

Extremwerttheorie

Andre Surenian

20. November 2008

Motivation

- Stichprobe X_1, \dots, X_n *iid*, (Schäden, Verluste aus Kreditportfolio)
- $M_1 = X_1, M_2 = \max(X_1, X_2), \dots, M_n = \max(M_1, \dots, M_n)$
- Wie ist $P(M_n \leq x)$ verteilt?

Motivation

Berechnung von $P(M_n \leq x)$

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = F^n(x)$$

Motivation

Zentraler Grenzwertungssatz

X_1, \dots, X_n iid, $S_n = X_1 + \dots + X_n$, $a_n = nE[X_1]$, $b_n = \sqrt{\text{var}[X_1]}$

$$\implies \lim_{n \rightarrow \infty} P\left(\frac{S_n - a_n}{b_n} \leq x\right) = \Phi(x), \quad x \in \mathbb{R}$$

Motivation

Ähnlich wie beim Zentralen Grenzwertungssatz: Versuche Aussagen über zentrierte und normierte Maxima zu finden.
Konkret: Suche reelle Folgen $c_n > 0$ und d_n , sodass:

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - d_n}{c_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H(x)$$

Definition

Verallgemeinerte Extremwertverteilung

$$H_{\xi}(x) = \begin{cases} \exp(-(1 + \xi x)^{-\frac{1}{\xi}}), & \xi \neq 0 \\ \exp(-e^{-x}), & \xi = 0 \end{cases}$$

$1 + \xi x > 0$. Dreiparametrische Familie bei Einführung eines Lageparameters $\mu \in \mathbb{R}$ und eines Skalenparameters $\sigma > 0$

$$H_{\xi, \mu, \sigma} := H_{\xi}\left(\frac{x - \mu}{\sigma}\right)$$

Für $\xi = 0$ Gumbelverteilung

$$\Lambda(x) = \exp(-e^{-x}), x \in \mathbb{R}$$

Für $\xi > 0$ Frechetverteilung

$$\Phi_\alpha(x) = \exp(-x^{-\alpha}), x \geq 0, \alpha > 0$$

Für $\xi < 0$ Weibullverteilung

$$\Psi_\alpha(x) = \exp(-(-x)^{-\alpha}), x \leq 0, \alpha < 0$$

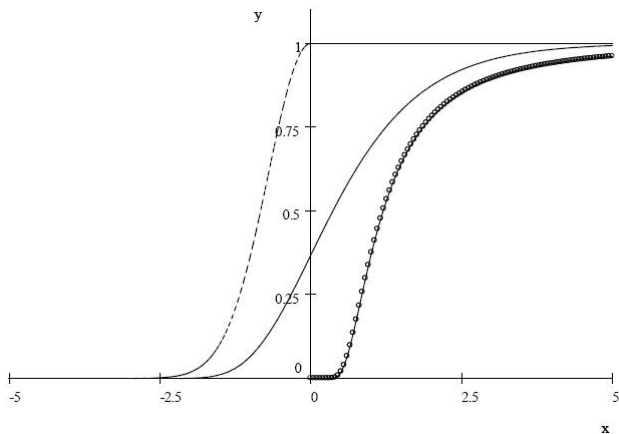


Abbildung: Gumbel, Frechet (gepunktet), Weibull (gestrichelt)

Bemerkung

Für festes x folgt, dass $\lim_{\xi \rightarrow 0} H_{\xi}(x) = H_0(x)$. Die verallgemeinerte Extremwertverteilung ist also stetig in ξ .

Maximum Domain of Attraction (MDA)

$$F \in MDA(H) \iff \exists c_n > 0, d_n \in \mathbb{R} : c_n^{-1}(M_n - d_n) \rightarrow H$$

Satz (Fisher-Tippet)

Ist $F \in MDA(H)$, dann folgt, dass H eine Verteilung vom Typ H_ξ ist.

Beispiel

Sei F exponentialverteilt mit $\beta > 0$

$\Rightarrow F(x) = 1 - \exp(-\beta x), x \geq 0$. Wähle $c_n := \frac{1}{\beta}$ und $d_n := \frac{\ln n}{\beta}$,

$$\begin{aligned} F^n(c_n x + d_n) &= (1 - \frac{1}{n} \exp(-x))^n, \\ \lim_{n \rightarrow \infty} F^n(c_n x + d_n) &= \exp(-e^{-x}), \end{aligned}$$

Heißt: $F \in MDA(H_0)$

Konvergenz von Minima

$$\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n)$$

$\tilde{F}(x) = 1 - F(-x) \Rightarrow$ Verteilung der Minima hat die Gestalt $1 - H_\xi(-x)$.

