# Onlinebeispiel 7.5. Einfluss der Körperkondition auf das Mortalitätsrisiko überwinternder Dunkelenten (*Anas rubripes*)

Kapitel 7.3.1.3 aus Henle, K., A. Grimm-Seyfarth & B. Gruber: Erfassung und Analyse von Tierpopulationen. Ulmer Verlag

Annegret Grimm-Seyfarth

07.04.2025

Conroy et al. (1989) untersuchten die Wintermortalität von Dunkelenten (*Anas rubripes*) in zwei Untersuchungsgebieten in New Jersey. Zur Illustration des Cox-Modells des relativen Risikos werden nur die Daten im ersten Untersuchungsjahr ausgewertet. Zwischen Enten im ersten Lebensjahr (Jungenten) und älteren Enten (Adulte) wurde unterschieden. Im Zeitraum vom 8. November bis 14. Dezember 1983 wurden 50 Enten besendert. Die besenderten Tiere wurden zu Fuß, mit Fahrzeugen, Boot und Flugzeug verfolgt, und ihr Status (lebend, tot, vermisst) wurde täglich bis zum 15. Februar 1984 kontrolliert. Die Analyse erfolgt für die 63 Tage ab 15. Dezember 1983, da zu diesem Zeitpunkt die Besenderung abgeschlossen war. In diesem Fall lebten zu diesem Zeitpunkt noch alle besenderten Enten. Die Körperkondition wurde beim Fang bestimmt. Pollock et al. (1989) verwendeten bei ihrer Auswertung der Daten von Conroy et al. (1989) als Index der Körperkondition das Verhältnis von Masse zu Flügellänge. Abweichend hierzu werden nachfolgend als Konditionsindex (KI) die Residuen – d.h. die Abweichungen vom Erwartungswert – aus einer Regressionsanalyse mit der dritten Wurzel der Masse als abhängiger und Flügellänge als unabhängiger Variablen benutzt. Die dritte Wurzel wird verwendet, da die Masse mit der dritten Potenz der Länge wachsen sollte. Der Index von Pollock et al. (1989) gewichtet die Länge stärker und stellt daher mehr einen Größenindex der relativen Flügellänge als ein Konditionsindex dar. Der hier verwendete Ansatz hat außerdem den Vorteil, dass sich das Basisrisiko auf Enten bezieht, deren Masse dem Erwartungswert für ihre Größe (Kondition = 0) entspricht und somit das relative Risiko bei gleicher Abweichung der Kondition nach oben und unten direkt miteinander verglichen werden kann.

Für diese beiden Beispiele nutzen wir das R-Paket survival (Therneau & Grambsch 2000, Therneau 2024). Für die Darstellung der Kaplan-Meier-Überlebensgrafiken nutzen wir das Paket ggsurvfit (Sjoberg et al. 2024).

```r
# check.packages function: install and load multiple R packages.
# Function from: https://gist.github.com/smithdanielle/9913897
check.packages <- function(pkg){
  new.pkg <- pkg[!(pkg %in% installed.packages()[, "Package"])]
  if (length(new.pkg))
    install.packages(new.pkg, dependencies = TRUE, type = "source")
  sapply(pkg, require, character.only = TRUE)
}

# benoetigte R pakete
pakete <- c("survival", "ggsurvfit", "emmeans")

# Pruefe und installiere
check.packages(pakete)
```

```
##  survival ggsurvfit   emmeans
##      TRUE      TRUE      TRUE
```

Weitere Informationen zur Nutzung des Paketes finden sich hier:

https://cran.r-project.org/web/packages/survival/survival.pdf

https://cran.r-project.org/web/packages/ggsurvfit/ggsurvfit.pdf

# Einlesen der Fangdaten

Wir laden zunächst die Daten ein. Dazu haben wir die Tabelle von Pollock et al. (1989) um den Index der Körperkondition ergänzt.

```
# Daten einlesen
UlmerBuch::beispiel.pfad() #Pfad zu den Beispieldaten
```

```
## Der Pfad zu den Beispiel Daten wurde gesetzt auf:
##  C:/Users/grimm/AppData/Local/Programs/R/R-4.4.3/library/UlmerBuch/extdata
```

```
ARData <- read.csv2("AR_Pollock_et_al_1989.csv")
head(ARData)
```

```
##   i ti Status Indikator Alter Masse  FL
## 1 1  2      H         1     1  1160 277
## 2 2  6      C         0     0  1140 266
## 3 3  6      C         0     1  1260 280
## 4 4  7      N         1     0  1160 264
## 5 5 13      H         1     1  1080 267
## 6 6 14      C         0     0  1120 262
```

Die Daten beinhalten die Kontrolldaten telemetrierter Dunkelenten (*Anas rubripes*) in New Jersey. `i`: Nummer des Individuums; `ti`: Anzahl der Beobachtungstage ab 15. Dezember 1983 für Exemplar `i`; `Status`: H - geschossen ("hunter"), C - gefunden ("censored"), N - natürlich verstorben; `Indikator`: 0 = lebend, 1 = tot; `Alter`: 0 = diesjährig, 1 = mindestens einjährig; `Masse` in g; `FL`: Flügellänge in mm. Quelle: Pollock et al. (1989).

