Inferência Estatística Comparada

AULA 9 - INFERÊNCIA BAYESIANA

Introdução

- A inferência mais completa sobre θ, sob a abordagem bayesiana, é a distribuição a posteriori para esse parâmetro.
- No entanto, a partir da posteriori, podemos extrair estimativas pontuais, intervalares (por região), testar hipóteses, etc.. Do mesmo modo, pré-experimentação, podemos construir estimadores, funções de testes, etc..
- Tais procedimentos são comumente derivados a partir da Teoria da Decisão.
- Relembremos, em linha gerais, o procedimento de obtenção de decisões (regra de decisões) ótimas sob a Teoria da Decisão.

Introdução

- A inferência mais completa sobre θ, sob a abordagem bayesiana, é a distribuição a posteriori para esse parâmetro.
- No entanto, a partir da posteriori, podemos extrair estimativas pontuais, intervalares (por região), testar hipóteses, etc.. Do mesmo modo, pré-experimentação, podemos construir estimadores, funções de testes, etc..
- Tais procedimentos são comumente derivados a partir da Teoria da Decisão.
- Relembremos, em linha gerais, o procedimento de obtenção de decisões (regra de decisões) ótimas sob a Teoria da Decisão.

Introdução

- A inferência mais completa sobre θ, sob a abordagem bayesiana, é a distribuição a posteriori para esse parâmetro.
- No entanto, a partir da posteriori, podemos extrair estimativas pontuais, intervalares (por região), testar hipóteses, etc.. Do mesmo modo, pré-experimentação, podemos construir estimadores, funções de testes, etc..
- Tais procedimentos são comumente derivados a partir da Teoria da Decisão.
- Relembremos, em linha gerais, o procedimento de obtenção de decisões (regra de decisões) ótimas sob a Teoria da Decisão.

Introdução

- A inferência mais completa sobre θ, sob a abordagem bayesiana, é a distribuição a posteriori para esse parâmetro.
- No entanto, a partir da posteriori, podemos extrair estimativas pontuais, intervalares (por região), testar hipóteses, etc.. Do mesmo modo, pré-experimentação, podemos construir estimadores, funções de testes, etc..
- Tais procedimentos são comumente derivados a partir da Teoria da Decisão.
- Relembremos, em linha gerais, o procedimento de obtenção de decisões (regra de decisões) ótimas sob a Teoria da Decisão.

Introdução

- A inferência mais completa sobre θ, sob a abordagem bayesiana, é a distribuição a posteriori para esse parâmetro.
- No entanto, a partir da posteriori, podemos extrair estimativas pontuais, intervalares (por região), testar hipóteses, etc.. Do mesmo modo, pré-experimentação, podemos construir estimadores, funções de testes, etc..
- Tais procedimentos são comumente derivados a partir da Teoria da Decisão.
- Relembremos, em linha gerais, o procedimento de obtenção de decisões (regra de decisões) ótimas sob a Teoria da Decisão.

Elementos do problema de decisão

- **Problema de decisão:** $(\mathcal{D}, \Theta, L, \mathbb{P})$, onde
- $\mathcal{D}$: espaço decisões (ações)
- Θ: "estado da natureza" desconhecido
- L: função de perda $L : \mathcal{D} \times \Theta \to \mathbb{R}$, onde $L(d, \theta)$ que "representa" a penalidade que o agente decisor incorre ao tomar a decisão $d \in \mathcal{D}$ sob o estado da natureza θ.
- $\mathbb{P}$: medida de probabilidade para o estado da natureza θ.
- Observação: Alternativamente à consideração de uma função de perda, é comum a adoção de uma função Utilidade, que representa "recompensas", ao invés de "penalidades".

Elementos do problema de decisão

- **Problema de decisão:** $(\mathcal{D}, \Theta, L, \mathbb{P})$, onde
- $\mathcal{D}$: espaço decisões (ações)
- Θ: "estado da natureza" desconhecido
- L: função de perda $L : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$, onde $L(d, \theta)$ que "representa" a penalidade que o agente decisor incorre ao tomar a decisão $d \in \mathcal{D}$ sob o estado da natureza θ.
- $\mathbb{P}$: medida de probabilidade para o estado da natureza θ.
- Observação: Alternativamente à consideração de uma função de perda, é comum a adoção de uma função Utilidade, que representa "recompensas", ao invés de "penalidades".