Bemerkung

Für symmetrische Verteilungen gilt $\tilde{F}(x) = F(x)$.

Maximum Domain of Attraction

- Welche Voraussetzungen an F garantieren, dass $F \in MDA(H_\xi)$?
- Welche Normierungskonstanten sind zu wählen?

MDA der Frechet Verteilung

Wie hat also eine Verteilungsfunktion F auszusehen, sodass gilt:

$$\lim_{n \rightarrow \infty} F^n(c_n x + d_n) = H_\xi(x), \quad \xi > 0$$

Definition

Eine positive, Lebesgue-messbare Funktion $L(x)$ auf \mathbb{R}^+ heißt von **langsamer Variation**, wenn gilt

$$\lim_{x \rightarrow \infty} \frac{L(cx)}{L(x)} = 1 \quad \forall c > 0$$

Satz

Für $\xi > 0$ und $L(x)$ von langsamer Variation gilt:

$$F \in MDA(H_\xi) \Leftrightarrow 1 - F(x) = x^{-\frac{1}{\xi}} L(x)$$

Für...

Verteilungen mit schweren Enden (heavy tailed distribution), z.B. Cauchyverteilung, Paretoverteilung, Burrverteilung

Beispiel

Paretoverteilung: $F(x) = 1 - \left(\frac{\kappa}{\kappa+x}\right)^\alpha$ $\alpha, \kappa > 0$

$$1 - F(x) = \left(\frac{\kappa}{\kappa+x}\right)^\alpha = x^{-\frac{1}{\xi}} L(x)$$

$$\Rightarrow \text{z.z.} \quad L(x) = \left(\frac{\kappa}{\kappa+x}\right)^\alpha x^{\frac{1}{\xi}}$$

$$\text{also } \lim_{x \rightarrow \infty} \frac{L(cx)}{L(x)} = \frac{\left(\frac{\kappa}{\kappa+xc}\right)^\alpha (xc)^{\frac{1}{\xi}}}{\left(\frac{\kappa}{\kappa+x}\right)^\alpha x^{\frac{1}{\xi}}} = 1$$

$$\exists \text{ für } \alpha = \frac{1}{\xi}$$

MDA der Gumbelverteilung

Satz

$x_F \leq \infty \Rightarrow F \in MDA(\Lambda) \iff \exists z < x_F$ und

$$1 - F(x) = c(x) \exp\left(-\int_z^x \frac{g(t)}{a(t)} dt\right),$$

wobei $g(x), c(x)$ messbar und $g(x) \rightarrow 1, c(x) \rightarrow c > 0$ und
 $\lim_{x \rightarrow x_F} a'(x) = 0$

Für...

Verteilungen mit leichten und mittelschweren Enden, z.B.
Exponentialverteilung, Normalverteilung, Loggammaverteilung usw.

MDA der Weibullverteilung

Wenig Interesse

Nicht von allzugroßen Interesse, weil $x_F < \infty$.

Jedoch...

Für $\xi < 0$ und $L(x)$ von langsamer Variation gilt:

$$F \in MDA(H_\xi) \Leftrightarrow x_F < \infty \text{ und } 1 - F(x_F - \frac{1}{x}) = x^{\frac{1}{\xi}} L(x)$$

Für...

Stetige Gleichverteilung und allgemein Verteilungen mit Potenzverhalten am rechten Rand, sodass: $1 - F(x) = \kappa(x_F - x)^\alpha$

Vorgehensweise

Beobachtung

X_1, \dots, X_n iid, Returns an Handelstagen, Schäden, etc.
Erzeuge m Blöcke (Perioden) zu n Beobachtungen, sodass,
Maxima sind $M_{nj}, j \in \{1, \dots, m\}$

Über...

$$h_{\xi, \mu, \sigma} = \frac{1}{\sigma} e^{-(1 + \xi \frac{x - \mu}{\sigma})^{-\frac{1}{\xi}}} (1 + \xi \frac{x - \mu}{\sigma})^{-\frac{1}{\xi} - 1}$$

Zur...

$$\sum_{i=1}^m \ln h_{\xi, \mu, \sigma}(M_{n,i})$$

Niveauüberschreitungen - threshold exceedances

Frage

Wie ist 'der hintere Bereich' $\{X \leq x + u | X > u\}$ einer Verteilung verteilt?