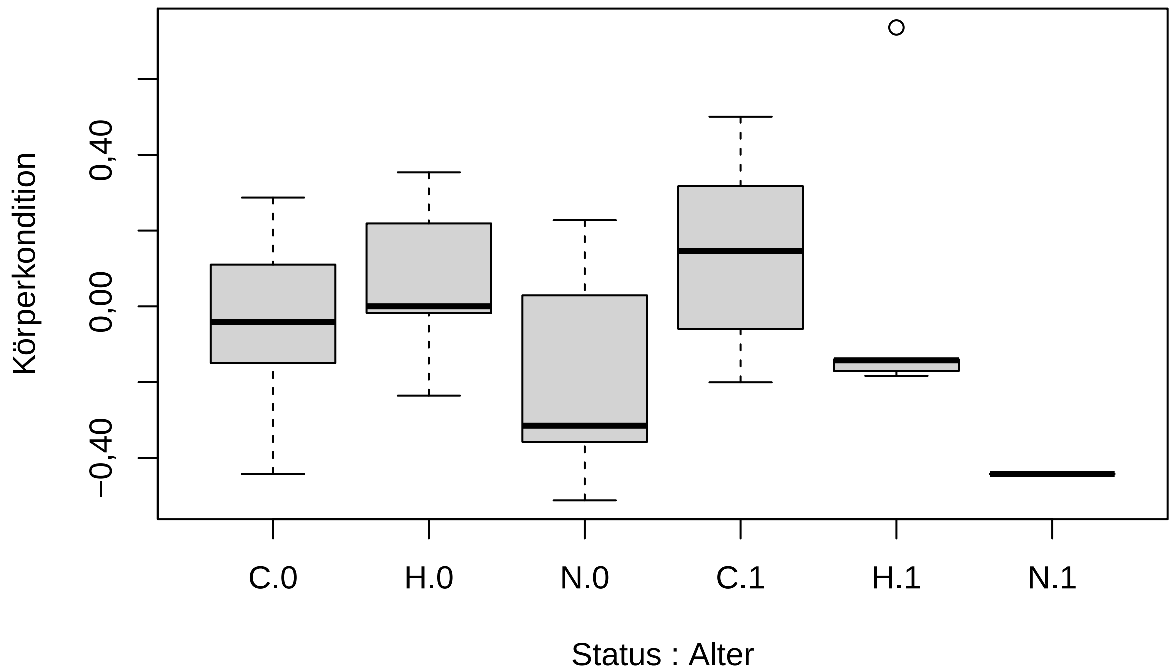
# Berechnung Konditionsindex

Zur Berechnung des Konditionsindex benötigen wir eine Regressionsanalyse mit der dritten Wurzel der Masse als abhängiger und Flügellänge als unabhängiger Variable.

```
# dritte Wurzel der Masse ziehen
ARData$MasseW <- ARData$Masse^(1/3)
# Regressionsgerade erstellen
KIfit <- lm(MasseW~FL, data=ARData)
# davon die Residuen ziehen
KI <- resid(KIfit)
# und zum Datensatz hinzufügen
ARData$KI <- KI
```

Schauen wir kurz einmal rein, ob die Tiere mit unterschiedlichem Status auch eine unterschiedliche Körperkondition aufweisen.

```
{boxplot(KI~Status+Alter, data=ARData, yaxt="n",
         ylab = "Körperkondition")
 axis(
   side = 2,
   at = pretty(ARData$KI),
   labels = format(pretty(ARData$KI), decimal.mark = ",", nsmall = 2))
}
```

