

# Die Varianzanalyse ohne Messwiederholung

Jonathan Harrington

Bitte noch einmal datasets.zip laden

## Variablen, Faktoren, Stufen

Eine Varianzanalyse ist die Erweiterung von einem t-test

t-test oder ANOVA (Analysis of Variance = Varianzanalyse)

Ein Faktor mit 2 Stufen

- Hat Gender einen Einfluss auf die Dauer?

## ANOVA

Ein Faktor mit mehr als 2 Stufen oder mehr als ein Faktor

- Es gibt 3 Altersgruppen, jung, mittel, alt. Hat die Altergruppe einen Einfluss auf die Dauer? (Ein Faktor mit 3 Stufen)
- Haben Gender und Dialekt einen Einfluss auf die Dauer? (2 Faktoren)

## Was ist die Varianzanalyse?

Mit der Varianzanalyse wird (durch einen F-Test) ein Verhältnis zwischen zwei Varianzen berechnet:  
**innerhalb von Stufen** und **zwischen Stufen**.

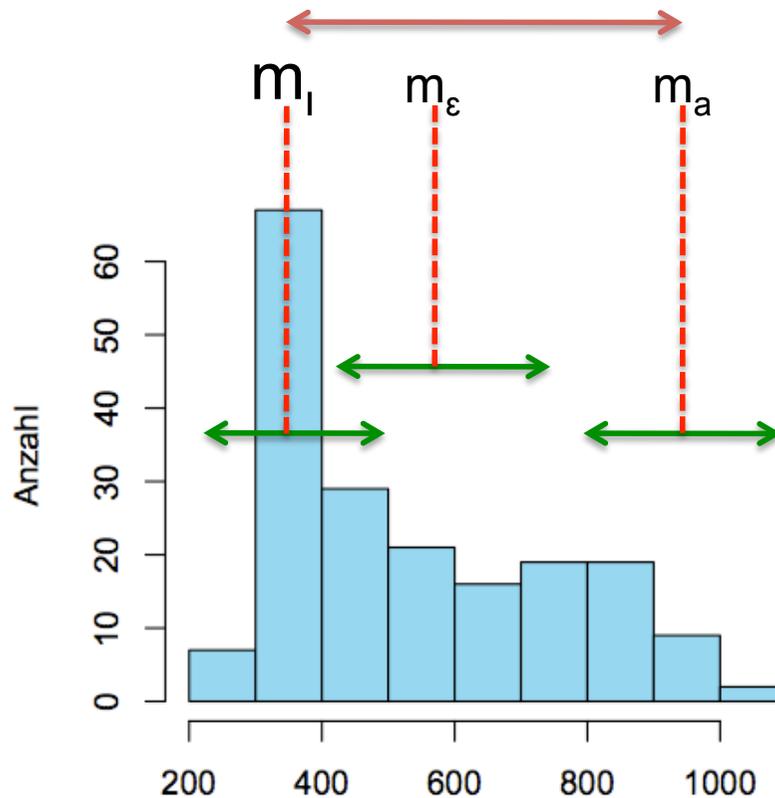
z.B. F1 von drei Vokalkategorien, /i,ε,a/.

**innerhalb:** Es gibt eine **randomisierte Variation von F1** innerhalb jeder Stufe (F1 von /i/ variiert, F1 von /ε/ variiert, F1 von /a/ variiert).

**zwischen:** F1 variiert, weil es eine **systematische Variation** zwischen den Verteilungen der Vokalkategorien gibt: die Werte von /i/, /ε/, und /a/ liegen in ganz unterschiedlichen F1-Bereichen, und je unterschiedlicher sie sind, **umso größer wird diese Varianz im Verhältnis zu der willkürlichen, randomisierten Varianz innerhalb der Stufen sein.**

# Was ist die Varianzanalyse?

## F1-Verteilung, drei Vokale



F

=

Varianz zwischen den Stufen

Varianz innerhalb der Stufen

Ist F signifikant größer als 1?

