Bruchpunktmodelle

Volker Schmid

17. Juli 2017

Bruchpunktmodelle

Bruchpunktmodelle

Bruchpunktmodelle

- Beim Bruchpunktmodell ändert sich ein Parameter eines Modells an einem Punkt sprungartig
- In der Regel Anwendung auf Zeitreihen
- Theoretisch diskrete oder stetige Zeit möglich

Beispiele:

- Aktienkurse
- Schlafdaten
- Krankheiten mit Schüben
- Stärke von Sportmannschaften
- Zeitreihen von Krankheitsfällen (Bruchpunkt: Ausbruch einer Epidemie)

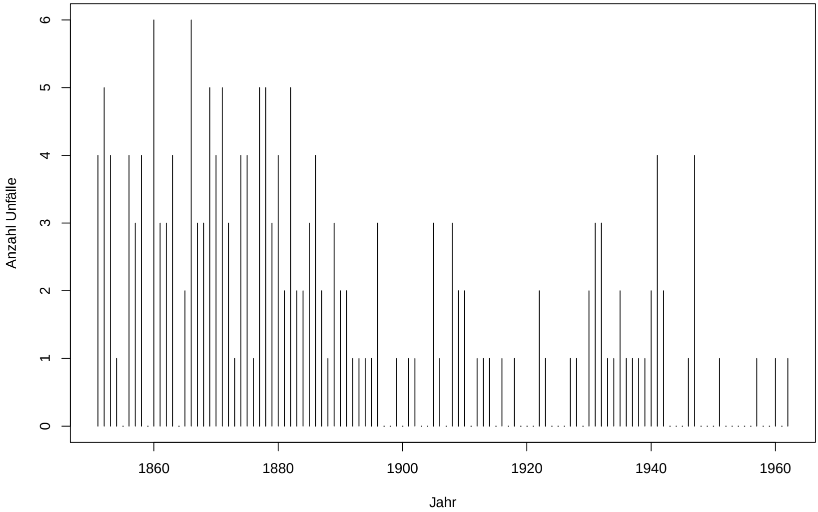
Beispiel: Unfälle in englischen Kohlebergwerken I

Der Datensatz {coal.txt} enthält die jährliche Anzahl von Unfällen in englischen Kohlebergwerken wärend der Jahre 1851-1962 (insgesamt 112 Jahre). Ein Plot der Daten zeigt einen deutlichen Rückgang der Unfälle ab etwa 1900.

```
coal <- read.table("../data/coal.txt", header = T)
```

```
plot(coal$year, coal$disasters, type = "h",
     xlab = "Jahr", ylab = "Anzahl Unfälle")
```

Beispiel: Unfälle in englischen Kohlebergwerken II



Bruckpunktmodell

Ein Bruchpunktmodell für diese Daten $\mathbf{y} = (y_1, \ldots, y_{112})$ hat folgende Form:

$$Y_i \sim \begin{cases} Po(\lambda_1), & i = 1, \ldots, \theta, \\ Po(\lambda_2), & i = \theta + 1, \ldots, 112, \end{cases}$$

Prioris:

- $\lambda_i \mid \alpha \sim Ga(3, \alpha)$ für $i = 1, 2$
- $\alpha \sim Ga(10, 10)$
- $\theta \sim U\{1, \ldots, 112\}$.

Posteriori

Die gemeinsame Posteriori-Dichte ist gegeben durch:

