Stochastiek 2

Inleiding in de Mathematische Statistiek

Maximum Likelihood-Schatters

Algemene definitie - 1

Gegegeven: waarneming X en model $\{P_\theta : \theta \in \Theta\}$.

Als X discreet is onder P_θ, definieer dan p_θ als de *kansmassa functie* van X. Dus als X waarden $x_1, x_2, \ldots$ aanneemt, dan

$$p_\theta(x_i) = P_\theta(X = x_i), \qquad i = 1, 2, \ldots.$$

Als X absoluut continu is onder P_θ, definieer dan p_θ als de *dichtheid* van X.

We noemen p_θ in beide gevallen de kansdichtheid van X.

Algemene definitie - 2

Gegegeven: waarneming X en model $\{P_\theta : \theta \in \Theta\}$. Zij p_θ de kansdichtheid van X onder P_θ.

Definitie.
Voor een vaste realisatie x van X definiëren we

$$L(\theta; x) = p_\theta(x), \quad \theta \in \Theta.$$

De functie $\theta \mapsto L(\theta; x)$ op Θ heet de likelihood-functie.

Algemene definitie - 3

Definitie.

Voor een realisatie x van X definiëren we de maximum likelihood-schatting als het punt $T(x)$ waar de likelihood $\theta \mapsto L(\theta; x)$ maximaal is. De maximum likelihood-schatter is de bijbehorende schatter $T(X)$.

Momentenschatters

Momenten

Definitie.
Zij X een stochastische variabele en $j \in \mathbb{N}$. Het $j$de moment van X is de verwachting EX^j.

Laat $X_1, \ldots, X_n$ o.o. zijn en gelijk verdeeld en $j \in \mathbb{N}$. Het $j$de steekproef moment is het gemiddelde

$$\bar{X}^j = \frac{1}{n} \sum_{i=1}^{n} X_i^j.$$

Momentenschatters

Idee: steekproefmomenten zijn goede schatters voor echte momenten (WGA). Schat parameter(s) z.d.d. ze *matchen*.

Definitie.
Laat $X_1, \ldots, X_n$ o.o. zijn en gelijk verdeeld volgens P_θ, $\theta \in \Theta$. Een momentenschatter $\hat{\theta}$ voor θ is de oplossing van een stelsel vergelijkingen

$$E_\theta X_1^j = \bar{X}^j, \qquad j \in J,$$

voor zekere $J \subset \mathbb{N}$.

N.B. De conventie is om J zo te kiezen dat de laagste momenten worden gebruikt en dat er een unieke oplossing is.

$\longrightarrow$ vb 3.27, 3.28

Bayes-schatters

Bayesiaanse paradigma

Tot nu toe: waarneming X heeft een verdeling die afhangt van een onbekende parameter θ. Deze heeft een vaste, deterministsche waarde die we proberen te schatten.

Bayesiaanse uitgangspunt: van onbekende grootheden zoals de parameter θ moet je niet aannemen dat ze een vaste, maar onbekende waarde hebben, maar die moet je opvatten als stochastische grootheden.

$\longrightarrow$ natuurconstanten

Bayesiaanse statistiek

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

- De ideeën/meningen die we hebben over de mogelijke waarden van de parameter **voordat** we de data gezien hebben kwantificeren we mbv een a-priori dichtheid $\pi : \Theta \to [0, \infty)$.
- De dichtheid p_θ van de data vatten we op als de **voorwaardelijke** dichtheid van X, gegeven dat de parameter de waarde θ heeft.
- **Nadat** we de data $X = x$ hebben gezien, bepalen we in deze setting de a-posteriori verdeling van de parameter: de verdeling van de parameter gegeven $X = x$.

Bepalen van de a-posteriori verdeling - 1

Dus voor de Bayesiaan:

- De parameter is een stochast $\bar{\Theta}$ met dichtheid $p_{\bar{\Theta}} = \pi$.
- Voor de voorwaardelijke dichtheid van X gegeven $\bar{\Theta} = \theta$ geldt $p_{X \mid \bar{\Theta}=\theta}(x) = p_\theta(x)$.

In deze setting heeft het paar $(X, \bar{\Theta})$ simultane dichtheid

$$p_{X,\bar{\Theta}}(x, \theta) = p_{X \mid \bar{\Theta}=\theta}(x) p_{\bar{\Theta}}(\theta) = p_\theta(x)\pi(\theta)$$

en X heeft marginale dichtheid

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

Bepalen van de a-posteriori verdeling - 2

Maar dan geldt voor de a-posteriori dichtheid $p_{\bar{\Theta} \,|\, X=x}$

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

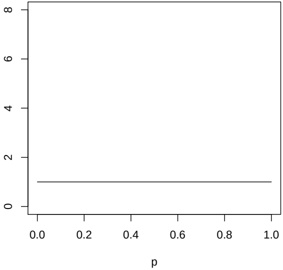
We gebruiken deze uitdrukking nu als **definitie** voor de a-posteriori verdeling.

Bayes' voorbeeld

- $X \sim \text{bin}(50, \theta)$, waarneming $X = 42$. Wat is θ?

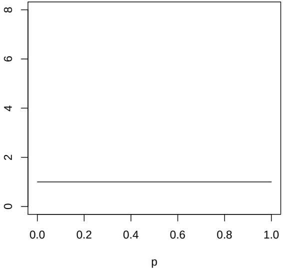
Bayes' voorbeeld

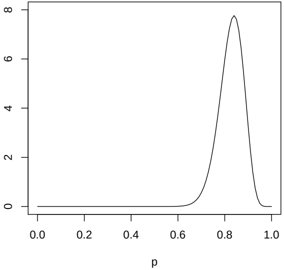
- $X \sim \text{bin}(50, \theta)$, waarneming $X = 42$. Wat is θ?
- Neem uniforme a-priori verdeling op $[0, 1]$.



Bayes' voorbeeld

- $X \sim \text{bin}(50, \theta)$, waarneming $X = 42$. Wat is θ?
- Neem uniforme a-priori verdeling op $[0, 1]$.





a-priori π $\quad \xrightarrow{\text{data } X = 42} \quad$ a-posteriori $p_{\bar{\Theta} \mid X=42}$

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

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

$\longrightarrow$ vb 3.38, 3.37.

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 $E((T(x) - \bar{\Theta})^2 \,|\, X = x)$ minimaal als

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

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