Kapitel 5

Erneuerungs- und Semi-Markov-Prozesse

Für den Poisson-Prozess und (reguläre) diskrete Markov-Prozesse impliziert die Markov-Eigenschaft, dass die Zwischenzeiten bzw. Verweildauern exponentialverteilt sind. Bei Erneuerungs- und Semi-Markov-Prozessen dürfen diese beliebig verteilt sein. Die Markov-Eigenschaft geht damit verloren bzw. gilt nur noch in eingeschränktem Sinn.

Ziel des Kapitels ist eine kurze Einführung in die Struktur dieser Prozesse. Details, insbesondere zur "klassischen" Erneuerungstheorie, und Beweise finden sich in FKO, Kapitel 4 und 5.

5.1 Erneuerungsprozesse

Die Pfade sind Treppenfunktionen mit Sprunghöhe +1 wie beim Poisson-Prozess; statt Exponentialverteilungen für die Verweildauern T_n sind nun beliebige Verteilungen zugelassen:

 T_n iid $\sim F(t)$, nichtnegativ. Interpretation und Anwendungsbeispiele analog zum Poisson-Prozess.

5.1.1 Definition und grundlegende Begriffe

Definition 5.1 Erneuerungsprozess

Sei $\{T_n, n \in \mathbb{N}\}$ eine Folge unabhängiger, nichtnegativer Zufallsvariablen mit gleicher Verteilungsfunktion F, für die F(0) < 1 gilt. Dann heißt die Folge $\{S_n, n \in \mathbb{N}_0\}$ mit

$$S_0 := 0$$
, $S_{n+1} = S_n + T_{n+1}$, $n = 0, 1, 2, ...$

Erneuerungsprozess (EP). Sei

$$N(t) = \max_{n \in \mathbb{N}_0} \{ n : S_n \le t \}$$

die Anzahl von Erneuerungen im Intervall [0,t]. Dann heißt $N=\{N(t),t\geq 0\}$ Erneuerungszählprozess, oft auch einfach Erneuerungsprozess.

Beispiele, Anwendungen

(a) Zuverlässigkeitstheorie

Ein bestimmter Bestandteil eines technischen Gerätes werde sofort nach seinem Ausfall durch ein neues Gerät ersetzt, usw. Die T_1, T_2, \ldots sind dann die Lebensdauern des ersten, zweiten,... Gerätes. N(t) gibt die Anzahl von Erneuerungen an.

Denkt man statt an ein technisches Gerät an andere Systeme, so sind viele weitere Interpretationen möglich.

(b) Marktforschung

Ein bestimmtes Produkt wird nach seinem Verbrauch erneut gekauft, usw. Der Erneuerungsprozess heißt in diesem Zusammenhang Wiederholkaufprozess.

(c) Sachversicherung

 S_1, S_2, \ldots sind die Zeitpunkte, zu denen Schadensfälle auftreten. N(t) ist die Anzahl der Schadensfälle.

(d) Warteschlangen, Lagerhaltung

Der Ankunftsprozess von Kunden kann in Verallgemeinerung zu den entsprechenden Beispielen

des letzten Kapitels einen Erneuerungsprozess bilden. Auch vom Bedarf in Lagerhaltungssystemen wird oft angenommen, dass er einen Erneuerungsprozess bildet; dabei ist T_n der Bedarf in der n-ten Periode.

(e) Klinische Studien, Survival-Analyse

In klinischen Studien sind T_1, T_2, \ldots z.B. die Zeitdauern nach denen epileptische Anfälle, Rückfälle, etc. auftreten. Ein wichtiger Spezialfall ist die Survival-oder Überlebenszeit-Analyse: Hier wird für (homogene) Gruppen von Patienten jeweils die Überlebenszeit T_1 bis zum Ende (Tod) oder auch nur zensiert beobachtet, d.h. das Ereignis Tod ist bis zum Ende der Studie noch nicht eingetreten (vgl. Fahrmeir, Hamerle und Tutz, 1996, Kap.7).

