Beispiel: Stationäre Verteilung einer MK mit 2 Zuständen $\{0, 1\}$

Berechnung: Löse zuerst $\tilde{\pi} = \tilde{\pi} P$ ohne Nebenbedingung $\pi \mathbb{1} = 1$, normiere anschließend durch $\pi = \frac{\tilde{\pi}}{\tilde{\pi}\mathbb{1}}$. (Eliminiere von den linearen Gleichungen ggf. zunächst die komplizierteste.)

$$P = \begin{bmatrix} 1-p_0 & p_0 \\ p_1 & 1-p_1 \end{bmatrix}, \quad \pi = [\pi_0, \pi_1]$$

$$(\pi_0, \pi_1) \begin{bmatrix} 1-p_0 & p_0 \\ p_1 & 1-p_1 \end{bmatrix} \stackrel{!}{=} (\pi_0, \pi_1)$$

$$\begin{aligned} \Leftrightarrow \qquad & \pi_0(1-p_0) + \pi_1 p_1 = \pi_0, \quad \Leftrightarrow \quad & \boxed{\pi_0 p_0 = \pi_1 p_1}, \\ & \pi_0 p_0 + \pi_1(1-p_1) = \pi_1 & \pi_0 p_0 = \pi_1 p_1 \end{aligned}$$

Lösung 1: $\tilde{\pi}_0, \tilde{\pi}_1$ unnormiert; keine Nebenbedingung

$$\widetilde{\pi}_0 = p_1 \Rightarrow \widetilde{\pi}_1 = p_0$$

$$\Leftrightarrow \boxed{\pi_0 = \frac{p_1}{p_0 + p_1}, \quad \pi_1 = \frac{p_0}{p_0 + p_1}}$$

Lösung 2: $\tilde{\pi}_0 + \tilde{\pi}_1 = 1$; Nebenbedingung $\pi_0 + \pi_1 = 1 \Leftrightarrow \pi_1 = 1 - \pi_0$

$$\begin{aligned}
& \pi_0(1 - p_0) + (1 - \pi_0)p_1 = \pi_0 \\
\Leftrightarrow \quad & -\pi_0 p_0 + p_1 - \pi_0 p_1 = 0 \\
\Leftrightarrow \quad & \pi_0(p_0 + p_1) = p_1 \\
\Leftrightarrow \quad & \boxed{\pi_0 = \frac{p_1}{p_0 + p_1}, \quad \pi_1 = \frac{p_0}{p_0 + p_1}}
\end{aligned}$$

Die Grenzverteilung π (mit $\pi = \pi P$) heißt auch stationäre Verteilung der Markovkette mit Übergangsmatrix P. Denn:
Wählt man als Startverteilung $p(0) = \pi$, so gilt:

$$\begin{aligned} & p(1) = p(0)P = \pi P = \pi \\ & \vdots \\ \Rightarrow \quad & p(t+1) = p(t)P = \pi P = \pi \quad \text{für alle } t. \end{aligned}$$

3.4 Instationäre und inhomogene Markov-Ketten

- Häufige Datensituation in Anwendungen:
  - Y_t diskrete bzw. kategoriale Ziel- (oder Response)variable.
  - $z = (z_1, \ldots, z_p)'$ Vektor von Kovariablen
  - $T = \mathbb{N}_0$
- Zeitreihendaten: (Y_t, Z_t), $t = 1, \ldots, T$, d.h. ein Pfad eines stochastischen Prozesses Y und eines Kovariablen-Prozesses werden von 1 bis T beobachtet.
- Longitudinaldaten: $\{(Y_{nt}, z_{nt}), t = 1, \ldots, T\}$, d.h. jeweils Pfade für $n = 1, \ldots, N$ Individuen / Objekte beobachtet.

- Mögliche Situationen:
  - N groß, T klein,
  - N klein, T groß (N=1: Zeitreihe),
  - N, T mittel bis groß.

- Datenstruktur der Pfade passt zu einer Markov-Kette für Y, aber
  - Zeitliche Stationarität fraglich $\Rightarrow$ Markov-Kette mit nichtstationären Übergangswahrscheinlichkeiten $p_{ij}(t)$
  - Homogenität der Population fraglich,

$$\Rightarrow P(Y_{n,t+1} = j | Y_{n,t} = i) = p_{ij,n}(t).$$

    von n abhängig. Bei homogener Population würde man

$$p_{ij,n}(t) \equiv p_{ij}(t), \quad n = 1, \ldots, N$$

    annehmen.
  - Auch Markov-Eigenschaft, zumindest 1. Ordnung, fraglich.

