruppen $(k > 2)$ systematisch voneinander unterscheiden
- **Formulierung der Hypothesen:**

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_\mathrm{k}$$

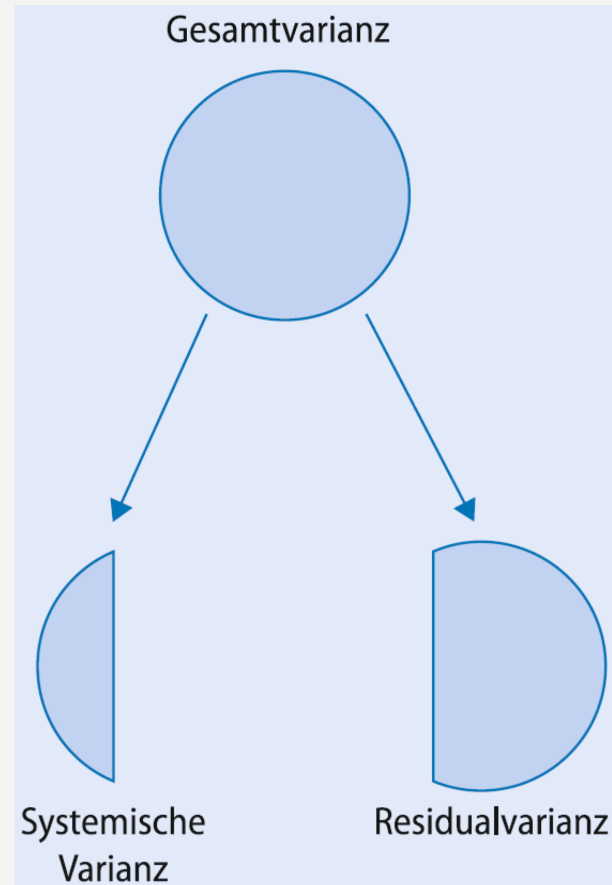
$$H_1: \neg H_0$$

Einfaktorielle Varianzanalyse: Grundprinzip

Sir Ronald Aylmer Fisher (1890-1962)

Das Prinzip der Varianzzerlegung:

- **Systematische Varianz** („Varianz zwischen den Gruppen“): verursacht durch unabhängige Variable
- **Residualvarianz** („Varianz innerhalb der Gruppen“): verursacht durch unbekannte Faktoren



Einfaktorielle Varianzanalyse

Terminologie:

- Y: abhängige Variable
- X: unabhängige Variable bzw. Faktor
- k: Anzahl der Faktorstufen
- j: Kennzeichnung der Faktorstufe (j = 1, 2, 3, … , k)
- i: Beobachtung (z.B. „Person“) innerhalb einer Faktorstufe
- N: Gesamtzahl der Beobachtungen
- n_j: Anzahl der Beobachtungen der j−ten Faktorstufe
- y_{ij}: i-te Beobachtung von Y in der j-ten Faktorstufe
- SY_j: Summe der beobachteten Werte unter der j-ten Faktorstufe ($SY_j = \sum_i y_{ij}$)
- $\overline{y}_j$: arithmetisches Mittel aller Messwerte unter der j-ten Faktorstufe ($\overline{y}_i = SY_j / n_j$)
- SY: Gesamtsumme aller Messwerte für Y ($SY = \sum_j \sum_i y_{ij}$)
- $\overline{y}$: arithmetisches Mittel der aller Messwerte für Y ($\overline{y} = SY / N$)

Quadratsummenzerlegung

- Totale Quadratsumme:

$$QS_{tot} = \sum_j \sum_i (y_{ij} - \bar{y})^2 \quad \text{mit} \quad df_{tot} = N - 1$$

- Treatmentquadratsumme (systematischer Anteil):

$$QS_X = \sum_j n_j \, (\bar{y}_j - \bar{y})^2 \quad \text{mit} \quad df_X = k - 1$$

- Fehlerquadratsumme (Anteil Residuum):

$$QS_e = \sum_j \sum_i (y_{ij} - \bar{y}_j)^2 \quad \text{mit} \quad df_e = N - k$$

⇒ Additive Beziehung:

$$QS_{tot} = QS_X + QS_e \quad \text{mit} \quad df_{tot} = df_X + df_e$$

Vereinfachte Berechnung

- Gleiche Stichprobengrößen (balanciertes Design):

$$(1) = \frac{SY^2}{N} \qquad (2) = \sum_j \sum_i y_{ij}^2 \qquad (3) = \frac{\sum_j SY_j^2}{n}$$