Elementos do problema de decisão

- **Problema de decisão:** $(\mathcal{D}, \Theta, L, \mathbb{P})$, onde
- $\mathcal{D}$: espaço decisões (ações)
- Θ: "estado da natureza" desconhecido
- L: função de perda $L : \mathcal{D} \times \Theta \to \mathbb{R}$, onde $L(d, \theta)$ que "representa" a penalidade que o agente decisor incorre ao tomar a decisão $d \in \mathcal{D}$ sob o estado da natureza θ.
- $\mathbb{P}$: medida de probabilidade para o estado da natureza θ.
- Observação: Alternativamente à consideração de uma função de perda, é comum a adoção de uma função Utilidade, que representa "recompensas", ao invés de "penalidades".

Elementos do problema de decisão

- **Problema de decisão:** $(\mathcal{D}, \Theta, L, \mathbb{P})$, onde
- $\mathcal{D}$: espaço decisões (ações)
- Θ: "estado da natureza" desconhecido
- L: função de perda $L : \mathcal{D} \times \Theta \to \mathbb{R}$, onde $L(d, \theta)$ que "representa" a penalidade que o agente decisor incorre ao tomar a decisão $d \in \mathcal{D}$ sob o estado da natureza θ.
- $\mathbb{P}$: medida de probabilidade para o estado da natureza θ.
- Observação: Alternativamente à consideração de uma função de perda, é comum a adoção de uma função Utilidade, que representa "recompensas", ao invés de "penalidades".

Elementos do problema de decisão

- **Problema de decisão:** $(\mathcal{D}, \Theta, L, \mathbb{P})$, onde
- $\mathcal{D}$: espaço decisões (ações)
- Θ: "estado da natureza" desconhecido
- L: função de perda $L : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$, onde $L(d, \theta)$ que "representa" a penalidade que o agente decisor incorre ao tomar a decisão $d \in \mathcal{D}$ sob o estado da natureza θ.
- $\mathbb{P}$: medida de probabilidade para o estado da natureza θ.
- Observação: Alternativamente à consideração de uma função de perda, é comum a adoção de uma função Utilidade, que representa "recompensas", ao invés de "penalidades".

Elementos do problema de decisão

- **Problema de decisão:** $(\mathcal{D}, \Theta, L, \mathbb{P})$, onde
- $\mathcal{D}$: espaço decisões (ações)
- Θ: "estado da natureza" desconhecido
- L: função de perda $L : \mathcal{D} \times \Theta \to \mathbb{R}$, onde $L(d, \theta)$ que "representa" a penalidade que o agente decisor incorre ao tomar a decisão $d \in \mathcal{D}$ sob o estado da natureza θ.
- $\mathbb{P}$: medida de probabilidade para o estado da natureza θ.
- Observação: Alternativamente à consideração de uma função de perda, é comum a adoção de uma função Utilidade, que representa "recompensas", ao invés de "penalidades".

Elementos do problema de decisão

- **Problema de decisão:** $(\mathcal{D}, \Theta, L, \mathbb{P})$, onde
- $\mathcal{D}$: espaço decisões (ações)
- Θ: "estado da natureza" desconhecido
- L: função de perda $L : \mathcal{D} \times \Theta \to \mathbb{R}$, onde $L(d, \theta)$ que "representa" a penalidade que o agente decisor incorre ao tomar a decisão $d \in \mathcal{D}$ sob o estado da natureza θ.
- $\mathbb{P}$: medida de probabilidade para o estado da natureza θ.
- Observação: Alternativamente à consideração de uma função de perda, é comum a adoção de uma função Utilidade, que representa "recompensas", ao invés de "penalidades".

Minimização de perda esperada

- Sob certas condições, assegura-se a representação numérica de preferência por funções de perda (utilidade) e o paradigma de minimização (maximização) de perda (utilidade) esperada.

- **Definição:** Uma decisão $d^* \in \mathcal{D}$ é uma decisão de Bayes contra $\mathbb{P}$, com relação à perda L, se, $\forall d \in \mathcal{D}$,

  $$\mathbb{E}(L(d^*, \theta)) \quad \leq \quad \mathbb{E}(L(d, \theta))$$

- $\mathbb{E}(L(d, \theta))$ é chamada **risco** da decisão d: $\rho(d, \mathbb{P}) = \mathbb{E}(L(d, \theta))$.

- $\mathbb{E}(L(d^*, \theta))$ é chamada **risco de Bayes**: $\rho^*(\mathbb{P}) = \rho(d^*, \mathbb{P}) = \mathbb{E}(L(d^*, \theta))$.
  (Em geral, $\rho^*(\mathbb{P}) = \inf\{\rho(d, \mathbb{P}) : d \in \mathcal{D}\}$)

Minimização de perda esperada

- Sob certas condições, assegura-se a representação numérica de preferência por funções de perda (utilidade) e o paradigma de minimização (maximização) de perda (utilidade) esperada.