3.4.1 Instationäre und inhomogene binäre Markov-Kette 1. Ordnung

- Modell für (binäre) Zeitreihendaten $\{Y_t,\ t \in \mathbb{N}_0\}$:

$$p_{ij}(t) = P(Y_{t+1} = j | Y_t = i), \quad i, j \in \{0, 1\}.$$

- Modell für Longitudinaldaten $\{Y_{nt}, t \in \mathbb{N}_0\}$, $n = 1, \ldots, N$ mit Kovariablen $\{z_{nt}, t \in \mathbb{N}_0\}$:

$$p_{ij,n}(t) = P(Y_{n,t+1} = j | Y_{nt} = i, z_{nt}).$$

  Die Kovariablen können dabei zeitabhängig oder auch zeitunabhängig sein.

- Es reicht die Übergangswahrscheinlichkeiten für $j = 1$ zu modellieren, da $p_{i0,n}(t) + p_{i1,n}(t) = 1$ gilt. Basis für die Modellierung sind binäre Regressionsmodelle, insbesondere Logit- oder Probitmodelle (vgl. Lineare Modelle, Generalisierte Regression).

- Separate Modellierung der Übergangswahrscheinlichkeiten:

$$p_{01,n}(t) = h(w'_{nt}\beta_0), \quad p_{11,n}(t) = h(z'_{nt}\beta_1).$$

h ist die sogenannte Responsefunktion $h : \mathbb{R} \to [0, 1]$. Für

$$h(x) = \exp(x)/(1 + \exp(x))$$

bzw.

$$h(x) = \Phi(x)$$

erhält man das Logit- bzw. das Probitmodell. Die Kovariablenvektoren w_{nt}, z_{nt} können auch identisch sein.

- Im reinen Zeitreihenfall (n=1) wird durch $\eta_t = w'_t\beta_0$ bzw. $\eta_t = z'_t\beta_1$ z.B. ein zeitlicher Trend parametrisch spezifiziert, z.B. ein linearer Trend

$$w'_t\beta_0 = \beta_{00} + \beta_{01}t$$

oder ein (stückweise) polynomialer Trend. Für Longitudinaldaten enthalten w_{nt}, z_{nt} in der Regel auch individuelle Kovariablen.

Probleme bei linearem Trend für inhomogene Markovketten:

$$P(y_{t+1} = 1 | y_t = 0) = \frac{\exp(\beta_{00} + \beta_{01} t)}{1 + \exp(\beta_{00} + \beta_{01} t)} \xrightarrow[t \to \infty]{} \begin{cases} 0, & \beta_{01} < 0 \\ 1, & \beta_{01} > 0 \end{cases}$$

$\Rightarrow$ Deterministisches Verhalten der Markovkette für $t \to \infty$.

- Lässt man statt einer parametrischen Trendfunktion $w_t' \beta_0$ eine nichtparametrisch spezifizierte Funktion $f_0(t)$ zu, so führt dies zu non- und semiparametrischen binären Regressionsmodellen (Generalisierte Regression, Computerintensive Verfahren).

- Konditionale Modellierung: Simultanes, autoregressives Modell, in dem Y_t als Regressor fungiert:

$$P(Y_{t+1} = 1 | Y_t, w_t).$$

- Beispiele:

$$\log\left(\frac{P(Y_{t+1}=1|Y_t,w_t)}{P(Y_{t+1}=0|Y_t,w_t)}\right) = w_t'\beta + Y_t\alpha \qquad \text{(additiv)}$$

$$\log\left(\frac{P(Y_{t+1}=1|Y_t,w_t)}{P(Y_{t+1}=0|Y_t,w_t)}\right) = w_t'\beta + Y_t w_t'\alpha \qquad \text{(mit Interaktion)}$$

- Das Modell mit Interaktion ist äquivalent zur separaten Modellierung (Einsetzen von $Y_t = 0$ bzw. 1 liefert die Zusammenhänge).