Vor allem im Kontext des ersten Beispiels nennt man die T_n Lebensdauern, die S_n Erneuerungszeitpunkte. Dabei wird S_0 nicht zu den Erneuerungszeitpunkten gezählt (in der Literatur ist dies gelegentlich der Fall). Für die Lebensdauer ist auch $P\{T_n = \infty\} > 0$, also $F(\infty) := \lim_{t \to \infty} F(t) < 1$, zugelassen. Die Annahme, dass T_1 dieselbe Verteilung wie T_2, T_3, \ldots besitzt, kann so gedeutet werden, dass zum Zeitpunkt t = 0 gerade eine Erneuerung stattgefunden hat (die aber vereinbarungsgemäß nicht gezählt wird). Besitzt T_1 eine von T_2, T_3, \ldots verschiedene Verteilung, so spricht man vom modifizierten Erneuerungsprozess.

Die T_n können diskrete oder stetige ZV sein. Im diesem Kapitel werden nur stetige ZV betrachtet, d.h. die Lebensdauerverteilung F besitzt eine Dichte f.

Beispiele:

- (a) Die Exponentialverteilung ergibt den Spezialfall des Poisson-Prozesses.
- (b) Weibull-Verteilung

Dichte:

$$f(t) = \lambda \alpha (\lambda t)^{\alpha - 1} \exp(-(\lambda t)^{\alpha}), \quad t \ge 0$$

wobei $\lambda > 0$ und $\alpha > 0$ Parameter der Verteilung sind.

Für $\alpha = 1$ ergibt sich die Exponentialverteilung als Spezialfall.

$$E(T) = \frac{1}{\lambda} \Gamma\left(\frac{1+\alpha}{\alpha}\right)$$

$$Var(T) = \frac{1}{\lambda^2} \left[\Gamma\left(\frac{\alpha+2}{\alpha}\right) - \left\{\Gamma\left(\frac{\alpha+1}{\alpha}\right)\right\}^2 \right]$$

(c) Gammaverteilung

Dichte

$$f(t) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \cdot t^{\alpha - 1} \exp(-\lambda t), \quad \alpha > 0, \lambda > 0, t > 0.$$

Für $\alpha = 1$ ergibt sich wiederum die Exponentialverteilung als Spezialfall.

$$\begin{array}{rcl} \mathrm{E}(T) & = & \frac{\alpha}{\lambda} \\ \mathrm{Var}(T) & = & \frac{\alpha}{\lambda^2} \\ \mathrm{Modus} & = & \frac{(\alpha-1)}{\lambda}, & \mathrm{f\"{u}r} \ \alpha > 1 \end{array}$$

(d) Die log-Normalverteilung

Sie ergibt sich aus einer normalverteilten ZV

 $X \sim N(\mu, \sigma^2)$ durch Exp-Transformation:

$$T = \exp(X), \quad X \sim N(\mu, \sigma^2)$$

Es gilt:

$$\begin{split} \mathbf{E}(T) &= \exp\left(\mu + \frac{1}{2}\sigma^2\right) \\ \mathbf{Var}(T) &= \exp(2\mu + \sigma^2) \cdot \left[\exp(\sigma^2) - 1\right] \\ f(t) &= \frac{1}{t} \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(\log(t) - \mu)^2}{\sigma^2}\right), \quad t > 0. \end{split}$$

Man definiert ferner die Survivorfunktion und insbesondere die Hazardrate:

Definition 5.2 Survivorfunktion und Hazardrate

Die Survivorfunktion

$$S(t) = 1 - F(t)$$

ist die unbedingte Wahrscheinlichkeit bis zum Zeitpunkt tzu überleben.

Die Hazardrate ist

$$\lambda(t) = \lim_{h \downarrow 0} \frac{P(t \le T \le t + h \mid T \ge t)}{h},$$

also die infinitesimale Rate für einen Ausfall zum Zeitpunkt t, gegeben Überleben bis zum Zeitpunkt t.

In o(h)-Schreibweise: Die Wahrscheinlichkeit für einen Ausfall im Intervall [t,t+h], gegeben Überleben bis zum Zeitpunkt t, ist

$$P(t < T < t + h \mid T > t) = \lambda(t)h + o(h).$$

Das Integral

$$\Lambda(t) = \int_{0}^{t} \lambda(u) du$$

wird als kumulierte Hazardrate bezeichnet.