- **Definição:** Uma decisão $d^* \in \mathcal{D}$ é uma decisão de Bayes contra $\mathbb{P}$, com relação à perda L, se, $\forall d \in \mathcal{D}$,

$$\mathbb{E}(L(d^*, \theta)) \quad \leq \quad \mathbb{E}(L(d, \theta))$$

- $\mathbb{E}(L(d, \theta))$ é chamada **risco** da decisão d: $\rho(d, \mathbb{P}) = \mathbb{E}(L(d, \theta))$.

- $\mathbb{E}(L(d^*, \theta))$ é chamada **risco de Bayes**: $\rho^*(\mathbb{P}) = \rho(d^*, \mathbb{P}) = \mathbb{E}(L(d^*, \theta))$. (Em geral, $\rho^*(\mathbb{P}) = \inf\{\rho(d, \mathbb{P}) : d \in \mathcal{D}\}$)

Minimização de perda esperada

- Sob certas condições, assegura-se a representação numérica de preferência por funções de perda (utilidade) e o paradigma de minimização (maximização) de perda (utilidade) esperada.

- **Definição:** Uma decisão $d^* \in \mathcal{D}$ é uma decisão de Bayes contra $\mathbb{P}$, com relação à perda L, se, $\forall d \in \mathcal{D}$,

$$\mathbb{E}(L(d^*, \theta)) \quad \leq \quad \mathbb{E}(L(d, \theta))$$

- $\mathbb{E}(L(d, \theta))$ é chamada **risco** da decisão d: $\rho(d, \mathbb{P}) = \mathbb{E}(L(d, \theta))$.

- $\mathbb{E}(L(d^*, \theta))$ é chamada **risco de Bayes**: $\rho^*(\mathbb{P}) = \rho(d^*, \mathbb{P}) = \mathbb{E}(L(d^*, \theta))$. (Em geral, $\rho^*(\mathbb{P}) = \inf\{\rho(d, \mathbb{P}) : d \in \mathcal{D}\}$)

Minimização de perda esperada

- Sob certas condições, assegura-se a representação numérica de preferência por funções de perda (utilidade) e o paradigma de minimização (maximização) de perda (utilidade) esperada.

- **Definição:** Uma decisão $d^* \in \mathcal{D}$ é uma decisão de Bayes contra $\mathbb{P}$, com relação à perda L, se, $\forall d \in \mathcal{D}$,

$$\mathbb{E}(L(d^*, \theta)) \quad \leq \quad \mathbb{E}(L(d, \theta))$$

- $\mathbb{E}(L(d, \theta))$ é chamada **risco** da decisão d: $\rho(d, \mathbb{P}) = \mathbb{E}(L(d, \theta))$.

- $\mathbb{E}(L(d^*, \theta))$ é chamada **risco de Bayes**: $\rho^*(\mathbb{P}) = \rho(d^*, \mathbb{P}) = \mathbb{E}(L(d^*, \theta))$. (Em geral, $\rho^*(\mathbb{P}) = \inf\{\rho(d, \mathbb{P}) : d \in \mathcal{D}\}$)

Minimização de perda esperada

- Sob certas condições, assegura-se a representação numérica de preferência por funções de perda (utilidade) e o paradigma de minimização (maximização) de perda (utilidade) esperada.

- **Definição:** Uma decisão $d^* \in \mathcal{D}$ é uma decisão de Bayes contra $\mathbb{P}$, com relação à perda L, se, $\forall d \in \mathcal{D}$,

$$\mathbb{E}(L(d^*, \theta)) \quad \leq \quad \mathbb{E}(L(d, \theta))$$

- $\mathbb{E}(L(d, \theta))$ é chamada **risco** da decisão d:
$\rho(d, \mathbb{P}) = \mathbb{E}(L(d, \theta))$.