3.4.2 Weitere inhomogene Markov-Ketten

- Eine inhomogene Markov-Kette 2. Ordnung besitzt die Übergangswahrscheinlichkeiten

$$P(Y_{t+1} = j | Y_t = i_1, Y_{t-1} = i_2) =: p_{i_1 i_2, j}(t).$$

- Im binären Fall $Y_t \in \{0, 1\}$ genügt es wieder, $p_{i_1\, i_2\, 1}(t)$ zu spezifizieren.

- Separate Modellierung: 4 separate binäre Regressionsmodelle für die Paare $(i_1, i_2) = (0,0), (0,1), (1,0)$ und $(1,1)$.

$$\log\left(\frac{p_{i_1 i_2,1}(t)}{p_{i_1 i_2,0}(t)}\right) = z_t' \beta_{i_1 i_2},$$

- Simultane (autoregressive) Logit-Modelle:

$$\log\left(\frac{P(Y_{t+1} = 1 | Y_t, Y_{t-1}, z_t)}{P(Y_{t+1} = 0 | Y_t, Y_{t-1}, z_t}\right) = z_t' \beta + \alpha_1 Y_t + \alpha_2 Y_{t-1}$$

oder

$$\log\left(\frac{P(Y_{t+1}=1|Y_t,Y_{t-1},z_t)}{P(Y_{t+1}=0|Y_t,Y_{t-1},z_t}\right) = z_t'\beta + Y_t z_t'\alpha_1 + Y_{t-1}z_t'\alpha_2 + Y_tY_{t-1}z_t'\alpha_3$$

- Einsetzen der verschiedenen 0/1-Kombinationen für Y_t, Y_{t-1} liefert

$$\beta_{00}=\beta, \quad \beta_{01}=\beta+\alpha_1, \quad \beta_{10}=\beta+\alpha_2, \quad \beta_{11}=\beta+\alpha_1+\alpha_2+\alpha_3.$$

  Durch Nullsetzen von Komponenten in $\alpha_1, \alpha_2, \alpha_3$ erhält man parametersparsamere Submodelle.

- Analoges Vorgehen bei Markov-Ketten höherer Ordnung möglich.

- Allgemeinere inhomogene Markov-Ketten (mit mehr als zwei Zuständen) lassen sich durch kategoriale autoregressive Regressionsmodelle behandeln.

3.5 Statistische Inferenz

3.5.1 Likelihood-Inferenz für stochastische Prozesse

- Beobachtung eines Pfades: Die Daten $(x_0), x_1, \ldots, x_T$ werden aufgefasst als Realisierungen

$$x_0 = X_0(\omega), x_1 = X_1(\omega), \ldots, x_t = X_t(\omega), \ldots, x_T = X_T(\omega)$$

eines stochastischen Prozesses $\{X_t, t \in \mathbb{N}_0\}$.

- Bei Markov-Ketten handelt es sich also um eine diskrete oder kategoriale Zeitreihe, die als bis T vorliegende Realisierung der Markov-Kette aufgefasst wird.

- Sei θ ein Vektor von unbekannten Parametern. Dann ist die Likelihoodfunktion

$$L(\theta|X) := f(x_0, \ldots, x_T|\theta)$$

die gemeinsame Dichte von $X_0, \ldots, X_T$, ausgewertet an $x_0, \ldots, x_T$ und als Funktion von θ betrachtet.

- Faktorisierung der Likelihood in das Produkt bedingter Dichten:

$$\begin{aligned} L(\theta|X) &= f(x_T, x_{T-1}, \ldots, x_0|\theta) \\ &= f_T(x_T|x_{T-1}, \ldots, x_0, \theta) \cdot f(x_{T-1}, \ldots, x_0|\theta) \\ &= \ldots \\ &= \prod_{t=1}^{T} f_t(x_t|x_{t-1}, \ldots, x_0, \theta) f_0(x_0|\theta). \end{aligned}$$

- Bei Markov-Ketten ergibt sich

$$L(\theta|X) = \prod_{t=1}^{T} f_t(x_t|x_{t-1}, \theta) f_0(x_0|\theta).$$

Bei homogenen Markov-Ketten kann der Index t bei f_t entfallen.

- Maximum-Likelihood-Prinzip: Wähle $\hat{\theta}_{ML}$ als Maximierer von $L(\theta|X)$.