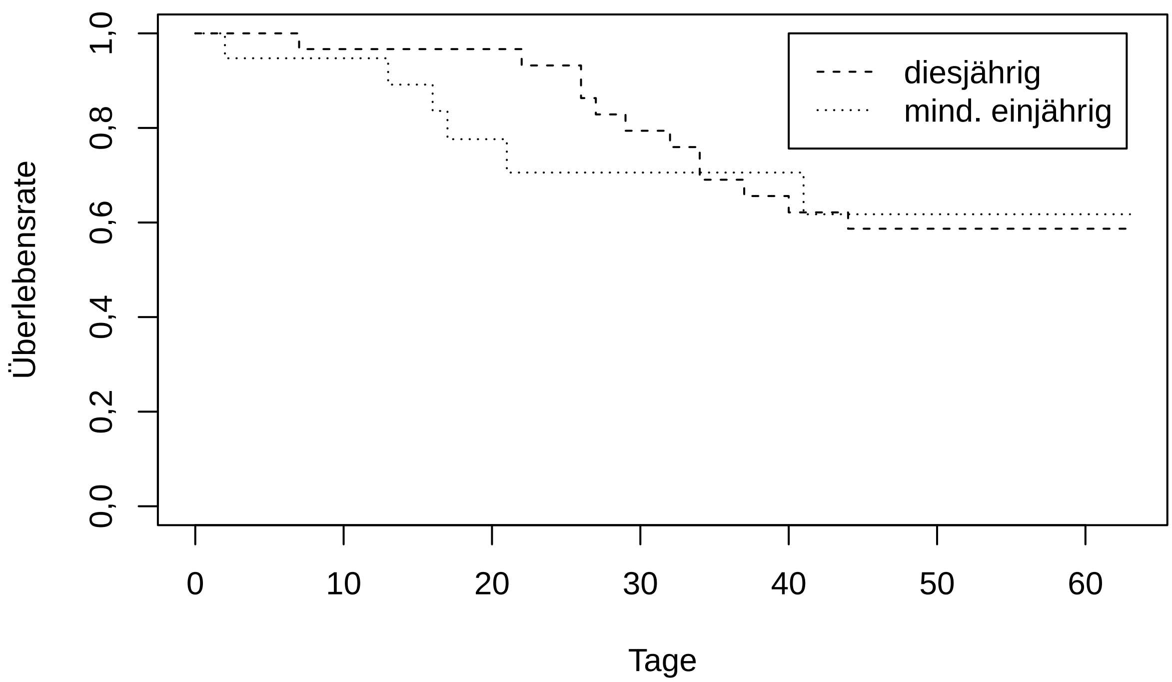



Analog dem Datensatz entspricht Status: H - geschossen ("hunter"), C - gefunden ("censored"), N - natürlich verstorben; und Alter: 0 = diesjährig, 1 = mindestens einjährig. Scheinbar sterben Tiere natürlichen Todes, wenn sie zu geringe Körperkondition aufweisen. Dies trifft für diesjährige zu wie für ältere Tiere. Im Gegensatz dazu werden von den Jägern auch Tiere mit besonders hoher Körperkondition geschossen, besonders bei den diesjährigen Tieren, wobei sich das nicht von den überlebenden Tieren statistisch unterscheidet. Bei den älteren Tieren wiederum sind solche mit hoher Kondition möglicher Weise zu schnell für Jäger, es werden eher (mit einer Ausnahme) Tiere mit mittlerer Kondition geschossen.

## Kaplan-Meier-Überlebensgrafik

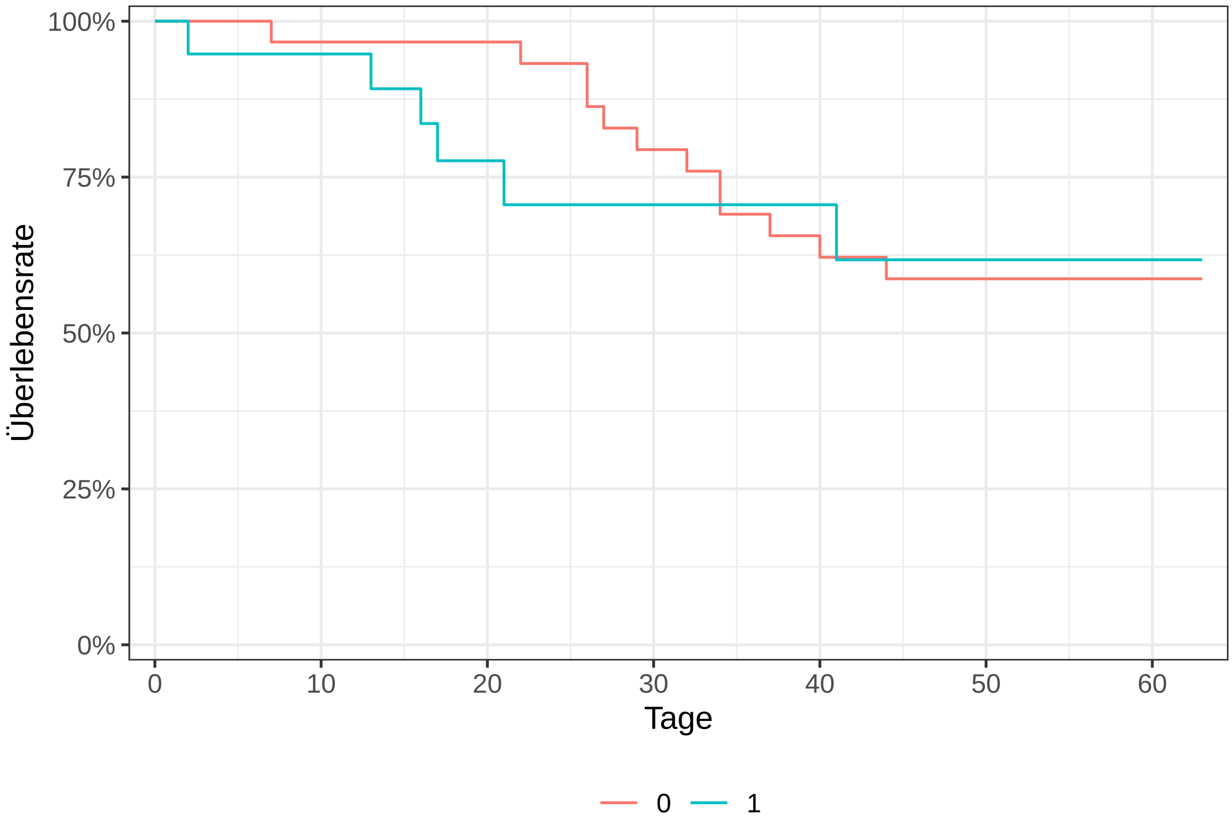
Um einen Überblick über die Daten zu bekommen, passen wir sie zunächst einem Kaplan-Meier-Modell an. Damit können die Überlebensgrafiken für diesjährige und mindestens einjährige Individuen dargestellt werden. Das einfachste Modell sieht folgendermaßen aus:

```r
fit <- survfit(Surv(ti, Indikator) ~ Alter, data = ARData)
{plot(fit, lty = 2:3, yaxt="n",
      xlab = "Tage",
      ylab = "Überlebensrate")
 axis(
   side = 2,
   at = seq(0,1,.2),
   labels = format(seq(0,1,.2), decimal.mark = ",", nsmall = 1))
legend(40, 1, c("diesjährig", "mind. einjährig"), lty = 2:3)
}
```