- $\mathbb{E}(L(d^*, \theta))$ é chamada **risco de Bayes**:
$\rho^*(\mathbb{P}) = \rho(d^*, \mathbb{P}) = \mathbb{E}(L(d^*, \theta))$.
(Em geral, $\rho^*(\mathbb{P}) = \inf\{\rho(d, \mathbb{P}) : d \in \mathcal{D}\}$)

Estimação

- Na sequência, vamos examinar problemas tradicionais em inferência estatística: estimação e testes de hipóteses.
- **Problema de estimação:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$, onde
- $\Delta = \{\delta : \mathcal{X} \to \Theta\}$: conjunto de estimadores para θ
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(\delta, (\theta, x)) = L(\delta(x), \theta)$, onde $L : \Theta \times \Theta \to \mathbb{R}$ associa a cada par $(d, \theta) \in \Theta^2$ a penalidade $L(d, \theta)$ para a estimativa d quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, vamos examinar problemas tradicionais em inferência estatística: estimação e testes de hipóteses.
- **Problema de estimação:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$, onde
- $\Delta = \{\delta : \mathcal{X} \to \Theta\}$: conjunto de estimadores para θ
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(\delta, (\theta, x)) = L(\delta(x), \theta)$, onde $L : \Theta \times \Theta \to \mathbb{R}$ associa a cada par $(d, \theta) \in \Theta^2$ a penalidade $L(d, \theta)$ para a estimativa d quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, vamos examinar problemas tradicionais em inferência estatística: estimação e testes de hipóteses.
- **Problema de estimação:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$, onde
- $\Delta = \{\delta : \mathcal{X} \to \Theta\}$: conjunto de estimadores para θ
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(\delta, (\theta, x)) = L(\delta(x), \theta)$, onde $L : \Theta \times \Theta \to \mathbb{R}$ associa a cada par $(d, \theta) \in \Theta^2$ a penalidade $L(d, \theta)$ para a estimativa d quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, vamos examinar problemas tradicionais em inferência estatística: estimação e testes de hipóteses.
- **Problema de estimação:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$, onde
- $\Delta = \{\delta : \mathcal{X} \to \Theta\}$: conjunto de estimadores para θ
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(\delta, (\theta, x)) = L(\delta(x), \theta)$, onde $L : \Theta \times \Theta \to \mathbb{R}$ associa a cada par $(d, \theta) \in \Theta^2$ a penalidade $L(d, \theta)$ para a estimativa d quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, vamos examinar problemas tradicionais em inferência estatística: estimação e testes de hipóteses.
- **Problema de estimação:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$, onde
- $\Delta = \{\delta : \mathcal{X} \to \Theta\}$: conjunto de estimadores para θ
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(\delta, (\theta, x)) = L(\delta(x), \theta)$, onde $L : \Theta \times \Theta \to \mathbb{R}$ associa a cada par $(d, \theta) \in \Theta^2$ a penalidade $L(d, \theta)$ para a estimativa d quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, vamos examinar problemas tradicionais em inferência estatística: estimação e testes de hipóteses.
- **Problema de estimação:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$, onde
- $\Delta = \{\delta : \mathcal{X} \to \Theta\}$: conjunto de estimadores para θ
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(\delta, (\theta, x)) = L(\delta(x), \theta)$, onde $L : \Theta \times \Theta \to \mathbb{R}$ associa a cada par $(d, \theta) \in \Theta^2$ a penalidade $L(d, \theta)$ para a estimativa d quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, vamos examinar problemas tradicionais em inferência estatística: estimação e testes de hipóteses.
- **Problema de estimação:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$, onde
- $\Delta = \{\delta : \mathcal{X} \to \Theta\}$: conjunto de estimadores para θ
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(\delta, (\theta, x)) = L(\delta(x), \theta)$, onde $L : \Theta \times \Theta \to \mathbb{R}$ associa a cada par $(d, \theta) \in \Theta^2$ a penalidade $L(d, \theta)$ para a estimativa d quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Consideremos $L \geq 0$ e (θ, X) contínuo. Nesse caso, para cada estimador $\delta \in \Delta$, o risco de δ é dado por

- $\rho(\delta, \bar{\mathbb{P}}) = \int_{\Theta \times \mathcal{X}} \bar{L}(\delta, (\theta, x))\, f(\theta, x)\, d\theta\, dx =$

$$= \int_\Theta \underbrace{\left[\int_{\mathcal{X}} L(\delta(x), \theta)\, f(x|\theta)\, dx \right]}_{RISCO\ FREQUENTISTA\ R(\delta,\theta)} f(\theta)\, d\theta =$$

$$= \int_{\mathcal{X}} \underbrace{\left[\int_\Theta L(\delta(x), \theta)\, f(\theta|x)\, d\theta \right]}_{\rho(\delta(x), \mathbb{P}_x)} f(x)\, dx =$$

$= \int_{\mathcal{X}} \rho(\delta(x), \mathbb{P}_x)\, f(x)\, dx \geq \int_{\mathcal{X}} \rho(d_x^*, \mathbb{P}_x)\, f(x)\, dx$, onde

para cada x, $d_x^* \in \Theta$ é decisão de Bayes de $(\Theta, \Theta, L, \mathbb{P}_x)$.