## Berechnung der Varianzen, innerhalb und zwischen

Diese Berechnung erfolgt über die sogenannte **Quadratsumme** oder **sum-of-squares**, die sich von der Varianz ableiten lässt (1) oder die durch die Formel (2) direkt berechnet werden kann

$$(1) \quad SS_x = (n - 1)s_x^2$$

(Quadratsumme von  $x$  gleicht die Varianz von  $x$  mal  $n-1$  ( $n$  ist die Anzahl der Stichproben).

$$(2) \quad SS_x = \sum (x - m_x)^2$$

Bestätigen

$x = 1:6$

$n = \text{length}(x)$

$v = \text{var}(x)$

$v * (n-1)$

$m = \text{mean}(x)$

$ssx = \text{sum}((x - m)^2)$

## Berechnung der Varianzen, innerhalb und zwischen

d.h. wenn wir die Quadratsummen wissen, gelangen wir zu den Varianzen, und wenn wir die Varianzen wissen, können wir den erwünschten F-Test durchführen.

Warum aber diese Schiene über die Quadratsummen?

Wegen einer Beziehung zwischen 3 Quantitäten, die auf eine sehr ähnliche Weise in der Regression vorkam.

$$\mathbf{SSY} = \mathbf{SSR} + \mathbf{SSE}$$

Die  
Quadratsumme  
über die  
gesamte  
Verteilung  
berechnet

=

Die  
Quadratsummen  
zwischen den  
Stufen

+

Die  
Quadratsummen  
innerhalb der  
Stufen

## Berechnung der Varianzen, innerhalb und zwischen

**vokal** 20 F2-Werte, 10 /I/, 10 /E/, ein Wert pro Person  
(also 20 Werte von 20 unterschiedlichen Personen)

```
vokal = read.table(file.path(pfad, "vokal.txt"))
```

```
attach(vokal)
```

```
table(V)
```

```
E  I  
10 10
```

## Berechnung der Varianzen, innerhalb und zwischen

Quadratsummen gesamt (SSY) =

$$SSY = \text{var}(F2) * (\text{length}(F2) - 1)$$

$$SS_y = (n - 1)s_y^2$$

Q-Summen innerhalb (SSE) +

$$SSE = SS_I + SS_E$$

Q-summen **innerhalb der Stufen** gleicht die  
Quadratsumme von /I/ plus die Q-Summe von /E/

Quadratsumme von /I/

Quadratsumme von /E/

`temp = V == "I"`

$$SSE = \text{var}(F2[temp]) * 9 + \text{var}(F2[!temp]) * 9$$

In einer Zeile

$$SSE = \text{sum}(tapply(F2, V, var) * 9)$$

Q-Summen zwischen (SSR)

$$SSR = SSY - SSE$$

## Berechnung der Varianzen, innerhalb und zwischen

$$\text{Fratio} = \frac{\text{Varianz zwischen den Stufen}}{\text{Varianz innerhalb der Stufen}} = \frac{\text{MSR}}{\text{MSE}}$$

Ist Fratio signifikant größer als 1?

$$\text{MSE} = \text{SSE} / 18$$

18 weil  $n-1 = 9$  pro Stufe

$$\text{MSR} = \text{SSR} / 1$$

weil 2 Stufen (/l, E/),  $2 - 1 = 1$

$$\text{Fratio} = \text{MSR} / \text{MSE}$$

[1] 7.22526

detach(vokal)

## ANOVA Berechnung in R

```
summary(aov(F2 ~ V, data=vokal))
```

		SSR		MSR = SSR/1								
Response: F2												
	Df	Sum Sq	Mean Sq	F value	Pr(>F)							
V	1	229837	229837	7.2253	0.01503	*						
Residuals	18	572583	31810									
---												
Signif. codes:		0	****	0.001	**	0.01	*	0.05	.	0.1	'	1

Annotations:

- SSR points to Sum Sq (229837)
- MSR = SSR/1 points to Mean Sq (229837)
- SSE points to Sum Sq (572583)
- MSE = SSE/18 points to Mean Sq (31810)
- Fratio = MSR/MSE points to F value (7.2253)
- 1 - pf(Fratio, 1, 18) points to Pr(>F) (0.01503)

F2 wird signifikant vom Vokal beeinflusst:  $F[1, 18] = 7.23, p < 0.05$ .