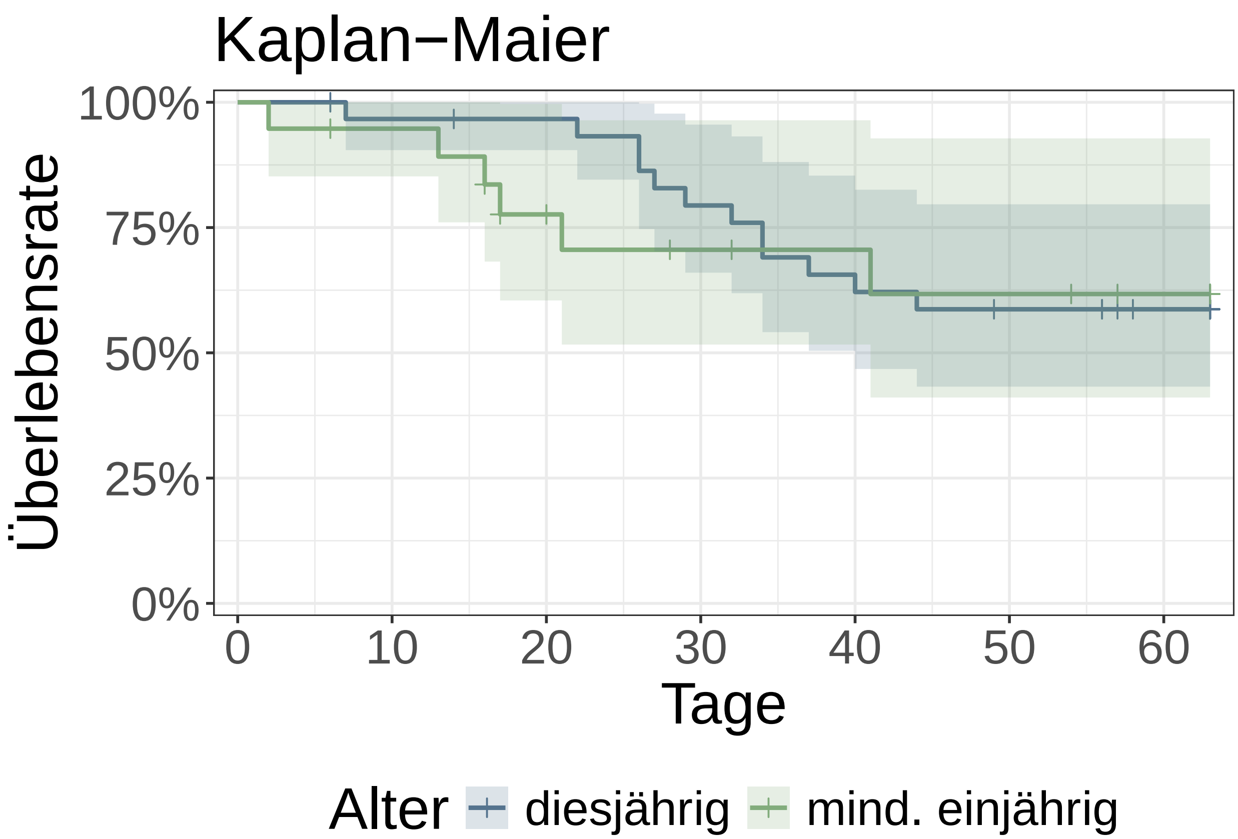

Dank des Paketes ggsurvfit können wir auf ggplot-Basis diesen Plot noch deutlich hübscher gestalten. Dies sieht folgendermaßen aus:

```r
survfit2(Surv(ti, Indikator) ~ Alter, data = ARData) %>%
ggsurvfit() +
  scale_ggsurvfit(x_scales = list(breaks = seq(0, 60, by = 10))) +
  labs(x = "Tage", y = "Überlebensrate")
```