- Ungleiche Stichprobengrößen:

$$(1) = \frac{SY^2}{N} \qquad (2) = \sum_j \sum_i y_{ij}^2 \qquad (3) = \sum_j \left(\frac{SY_j^2}{n_j}\right)$$

- Berechnung der Quadratsummen:

$QS_X = (3) - (1)$ mit $df_X = k - 1$

$QS_e = (2) - (3)$ mit $df_e = N - k$

$QS_{tot} = (2) - (1)$ mit $df_{tot} = N - 1$

Signifikanztest (F-Test)

- Die Prüfgröße F ergibt sich aus dem Verhältnis der mittleren Quadrate, die auf den Einfluss der unabhängigen Variable und den Fehler zurückzuführen sind
  - Mittlere Quadrate (MQ): $MQ_X = \frac{QS_X}{df_X}$ sowie $MQ_e = \frac{QS_e}{df_e}$
- Berechnung der Prüfgröße F:

$$F = \frac{MQ_X}{MQ_e}$$

- Testentscheidung nach festgelegtem Signifikanzniveau:
  - F_{krit} aus Tabelle der F-Verteilung ablesen, unter df_X und df_e
  - H_0 wird abgelehnt, wenn $F > F_{krit}$

Effektstärke: η^2

- Maß der „Varianzaufklärung“ in den beobachteten Werten: beziffert den Anteil der beobachteten Variation in der abhängigen Variable, die auf den Einfluss der unabhängigen Variable (d.h. dem Faktor) zurückzuführen ist
- Wertebereich: 0 … 1,0 (bzw. 0 … 100%)

$$\eta^2 = \frac{QS_X}{QS_{tot}} = \frac{QS_X}{QS_X + QS_e}$$

Einfaktorielle Varianzanalyse: Beispiel

- Ein Forscherteam der Universität von Transsilvanien vermutet einen Einfluss von Filmen auf den Blutdurst von Vampiren. Hierzu führen sie ein Experiment mit drei Gruppen von je 5 Teilnehmern durch, die sich jeweils einen Film ansehen und dann ihren Blutdurst auf einer 10-stufigen Skala angeben sollen. Das Signifikanzniveau wird auf 5% festgelegt.

Einfaktorielle Varianzanalyse: Beispiel

					$N = 15n = 5$
$i = 1$	3	3	8		
$i = 2$	1	4	7		
$i = 3$	3	3	9		
$i = 4$	2	2	4		
$i = 5$	1	3	7		
$\boldsymbol{SY_j}$	10	15	35	$\boldsymbol{SY}$	60
$\boldsymbol{\bar{y}_j}$	2	3	7	$\boldsymbol{\bar{y}}$	4

Einfaktorielle Varianzanalyse: Beispiel

Einfaktorielle Varianzanalyse: Beispiel

1. Gleiche Stichprobengrößen balanciertes Design:

$$(1) = \frac{SY^2}{N} = \frac{60^2}{15} = 240$$

$$(2) = \sum_j \sum_i y_{ij}^2 = 3^2 + 1^2 + 3^2 + \ldots + 7^2 = 330$$

$$(3) = \frac{\sum_j SY_j^2}{n} = \frac{10^2 + 15^2 + 35^2}{5} = 310$$

2. Berechnung der Quadratsummen:

$$QS_X = (3) - (1) = 310 - 240 = 70 \qquad df_X = k - 1 = 2$$

$$QS_e = (2) - (3) = 330 - 310 = 20 \qquad df_e = N - k = 12$$

$$QS_{tot} = (2) - (1) = 330 - 240 = 90 \qquad df_{tot} = N - 1 = 14$$

Einfaktorielle Varianzanalyse: Beispiel

3. Signifikanztest:

- Berechnung der Mittleren Quadrate:

  $MQ_X = \frac{QS_X}{df_X} = \frac{70}{2} = 35$ sowie $MQ_e = \frac{QS_e}{df_e} = \frac{20}{12} = 1{,}67$

- Berechnung der Prüfgröße F:

  $$F = \frac{MQ_X}{MQ_e} = \frac{35}{1{,}67} = \mathbf{20,96}$$

- Kritischer F-Wert bei $\alpha = 0{,}05$:
  - für $df_X = 2$ und $df_e = 12$: $F_{krit} = \mathbf{3,89}$

