Stochastiek 2

Inleiding in de Mathematische Statistiek

Bayes-schatters

Bayesiaanse statistiek

Setting: waarneming X met dichtheid p_θ, $\theta \in \Theta$.

- Vat de parameter op als een stochast $\bar{\Theta}$ met a-priori dichtheid π (op Θ).
- Vat de dichtheid p_θ op als de voorwaardelijke dichtheid van X, gegeven $\bar{\Theta} = \theta$.
- Bepaal de a-posteriori verdeling van de parameter: dichtheid

$$p_{\bar{\Theta}\,|\,X=x}(\theta) = \frac{p_\theta(x)\pi(\theta)}{\int p_\theta(x)\pi(\theta)\,d\theta}, \qquad \theta \in \Theta.$$

Bayes-schatters

We kunnen nu de a-posteriori verdeling gebruiken om schatters te construeren. Standaard schatter: de verwachting.

Definitie.
De Bayes-schatter is de verwachting van de a-posteriori verdeling. Dus voor de realisatie $X = x$ is de schatting

$$T(x) = \int \theta p_{\bar{\Theta}\,|\,X=x}(\theta)\,d\theta = \frac{\int \theta p_\theta(x)\pi(\theta)\,d\theta}{\int p_\theta(x)\pi(\theta)\,d\theta}.$$

Karakterisering Bayes-schatters

Definitie.

Het Bayes-risico van de schatter $T = T(X)$ is

$$R(\pi; T) = \int E_\theta (T - \theta)^2 \pi(\theta)\, d\theta.$$

Stelling.

De Bayes-schatter minimaliseert het Bayes-risico

$$T \mapsto R(\pi; T)$$

over alle schatters.

Karakterisering Bayes-schatters - bewijs - 1

In de Bayesiaanse setup:

$$\bar{\Theta} \sim \pi$$
$$X \mid \bar{\Theta} = \theta \sim p_\theta.$$

Volgt dat

$$\begin{aligned}
R(\pi; T) &= \int E((T(X) - \theta)^2 \mid \bar{\Theta} = \theta)\pi(\theta)\, d\theta \\
&= E(T(X) - \bar{\Theta})^2 \\
&= \int E((T(x) - \bar{\Theta})^2 \mid X = x)\, p_X(x)\, dx.
\end{aligned}$$

Karakterisering Bayes-schatters - bewijs - 2

Lemma.
Zij Y een stochastische variabele met $EY^2 < \infty$. Dan is $t \mapsto E(t - Y)^2$ minimaal in $t = EY$.

Bewijs.
Zelf doen! $\square$

Dus voor iedere vaste x is $E((T(x) - \bar{\Theta})^2 \mid X = x)$ minimaal als

$T(x) = E(\bar{\Theta} \mid X = x) =$ "a-posteriori verwachting bij realisatie $X = x$".

Maar dan minimaliseert de a-posteriori verwachting ook het risico $R(\pi; T)$. $\square$

$\rightarrow$ vb 3.37

Optimaliteitstheorie

Voldoende statistieken - idee

Terminologie: een functie van de waarneming X noemen we ook wel een statistiek.

Idee: als we in plaats van de waarneming X alleen een statistiek $V(X)$ zien, gooien we i.h.a. informatie over de parameter θ weg. Als $V(X)$ toch nog alle informatie over θ bevat, is $V(X)$ voldoende voor θ.

$\rightarrow$ vb 6.1

Voldoende statistieken - definitie discrete geval

Definitie. Stel dat X voor ieder P_θ in het model een discrete verdeling heeft. Dan heet de statistiek $V(X)$ voldoende voor θ als voor alle mogelijke x en v,

$$P_\theta(X = x \mid V = v)$$

niet van θ afhangt.

$\rightarrow$ vb 6.3

Voldoende statistieken - factorisatiestelling

Stelling.

Stel dat X voor ieder P_θ in het model een discrete verdeling heeft en zij p_θ de corresponderende kansdichtheid. Dan is de statistiek $V(X)$ voldoende voor θ dan en slechts dan als er functies g_θ en h bestaan z.d.d. voor alle x en θ,

$$p_\theta(x) = g_\theta(V(x))h(x).$$

$\rightarrow$ vb 6.5

Voldoende statistieken - definitie continue geval

Definitie. Zij p_θ de kansdichtheid van X onder P_θ (discreet of continu). Een statistiek $V(X)$ heet voldoende voor θ als er functies g_θ en h bestaan z.d.d. voor alle x en θ,

$$p_\theta(x) = g_\theta(V(x))h(x).$$

Lemma.

