Das Boolesche Modell

Lukas Hägele

3.12.2009

Ziele des heutigen Seminars:

Das Boolesche Modell ist sehr anwendungsbezogen. Daher ist unser Ziel, am Ende die folgenden statistischen Fragen zu beantworten:

- Wann ist das Boolesche Modell geeignet für die Modellierung eines Datensatzes?
- Wie passe ich das Boolesche Modell an einen Datensatz an?

Definition Boolesches Modell:

Keime:
$\{S_n\}$; $n \in \mathbb{N}$ mit $S_n : \Omega \rightarrow \mathbb{R}^d$ homogener Poisson-Prozess in $\mathbb{R}^d$ mit Intensität $\lambda > 0$

Körner:
$\{\Xi_n\}$; $n \in \mathbb{N}$ mit $\Xi_n : \Omega \rightarrow \mathcal{C}$ Folge zufälliger kompakter Mengen unabhängig und identisch verteilt mit Verteilung Q

mit den folgenden Bedingungen:

- $\{S_n\}$ und $\{\Xi_n\}$ unabhängig
- $\mathbb{E}(\nu_d(\Xi_1 \oplus C)) < \infty$, $\forall C \in \mathcal{C}$
  $\mathcal{C}$ Familie aller kompakter Teilmengen des $\mathbb{R}^d$

Definition Boolesches Modell:

Das Boolesche Modell Ξ ist ein spezielles Keim-Korn-Modell:

$$\Xi = \bigcup_{n=1}^{\infty} (\Xi_n + S_n)$$

$$\text{mit } \Xi_n + S_n = \{\xi + S_n : \xi \in \Xi_n\}$$

Insbesondere ist Ξ eine ZAM (folgt aus Bedingung 2)

Veranschaulichung Boolesches Modell:

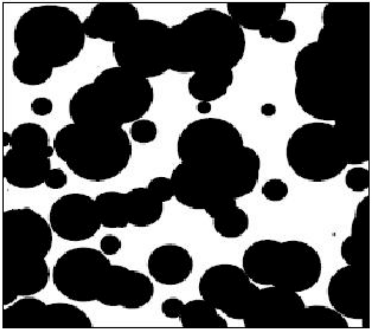


Abbildung 3: Boolesches (Keim–Korn–) Modell

Wiederholung:

Definition Kapazitätsfunktional:

$\mathbf{T}_{\Xi} : \mathcal{C} \to [0,1]$ mit $\mathbf{T}_{\Xi}(C) = \mathbb{P}(\Xi \cap C \neq \emptyset) \quad \forall C \in \mathcal{C}$

Satz von Choquet:

Seien Ξ_1, Ξ_2 zwei beliebige ZAM. Es gilt $\mathbb{P}_{\Xi_1} = \mathbb{P}_{\Xi_2}$ genau dann, wenn $\mathbf{T}_{\Xi_1} = \mathbf{T}_{\Xi_2}$

Eigenschaften Boolesches Modell:

Das Boolesche Modell $\Xi = \bigcup_{n=1}^{\infty} (\Xi_n + S_n)$ ist stationär

$$
\begin{aligned}
T_{\Xi}(C) = \mathbb{P}(\Xi \cap C \neq \emptyset) &= \mathbb{P}(\bigcup_{n=1}^{\infty} (\Xi_n + S_n) \cap C \neq \emptyset) \\
&= \mathbb{P}(\bigcup_{n=1}^{\infty} (\Xi_n + (S_n + x)) \cap C \neq \emptyset) \\
&= \mathbb{P}(\bigcup_{n=1}^{\infty} ((\Xi_n + S_n) + x) \cap C \neq \emptyset) \\
&= \mathbb{P}((\Xi + x) \cap C \neq \emptyset) = T_{\Xi + x}(C)
\end{aligned}
$$

Eigenschaften Boolesches Modell:

$\Xi = \bigcup_{n=1}^{\infty} (\Xi_n + x_n)$ ist isotrop, wenn Ξ_1 isotrop ist.

$$
\begin{aligned}
T_{\Xi}(C) = \mathbb{P}(\Xi \cap C \neq \emptyset) \quad & = \mathbb{P}\Big(\bigcup_{n=1}^{\infty} (\Xi_n + S_n) \cap C \neq \emptyset\Big) \\
& = \mathbb{P}\Big(\bigcup_{n=1}^{\infty} (\delta(\Xi_n) + \delta(S_n)) \cap C \neq \emptyset\Big) \\
& = \mathbb{P}\Big(\bigcup_{n=1}^{\infty} (\delta(\Xi_n + S_n)) \cap C \neq \emptyset\Big) \\
& = \mathbb{P}(\delta(\Xi) \cap C \neq \emptyset) = T_{\delta(\Xi)}(C)
\end{aligned}
$$