Beispiel:

(a) Exponential verteilung: $\lambda(t) = \lambda$

(b) Weibull-Verteilung: $\lambda(t) = \lambda \alpha (\lambda t)^{\alpha-1}$

Es gelten folgende

Zusammenhänge

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}$$

$$S(t) = \exp\left[-\int_0^t \lambda(u)du\right] = \exp(-\Lambda(t))$$

$$f(t) = \lambda(t) \cdot S(t) = \lambda(t) \exp\left[-\int_0^t \lambda(u)du\right]$$

Beweis:

Die erste Beziehung erfolgt direkt aus der Definition der Hazardrate. Die zweite Gleichung folgt wegen

$$\int_{0}^{t} \lambda(u)du = \int_{0}^{t} \frac{f(u)}{1 - F(u)} du = [-\ln(1 - F(u))]_{0}^{t} = -\ln(1 - F(t)).$$

5.1.2 Zur Theorie der Erneuerungsprozesse

Die klassische Theorie der EP ist insbesondere interessiert an der Verteilung von S_n und N(t), sowie der erwarteten Zahl R(t) = E(N(t)) von Erneuerungen bis t. Es gelten folgende Sätze (FKO S131-133):

Satz 5.1 Verteilung von S_n und N(t)

$$F_n(t) := P\{S_n \le t\} = \int_{[0,t]} F_{n-1}(t-x)dF(x), \quad t \ge 0,$$

mit $F_1(t) := F(t)$; in der Schreibweise mit dem Faltungssymbol also

$$F_n = F * F_{n-1} = F^{n*} = \underbrace{F * \dots * F}_{n-mal}.$$

$$P{N(t) = n} = F_n(t) - F_{n+1}(t), \quad n = 0, 1, ...; \quad F_0(t) \equiv 1.$$

Für die Dichte $f_n(t)$ von $F_n(t)$ gilt

$$f_n(t) = \int_0^t f_{n-1}(t-x)f(x)dx, \quad f_1(t) := f(t)$$

Es ist

$$E(N(t)) = R(t) = \sum_{n=1}^{\infty} F_n(t).$$

Analytisch ist die Faltung nur in Spezialfällen lösbar. Da die numerische Integration sehr aufwendig wird, sind folgende asymptotischen Ergebnisse (für $t \to \infty$) nützlich. Dabei werden rekurrente EP vorausgesetzt, d.h. es muss gelten

$$\lim_{t \to \infty} F(t) = 1.$$

Ansonsten treten mit positiver Wahrscheinlichkeit ab einem gewissen Erneuerungszeitpunkt keine weiteren Erneuerungen mehr auf. Im folgenden sei für rekurrente EP mit stetigen T_n

$$\mu = E(T_n) = \int_0^\infty x f(x) dx = \int_0^\infty (1 - F(x)) dx.$$

Dann gilt

Satz 5.2 Grenzwertsätze

Für einen rekurrenten EP mit stetigen T_n ,

$$\mu = \mathrm{E}(T_n), \quad \sigma^2 = \mathrm{Var}(T_n) < \infty \text{ gelten } :$$

- 1. $\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\mu}$ m. Wkeit 1
- 2. $\lim_{t\to\infty} \frac{R(t)}{t} = \frac{1}{\mu}$, elementares Erneuerungstheorem

$$\lim_{t\to\infty} (R(t) - \frac{t}{\mu}) = \frac{\sigma^2 - \mu^2}{2\mu^2},$$

3. $\lim_{t\to\infty} P\left(\frac{N(t)-\frac{t}{\mu}}{\frac{\sigma}{\mu}\sqrt{\frac{t}{\mu}}} \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{t^2}{2}) dt$, d.h. N(t) ist für große t approximativ normalverteilt mit Erwartungswert $\frac{t}{\mu}$ und der Varianz $\frac{\sigma^2 t}{\mu^3}$.

Aussage 1. bzw. 2. enthalten das plausible Resultat, dass auf lange Sicht die (durchschnittliche) Anzahl von Erneuerungen pro Zeiteinheit gleich dem Kehrwert der erwarteten Dauer $\mu = E(T_n)$ ist.