Maximierung üblicherweise durch Übergang zur Log-Likelihood

$$l(\theta|X) := \ln L(\theta|X) = \sum_{t=1}^{T} l_t(x_t|x_{t-1}, \theta) + l_0(x_0|\theta)$$

mit den log-Likelihood Beiträgen $l_t(x_t|.) = \ln f_t(x_t|.)$.

- Falls die X_t unabhängig sind, ergibt sich wie üblich

$$L(\theta|X) = f(x_0, \ldots, x_T|\theta) = \prod_{t=0}^{T} f_t(x_t|\theta).$$

- Beobachtung mehrerer unabhängiger Pfade eines stochastischen Prozesses: Die Daten $x_{n0}, \ldots, x_{nT}$, $n = 1, \ldots, N$ werden aufgefasst als unabhängige Realisierungen

$$x_{n0} = X_0(\omega_n), \ldots, x_{nT} = X_T(\omega_n), \quad n = 1, \ldots, N$$

des Prozesses $\{X_t, t \in \mathbb{N}_0\}$. Die gemeinsame Dichte und die Likelihood ergeben sich als Produkt

$$\begin{aligned} L(\theta|X) &= \prod_{n=1}^{N} f(x_{nT}, \ldots, x_{n0}|\theta) \\ &= \prod_{n=1}^{N} \prod_{t=1}^{T} f_t(x_{nt}|x_{n,t-1}, \theta) f_0(x_{n0}|\theta) \end{aligned}$$

der beteiligten Likelihoodfunktionen für die Realisierungen $n = 1, \ldots, N$.

- Longitudinaldaten mit Kovariablen: Beobachtungen zu einer Zielvariablen Y und einem Vektor z von Kovariablen.

  Die Daten werden aufgefasst als unabhängige Realisierungen von stochastischen Prozessen Y_n, $n = 1, \ldots, N$, gegeben die Kovariablen.

  Die Likelihood für Beobachtungseinheit n ist gegeben durch

$$L_n(\theta|Y_n) = \prod_{t=1}^{T} f(y_{nt}|y_{n,t-1}, z_{nt}, \theta) f(y_{n0}).$$

  Wegen der Unabhängigkeit der verschiedenen Pfade ergibt sich die gesamte Likelihood zu

$$L(\theta|Y) = \prod_{n=1}^{N} L_n(\theta|Y_n).$$

3.5.2 Inferenz bei homogenen Markov-Ketten

- Sei X eine homogene Markov-Kette mit endlichem Zustandsraum $S = \{1, \ldots, m\}$.

- Unbekannte Parameter θ:

$P = (p_{ij})$	Übergangsmatrix
$p(0)$	Anfangsverteilung

- Für einen beobachteten Pfad

$$X_0(\omega) = i_0, X_1(\omega) = i_1, \ldots, X_T(\omega) = i_T$$

ergibt sich die Likelihoodfunktion

$$\begin{aligned} L(p(0), P) &= P\{X_0 = i_0, \ldots, X_T = i_T\} \\ &= p_i(0) p_{i_0 i_1} \cdot \ldots \cdot p_{i_{T-1} i_T}. \end{aligned}$$

- Problem: Es gibt nur eine Beobachtung zur Schätzung von $p(0)$.

$$\Rightarrow \widehat{p}_{i_0}(0) = 1, \quad \widehat{p}_i(0) = 0, i \neq i_0.$$

Betrachte deshalb nur die (auf $X_0 = x_0$) bedingte Likelihood

$$L(P) = \prod_{i,j} p_{ij}^{n_{ij}}$$

mit $n_{ij} =$ Anzahl der beobachteten Übergänge von i nach j.

- Log-Likelihood:

$$\log L(P) = \sum_{i,j} n_{ij} \log(p_{ij}).$$

- Maximieren von $\log L(P)$ unter der Nebenbedingung, dass die Zeilensummen $=1$ sind, ergibt die ML-Schätzer (mit $n_i =$ Anzahl der beobachteten Besuche in i)

$$\widehat{p}_{ij} = \frac{n_{ij}}{n_i}.$$

Beweis: Nach der Lagrange-Methode ist die Maximierung mit Nebenbedingung äquivalent zur Maximierung von

$$l^* = \sum_{i,j} n_{ij} \log p_{ij} - \sum_i \lambda_i \left(\sum_{j \in S} p_{ij} - 1 \right)$$

ohne Nebenbedingungen. [Einschub: Lagrange-Methode mit einer Nebenbedingung: Für jede Lösung a des Maximierungsproblems

