Estimadores de Bayes

- Estimador e estimativa;
- Função de perda;
- Estimador de Bayes;
- Consistência do estimador de Bayes;
- Estimador de Bayes para grandes amostras;
- Limitações.

Definição 15 (Priori imprópria)

Seja $\xi : \Lambda \to (0, \infty)$, $\Omega \subseteq \Lambda$, *uma função tal que* $\int_\Omega \xi(\theta)\, d\theta = \infty$. *Se utilizamos* ξ *como uma p.d.f. para* θ, *dizemos que* ξ *é uma* ***priori imprópria*** *para* θ.

Exemplo 3 (Priori imprópria para a taxa de uma Poisson)

Suponha que $X_1, X_2, \ldots, X_n$ *formam uma amostra aleatória com distribuição Poisson com taxa* $\theta > 0$, *desconhecida. Desta vez, fazemos a escolha de hiperparâmetros* $\alpha = \beta = 0$, *o que leva a*

$$\xi(\theta) = \frac{1}{\theta}.$$

A posteriori passa a ser

$$\xi(\theta \mid \boldsymbol{x}) = \frac{n^S}{\Gamma(S)} \theta^{n-1} e^{-S\theta},$$

onde $S = \sum_{i=1}^n x_i$.

Estimadores (de Bayes)

Definição 16 (Estimador)

Sejam $X_1, X_2, \ldots, X_n$ *variáveis aleatórias com distribuição conjunta indexada por* θ. *Um* ***estimador*** *de* θ *é qualquer função real* $\delta : X_1, X_2, \ldots, X_n \to \mathbb{R}^d$, $d \geq 1$.

Definição 17 (Estimativa)

Dizemos que o valor de δ *avaliado nas realizações de* $X_1, X_2, \ldots, X_n$, $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$, $\delta(\boldsymbol{x})$ *é uma* ***estimativa*** *de* θ.

Definição 18 (Função de perda)

Uma função de perda é uma função real em duas variáveis

$$L : \Omega \times \mathbb{R}^d \to \mathbb{R},$$

em que dizemos que o estatístico <u>*perde*</u> $L(\theta, a)$ *se o parâmetro vale* θ *e a estimativa dada vale* a.

Exemplos de funções de perda são $L(\theta, a) = (\theta - a)^2$ e $L(\theta, a) = |\theta - a|$.

Observação 2 (Perda esperada *a priori*)

Se escolhemos uma priori $\xi(\theta)$, nossa perda esperada, ***antes*** *de observar os dados é*

$$E_\xi[L(\theta, a)] = \int_\Omega L(\theta, a)\xi(\theta)\, d\theta.$$

Vemos então que a escolha da distribuição a priori está inextrincavelmente ligada à função de perda.

Definição 19 (Estimador de Bayes)

Considere a perda esperada a posteriori:

$$E_{\theta|\boldsymbol{x}}\left[L(\theta, a)\right] = E[L(\theta, a) \mid \boldsymbol{x}] = \int_{\Omega} L(\theta, a)\xi(\theta \mid \boldsymbol{x})\, d\theta.$$

*Dizemos que δ^* é um **estimador de Bayes** se, para toda realização $\boldsymbol{X} = \boldsymbol{x}$,*

$$E[L(\theta, \delta^*(\boldsymbol{x})) \mid \boldsymbol{x}] = \min_{a \in \mathcal{A}} E[L(\theta, a) \mid \boldsymbol{x}].$$

- Em outras palavras, um estimador de Bayes é uma função real dos dados que minimiza a perda esperada com respeito à posteriori dos parâmetros.

Estimador de Bayes sob perda quadrática

Suponha que a função de perda seja

$$L(\theta, \delta^*) = (\theta - \delta^*)^2.$$

Dizemos que a função de perda é **quadrática**. Temos o seguinte resultado:

Teorema 9 (δ^* sob perda quadrática)

(De Groot, Corolário 7.4.1)
Seja θ um parâmetro tomando valores reais. Sob perda quadrática,

$$\delta^*(\boldsymbol{x}) = E[\theta|\boldsymbol{X} = \boldsymbol{x}] = \int_{\Omega} \theta \xi(\theta \mid \boldsymbol{x}) \, d\theta.$$

Prova: Escrever a perda esperada *a posteriori* explicitamente, usar a lei de esperanças e minimizar a expressão resultante com respeito ao estimador (ex. diferenciar e igualar a derivada a zero).