$$\begin{aligned}
p(\lambda_1, \lambda_2, \alpha, \theta \mid \mathbf{y}) &\propto p(\mathbf{y} \mid \lambda_1, \lambda_2, \alpha, \theta)\, p(\lambda_1, \lambda_2, \alpha, \theta) \\
&= p(\mathbf{y} \mid \lambda_1, \lambda_2, \alpha, \theta)\, p(\lambda_1 \mid \lambda_2, \alpha, \theta) \\
&\qquad \cdot p(\lambda_2 \mid \alpha, \theta)\, p(\alpha \mid \theta)\, p(\theta) \\
&= p(\mathbf{y} \mid \lambda_1, \lambda_2, \theta)\, p(\lambda_1 \mid \alpha)\, p(\lambda_2 \mid \alpha)\, p(\alpha) p\,(\theta) \\
&\propto \left(\prod_{i=1}^{\theta} \lambda_1^{y_i} \exp(-\lambda_1) \right) \left(\prod_{i=\theta+1}^{n} \lambda_2^{y_i} \exp(-\lambda_2) \right) \\
&\qquad \cdot \underbrace{\alpha^3 \lambda_1^{3-1} \exp(-\alpha\lambda_1)}_{\lambda_1 \mid \alpha \sim Ga(3,\alpha)} \underbrace{\alpha^3 \lambda_2^{3-1} \exp(-\alpha\lambda_2)}_{\lambda_2 \mid \alpha \sim Ga(3,\alpha)} \\
&\qquad \cdot \underbrace{\alpha^{10-1} \exp(-10\alpha)}_{\alpha \sim Ga(10,10)} \underbrace{\mathrm{I}(1 \le \theta \le 112)}_{\theta \sim \mathrm{U}\{1,\ldots,112\}} .
\end{aligned}$$

Full Conditionals I

$$
\begin{aligned}
\lambda_1 \mid \lambda_2, \alpha, \theta, \mathbf{y} &\propto \left(\prod_{i=1}^{\theta} \lambda_1^{y_i} \exp(-\lambda_1) \right) \lambda_1^{3-1} \exp(-\alpha\lambda_1) \\
&= \lambda_1^{3+\sum_{i=1}^{\theta} y_i - 1} \exp\left(-\lambda_1(\theta + \alpha) \right)
\end{aligned}
$$

Dies ist der Kern einer $Ga(3 + \sum_{i=1}^{\theta} y_i, \theta + \alpha)$-Verteilung.

$$
\begin{aligned}
\lambda_2 \mid \lambda_1, \alpha, \theta, \mathbf{y} &\propto \left(\prod_{i=\theta+1}^{n} \lambda_2^{y_i} \exp(-\lambda_2) \right) \lambda_2^{3-1} \exp(-\alpha\lambda_2) \\
&= \lambda_2^{3+(\sum_{i=\theta+1}^{n} y_i)-1} \exp\left(-\lambda_2(\alpha + [n - \theta]) \right)
\end{aligned}
$$

Dies ist der Kern einer $Ga(3 + \sum_{\theta+1}^{112} y_i, \alpha + 112 - \theta)$, da $n = 112$.

Full Conditionals II

$$\begin{aligned}\alpha \mid \lambda_1, \lambda_2, \theta, \mathbf{y} &\propto \alpha^3 \exp(-\alpha\lambda_1)\alpha^3 \exp(-\alpha\lambda_2)\alpha^{10-1} \exp(-10\alpha) \\ &= \alpha^{16-1} \exp\big(-\alpha(10 + \lambda_1 + \lambda_2)\big)\end{aligned}$$

Dies ist der Kern einer $Ga(16, 10 + \lambda_1 + \lambda_2)$.

Full Conditionals III

$$
\begin{aligned}
\theta \mid \lambda_1, \lambda_2, \alpha, \mathbf{y} &\propto \left(\prod_{i=1}^{\theta} \lambda_1^{y_i} \exp(-\lambda_1) \right) \left(\prod_{i=\theta+1}^{n} \lambda_2^{y_i} \exp(-\lambda_2) \right) \\
&\quad \cdot \mathrm{I}(1 \leq \theta \leq 112) \\
&\propto \left(\prod_{i=1}^{\theta} \lambda_1^{y_i} \exp(-\lambda_1) \right) \left(\prod_{i=\theta+1}^{n} \lambda_2^{y_i} \exp(-\lambda_2) \right) \\
&\quad \cdot \frac{\left(\prod_{i=1}^{\theta} \lambda_2^{y_i} \exp(-\lambda_2) \right)}{\left(\prod_{i=1}^{\theta} \lambda_2^{y_i} \exp(-\lambda_2) \right)} \mathrm{I}(1 \leq \theta \leq 112) \\
&\propto \left(\prod_{i=1}^{\theta} \lambda_1^{y_i} \lambda_2^{-y_i} \exp(-\lambda_1) \exp(\lambda_2) \right) \mathrm{I}(1 \leq \theta \leq 112) \\
&= \exp\left(\theta(\lambda_2 - \lambda_1) \right) \left(\frac{\lambda_1}{\lambda_2} \right)^{\sum_{i=1}^{\theta} y_i} \mathrm{I}(1 \leq \theta \leq 112)
\end{aligned}
$$

