

## Lösung zu Kapitel 10: Beispiel 6

`leben-nach-tod.csv` enthält Angaben zum Glauben an ein Leben nach dem Tod. Die abhängige Variable (`leben.nt`) soll zu den Prädiktoren Geschlecht (`sex`) und Alter (`alter`) in einem möglichst guten Modell in Beziehung gesetzt werden.

- Wir lesen `leben-nach-tod.csv` ein ermöglichen mit `attach` den direkten Zugriff auf die Variablen des Datensatzes.

R

```
> lnt <- read.csv2("leben-nach-tod.csv", header = TRUE)
> attach(lnt)
```

- In `leben.1` speichern wir ein volles Modell mit den Haupteffekten von `sex` und `alter`, sowie deren Interaktion `sex:alter`. Da wir eine logistische Regression berechnen müssen wir in `glm` das Argument `family=binomial` setzen.  
In `leben.2` speichern wir ein reduziertes Modell aus dem wir mit `update` die Interaktion `sex:alter` entfernt haben. Wenn wir beide Modelle einem Likelihood-Ratio-Test (mit angefordertem  $\chi^2$ -Test durch `test="Chisq"`) unterziehen sehen wir, dass das reduzierte Modell mit  $p = .372$  nicht signifikant schlechter passt als das volle.

R

```
> leben.1 <- glm(leben.nt ~ sex * alter, family = binomial)
> leben.2 <- update(leben.1, . ~ . - sex:alter)
> anova(leben.2, leben.1, test = "Chisq")
```

### Analysis of Deviance Table

```
Model 1: leben.nt ~ sex + alter
Model 2: leben.nt ~ sex * alter
      Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1          380      476.97
2          379      476.18  1   0.79589   0.3723
```

- Wir können mit der Modellselektion fortfahren und eliminieren in `leben.3` und `leben.4` jeweils die Effekte für `sex` bzw. `alter` und unterziehen diese Modell wieder einem Likelihood-Ratio-Test. Nun zeigt sich, dass zwar `leben.3` signifikant schlechter ( $p = .001$ ) passt, aber `leben.4` hingegen mit  $p = .206$  noch akzeptabel ist.  
Wir gehen daher von `leben.4` aus und versuchen hier `sex` zu entfernen, was jedoch zu einem signifikanten Ergebnis ( $p < .001$ ) führt, wodurch `leben.4` zum Modell unserer Wahl wird das wir uns mittels `summary` detailliert ausgeben lassen.

R

```
> leben.3 <- update(leben.2, . ~ . - sex)
> leben.4 <- update(leben.2, . ~ . - alter)
> anova(leben.3, leben.1, test = "Chisq")
```

### Analysis of Deviance Table

```
Model 1: leben.nt ~ alter
```

```

Model 2: leben.nt ~ sex * alter
      Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1          381      489.75
2          379      476.18  2    13.577  0.001127 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

**R**

```
> anova(leben.4, leben.1, test = "Chisq")
```

#### Analysis of Deviance Table

```

Model 1: leben.nt ~ sex
Model 2: leben.nt ~ sex * alter
      Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1          381      479.33
2          379      476.18  2     3.157  0.2063

```

**R**

```

> leben.5 <- update(leben.4, . ~ . - sex)
> anova(leben.5, leben.4, test = "Chisq")

```

#### Analysis of Deviance Table

```

Model 1: leben.nt ~ 1
Model 2: leben.nt ~ sex
      Resid. Df Resid. Dev Df Deviance P(>|Chi|)
1          382      492.06
2          381      479.33  1    12.727  0.0003604 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

**R**

```
> summary(leben.4)
```

```

Call:
glm(formula = leben.nt ~ sex, family = binomial)

```

#### Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.0715	-1.0715	-0.7815	1.2871	1.6340

#### Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-1.0296	0.1571	-6.555	5.58e-11 ***
sexMann	0.7754	0.2192	3.538	0.000404 ***

---

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

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 492.06 on 382 degrees of freedom  
Residual deviance: 479.33 on 381 degrees of freedom  
AIC: 483.33

Number of Fisher Scoring iterations: 4

R

```
> detach(lnt)
```