- Testentscheidung:
  - H_0 wird abgelehnt, da $F = 20{,}96 > F_{krit}$

Einfaktorielle Varianzanalyse: Beispiel

4. Effektstärke:

$$\eta^2 = \frac{QS_X}{QS_{tot}} = \frac{70}{90} = 0{,}778$$

5. Bericht:

Die Hypothese konnte bestätigt werden. Demnach hat der Inhalt von Filmen tatsächlich einen Einfluss auf den Blutdurst von Vampiren. In den untersuchten Gruppen unterschieden sich die Testvampire im Hinblick auf ihren Blutdurst in Abhängigkeit von dem Film, den sie zuvor jeweils gesehen hatten („Tatort“: $\bar{x} = 2$, „Alarm für Cobra 11“: $\bar{x} = 3$ oder „Machete Kills“: $\bar{x} = 7$). Der Faktor „Film“ erklärt insgesamt 77,8 Prozent der Varianz im Blutdurst von Vampiren ($F(2{,}12) = 20{,}96$; $p < 0{,}05$).

Post hoc-Verfahren

- Der F-Test ist ein globaler Test, d.h. ein signifikantes Testergebnis belegt Unterschiede zwischen mindestens zwei der untersuchten Gruppen (Faktorstufen)
- Oft will der Forscher dann wissen, welche Gruppen sich konkret signifikant von anderen Gruppen unterscheiden
- Die Analyse erfolgt mittels besonderer Verfahren, den *Post hoc*-Tests
- R bietet einer Reihe dieser Verfahren an

Post hoc-Verfahren

- Wenn davon ausgegangen werden kann, dass die Varianzen in der Population ähnlich sind:
  - Bonferroni: liefert die beste Kontrolle über den α-Fehler
  - REGWQ / Turkey: gute Teststärke und strenge Kontrolle über den α-Fehler
  - Gabriel: bei leicht abweichenden Stichprobengrößen
  - GT2 n. Hochberg: bei stark abweichenden Stichprobengrößen
  - Dunnett: wenn Gruppen gegen einen Kontroll-Mittelwert verglichen werden sollen
- Wenn es Zweifel an der Ähnlichkeit der Varianzen gibt:
  - Games-Howell: liefert allgemein die besten Ergebnisse

Post hoc-Verfahren

Post-Hoc-Tests

	Group_Var	contrast	Delta_M	conf_lower	conf_upper	p	d	se	t	df
2	Gruppe	Tatort-Cobra 11	1	-1.178302	3.178302	0.4618859342	1.154701	0.5477226	-1.825742	8
1	Gruppe	Tatort-Machete	5	2.821698	7.178302	0.0001409786	3.333333	0.9486833	-5.270463	8
3	Gruppe	Cobra 11-Machete	4	1.821698	6.178302	0.0009867809	2.828427	0.8944272	-4.472136	8

Bonferroni	Tatort	Cobra 11	-1,000	,816	,733	-3,27	1,27
		Machete	-5,000*	,816	<,001	-7,27	-2,73
	Cobra 11	Tatort	1,000	,816	,733	-1,27	3,27
		Machete	-4,000*	,816	,001	-6,27	-1,73
	Machete	Tatort	5,000*	,816	<,001	2,73	7,27
		Cobra 11	4,000*	,816	,001	1,73	6,27
Games-Howell	Tatort	Cobra 11	-1,000	,548	,228	-2,60	,60
		Machete	-5,000*	,949	,004	-7,90	-2,10
	Cobra 11	Tatort	1,000	,548	,228	-,60	2,60
		Machete	-4,000*	,894	,014	-6,89	-1,11
	Machete	Tatort	5,000*	,949	,004	2,10	7,90
		Cobra 11	4,000*	,894	,014	1,11	6,89

*. Die Mittelwertdifferenz ist in Stufe 0.05 signifikant.

Post hoc-Verfahren

	Group_Var	contrast	Delta_M	conf_lower	conf_upper	p	d	se	t	df
2	Gruppe	Tatort-Cobra 11	1	-1.178302	3.178302	0.4618859342	1.154701	0.5477226	-1.825742	8
1	Gruppe	Tatort-Machete	5	2.821698	7.178302	0.0001409786	3.333333	0.9486833	-5.270463	8
3	Gruppe	Cobra 11-Machete	4	1.821698	6.178302	0.0009867809	2.828427	0.8944272	-4.472136	8