5.2 Semi-Markov-Prozesse

5.2.1 Definition und grundlegende Eigenschaften

Die Struktur eines typischen Pfades eines Semi-Markov-Prozesses gleicht der eines Markov-Prozesses, mit der Ausnahme, dass die Verweilzeiten nicht notwendig exponentialverteilt sein müssen.

 Y_n Zustand nach dem n-ten Übergang, $n = 0, 1, \ldots$

 T_{n+1} Verweildauer im Zustand Y_n

 S_n Zeitpunkt des n-ten Übergangs.

Die Folge $\{Y_n, n \in \mathbb{N}_0\}$ der Zustände soll eine Markov-Kette bilden, während die Zeitpunkte S_n , zu denen ein Übergang erfolgt, und damit die Verweildauer durch einen gesonderten Wahrscheinlichkeitsmechanismus beschrieben werden. Dies führt zur

Definition 5.3 Markov-EP, Semi-Markov-Prozess

 $\{Y_n, n \in \mathbb{N}_0\}$ sei eine Folge von ZV mit abzählbarem Zustandsraum $\{S_n, n \in \mathbb{N}_0\}$ eine Folge nichtnegativer ZV mit $0 =: S_0 \le S_1 \le S_2 \le \dots$ Der SP $(Y, S) = \{(Y_n, S_n), n \in \mathbb{N}_0\}$ heißt Markov-Erneuerungsprozess (MEP) : $\Leftrightarrow t \ge s \ge s_{n-1} \ge \dots \ge 0, \ \forall \ j, i_n, \dots, i_0 \in S$

(SM)
$$P\{Y_{n+1} = j, S_{n+1} - S_n \le t \mid Y_n = i, Y_{n-1} = i_{n-1}, \dots, Y_0 = i_0, S_n = s, S_{n-1} = s_{n-1}, \dots\}$$

= $P\{Y_{n+1} = j, S_{n+1} - S_n \le t \mid Y_n = i\}.$

Ist (Y,S) ein MEP, so heißt $X=\{X(t),t\geq 0\}$ der zu (Y,S) gehörige Semi-Markov-Prozess (SMP): \Leftrightarrow

$$X(t) = Y_n$$
 für $S_n \le t < S_{n+1}$, $n = 0, 1, 2, ...$

Interpretation: Die Markov-Eigenschaft gilt nur noch bei Bedingen auf $X(S_n) = i$, d.h. bedingt auf $t = S_n$; nicht jedoch für beliebige Zeitpunkte. Die Eigenschaft (SM) heißt deshalb Semi-Markov-Eigenschaft.

Definition 5.4 Homogener SMP

Gilt außerdem

$$P\{Y_{n+1} = j, S_{n+1} - S_n \le t \mid Y_n = i\} = P\{Y_1 = j, S_1 \le t \mid Y_0 = i\} =: Q_{ij}(t)$$

unabhängig von n, so heißt der MEP bzw. SMP homogen. Die bedingten Wahrscheinlichkeiten $Q_{ij}(t)$ heißen Semi-Markov-Übergangswahrscheinlichkeiten (SÜW), $Q(t) = (Q_{ij}(t))$ Semi-Markov-Übergangsmatrix (SÜM).

Vereinbarung: Im weiteren werden nur homogene SMP betrachtet und zusätzlich gefordert, dass für die Verteilung

$$Q_i(t) := \sum_{j \in S} Q_{ij}(t) = P\{T_{n+1} \le t \mid Y_n = i\} = P\{S_1 \le t \mid Y_0 = i\}$$

der Verweilzeit in i gilt:

$$Q_i(0) < 1$$
 und $\lim_{t \to \infty} Q_i(t) = 1$.

Also: Der SMP verbleibt eine positive Zeitspanne in i (dies entspricht der Forderung F(0) < 1 bei EP)

und verlässt m. Wkeit 1 nach endlicher Zeit i wieder (dies entspricht $F(\infty) = 1$ für einen rekurrenten EP).