Das gleiche

```
reg = lm(F2 ~ V, data = vokal)
anova(reg)
```

## Beziehung: t-test und ANOVA

Da wir in diesem Fall mit einem Faktor und 2 Stufen zu tun haben, hätten wir das gleiche Ergebnis mit einem t-test bekommen können

```
t.test(F2 ~ V, var.equal=T, data = vokal)
```

```
t = -2.688, df = 18, p-value = 0.01503
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -381.97452  -46.82548
sample estimates:
mean in group E mean in group I
      1638.8      1853.2
```



Die t-Statistik ist die Wurzel vom F-Ratio aus der ANOVA

## ANOVA: einige Voraussetzungen

### 1. ähnlich stark besetzte Stufen und Faktoren

zB 20 initiale, 20 mediale, 20 finale /t/s, um zu messen, ob die Silbenposition (= Faktor) einen Einfluss auf die Dauer hat. Um zusätzlich zu messen, ob Dialekt (Bayern, Hessen) einen Einfluss ausübt: 30 aus Bayern, 30 aus Hessen, jeweils 10 pro Silbenposition.

<u>Silbenposition</u>	<u>initial</u>	<u>medial</u>	<u>final</u>
<u>Dialekt</u>			
Bayern	10	10	10
Hessen	10	10	10

## ANOVA: einige Voraussetzungen

2. Alle Werte sind von unterschiedlichen Sprechern (60 Vpn., ein Wert pro Vpn für dieses Beispiel).

Ein Sprecher produzierte eine Silbe in initialer, medialer, und finaler Position = **ANOVA mit Messwiederholungen** (Repeated-Measures ANOVA).

<u>Silbenposition</u>	<u>initial</u>	<u>medial</u>	<u>final</u>
<u>Dialekt</u>			
Bayern	10	10	10
Hessen	10	10	10

### 3. Generell für ANOVA: keine Wiederholungen **innerhalb einer Zelle**

z.B. Die Vpn. produzierten 10 x dieselben initialen Silben. Diese Werte müssen gemittelt werden, sodass man nur einen Wert pro Vpn. pro Zelle hat.

<u>Silbenposition</u> <u>Dialekt</u>	<u>initial</u>	<u>medial</u>	<u>final</u>
Bayern	10	10	10
Hessen	10	10	10

## ANOVA: Voraussetzungen

4. Ein Anova kann eigentlich nie auf die Daten einer einzigen Versuchsperson angewandt werden

Abgesehen von einem trivialen Fall wie:

Eine Vpn. produzierte *heed*, *hid*, *hard*, *had* (jeweils einmal).

Unterscheiden sich die Wörter in der Dauer?

## Zwei Faktoren

```
vok = read.table(file.path(pfad, "vok.txt"))
```

```
head(vok)
```

```
with(vok, table(Vokal, Gen))
```

```
"F2"      "Vokal" "Gen"
```

F2 Daten, 60 Sprecher, 30 m, 30

```
table(Vokal, Gen)
```

w, drei Vokale

Gen

Vokal	m	w
E	10	10
I	10	10
a	10	10

Hat Vokal einen Einfluss auf F2?

Hat Gender einen Einfluss auf F2?

Gibt es eine Interaktion zwischen Vokal und Gender?