Einfaktorielle Varianzanalyse

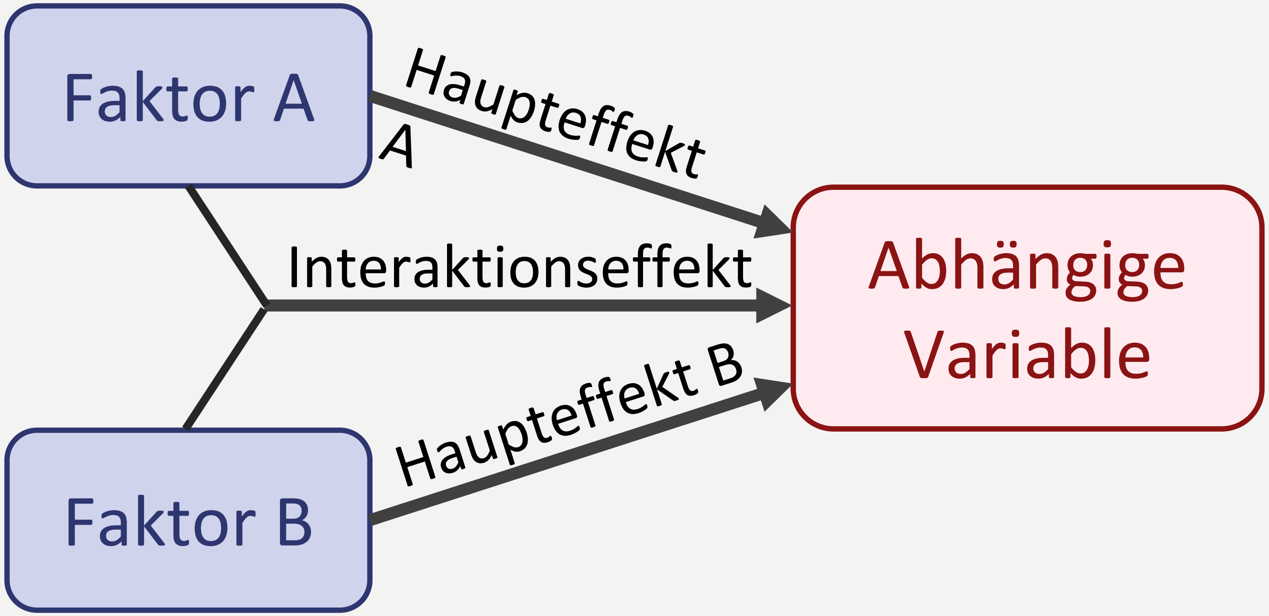
Voraussetzungen für den Signifikanztest:

- Die abhängige Variable ist *intervallskaliert*
- *Varianzhomogenität*: die Varianzen der Populationen der untersuchten Gruppen sind gleich
- Die Messwerte in allen Bedingungen sind voneinander *unabhängig*
- Die Fehlerkomponenten sind *normalverteilt* und voneinander *unabhängig*

- *Aber:* Voraussetzungen verlieren bei wachsendem Stichprobenumfang ihre Bedeutung

Zweifaktorielle Varianzanalyse

- Problemformulierung:



Zweifaktorielle Varianzanalyse: SPSS

Deskriptive Statistiken

Abhängige Variable: Blutdurst

Film	Geschlecht	Mittelwert	Standardabweichung	N
Tatort	weiblich	3,00	,000	2
	männlich	1,33	,577	3
	Gesamt	2,00	1,000	5
Cobra 11	weiblich	3,33	,577	3
	männlich	2,50	,707	2
	Gesamt	3,00	,707	5
Machete	weiblich	5,50	2,121	2
	männlich	8,00	1,000	3
	Gesamt	7,00	1,871	5
Gesamt	weiblich	3,86	1,464	7
	männlich	4,13	3,314	8
	Gesamt	4,00	2,535	15