Bemerkung: Alternativ können die Verweildauerverteilungen auch durch Hazardraten charakterisiert werden:

$$\lambda_i(t) = \frac{q_i(t)}{1 - Q_i(t)}, \quad q_i(t) \text{ Dichte.}$$

Satz 5.3 Struktur von SMP

 $\{Y_n, n \in \mathbb{N}_0\}$ ist eine homogene MK mit Übergangsmatrix $Q = (q_{ij})$, wobei

$$q_{ij} = \lim_{t \to \infty} Q_{ij}(t).$$

Die bedingte Verteilungsfunktion der Verweildauer T_{n+1} bedingt auf den Zustand $Y_n = i$ und $Y_{n+1} = j$ ist $G_{ij}(t) = P(T_{n+1} \le t \mid Y_n = i, Y_{n+1} = j) = \frac{Q_{ij}(t)}{q_{ij}}$. (Für $q_{ij} = 0$ setze $G_{ij}(t) = 0$.)

Daraus kann man leicht einen Algorithmus zur Simulation von SMP herleiten:

- 1. Simuliere die MK Y gemäß ÜM (q_{ij})
- 2. Simuliere dann die Verweildauer T gemäß der Verteilung $G_{ij}(t)$.

Bemerkung: Alternativ kann ein SMP auch über die (abgangsspezifischen) Hazardraten

$$\lambda_{ij}(t) = \lim_{h \downarrow 0} \frac{P(Y_{n+1} = j, t \le T_{n+1} \le t + h \mid Y_n = i, T_{n+1} \ge t)}{h}$$

$$= \lim_{h \downarrow 0} \frac{P(Y_{n+1} = j, t \le T_{n+1} \le t + h \mid Y_n = i)}{h \cdot P(T_{n+1} \ge t \mid Y_n = i)}$$

$$= \frac{q_{ij}(t)}{1 - Q_i(t)}, \quad \text{mit } q_{ij}(t) : \text{ "Dichte" zu } Q_{ij}(t), \text{ d.h.}$$

$$q_{ij}(t) = \frac{\partial Q_{ij}(t)}{\partial t}$$

charakterisiert werden.

Beispiel:

MEP mit

$$Q(t) = (Q_{ij}(t)) = \begin{pmatrix} 0.6(1 - \exp(-5t)) & 0.4(1 - \exp(-2t)) \\ 0.5 - 0.2 \exp(-3t) - 0.3 \exp(-5t) & 0.5 - 0.5 \exp(-2t) - t \exp(-2t) \end{pmatrix}$$

1. Verteilungsfunktion der Verweildauer in i

$$Q_1(t) = 1 - 0.6 \exp(-5t) - 0.4 \exp(-2t)$$

$$Q_2(t) = 1 - 0.2 \exp(-3t) - 0.5 \exp(-2t)$$

$$-0.3 \exp(-5t) - t \exp(-2t)$$

2. Übergangsmatrix der eingebetten MK

$$Q = \lim_{t \to \infty} Q(t) = \begin{pmatrix} 0.6 & 0.4 \\ 0.5 & 0.5 \end{pmatrix}$$

3. Bedingte Verteilungsfunktionen

$$G_{ij}(t) = P(T_{n+1} \le t \mid Y_n = i, Y_{n+1} = j)$$

$$= \frac{Q_{ij}(t)}{q_{ij}}$$

$$G_{11}(t) = 1 - \exp(-5t)$$

$$G_{12}(t) = 1 - \exp(-2t)$$

$$G_{21}(t) = 1 - \frac{2}{5}\exp(-3t) - \frac{3}{5}\exp(-5t)$$

$$G_{22}(t) = 1 - \exp(-2t) - 2t\exp(-2t)$$

4. Erwartete Verweildauern

$$\mu_i = \int_{0}^{\infty} [1 - Q_i(t)] dt$$

$$\mu_1 = \int_0^\infty 0.6 \exp(-5t) dt + \int_0^\infty 0.4 \exp(-2t) dt$$

$$= 0.6 \left[-\frac{1}{5} \exp(-5t) \right]_0^\infty + 0.4 \left[-\frac{1}{2} \exp(-2t) \right]_0^\infty$$

$$= 0.6 \cdot \frac{1}{5} + 0.4 \cdot \frac{1}{2}$$

$$= \frac{6}{50} + \frac{4}{20}$$

$$= \frac{12}{100} + \frac{20}{100} = \frac{32}{100} = 0.32$$

$$\mu_2 = 0.626$$

5. Hazardrate

$$\lambda_{ij}(t) = \frac{q_{ij}(t)}{1 - Q_i(t)}$$

z.B.