Definition - Excess Verteilung

Sei $X \sim F$. Die Excess Verteilung über die Schranke u ist:

$$F_u(x) = P(X - u \leq x | X > u) = \frac{F(x+u) - F(u)}{1 - F(u)}$$

für $x \in [0, x_F - u)$, wobei $x_F \leq \infty$ der rechte Endpunkt von F .

Mean excess function

$X \sim F$, $E[X] < \infty$, dann ist die mean excess function durch

$$e(u) = E[X - u | X > u]$$

gegeben

Definition - Verallgemeinerte Paretoverteilung

Für $\beta > 0$ und $x \geq 0$ wenn $\xi \geq 0$, und $x \in [0, -\frac{\beta}{\xi}]$ wenn $\xi < 0$ ist

$$G_{\xi, \beta}(x) = \begin{cases} 1 - (1 + \frac{\xi x}{\beta})^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp(-\frac{x}{\beta}), & \xi = 0 \end{cases}$$

die **verallgemeinerte Paretoverteilung**. ξ ist der Shape bzw. Tailparameter und β der Skalenparameter.

$$G_{\xi, \beta}(x) \in MDA(H_{\xi}) \quad \forall \xi \in \mathbb{R}$$

für $\xi < 1$ ist der Erwartungswert: $E(X) = \frac{\beta}{1-\xi}$

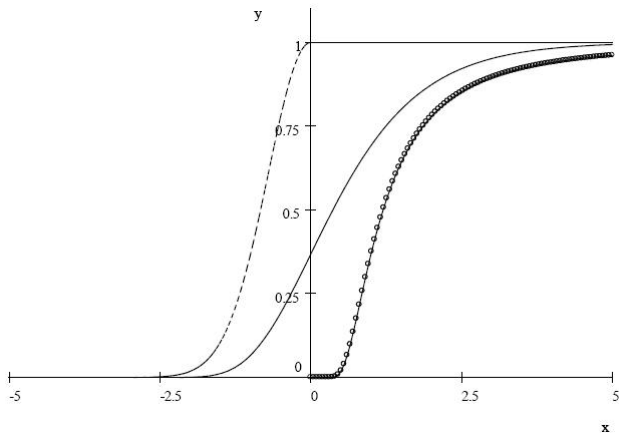


Abbildung: $\xi = 0$, $\xi < 0$ (gepunktet), $\xi > 0$ (gestrichelt)

Satz

Es existiert eine positive, messbare Funktion $\beta(u)$, sodass:

$$\lim_{u \rightarrow x_F} \sup_{x \in [0, x_F - u)} |F_u(x) - G_{\xi, \beta(u)}(x)| = 0 \iff F \in MDA(H_\xi), \xi \in \mathbb{R}$$

→ Annahme

F Verlustverteilung, sodass $F \in MDA(H_\xi), \xi \in \mathbb{R} \Rightarrow$ Bei vorgegebener Schranke $u \in \mathbb{R}$ gilt:

$$F_u(x) = G_{\xi, \beta}(x) \text{ für } \beta > 0 \text{ und } x \in [0, x_F - u).$$

Umgang mit Daten

n unabhängigen Beobachtungen (Verluste) $X_1, \dots, X_n \sim F$.
 Betrachte nur Daten mit $X_i \geq u \Rightarrow$ Anzahl der Beobachtungen,
 die u überschritten haben sei N_u .

Neue, gefilterten Daten seien nun X'_1, \dots, X'_{N_u} . Die Höhe des
 überschrittenen Verlustes (excess loss) ergibt sich durch
 $Y_j = X'_j - u$.

$$\text{Über } g_{\xi, \beta} = \beta \left(1 + \frac{\xi x}{\beta}\right)^{-(1 + \frac{1}{\xi})} \text{ zu}$$

$$\ln L(\xi, \beta; Y_1, \dots, Y_{N_u}) = \sum_{j=1}^{N_u} \ln g_{\xi, \beta}(Y_j)$$

Lösungen der Maximum Likelihood Gleichungen liefern uns das
 Modell einer verallgemeinerten Paretoverteilung des Types $G_{\xi, \beta}^{\wedge, \wedge}$ für
 die Excess Funktion F_u .