Full Conditionals IV

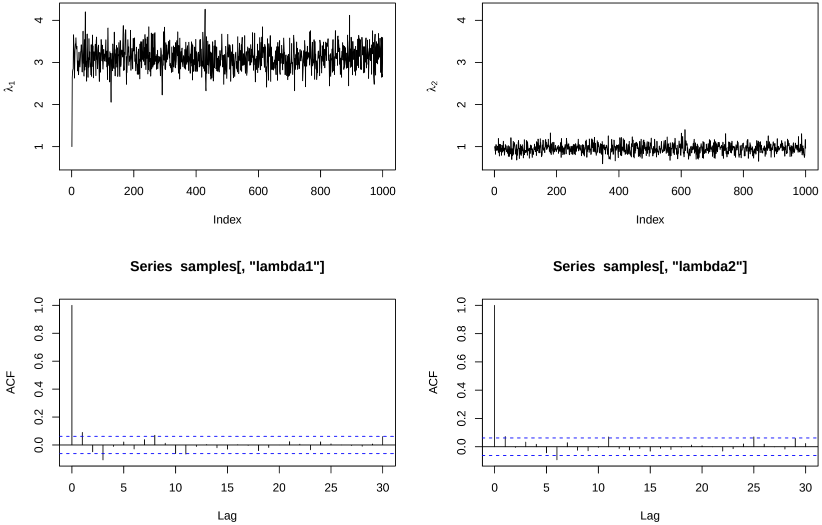
$$\theta \mid \lambda_1, \lambda_2, \alpha, \mathbf{y} = \exp\left(\theta(\lambda_2 - \lambda_1)\right) \exp\left\{ \left(\sum_{i=1}^{\theta} y_i \right) \log\left(\frac{\lambda_1}{\lambda_2} \right) \right\}$$
$$\cdot \, \mathrm{I}(1 \leq \theta \leq 112)$$

Ergebnisse I

```r
samples <- breakpoint.gibbs(1000, y = coal$disaster)

# plots
par(mfrow = c(2,2))
plot(samples[,"lambda1"], type = "l",
     ylab = expression(lambda[1]),
     ylim = c(min(samples[,1:2]), max(samples[,1:2])))
plot(samples[,"lambda2"], type = "l",
     ylab = expression(lambda[2]),
     ylim = c(min(samples[,1:2]), max(samples[,1:2])))
acf(samples[,"lambda1"])
acf(samples[,"lambda2"])
```

Ergebnisse II

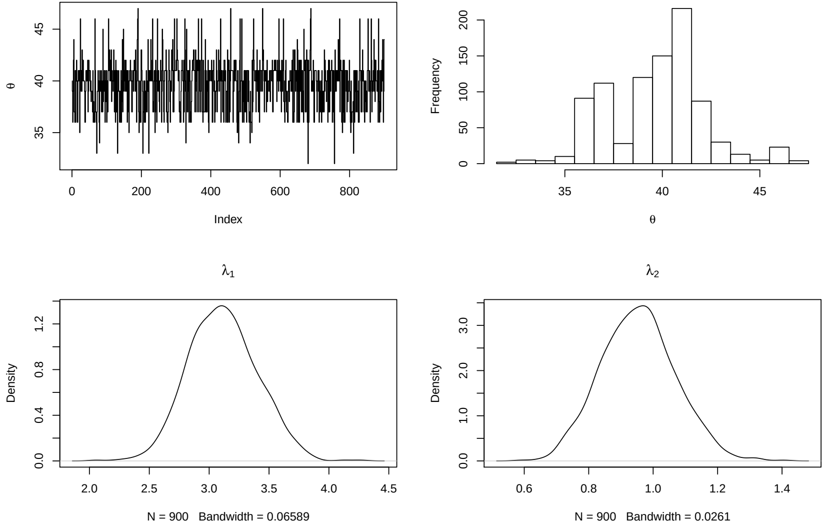


Betrachtung der Pfade: 1000 Realisationen reichen aus, fast kein Burn-in notwendig, kaum Autokorrelation.