$$\max f(x), \text{ so dass } g(x) = 0$$

gibt es ein $\lambda_0 \in \mathbb{R}$, so dass die Gradienten parallel sind

$$\nabla_x f(a) = \lambda_0 \nabla_x g(a),$$

d.h. a löst für $\lambda = \lambda_0$

$$\max f(x) - \lambda g(x). \quad]$$

Nullsetzen der partiellen Ableitungen von

$$l^* = \sum_{i,j} n_{ij} \, \log p_{ij} - \sum_i \lambda_i \left(\sum_{j \in S} p_{ij} - 1 \right)$$

liefert

$$\frac{\partial l^*}{\partial p_{ij}} = \frac{n_{ij}}{p_{ij}} - \lambda_i = 0,$$

also

$$n_{ij} = \lambda_i p_{ij}.$$

Summation über j liefert wegen $\sum_j p_{ij} = 1$

$$\lambda_i = \sum_{j \in S} n_{ij} = n_i,$$

und daraus schließlich

$$\widehat{p}_{ij} = \frac{n_{ij}}{n_i}.$$

- Bei ergodischen Markov-Ketten gilt für $T \to \infty$

$$\widehat{p}_{ij} \overset{a}{\sim} N\left(p_{ij}, \frac{\widehat{p}_{ij}(1-\widehat{p}_{ij})}{n_i}\right).$$

- Im homogenen Fall lassen sich die Übergangswahrscheinlichkeiten bei mehreren beobachteten Pfaden wie bei einem Pfad schätzen.

- Prinzipiell lassen sich bei mehreren beobachteten Pfaden auch Übergangswahrscheinlichkeiten $p_{ij}(t)$ im inhomogenen Fall schätzen:

$$\widehat{p}_{ij}(t) = \frac{n_{ij}(t)}{n_i(t)}$$

mit

$n_{ij}(t)$	Anzahl der Übergänge von i nach j mit $X_t = i$ und $X_{t+1} = j$
$n_i(t)$	Anzahl der Beobachtungen mit $X_t = i$.

Problem: ohne Modellannahmen hohe Varianz von $\widehat{p}_{ij}(t)$, falls $n_i(t)$ klein.
Alternative: parametrische oder nonparametrische Modellierung von $p_{ij}(t)$.

- Likelihood-Quotienten-Tests für homogene Markov-Ketten:

  Zum Test von Hypothesen H_0 versus H_1, z.B.

$\mathsf{H}_0 : \{X_t\}$ ist i.i.d.,	$\mathsf{H}_1 : \{X_t\}$ ist Markov-Kette 1. Ordnung
$\mathsf{H}_0 : \{X_t\}$ ist Markov-Kette 1. Ordnung,	$\mathsf{H}_1 : \{X_t\}$ ist Markov-Kette 2. Ordnung

  kann der Likelihood-Quotienten-Test verwendet werden:

  $l(\hat{\theta}_1)$ maximierte Loglikelihood im H_1-Modell

  $l(\hat{\theta}_0)$ maximierte Loglikelihood im H_0-Modell

  Unter H_0 gilt:

  $$2(l(\hat{\theta}_1) - l(\hat{\theta}_0)) \overset{a}{\sim} \chi^2(r), \quad T \to \infty$$

  mit r=Differenz der Anzahl der geschätzten Parameter in H_0- und H_1-Modell.

Fallstudie: Niederschlag bei den Snoqualmie-Wasserfällen

Daten: 36 Jahre $n = 1948, \ldots, 1983$; jeweils Januar $t = 1, \ldots, 31$. (Quelle: Guttorp)

$$X_{nt} = \begin{cases} 1 & \text{Regen am Tag } t \text{ in Jahr } n \\ 0 & \text{kein Regen am Tag } t \text{ im Jahr } n. \end{cases}$$

Annahme: Unabhängige Pfade $n = 1, \ldots, 36$ desselben Prozesses X werden beobachtet.