Anwendung im Risikomanagement

Bemerkung

$\bar{F}(x) = 1 - F(x) = P(X > x)$ bezeichne die Wahrscheinlichkeit des Tails einer Verteilung.

Beziehung

Bei vorgegebener Schranke u , gilt für $x \geq u$ die Wahrscheinlichkeit des Tails:

$$\begin{aligned}\bar{F}(x) &= \bar{F}(u)\bar{F}_u(x-u) \\ &= \bar{F}(u)\left(1 + \xi\frac{x-u}{\beta}\right)^{-\frac{1}{\xi}}\end{aligned}$$

Über...

$$F(x) = -\bar{F}(u) \left(1 + \xi \frac{x-u}{\beta}\right)^{-\frac{1}{\xi}} + 1$$

...zum V@R

Für $\alpha \geq F(u)$ erhalten wir

$$VaR_\alpha = u + \frac{\beta}{\xi} \left(\left(\frac{1-\alpha}{\bar{F}(u)} \right)^{\frac{1}{\xi}} - 1 \right)$$

und expected short fall

$$ES_\alpha = VaR_\alpha + e(VaR)$$

Vorgehensweise

$$\bar{F}(x) = \bar{F}(u) \left(1 + \xi \frac{x-u}{\beta}\right)^{-\frac{1}{\xi}}$$

Schätzer $\hat{\beta}$ und $\hat{\xi}$ über **Annahme 1**.

$$\hat{\bar{F}}(x) = \frac{N_u}{n} \left(1 + \hat{\xi} \frac{x-u}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}}$$

Der Hill Schätzer

Schätzung...

...des Tail-Index $\alpha > 0$ für eine Verteilung $F \in MDA(H_\xi)$ für $\xi > 0$, sodass: $\bar{F}(x) = L(x)x^{-\alpha}$,

...über mean excess function von $\ln X$.

$$\begin{aligned}
 e^*(\ln u) &= E(\ln X - \ln u | \ln X > \ln u) \\
 &= \frac{1}{\bar{F}(u)} \int_u^\infty (\ln x - \ln u) dF(x) \\
 &= \frac{1}{\bar{F}(u)} \int_u^\infty \frac{\bar{F}(x)}{x} dx \\
 &= \frac{1}{\bar{F}(u)} \int_u^\infty L(x)x^{-(\alpha+1)} dx
 \end{aligned}$$

Für $u \rightarrow \infty$ folgt

$$e^*(\ln u) \sim \frac{L(u)u^{-\alpha}\alpha^{-1}}{\bar{F}(u)} = \alpha^{-1}$$

Hill-Schätzer

Für beobachtete und absteigend geordnete Werte $X_{n,n} \leq \dots \leq X_{1,n}$ ist $e^*(\ln X_{k,n}) \approx \alpha^{-1}$, wobei $X_{k,n} = u$ die Schranke und der **Hill Schätzer**:

$$\hat{\alpha}_{k,n}^{(H)} = \left(\frac{1}{k} \sum_{j=1}^k \ln X_{j,n} - \ln X_{k,n} \right)^{-1}, \quad 2 \leq k \leq n$$

Schätzer für $\bar{F}(x)$ über den Hill Schätzer

- $\bar{F}(x) = Cx^{-\alpha}$ $x \geq u > 0$ für eine hohe Schranke u
- Tail-Index α über den Hill Schätzer $\hat{\alpha}_{k,n}^{(H)}$
- u ist k -ter Wert der absteigend geordneten Teststatistik $X_{k,n}$
- C

- Schreibe C als $u^\alpha \bar{F}(u)$ ($F(u) = L(u)u^{-\alpha}$)
- $\Rightarrow \bar{F}(x) = u^\alpha \bar{F}(u)x^{-\alpha}$

$$\hat{\bar{F}}(x) = \frac{k}{n} \left(\frac{x}{X_{k,n}} \right)^{-\hat{\alpha}_{k,n}^{(H)}}$$