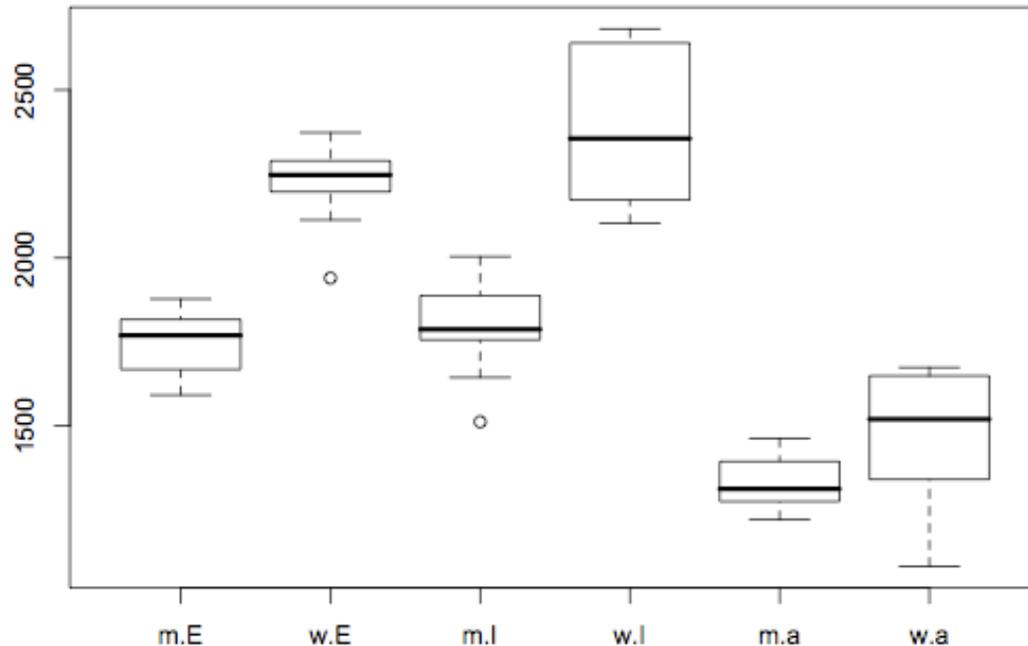
= ist der F2-Unterschied zwischen /I, E, a/ ähnlich für M und W?

N.B. Die Annahme ist wir haben 60 Vpn, einen Wert pro Sprecher

## Zwei Faktoren

### Boxplot Abbildung

```
boxplot(F2 ~ Gen * Vokal,  
data = vok)
```

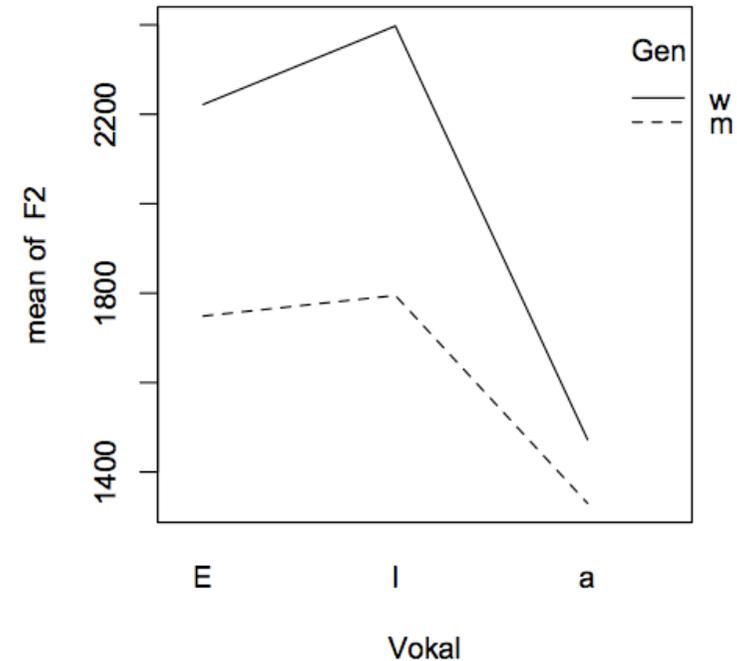


Hat Vokal einen Einfluss auf F2?