Es lassen sich auch Konfidenzintervalle hinzufügen und weitere Sachen anpassen:

```
survfit2(Surv(ti, Indikator) ~ Alter, data = ARData) %>%
ggsurvfit(linewidth = 0.8) +
  add_censor_mark() +
  add_confidence_interval() +
  add_quantile() +
  #add_risktable() +
  add_legend_title() +
  scale_ggsurvfit() +
  scale_color_manual(values = c('#54738E', '#82AC7C'),
                     labels = c("diesjährig", "mind. einjährig")) +
  scale_fill_manual(values = c('#54738E', '#82AC7C'),
                    labels = c("diesjährig", "mind. einjährig")) +
  theme(plot.title = element_text(size=24),
        axis.title.x = element_text(size=22),
        axis.title.y = element_text(size=22),
        axis.text = element_text(size=18),
        legend.text = element_text(size = 18),
        legend.title = element_text(size = 22)) +
  labs(
    title = "Kaplan-Maier",
    y = "Überlebensrate",
    x = "Tage"
  )
```




## Cox-Modell mit Kovariablen

Als nächstes rechnen wir das Cox-Modell des relativen Risikos. Dies schauen wir uns in allen Teilschritten an.

### Modell erstellen

Wir gehen davon aus, dass es Unterschiede im Überleben zwischen diesjährigen und mindestens einjährigen Individuen gibt. Außerdem haben wir anhand des obigen Boxplots eine Idee über den Einfluss der Körperkondition bekommen: es könnte sich um einen quadratischen Zusammenhang handeln, der sich jedoch je nach Alter unterschiedlich verhält. Dies beschreiben wir als Interaktion. Wir nehmen also an, dass eine geringe Kondition das Überleben verschlechtert, aber eine zu hohe Kondition dazu führt, wahrscheinlicher geschossen zu werden. Für Adulte sollte der quadratische Term jedoch 0 sein.

```
fit2 <- coxph(Surv(ti, Indikator) ~ Alter * poly(KI,2), data = ARData)
drop1(fit2, test="Chi")
```

```
## Single term deletions
##
## Model:
## Surv(ti, Indikator) ~ Alter * poly(KI, 2)
##                     Df    AIC    LRT Pr(>Chi)
## <none>                 121.48
```

```
## Alter:poly(KI, 2)  2 128.26 10.772  0.00458 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# der Interaktionsterm ist signifikant
summary(fit2)
```

```
## Call:
## coxph(formula = Surv(ti, Indikator) ~ Alter * poly(KI, 2), data = ARData)
##
##   n= 50, number of events= 18
##
##                          coef  exp(coef)   se(coef)      z Pr(>|z|)
## Alter              -2.037e+00  1.305e-01  2.476e+00 -0.822   0.4108
## poly(KI, 2)1        1.511e+00  4.532e+00  2.991e+00  0.505   0.6134
## poly(KI, 2)2        3.908e+00  4.981e+01  3.259e+00  1.199   0.2304
## Alter:poly(KI, 2)1 -3.338e+01  3.183e-15  2.273e+01 -1.469   0.1419
## Alter:poly(KI, 2)2 -2.378e+01  4.724e-11  1.337e+01 -1.778   0.0754 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## Alter              1.305e-01  7.664e+00 1.018e-03     16.72
## poly(KI, 2)1       4.532e+00  2.207e-01 1.289e-02   1593.22
## poly(KI, 2)2       4.981e+01  2.008e-02 8.381e-02  29603.28
## Alter:poly(KI, 2)1 3.183e-15  3.142e+14 1.436e-34  70556.76
## Alter:poly(KI, 2)2 4.724e-11  2.117e+10 1.953e-22     11.43
##
## Concordance= 0.76  (se = 0.057 )
## Likelihood ratio test= 16.52  on 5 df,   p=0.006
## Wald test            = 10.85  on 5 df,   p=0.05
## Score (logrank) test = 11.23  on 5 df,   p=0.05
```

```r
emmeans(fit2, ~ Alter*poly(KI,2))
```

```
##  Alter       KI  emmean   SE  df asymp.LCL asymp.UCL
##      0 -2.5e-18 -0.4082 0.34 Inf     -1.08     0.259
##      1 -2.5e-18  0.0386 1.31 Inf     -2.52     2.599
##
## Results are given on the log (not the response) scale.
## Confidence level used: 0.95
```

Die Interaktion ist signifikant (p = 0,005). Nach emmeans (dabei werden die Interaktionsterme interpretierbar dargestellt) hat das Alter einen positiven Effekt auf das Überleben (ältere Tiere hatten eine höhere Überlebenswahrscheinlichkeit), Körperkondition schauen wir uns gleich noch genauer an.