Estimador de Bayes sob perda absoluta

Teorema 10 (δ^* sob perda absoluta)

(De Groot, Corolário 7.4.2)
Suponha que a função de perda é dada por

$$L(\theta, \delta^*) = |\theta - \delta^*|.$$

*Dizemos que a função de perda é **absoluta**.*
*Seja θ um parâmetro tomando valores na reta. Sob perda absoluta, $\delta^*(\boldsymbol{x})$ é a **mediana** a posteriori, isto é,*

$$\int_{-\infty}^{\delta^*(\boldsymbol{x})} \xi(\theta \mid \boldsymbol{x})\, d\theta = \frac{1}{2}.$$

Prova: Decompor a perda esperada em duas integrais de funções não-negativas utilizando as propriedades da função valor absoluto e aplicar a regra de Leibnitz duas vezes para encontrar o ponto de mínimo.

O estimador de Bayes em grandes amostras

- Sob condições brandas de regularidade, à medida que o tamanho de amostra cresce, a influência da priori diminui.

Exemplo 4 (Proporção de itens defeituosos)

Suponha que estamos interessados na proporção θ de itens defeituosos em uma linha de produção. Suponha ainda que

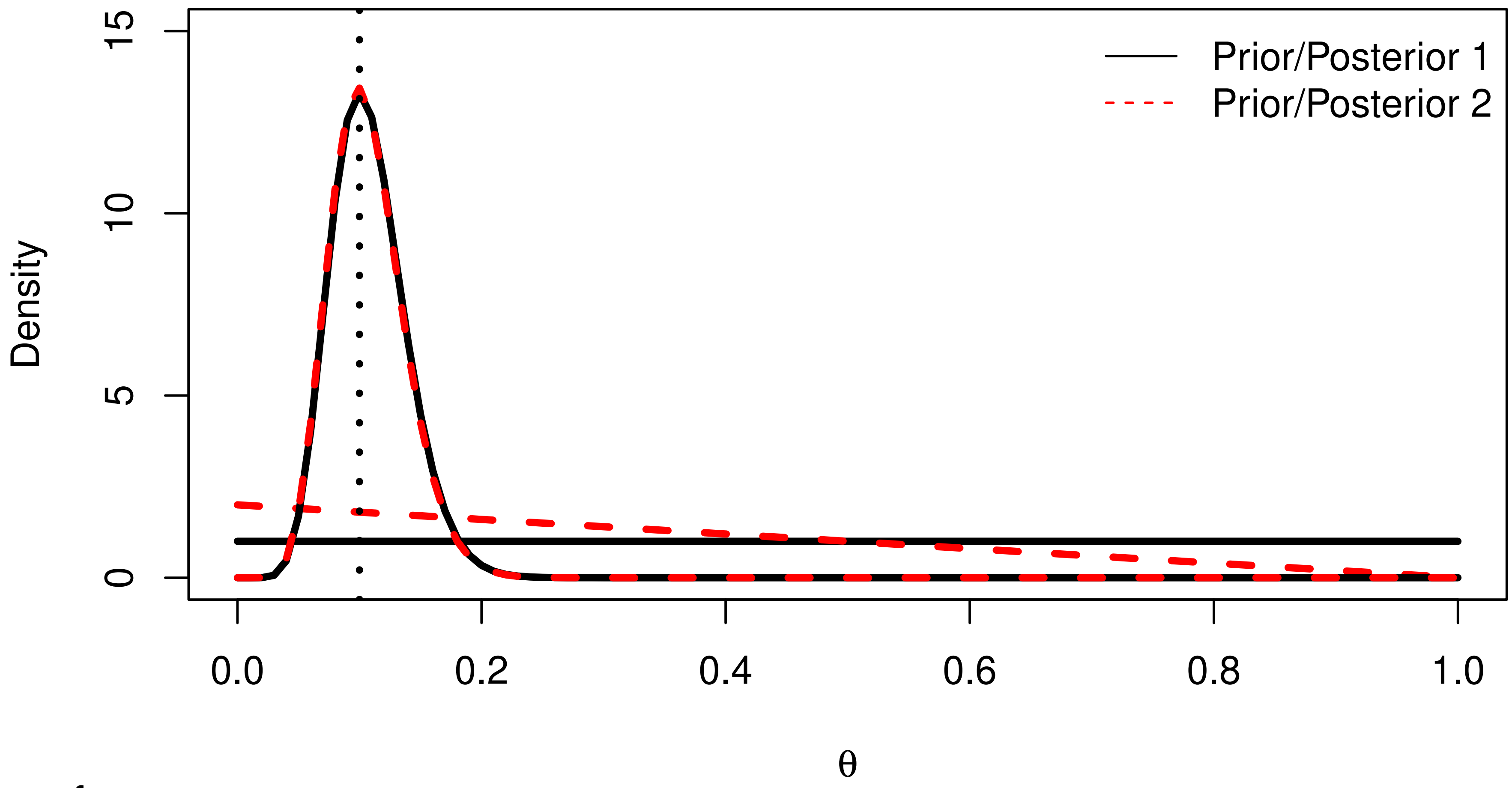
- *Priori 1:* $\xi_1(\theta) = 1,\ 0 < \theta < 1$;
- *Priori 2:* $\xi_2(\theta) = 2(1 - \theta),\ 0 < \theta < 1$;
- *Dados: de* $n = 100$ *itens observados,* $y = 10$ *apresentaram defeito.*