Bemerkung

Der Schätzer über die verallgemeinerte Paretoverteilung

$\hat{F}(x) = \frac{N_u}{n} \left(1 + \xi \frac{x-u}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}}$ kann mit dem Hill Schätzer kombiniert

werden durch geeignete Substitution. $\hat{\xi}^{(H)} = \frac{1}{\hat{\alpha}_{k,n}^{(H)}}$, $X_{k,n} = N_u$ so erhalten wir:

$$\hat{F}(x) = \frac{N_u}{n} \left(1 + \hat{\xi}^{(H)} \frac{x-u}{\hat{\xi}^{(H)} u}\right)^{-\frac{1}{\hat{\xi}^{(H)}}}$$

Bisher

Finde und schätze geeignete Verteilung und berechne Wahrscheinlichkeiten für (Höhe der) Grenzüberschreitungen.

Nun

Betrachte und zähle das 'Ereignis' Grenzüberschreitung. Höhe der Überschreitung zunächst irrelevant.

Definition - White noise Prozess

Ein stochastischer Prozess $(\epsilon_t)_{t \in \mathbb{N}}$ ist ein white noise prozess, wenn für alle $t \in \mathbb{N}$ gilt:

- $E(\epsilon_t) = 0$
- $E(\epsilon_t^2) =: \sigma^2 < \infty$
- $\forall h > 0 : \text{Var}(\epsilon_{t+h}, \epsilon_t) = 0$

Betrachte...

- white noise Prozess $(X_t)_{t \in \mathbb{N}}$, gemeinsame Verteilung $F \in MDA(H_\xi)$ liegt für $\xi \in \mathbb{R}$.
- Für festes x sei $u_n(x) := c_n x + d_n$ eine Folge von Schranken.

Folgerung für einzelne Beobachtung

Das Ereignis Niveauüberschreitung der i – ten Beobachtung
 $\{X_i > u_n(x)\} \sim Ber(\bar{F}(u_n(x)))$

Folgerung für Stichprobe

Anzahl der Grenzüberschreitungen $N_{u_n(x)} \sim Bin(n, \bar{F}(u_n(x)))$ mit
Erwartungswert $n\bar{F}(u_n(x))$

$n \rightarrow \infty$

Grenzübergang

$$\begin{aligned} \lim_{n \rightarrow \infty} n\bar{F}(c_n x + d_n) &= n(1 - F(c_n x + d_n)) \text{ Wegen } -\ln(1 - y) \sim y \\ &\text{für } y \rightarrow 0 \text{ und } y := 1 - F(c_n x + d_n) \\ &= \lim_{n \rightarrow \infty} -n \ln(1 - (1 - F(c_n x + d_n))) \\ &= \lim_{n \rightarrow \infty} -n \ln(F(c_n x + d_n)) \\ &= \lim_{n \rightarrow \infty} -\ln(F^n(c_n x + d_n)) \\ &= -\ln H_\xi(x) \end{aligned}$$

Folgerung

Die Anzahl der Grenzüberschreitungen $N_{u_n(x)}$ einer Stichprobe konvergiert gegen eine Poissonverteilung mit dem Parameter $\lambda(x) = -\ln H_\xi(x)$.

Definition - Zählprozess

Sei Y_1, \dots, Y_n Folge von ZV im Zustandsraum Y (z.B. \mathbb{R} oder \mathbb{R}^2)
dann ist der Zählprozess über der Menge $A \subset Y$ definiert als

$$N(A) = \sum_{i=1}^n I_{\{Y_i \in A\}}$$

Definition - Poissonprozess

Der Zählprozess $N(\cdot)$ ist ein Poissonprozess auf Y mit der Mittelwertfunktion Λ wenn:

- Für $A \subset Y$ und $k \geq 0$,

$$P(N(A) = k) = \begin{cases} e^{-\Lambda(A)} \frac{\Lambda(A)^k}{k!}, & \Lambda(A) < \infty \\ 0, & \Lambda(A) = \infty \end{cases}$$

- Für $m \geq 1$ und für paarweise disjunkte Mengen $A_1, \dots, A_m \subset Y$ sind die Zufallsvariablen $N(A_1), \dots, N(A_m)$ unabhängig.

Für $\Lambda(\cdot)$ gilt, dass $E(N(A)) = \Lambda(A)$. $\lambda(x)$ bezeichnet die Intensität des Poissonprozesses.

Folgerung

Habe...