$$q_{12}(t) = \frac{\partial Q_{12}(t)}{\partial t}$$
$$= 0.4 \cdot 2 \exp(-2t)$$
$$= 0.8 \exp(-2t).$$

$$Q_1(t) = 1 - 0.6 \cdot \exp(-5t) - 0.4 \cdot \exp(-2t).$$

$$\Rightarrow \lambda_{12}(t) = \frac{0.8 \exp(-2t)}{0.6 \exp(-5t) + 0.4 \exp(-2t)}$$
$$= \frac{2}{1 + 1.5 \exp(-3t)}.$$

Bemerkung MK, MP, EP als Spezialfälle

- (a) Ein Erneuerungsprozess ergibt sich mit $S = \{1\}$ und daher $Q_{11}(t) = F(t)$.
- (b) Eine Markov-Kette erhält man durch die Wahl

$$Q_{ij}(t) = \begin{cases} 0 & \text{für } t < 1\\ q_{ij} & \text{für } t \ge 1 \end{cases}$$

(da $q_{ij} = \lim_{t\to\infty} Q_{ij}(t)$) $\Rightarrow T_1 = T_2 = \ldots = 1$ deterministische Verweildauern.

(c) Mit $G_{ij}(t) = 1 - \exp(-\lambda_i t), \quad t \ge 0$

$$q_{ii} = \begin{cases} 0 & , i \text{ stabil} \\ 1 & , i \text{ absorbierend} \end{cases}$$

erhält man einen MP.

5.2.2 Grenzwertsätze

Rekurrenzeigenschaften und Irreduzibilität eines Semi-Markov-Prozesses leiten sich aus den entsprechenden Eigenschaften der Markovkette Y ab.

Der SMP heißt aperiodisch, wenn für alle i die Zwischenzeiten T_{ii} zwischen aufeinanderfolgenden Besuchen in i nicht gitterförmig sind, d.h. es gibt kein $d \geq 0$ mit $\sum_{n=0}^{\infty} P(T_{ii} = nd) = 1$. (Dies ist insbesondere bei stetigen Verweildauern erfüllt.)

Es gelten ähnliche Aussagen wie bei MP, wobei der Erwartungswert $\frac{1}{\lambda_i}$ von exponentialverteilten Verweildauern durch den Erwartungswert

$$\mu_i = \int_{0}^{\infty} t q_i(t) dt = \int_{0}^{\infty} (1 - Q_i(t)) dt$$

ersetzt wird. Zum Beispiel gilt

Satz 5.4 Grenzwertsatz

Der SMP sei aperiodisch, irreduzibel und positiv-rekurrent. Dann gilt für die Zustandswahrscheinlichkeit $p_j(t) = P(X(t) = j)$

$$\pi_j := \lim_{t \to \infty} p_j(t) = \frac{\nu_j \mu_j}{\sum_{i \in S} \nu_i \mu_i}$$

(unabhängig von der Anfangsverteilung), wobei ν mit

$$\nu = \nu Q$$

ein zur ÜM Q von Y gehöriges strikt positives invariantes Maß (d.h. bis auf Normierung die stationäre Grenzverteilung von Y) ist.

Eine Plausibilitätserklärung verläuft wie für MP.

5.3 Bemerkung zur statistischen Inferenz

Im Vergleich zu MP sind statt Parametern λ_i der Exponentialverteilungen die entsprechenden Verweildauern parametrisch oder nichtparametrisch zu schätzen sowie die ÜM Q der eingebetteten MK. Die moderne Inferenz benutzt das Instrumentarium von Zählprozessen (vgl. Kap.7) um die (Übergangsspezifischen) Hazardraten $\lambda_{ij}(t)$ als Funktion von t zu schätzen.