Ergebnisse III

```
burnin <- 1:100 # hier großzügiger burn-in angewendet
par(mfrow = c(2,2))
# theta
plot(samples[-burnin, "theta"], type = "S",
     ylab = expression(theta))
hist(samples[-burnin, "theta"], xlab = expression(theta),
     breaks = ((min(samples[-burnin, "theta"])-1):
                 (max(samples[-burnin, "theta"])))+0.5,
     main = "" )
# lambda_1
plot(density(samples[-burnin, "lambda1"]), main =
       expression(lambda[1]))
# lambda_2
plot(density(samples[-burnin, "lambda2"]), main =
       expression(lambda[2]))
```

Ergebnisse IV



- Mittelwert und Median für alle Parameter:

Ergebnisse V

```r
apply(samples[-burnin,], MAR = 2, mean)
```

```
##    lambda1    lambda2      alpha      theta
##  3.0959415  0.9442205  1.1388692 39.9166667
```

```r
apply(samples[-burnin,], MAR = 2, median)
```

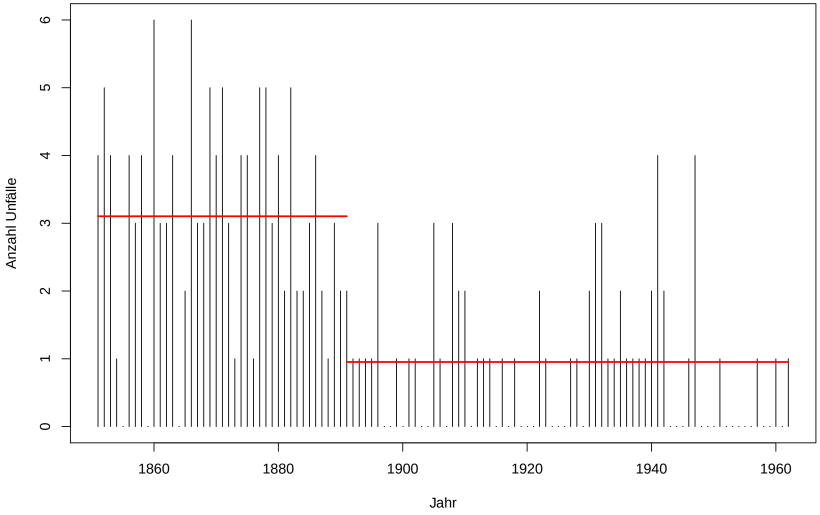
```
##    lambda1    lambda2      alpha      theta
##  3.0870221  0.9445069  1.1162097 40.0000000
```

Median-Modell im Histogramm:

Ergebnisse VI

```r
plot(coal$year, coal$disasters, type = "h",
     xlab = "Jahr", ylab = "Anzahl Unfälle")
medianmodel <- apply(samples[-burnin,], MAR = 2, median)
jahr1 <- min(coal$year)
lines(c(jahr1, jahr1 + medianmodel["theta"]),
      rep(medianmodel["lambda1"], 2), col=2, lwd = 2)
lines(c(jahr1 + medianmodel["theta"], max(coal$year)),
      rep(medianmodel["lambda2"], 2), col=2, lwd = 2)
```

Ergebnisse VII



Ergebnisse VIII

```r
library(coda)
samples2 <- as.mcmc(samples)
summary(samples2)
```

```
##
## Iterations = 1:1000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 1000
##
## 1. Empirical mean and standard deviation for each variab
##    plus standard error of the mean:
##
##            Mean     SD Naive SE Time-series SE
## lambda1  3.0996 0.3059 0.009672       0.010393
## lambda2  0.9447 0.1154 0.003650       0.003837
```

Ergebnisse IX

```
## alpha    1.1349 0.2861 0.009047       0.009047
## theta   39.9040 2.4517 0.077529       0.080051
##
## 2. Quantiles for each variable:
##
##            2.5%     25%    50%    75%  97.5%
## lambda1  2.5331  2.8983  3.087  3.299  3.722
## lambda2  0.7224  0.8665  0.944  1.018  1.180
## alpha    0.6491  0.9275  1.112  1.326  1.734
## theta   36.0000 39.0000 40.000 41.000 46.000
```

```r
plot(samples2)
```

Ergebnisse X