- white noise Prozess $(X_i)_{i \in \mathbb{N}}$
- $u_n(x) = c_n x + d_n$ für festes x . $n \in \mathbb{N}$ und $1 \leq i \leq n$
- Sei $Y_{i,n} = \frac{i}{n} I_{\{X_i > u_n(x)\}}$ für $n \in \mathbb{N}$ und $1 \leq i \leq n$

Also

Der Prozess, der nun unsere Grenzüberschreitungen aus einer Stichprobe mit Umfang n zählt ist durch

$$N_n(A) = \sum_{i=1}^n I_{\{Y_{i,n} \in A\}}$$

gegeben. Für $n \rightarrow \infty$ ein homogener Poissonprozess.

peak over threshold

Annahmen

- Der Zeitpunkt des Eintreffens des Ereignisses 'Grenzüberschreitung' gehorcht einem homogenen Poissonprozess
- Die Höhe der Grenzüberschreitung gehorcht einer verallgemeinerten Paretoverteilung
- Zeitpunkt und Höhe einer Grenzüberschreitung sind unabhängig

Erweiterung zu...

Zweidimensionalen Poissonprozess über (t, x) für Zeitpunkt und Höhe einer Grenzüberschreitung.

⇒

$Y = U \times V$ erweitert, $U = (0, 1]$ und $V = (u, \infty)$ darstellt. Für $A \subset Y$ ist Poissonprozess:

$$N(A) = \sum_{i=1}^n I_{\{(i/n, X_i) \in A\}}$$

$$\lambda(t, x) = \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi} - 1}$$

$$\Lambda(A) = \int_{t_1}^{t_2} \int_x^{\infty} \lambda(y) dy dt = -(t_2 - t_1) \ln H_{\xi, \mu, \sigma}(x)$$

Kreise schließen sich

Denn...

Für eine Schranke $x \geq u$ folgt nun, dass der eindimensionale Prozess der Höhe der Grenzüberschreitung Intensität $\tau(x) := -\ln H_{\xi, \mu, \sigma}(x)$ ist. Wahrscheinlichkeit $\bar{F}_u(x)$ für Grenzüberschreitungen über

$$\bar{F}_u(x) = \frac{\tau(x+u)}{\tau(u)} = \left(1 + \frac{\xi x}{\sigma + \xi(u - \mu)}\right)^{-\frac{1}{\xi}} = \bar{G}_{\xi, \beta}(x)$$

Und...

Betrachte das Ereignis $\{M_n \leq x\}$ für ein $x \geq u$.
 Zählprozesssprache $P(N(A) = 0)$

$$P(N(A) = 0) = \exp(-\Lambda(A)) = H_{\xi, \mu, \sigma}(x), x \geq u$$

self-exciting POT-Modell

Bisher

- Homogener Poissonprozess

Nun

- Versuche vergangene Grenzüberschreitungen (Höhe und Zeitpunkt) miteinzubeziehen über **bedingte Intensität**

- Daten X_1, \dots, X_n
- Schranke u
- N_u Grenzüberschreitungen
- $\{(i, X_i) : 1 \leq i \leq n, X_i > u\}$ (kein relativen Zeitpunkt)
- oder $\{(T_j, \tilde{X}_j) : j = 1, \dots, N_u\}$

Betrachte $Y_i = iI_{\{X_i > u\}}$ liefert Zeitpunkt der Grenzüberschreitung, falls einer Auftritt, Andernfalls 0.

Zählprozess mit bedingter Intensität

Zählprozess aller Überschreitungen $N(\cdot)$ im Zustandsraum $Y = (0, n]$ ist

$$N(A) = \sum_{i=1}^n I_{\{Y_i \in A\}}$$

mit der *bedingten Intensität*

$$\lambda^*(t) = \tau + \psi \sum_{j: 0 < T_j < t} h(t - T_j, \tilde{X}_j - u)$$

ist, wobei $\tau > 0$, $\psi \geq 0$ und h eine nichtnegative Funktion ist.

Bemerkung

Beispiele für h

- $h(s, x) = \exp(\delta x - \gamma s)$, wobei $\delta, \gamma > 0$ oder
- $h(s, x) = \exp(\delta x)(s + \gamma)^{-(\rho+1)}$, wobei $\delta, \gamma, \rho > 0$.

D a n k e