Das Kapazitätsfunktional im Booleschen Modell:

Für das Boolesche Modell gilt:

$\mathbf{T}_{\Xi}(C) = 1 - exp(-\lambda \mathbb{E} \nu_d(\Xi_1 \oplus \check{C}))$
mit $\check{C}$ Spiegelung von C am Ursprung

Beweis an Tafel

Volumenanteil im Booleschen Modell:

Volumenanteil

$p = \mathbb{E}\nu_d(\Xi \cap B) \quad , \quad \nu_d(B) = 1$

Für das Boolesche Modell gilt:

$$
\begin{aligned}
p &= \mathbb{E}\nu_d(\Xi \cap B) \\
&= \mathbb{E} \int_B \mathbb{1}(x \in \Xi) dx \\
&= \int_B \mathbb{E}\mathbb{1}(x \in \Xi) dx \\
&= \int_B \mathbb{P}(x \in \Xi) dx \\
&= \int_B \mathbb{P}(0 \in \Xi) dx \\
&= \mathbb{P}(0 \in \Xi) \\
&= \mathbb{P}(\Xi \cap \{0\} \neq \emptyset) \\
&= \mathbf{T}_\Xi(\{0\}) \\
&= 1 - exp(-\mathbb{E}\nu_d(\Xi_1))
\end{aligned}
$$

Kontaktverteilungsfunktion im Booleschen Modell:

Kontaktverteilungsfunktion

$H_B(r) = \mathbb{P}(rB \cap \Xi \neq \emptyset \mid o \notin \Xi) \quad \forall r \geq 0$

- $B \in \mathcal{C}$ nennt sich das "strukturierende Element" von H_B
- B ist sternförmig bezüglich dem Ursprung

Für das Boolesche Modell gilt:

$H_B(r) = 1 - exp(-\lambda(\mathbb{E}\nu_d(\Xi_1 \oplus r\check{B}) - \mathbb{E}\nu_d(\Xi_1)))$

Beweis Tafel

Verallgemeinerter Satz von Steiner

Sei Ξ eine isotrope, konvexe, zufällige, abgeschlossene Menge, mit $\mathbb{E}\nu_d(\Xi \oplus B(o,r)) < \infty\,,\ \forall r > 0$. Dann gilt

$$\mathbb{E}\nu_d(\Xi \oplus B) = \sum_{k=0}^{d} \alpha_{dk} \mathbb{E} V_{d-k}(\Xi) V_k(B),\ \forall B \in \mathcal{K}$$

mit $\mathcal{K}$ Familie aller konvexer Körper in $\mathbb{R}^d$ und $\alpha_{dk} = \frac{k!\kappa_k (d-k)!\kappa_{d-k}}{d!\kappa_d}$ mit $\kappa_i =$ i-dimensionales Volumen der Einheitskugel in $\mathbb{R}^i$.

- insbesondere gilt für V_{d-1} und V_1:

$$2V_{d-1}(B) = \text{"Oberfläche von B"}$$

$$V_1(B) = \frac{d\kappa_d}{2\kappa_{d-1}} \beta(B),$$

mit der mittleren Breite $\beta(B)$

$\mathbf{T}_{\Xi}$ und H_B für konvexe isotrope Körner

Für das Kapazitätsfunktional gilt:

$$\mathbf{T}_{\Xi}(C) = 1 - exp(-\lambda \sum_{k=0}^{d} \alpha_{dk} \mathbb{E} V_{d-k}(\Xi_1) V_k(C)), \quad \forall C \in \mathcal{K}$$

Für die Kontaktverteilungsfunktion mit $B \in \mathcal{K}$ gilt:

$$H_B(r) = 1 - exp(-\lambda \sum_{k=1}^{d} \alpha_{dk} r^k \mathbb{E} V_{d-k}(\Xi_1) V_k(B)) \quad \forall r \geq 0$$

Lineare Kontaktverteilungsfunktion:

- Sei d $= 2$
- Sei B die Strecke vom Ursprung zu einem beliebigen Punkt $u \in \partial B(o, 1)$

$H_{B,l}$ lineare Verteilungsfunktion in Richtung u:

$H_{B,l}(r) = 1 - exp(-\lambda \frac{\mathbb{E}2V_1(\Xi_1)}{\pi} r)\,,\ r \geq 0$

$\Rightarrow -\frac{1}{r} ln(1 - H_{B,l}(r)) = \lambda \frac{\mathbb{E}2V_1(\Xi_1)}{\pi}$ ist konstant in r

Quadratische Kontaktverteilungsfunktion:

- Sei $B = [0,1]^2$
- Sei Ξ_1 isotrope ZAM mit $\mathbb{P}(\Xi_1 \neq \emptyset) = 1$ und $\mathbb{E}max\{| x - y |: x, y \in \Xi_1\}^2 < \infty$
- Sei $\mathbb{E}l(\Xi_1)$ die mittlere Randlänge des Korns Ξ_1

$H_{B,q}$ quadratische Verteilungsfunktion:

$H_{B,q}(r) = 1 - exp(-\lambda(\frac{2\mathbb{E}l(\Xi_1)}{\pi}r + r^2))$

$\Rightarrow -\frac{1}{r}ln(1 - H_{B,q}(r)) = \lambda\frac{2\mathbb{E}l(\Xi_1)}{\pi} + r$ ist lineare Funktion in r

Umformung Kontaktverteilungsfunktion:

Kontaktverteilungsfunktion

$H_B(r) = \mathbb{P}(rB \cap \Xi \neq \emptyset \mid o \notin \Xi) \quad \forall r \geq 0$

$= \mathbb{P}(o \in \Xi \oplus r\check{B} \mid o \notin \Xi)$, da $\{rB \cap \Xi \neq \emptyset\} = \{o \in \Xi \oplus r\check{B}\}$

$= 1 - \mathbb{P}(o \notin \Xi \oplus r\check{B} \mid o \notin \Xi)$

$= 1 - \frac{\mathbb{P}(o \notin \Xi \oplus r\check{B})}{\mathbb{P}(o \notin \Xi)}$

$= 1 - \frac{1 - \mathbb{P}(o \in \Xi \oplus r\check{B})}{1 - \mathbb{P}(o \in \Xi)}$

Schätzer der Kontaktverteilungsfunktion:

Wir betrachten den folgenden Schätzer für H_B

$$\hat{H}_{B,W}(r_i) = 1 - \frac{1-\hat{p}_{B,W}(r_i)}{1-\hat{p}_W} \quad \forall i = 1, \ldots, m$$

mit

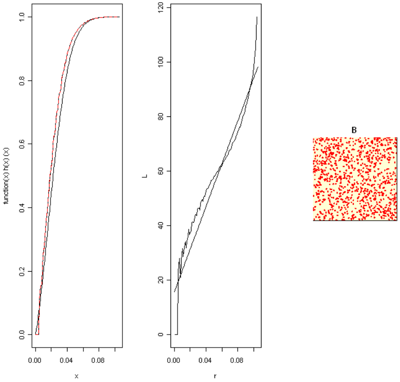
$$\hat{p}_{B,W}(r) = \frac{\nu_d((W \ominus r\check{B}) \cap (\Xi \oplus r\check{B}))}{\nu_d(W \ominus r\check{B})}$$

und

$$\hat{p}_W = \frac{\nu_d(W \cap \Xi)}{\nu_d(W)}$$

Schätzer der Kontaktverteilungsfunktion:

- Wann ist das Boolesche Modell geeignet für die Modellierung eines Datensatzes?



- rechts: Ausschnitt Boolesches Modell B mit $\nu_2(W) = 1$
- links:
  $H_{B,s}(r)$ (sphärisch) und
  $\hat{H}_{B,W,s}(r_i)$ in rot
- mitte:
  $-\frac{1}{r} ln(1 - H_{B,s}(r))$ und
  $-\frac{1}{r} ln(1 - \hat{H}_{B,W,s}(r_i))$

Minimum-Kontrast-Schätzer:

- Wie passe ich das Boolesche Modell an einen Datensatz an?

Wir betrachten sphärische Kontaktverteilungsfunktion:

$$H_B(r) = 1 - exp(-\lambda \sum_{j=0}^{d-1} \kappa_{d-j} \mathbb{E} V_j(\Xi_1) r^{d-j})$$

und suchen den Minimum-Kontrast-Schätzer $\hat{\theta}$ für die Intensität und die mittleren inneren Volumina mit:

$$\hat{\theta}(\xi) = argmin_{\theta \in \Theta} \Delta_W(\xi; \theta) \quad \theta = (\lambda, \delta_1, ..., \delta_{d-1})$$

wobei:

$$\Delta_W(\xi;\theta) = \sum_{i=1}^{m} (log(1 - \hat{H}_{B,W}(r_i, \xi)) - \lambda(\kappa_d r_i^d + \sum_{j=1}^{d-1} \kappa_{d-j} \delta_j r_j^{d-j}))^2$$

Quellen:

- Stoyan, D., Kendall, W.S., Mecke, J. (1995): Stochastic Geometry and its Applications, Wiley
- Schmidt, V. (2008): Räumliche Statistik (Vorlesungsskript)
- R-Simulation Olaf Wied

"Vielen Dank für eure Aufmerksamkeit!"