Hat Gender einen Einfluss auf F2?

### Interaktion-Abbildung

```
with(vok, interaction.plot(Vokal,  
Gen, F2))
```



Gibt es eine Interaktion zwischen Vokal und Gender?

## Zwei Faktoren

```
vok.aov = aov(F2 ~ Vokal * Gen, data = vok)
```

das gleiche

```
vok.aov = aov(F2 ~ Vokal + Gen + Vokal:Gen, data = vok)
```



Interaktion

```
summary(vok.aov)      oder      anova(vok.aov)
```

```
oder:
```

```
reg = lm(F2 ~ Vokal * Gen, data = vok)
```

```
anova(reg)
```

### Analysis of Variance Table

Response: F2

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Vokal	2	5578128	2789064	119.637	< 2.2e-16	***
Gen	1	2474570	2474570	106.147	2.354e-14	***
Vokal:Gen	2	563391	281696	12.083	4.603e-05	***
Residuals	54	1258885	23313			

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.'  
0.1 ' ' 1

## post-hoc Tukey-Test

Wenn eine Interaktion vorliegt, muss geprüft werden, ob sich die Stufen-Kombination in dem Parameter (hier F2) unterscheiden.

Tukey Tests normieren die Wahrscheinlichkeiten für die Anzahl der Tests (wenn 100 Tests post-hoc durchgeführt werden, kann es vorkommen dass einige per Zufall signifikant sein werden – dafür normiert der Tukey Test).

## post-hoc tests

tk = TukeyHSD(vok.aov)

tk

Vergleicht die Stufen des Vokals paarweise

```
$Vokal
      diff      lwr      upr      p adj
I-E  110.80  -5.561759  227.1618 0.0650875
a-E -584.25 -700.611759 -467.8882 0.0000000
a-I -695.05 -811.411759 -578.6882 0.0000000
```

Vergleicht die Stufen von Gender paarweise (trägt nichts neues bei, da wir schon aus dem Haupttest wissen, dass es signifikante Unterschiede innerhalb von Gender gibt).

```
$Gen
      diff      lwr      upr      p adj
w-m 406.1667 327.1282 485.2052      0
```

Vergleicht alle Vokal \* Gender Stufenkombinationen paarweise

```
$`Vokal:Gen`
      diff      lwr      upr      p adj
I:m-E:m  46.0  -155.74006  247.74006 0.9841188
usw...
```

## post-hoc tests

\$` Vokal : Gen`

Wir benötigen aber nur Vergleiche zwischen Stufen von **einem** Faktor, wenn die Stufen aller anderen Faktoren **konstant** sind.

1. Unterscheiden sich die Vokale desselben Geschlechts?  
(Vokal variiert, Geschlecht ist konstant).

I vs E in Männern

I vs E in Frauen

I vs a in Männern

I vs a in Frauen

E vs a in Männern

E vs a in Frauen

2. Unterscheiden sich Männer und Frauen derselben Vokale?  
(Geschlecht variiert, Vokal ist konstant)

m vs. w in I

m vs. w in E

m vs. w in a

Aber *nicht* wenn beide Faktoren variieren.

m-E vs. w-a, m-a vs w-I usw.

# 1. Unterscheiden sich die Vokale desselben Geschlechts (Vokal variiert, Geschlecht ist konstant)?

\$`Vokal:Gen`	diff	lwr	upr	p adj
I:m-E:m	46.0	-155.74006	247.74006	0.9841188
a:m-E:m	-418.9	-620.64006	-217.15994	0.0000015
E:w-E:m	473.2	271.45994	674.94006	0.0000001
I:w-E:m	648.8	447.05994	850.54006	0.0000000
a:w-E:m	-276.4	-478.14006	-74.65994	0.0021912
a:m-I:m	-464.9	-666.64006	-263.15994	0.0000001
E:w-I:m	427.2	225.45994	628.94006	0.0000010
I:w-I:m	602.8	401.05994	804.54006	0.0000000
a:w-I:m	-322.4	-524.14006	-120.65994	0.0002373
E:w-a:m	892.1	690.35994	1093.84006	0.0000000
I:w-a:m	1067.7	865.95994	1269.44006	0.0000000
a:w-a:m	142.5	-59.24006	344.24006	0.3094441
I:w-E:w	175.6	-26.14006	377.34006	0.1221478
a:w-E:w	-749.6	-951.34006	-547.85994	0.0000000
a:w-I:w	-925.2	-1126.94006	-723.45994	0.0000000