## Modelle vergleichen

Wir können auch hier verschiedene Modelle mittels AIC vergleichen (siehe Kapitel 9.3).

```r
AIC(fit2)
```

```
## [1] 121.4833
```

```r
AIC(coxph(Surv(ti, Indikator) ~ Alter * KI, data = ARData))
```

```
## [1] 125.786
```

```r
AIC(coxph(Surv(ti, Indikator) ~ Alter + poly(KI,2), data = ARData))
```

```
## [1] 128.2556
```

```r
AIC(coxph(Surv(ti, Indikator) ~ Alter, data = ARData))
```

```
## [1] 129.9209
```

```r
AIC(coxph(Surv(ti, Indikator) ~ KI, data = ARData))
```

```
## [1] 125.605
```

```r
AIC(coxph(Surv(ti, Indikator) ~ poly(KI,2), data = ARData))
```

```
## [1] 127.3074
```

Wir sehen, dass unser ursprünglich gebautes Modell (basierend auf biologischen Annahmen und einer guten Datenexploration zu Beginn) den deutlich niedrigsten AIC Wert aufweist. Wir haben also intiutiv das beste Modell gebaut. Schauen wir uns noch die p-Werte an. Dazu schauen wir uns unser gebautes Modell sowie eins ohne Interaktion an.

```r
drop1(fit2, test="Chi")
```

```
## Single term deletions
##
## Model:
## Surv(ti, Indikator) ~ Alter * poly(KI, 2)
##                    Df    AIC    LRT Pr(>Chi)
## <none>                121.48
## Alter:poly(KI, 2)  2 128.26 10.772  0.00458 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
drop1(coxph(Surv(ti, Indikator) ~ Alter + poly(KI,2), data = ARData),
      test="Chi")
```

```
## Single term deletions
##
## Model:
## Surv(ti, Indikator) ~ Alter + poly(KI, 2)
```

```
##              Df    AIC    LRT Pr(>Chi)
## <none>          128.26
## Alter         1 127.31 1.0518  0.30508
## poly(KI, 2)   2 129.92 5.6654  0.05886 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Der Gesamtterm (Alter * poly(KI,2)) ist signifikant. Bei einem Modell ohne Interaktion wäre nur noch die quadratische Körperkondition signifikant.

## Modellüberprüfung

Auch dieses Modell können wir auf Einhaltung der Annahmen prüfen. Dazu schauen wir uns den Waldtest an, der angibt, ob die finalen Koeffizienten vom Ausgangswert abweichen. Außerdem testen wir die Annahme des proportionalen Risikos für die Anpassung eines Cox-Regressionsmodells (`coxph`).

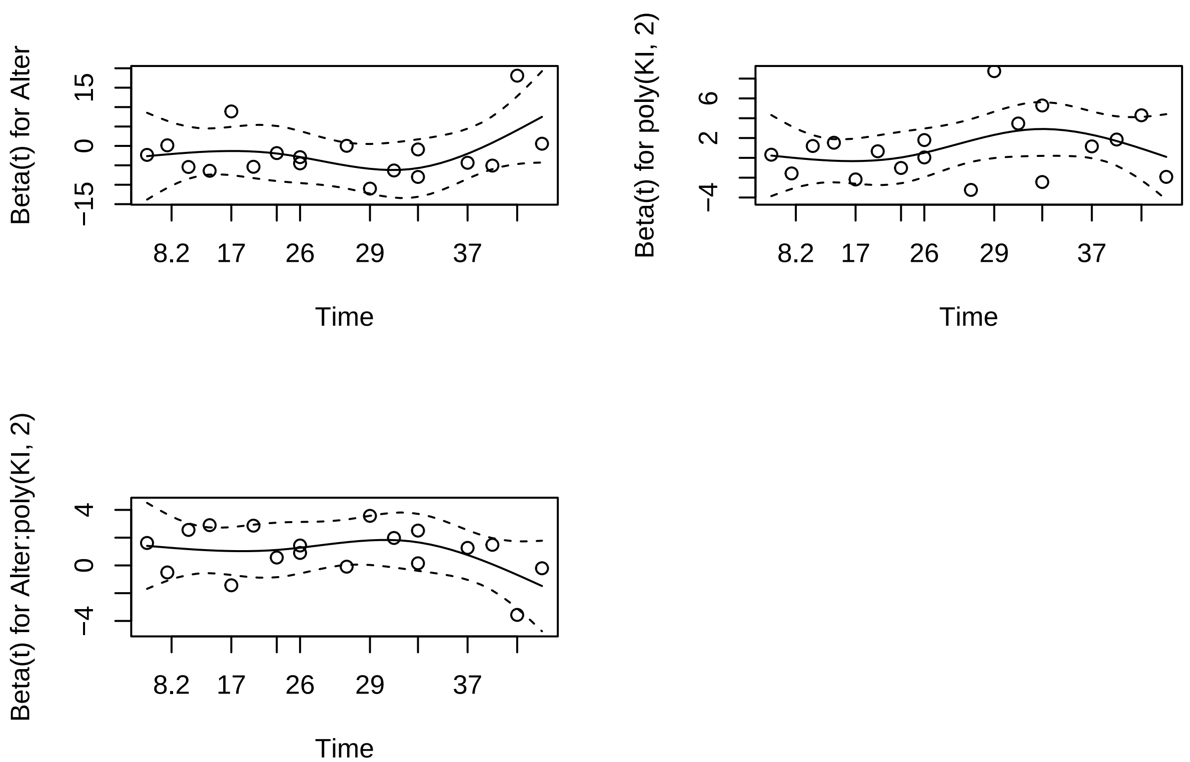
```
summary(fit2)
```