Perguntas:

- $\bar{x}_n = ?$
- $E_1[\theta \mid \mathbf{x}] = \int_0^1 \theta \xi_1(\theta \mid \mathbf{x})\, d\theta = ?$;
- $E_2[\theta \mid \mathbf{x}] = \int_0^1 \theta \xi_2(\theta \mid \mathbf{x})\, d\theta = ?$;

Proporção de itens defeituosos: prioris e posterioris

Ver também exemplo 7.3.3 de De Groot.



Definição 20 (Estimador consistente)

Seja $\delta_1, \delta_2, \ldots, \delta_n$ uma sequência de estimadores de θ. Se quando $n \to \infty$ a sequência converge para θ, dizemos que esta é uma sequência consistente de estimadores.

Observação 3 (A média amostral é consistente para o caso Bernoulli)

Se $X_1, X_2, \ldots, X_n$ são i.i.d. Bernoulli com parâmetro θ condicional a θ, temos pela LGN: $\bar{X}_n \xrightarrow{p} \theta$.

Observação 4 (O estimador de Bayes é consistente para o caso Bernoulli)

Para $\alpha > 0$ e $\beta > 0$ fixos, a média a posteriori vale

$$\delta^*(\boldsymbol{x}) = E[\theta \mid \boldsymbol{x}] = \frac{\alpha + y}{\alpha + \beta + n},$$

onde $y = \sum_{i=1}^{n} x_i$. É fácil ver que $\delta^*(\boldsymbol{x}) \xrightarrow{p} \bar{x}_n \xrightarrow{p} \theta$.

O que aprendemos?

- Estimador;
  "Um estimador é qualquer função real dos dados"
- Função de perda;
  "Uma função real que quantifica a perda incorrida por uma estimativa incorreta"
- Estimador de Bayes;
  "Um estimador que minimiza a perda esperada *a posteriori*"
- Propriedades e limitações do estimador de Bayes;
  "À medida que o tamanho da amostra cresce, o estimador se aproxima do valor verdadeiro, a influência da priori diminui, mas precisamos sempre de uma função de perda bem especificada"

Leitura recommendada

- De Groot seção 7.4;
- * Casella & Berger, seção 7.3.
- **Exercícios recomendados**
  - De Groot, seção 7.4: exercícios 2, 4, 7, 11 e 14.