- Als V voldoende is en $V = f(V^*)$ voor een statistiek V^* en een functie f, dan is V^* ook voldoende.
- Als f een bijectie is, dan is V voldoende dan en slechts dan als $f(V)$ voldoende is.

$\rightarrow$ vb 6.8, 6.9

Schatters met een minimale kwadratische fout

Willen: schatter T met een zo klein mogelijke $E_\theta(T-\theta)^2$.

Maar: er bestaat i.h.a. *geen* T_0 zdd $E_\theta(T_0-\theta)^2 \leq E_\theta(T-\theta)^2$ voor *alle* T en θ (check!).

Mogelijke "oplossingen":

- Bayes criterium: kies een prior π en minimaliseer $T \mapsto \int MSE(\theta; T)\pi(\theta)\, d\theta$. Geeft de Bayes-schatter.
- Minimaliseer $T \mapsto \sup_{\theta\in\Theta} MSE(\theta; T)$. Oplossing heet een minimax schatter. (vb 6.14).
- Beperk tot zuivere schatters, i.e. minimaliseer $T \mapsto Var_\theta T$ over de klasse van zuivere schatters. (Voor "grote n" een redelijke restrictie?)

UMVZ schatters

Definitie.
Een schatter T heet UMVZ voor θ als T zuiver is en voor alle zuivere schatters S en $\theta \in \Theta$: $Var_\theta T \leq Var_\theta S$.

Bij het zoeken naar UMVZ schatter kunnen we ons beperken tot schatters op basis van een voldoende statistiek:

Stelling. [Rao-Blackwell]
Zij T een schatter en V een voldoende statistiek. Dan bestaat er een schatter $T^* = T^*(V)$ zodanig dat voor alle $\theta \in \Theta$,

$$E_\theta T^* = E_\theta T, \qquad Var_\theta T^* \leq Var_\theta T$$

Bewijs Rao-Blackwell

Extra aanname: X is discreet verdeeld.

Definieer $T^* = T^*(V) = E_\theta(T \mid V)$. Dan:

- T^* is een schatter:

$$T^*(v) = E_\theta(T(X) \mid V = v) = \sum_x T(x) \underbrace{P_\theta(X = x \mid V = v)}_{\text{onafh. van } \theta}.$$

  Dus T^* is een functie van v die niet van θ afhangt.

- Standaard eigenschappen van de voorwaardelijke verwachting geven $E_\theta T^* = E_\theta T$ en $Var_\theta T^* \leq Var_\theta T$. (check!)

(Als X niet discreet is, algemene theorie van voorwaardelijke verwachtingen nodig).

□

Volledigheid

Vraag: wanneer is een zuivere schatter op basis van een voldoende statistiek UMVZ?

Definitie.
Een statistiek V heet volledig als voor "alle" functies f:

$$\forall\theta : E_\theta f(V) = 0 \qquad \Longrightarrow \qquad \forall\theta : P_\theta(f(V) = 0) = 1.$$

Stelling.
Als V voldoende en volledig is en $T = T(V)$ is een zuiver schatter op basis van V, dan is T UMVZ.

Bewijs

Zij S een zuivere schatter. Vanwege Rao-Blackwell is er een zuivere schatter $S^* = S^*(V)$ met een kleinere variantie. Omdat $T = T(V)$ ook zuiver is, geldt

$$\forall\theta : E_\theta(S^*(V) - T(V)) = \theta - \theta = 0.$$

Vanwege volledigheid (neem $f = S^* - T$) hebben we dan dus

$$\forall\theta : S^*(V) = T(V)$$

met P_θ-kans 1. Volgt dat voor alle θ,

$$Var_\theta T = Var_\theta S^* \leq Var_\theta S.$$

$\square$

→ vb. 6.19

Exponentiële families

Definitie. Een familie kansdichtheden p_θ, $\theta \in \Theta$, heet een k-dimensionale exponentiële familie als er functies c, h, Q_j en V_j bestaan z.d.d. voor alle x,

$$p_\theta(x) = c(\theta)h(x)e^{\sum_{j=1}^k Q_j(\theta)V_j(x)}.$$

N.B. Als $X \sim p_\theta$ met p_θ uit een k-dimensionale exponentiële familie, dan is $(V_1, \ldots, V_k)$ een voldoende statistiek.

Stelling.

$X \sim p_\theta$ met p_θ uit een k-dimensionale exponentiële familie.
Als

$$\Big\{(Q_1(\theta), \ldots, Q_k(\theta)) : \theta \in \Theta\Big\} \subset \mathbb{R}^k$$

een inwendig punt bevat, dan is $(V_1, \ldots, V_k)$ voldoende en volledig.

$\rightarrow$ vb. 6.22 t/m 6.25