```
## Call:
## coxph(formula = Surv(ti, Indikator) ~ Alter * poly(KI, 2), data = ARData)
##
##   n= 50, number of events= 18
##
##                          coef  exp(coef)   se(coef)      z Pr(>|z|)
## Alter              -2.037e+00  1.305e-01  2.476e+00 -0.822   0.4108
## poly(KI, 2)1        1.511e+00  4.532e+00  2.991e+00  0.505   0.6134
## poly(KI, 2)2        3.908e+00  4.981e+01  3.259e+00  1.199   0.2304
## Alter:poly(KI, 2)1 -3.338e+01  3.183e-15  2.273e+01 -1.469   0.1419
## Alter:poly(KI, 2)2 -2.378e+01  4.724e-11  1.337e+01 -1.778   0.0754 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                    exp(coef) exp(-coef) lower .95 upper .95
## Alter              1.305e-01  7.664e+00 1.018e-03     16.72
## poly(KI, 2)1       4.532e+00  2.207e-01 1.289e-02   1593.22
## poly(KI, 2)2       4.981e+01  2.008e-02 8.381e-02  29603.28
## Alter:poly(KI, 2)1 3.183e-15  3.142e+14 1.436e-34  70556.76
## Alter:poly(KI, 2)2 4.724e-11  2.117e+10 1.953e-22     11.43
##
## Concordance= 0.76  (se = 0.057 )
## Likelihood ratio test= 16.52  on 5 df,   p=0.006
## Wald test            = 10.85  on 5 df,   p=0.05
## Score (logrank) test = 11.23  on 5 df,   p=0.05
```

```
temp <- cox.zph(fit2)
print(temp) # Ergebnisse darstellen
```

```
##                   chisq df     p
## Alter              1.36  1 0.243
## poly(KI, 2)        3.56  2 0.169
## Alter:poly(KI, 2)  5.18  2 0.075
## GLOBAL             5.66  5 0.341
```