Modell 1: $\{X_{nt}\}$ sind i.i.d. B(1,p) verteilt.
Häufigkeitstabelle:

	$X_t = 0$	$X_t = 1$	
	325	791	1116

relative Häufigkeiten für $X_t = 0 : \widehat{p} = \frac{325}{791+325} = 0.291$

relative Häufigkeiten für $X_t = 1 : 1 - \widehat{p} = 0.709$

Likelihood: $L(p) \propto p^{325}(1-p)^{791}$ (1 unbekannter Parameter)

Modell 2: X_t ist homogene Markov-Kette 1. Ordnung mit (jahres–unabhängiger) Übergangsmatrix P

Kontingenztafel:

	$X_t = 0$	$X_t = 1$	
$X_{t-1} = 0$	186	123	309
$X_{t-1} = 1$	128	643	771
	314	766	1080

Beachte: Hier gehen nur 30·36=1080 Beobachtungen ein.

Geschätzte Übergangsmatrix (z.B $\widehat{p}_{01} = \frac{123}{309} = 0.398$):

$$\widehat{P} = \begin{pmatrix} 0.602 & 0.398 \\ 0.166 & 0.834 \end{pmatrix}$$

Likelihood (zwei unbekannte Parameter):

$$L(P) = \prod_{n=1}^{36} P(X_{n0} = x_{n0}) \cdot p_{00}^{186} p_{01}^{123} p_{10}^{128} p_{11}^{643}$$

LQ–Test zu

$$\mathsf{H}_0 : \{X_t\} \text{ i.i.d} \qquad \text{versus} \qquad \mathsf{H}_1 : \{X_t\} \text{ MK 1. Ordnung}$$

Beachte: unter H_0 und H_1 werden "gleiche" Daten benötigt.

$\Rightarrow$ Lasse bei Modell 1 die Startwerte $(n = 1, ..., 36)$ weg.

$\Rightarrow \widehat{p} = \frac{314}{766+314} = 0.291$

LQ–Statistik ($\overset{a}{\sim} \chi^2_1$ unter H_0):

$$\begin{aligned} 2\{\log(L(\widehat{P})) - \log(L(\widehat{p}))\} &= 2\{[186 \cdot \log(0.602) + 123 \cdot \log(0.398) \\ &\quad +128 \cdot \log(0.166) + 643 \cdot \log(0.834)] \\ &\quad -[314 \cdot \log(0.291) + 766 \cdot \log(0.709)]\} \\ &= 193.49 \end{aligned}$$

$\Rightarrow H_0$ (i.i.d. Modell 1) ablehnen

Modell 3: Homogene Markov-Kette 2. Ordnung

Kontingenztafel:

		$X_t = 0$	$X_t = 1$	
$X_{t-2} = 0$	$X_{t-1} = 0$	109	67	176
$X_{t-2} = 0$	$X_{t-1} = 1$	25	94	119
$X_{t-2} = 1$	$X_{t-1} = 0$	70	52	122
$X_{t-2} = 1$	$X_{t-1} = 1$	100	527	627
		304	740	1044

$\Rightarrow$ geschätzte ÜM für MK 2. Ordnung, $p_{ij,k} := P(X_t = k | X_{t-1} = j, X_{t-2} = i)$,

$$\widehat{P}_2 = (\widehat{p}_{ij,k})_{ij;k} = \begin{pmatrix} 0.619 & 0.381 \\ 0.210 & 0.790 \\ 0.574 & 0.426 \\ 0.159 & 0.841 \end{pmatrix}$$

LQ–Test zu

$$\text{H}_0 : \{X_t\} \text{ MK 1. Ordnung} \qquad \text{versus} \qquad \text{H}_1 : \{X_t\} \text{ MK 2. Ordnung}$$

$$\begin{aligned}\log(L(\widehat{P}_2)) &= 109 \cdot \log(0.619) + ... + 527 \cdot \log(0.841) \\ &= -536.48 \\ \log(L(\widehat{P})) &= -537.67\end{aligned}$$

$\Rightarrow 2\{\log(L(\widehat{P}_2)) - \log(L(\widehat{P}))\} = 2.37$ bei 2 Freiheitsgraden

$\Rightarrow p$–Wert 0.15

$\Rightarrow$ H$_0$ wird nicht abgelehnt.

Achtung: Auch hier muss wieder die Datengrundlage angepasst werden!

Das Markov-Modell impliziert eine geometrische Verteilung für die Dauer von trockenen / nassen Perioden, die sich nicht mit der Realität deckt. Längere Perioden werden beobachtet, als vom Modell vorhergesagt werden. $\rightarrow$ 3.6 Hidden Markov-Modelle.