Zweifaktorielle Varianzanalyse: SPSS

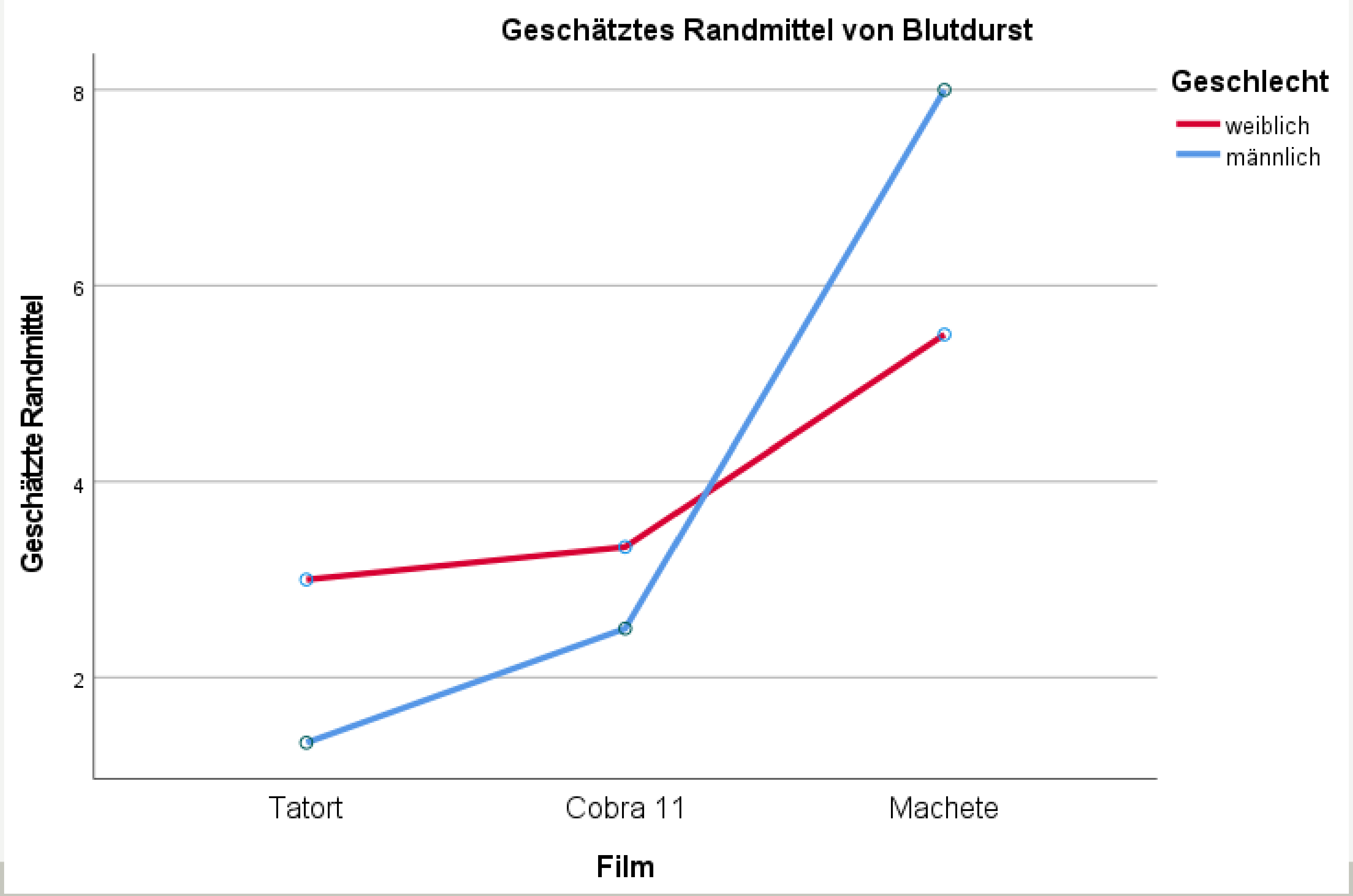
Tests der Zwischensubjekteffekte

Abhängige Variable: Blutdurst

Quelle	Typ III Quadratsumme	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Korrigiertes Modell	81,667[a]	5	16,333	17,640	<,001	,907
Konstanter Term	224,044	1	224,044	241,968	<,001	,964
Gruppe	58,022	2	29,011	31,332	<,001	,874
Geschlecht	,000	1	,000	,000	1,000	,000
Gruppe * Geschlecht	11,667	2	5,833	6,300	,019	,583
Fehler	8,333	9	,926			
Gesamt	330,000	15				
Korrigierte Gesamtvariation	90,000	14				

a. R-Quadrat = ,907 (korrigiertes R-Quadrat = ,856)

Zweifaktorielle Varianzanalyse: SPSS



Grafische Darstellung des Interaktions-effekts

Aufgabe 1

Freiheitsgrade und F-Werte

Bestimmen Sie für folgende Fragestellungen jeweils die Freiheitsgrade und den kritischen F-Wert.

a. Es wurde ein Experiment zur Werbewirkung durchgeführt. Die Versuchspersonen sahen unterschiedliche Versionen eines Werbespots. Insgesamt gab es vier Gruppen: Werbespot mit Humor, mit Erotik, mit Prominenten und mit Tieren. In jeder Gruppe waren 31 Versuchspersonen. (Signifikanzniveau von 1%)

b. Eine Studie hat die durchschnittliche Vor- und Nachbereitungszeit von Studierenden an der LMU untersucht. Untersucht wurden 50 BWLer, 40 KWler, 60 Mediziner, 40 Soziologen und 40 Juristen. (Signifikanzniveau von 5%)