# 2. Unterscheiden sich Männer und Frauen derselben Vokale? (Geschlecht variiert, Vokal ist konstant)?

alles andere ist unwichtig

\$Vokal

	diff	lwr	upr	p adj
I-E	110.80	-5.561759	227.1618	0.0650875
a-E	-584.25	-700.611759	-467.8882	0.0000000
a-I	-695.05	-811.411759	-578.6882	0.0000000

/I/ vs /E/ nicht signifikant

/a/ vs /E/ sig: p < 0.001

/a/ vs /I/ sig: p < 0.001

\$Gen

	diff	lwr	upr	p adj
w-m	406.1667	327.1282	485.2052	0

Gender: sig. p < 0.001

abgesehen von /a/

\$`Vokal:Gen`

	diff	lwr	upr	p adj
I:m-E:m	46.0	-155.74006	247.74006	0.9841188
a:m-E:m	-418.9	-620.64006	-217.15994	0.0000015
E:w-E:m	473.2	271.45994	674.94006	0.0000001
a:m-I:m	-464.9	-666.64006	-263.15994	0.0000001
I:w-I:m	602.8	401.05994	804.54006	0.0000000
a:w-a:m	142.5	-59.24006	344.24006	0.3094441
I:w-E:w	175.6	-26.14006	377.34006	0.1221478
a:w-E:w	-749.6	-951.34006	-547.85994	0.0000000
a:w-I:w	-925.2	-1126.94006	-723.45994	0.0000000

## Haupteffekte

Response: F2

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Vokal	2	5578128	2789064	119.637	< 2.2e-16	***
Gen	1	2474570	2474570	106.147	2.354e-14	***
Vokal:Gen	2	563391	281696	12.083	4.603e-05	***
Residuals	54	1258885	23313			

## Post-hoc

/l/ vs /E/ nicht signifikant     /a/ vs /E/ sig:  $p < 0.001$

/a/ vs /l/ sig:  $p < 0.001$      Gender: sig.  $p < 0.001$   
abgesehen von /a/

Vokal ( $F[2,54] = 119.6$ ,  $p < 0.001$ ) sowie Gender ( $F[1,54] = 106.1$ ,  $p < 0.001$ ) hatten einen signifikanten Einfluss auf F2 und es gab eine signifikante Interaction ( $F[2,54] = 12.1$ ,  $p < 0.001$ ) zwischen diesen Faktoren. Post-hoc Tukey-Tests zeigten signifikante **F2-Unterschiede zwischen /a/ vs /E/ ( $p < 0.001$ )** und **zwischen /a/ vs /l/ ( $p < 0.001$ )** jedoch nicht zwischen /l/ vs. /E/. F2 von Männern und Frauen unterschieden sich signifikant für alle Vokale außer /a/.

## tksel() Funktion

Die benötigten Kombinationen bekommt man auch mit **tksel()**

```
source(file.path(pfad, "tksel.txt"))
```

```
names(tk)
```

```
[1] "Vokal"
```

```
"Gen"
```

```
"Vokal:Gen"
```

Interaktion an  
dritter Stelle

Vokal variiert

```
tksel(tk[[3]], 1)
```

Faktor 1 variiert

Geschlecht variiert

```
tksel(tk[[3]], 2)
```

Faktor 2 variiert

```
vok.aov = aov(F2 ~ Vokal * Gen, data = vok)
```