```r
# Modellvalidierung als Plot darstellen
# Achsenbeschriftungen werden automatisch erstellt und sind daher Englisch
par(mfrow=c(2,2))
plot(temp)
```




Der Wald-Test ist signifikant, die Parameter haben also einen Einfluss auf das Überleben. Der Anpassungs-test wiederum zeigt keine signifikanten Abweichungen von den Modellannahmen. Unser Modell passt also gut.

### Überlebensgrafiken

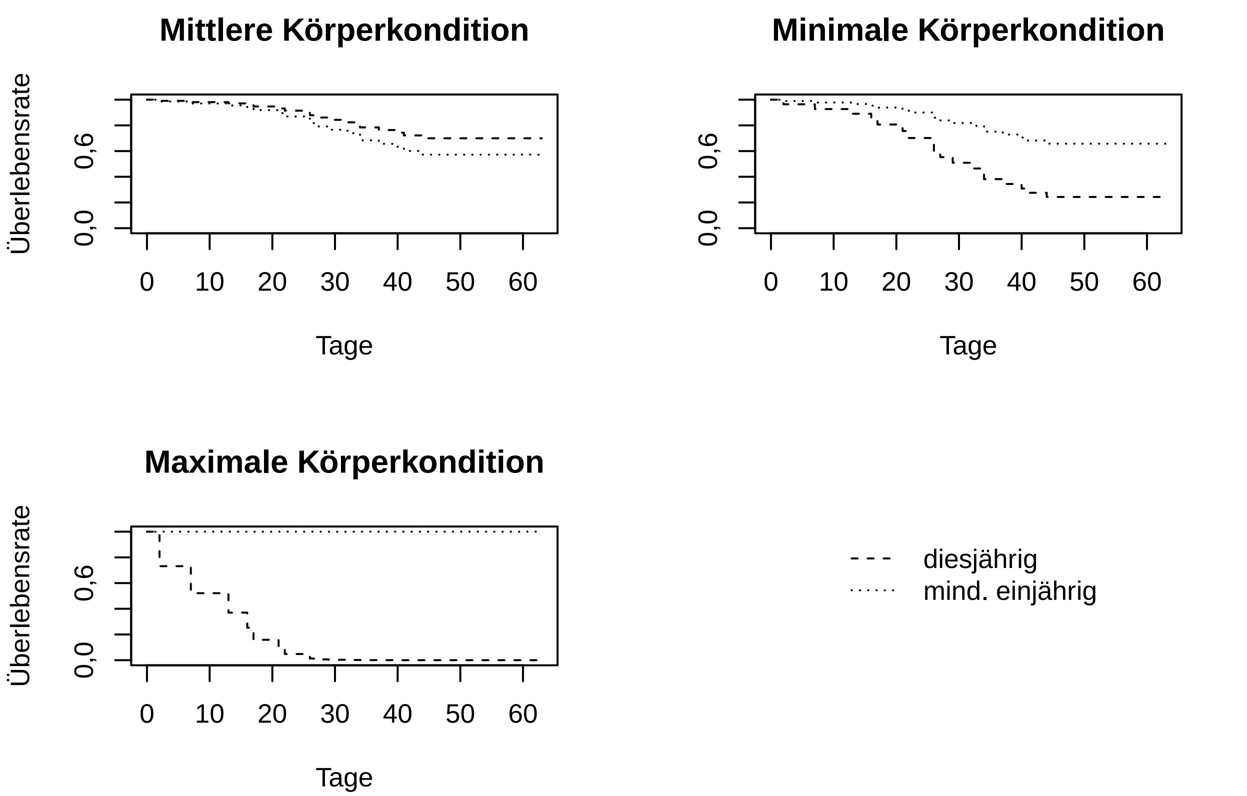
```r
par(mfrow=c(2,2))
plot(survfit(fit2, newdata=data.frame(Alter=c(0,1), KI=mean(ARData$KI))),
     xlab = "Tage", ylab="Überlebensrate", lty=c(2,3), yaxt="n",
     main="Mittlere Körperkondition")
 axis(
   side = 2,
   at = seq(0,1,.2),
   labels = format(seq(0,1,.2), decimal.mark = ",", nsmall = 1))
#legend(40, 1, c("diesjährig", "mind. einjährig"), lty = 2:3)

plot(survfit(fit2, newdata=data.frame(Alter=c(0,1), KI=min(ARData$KI))),
     xlab = "Tage", ylab="", lty=c(2,3), yaxt="n",
     main="Minimale Körperkondition")
```

```
 axis(
   side = 2,
   at = seq(0,1,.2),
   labels = format(seq(0,1,.2), decimal.mark = ",", nsmall = 1))
#legend(40, 1, c("diesjährig", "mind. einjährig"), lty = 2:3)

plot(survfit(fit2, newdata=data.frame(Alter=c(0,1), KI=max(ARData$KI))),
     xlab = "Tage", ylab="Überlebensrate", lty=c(2,3), yaxt="n",
     main="Maximale Körperkondition")
 axis(
   side = 2,
   at = seq(0,1,.2),
   labels = format(seq(0,1,.2), decimal.mark = ",", nsmall = 1))
#legend(20, 1, c("diesjährig", "mind. einjährig"), lty = 2:3, bty="n")

plot(0, xaxt = 'n', yaxt = 'n', bty = 'n', pch = '', ylab = '', xlab = '')
legend("top", c("diesjährig", "mind. einjährig"), lty = 2:3, bty="n")
```




Jungtiere haben bei mittlerer Körperkondition die höchste Überlebensrate, bei hohen sterben sie fast sicher. Bei adulten ist es umgekehrt, sie haben bei hoher Körperkondition die höchsten Überlebensraten, wobei der Unterschied zwischen minimaler Körperkondition und mittlerer Körperkondition bei ihnen sehr gering ist.

# Literaturverzeichnis


Conroy, M.J., Costanzo, G.R., Stotts, D.B. (1989): Winter survival of female American black ducks on the Atlantic coast. — The Journal of Wildlife Management 53: 99–109.

Pollock, K.H., Winterstein, S.R., Conroy, M.J. (1989): Estimation and analysis of survival distributions for radio-tagged animals. — Biometrics 45: 99–109.

Sjoberg, D., Baillie, M., Fruechtenicht, C., Haesendonckx, S., Treis, T. (2024): ggsurvfit: Flexible time-to-event figures. R package version 1.1.0, https://CRAN.R-project.org/package=ggsurvfit.

Therneau, T. (2024): A package for survival analysis in R. R package version 3.8-3, https://CRAN.R-project.org/package=survival

Therneau, T.M., Grambsch, P.M. (2000): Modeling survival data: Extending the cox model. — Springer, New York.