Estimação

- Consideremos $L \geq 0$ e (θ, X) contínuo. Nesse caso, para cada estimador $\delta \in \Delta$, o risco de δ é dado por
- $\rho(\delta, \bar{\mathbb{P}}) = \int_{\Theta \times \mathcal{X}} \bar{L}(\delta, (\theta, x)) \; f(\theta, x) \; d\theta \; dx =$

$$= \int_\Theta \underbrace{\left[\int_{\mathcal{X}} L(\delta(x), \theta) \; f(x|\theta) \; dx \right]}_{RISCO\ FREQUENTISTA\ R(\delta,\theta)} f(\theta) \; d\theta =$$

$$= \int_{\mathcal{X}} \underbrace{\left[\int_\Theta L(\delta(x), \theta) \; f(\theta|x) \; d\theta \right]}_{\rho(\delta(x), \mathbb{P}_x)} f(x) \; dx =$$

$= \int_{\mathcal{X}} \rho(\delta(x), \mathbb{P}_x) \; f(x) \; dx \geq \int_{\mathcal{X}} \rho(d_x^*, \mathbb{P}_x) \; f(x) \; dx$, onde

para cada x, $d_x^* \in \Theta$ é decisão de Bayes de $(\Theta, \Theta, L, \mathbb{P}_x)$.

Estimação

- Desse modo, definindo $\delta^* : \mathcal{X} \to \Theta$ por $\delta^*(x) = d_x^*$, obtemos que, para todo $\delta \in \Delta$

  $$\rho(\delta^*, \bar{\mathbb{P}}) \;\leq\; \rho(\delta, \bar{\mathbb{P}})$$

  Assim, δ^* é estimador de Bayes considerando $\bar{L}$ contra $\bar{\mathbb{P}}$.

- Note que a obtenção da estimativa para θ considerando perda $\bar{L}$ ao observar $x \in \mathcal{X}$ envolve APENAS a distribuição a posteriori de θ dado $X = x$ (minimização da integral "interna" em vermelho). Ou ainda, envolve a distribuição a priori para θ e APENAS a função de verossimilhança para θ gerada por x, sem fazer menção (considerar) funções de verossimilhança geradas por outros pontos do espaço amostral (note que não é necessário construir o estimador δ^* para então obter a estimativa $\delta^*(x)$).

Estimação

- Desse modo, definindo $\delta^* : \mathcal{X} \to \Theta$ por $\delta^*(x) = d_x^*$, obtemos que, para todo $\delta \in \Delta$

  $$\rho(\delta^*, \bar{\mathbb{P}}) \;\leq\; \rho(\delta, \bar{\mathbb{P}})$$

  Assim, δ^* é estimador de Bayes considerando $\bar{L}$ contra $\bar{\mathbb{P}}$.

- Note que a obtenção da estimativa para θ considerando perda $\bar{L}$ ao observar $x \in \mathcal{X}$ envolve APENAS a distribuição a posteriori de θ dado $X = x$ (minimização da integral "interna" em vermelho). Ou ainda, envolve a distribuição a priori para θ e APENAS a função de verossimilhança para θ gerada por x, sem fazer menção (considerar) funções de verossimilhança geradas por outros pontos do espaço amostral (note que não é necessário construir o estimador δ^* para então obter a estimativa $\delta^*(x)$).

Estimação

- Sob certas condições, decisões ótimas (de Bayes), tais como estimativas pontuais ou intervalares, decisões em testes de hipóteses, são obtidas de modo similar: fixada $L : \mathcal{D} \times \Theta \to \mathbb{R}$, ao observar $x \in \mathcal{X}$, obtem-se a posteriori de θ dado $X = x$ e então toma-se $d^* \in \mathcal{D}$ solução do problema $(\mathcal{D}, \Theta.L, \mathbb{P}_x)$.

- Funções de perda comumente adotadas em estimação $(\mathcal{D} = \Theta \subset \mathbb{R})$:

- 1) $L(d, \theta) = (d - \theta)^2$ **(perda quadrática)**: se para todo $x \in \mathcal{X}$, $\mathbb{E}(\theta|X = x) < \infty$ e $\mathbb{E}(\theta|X = x) \in \Theta$, temos

  $\delta^*(x) = \mathbb{E}(\theta|X = x)$ é estimativa de Bayes com relação à perda quadrática. Nesse caso, o estimador é dado por

Estimação

- Sob certas condições, decisões ótimas (de Bayes), tais como estimativas pontuais ou intervalares, decisões em testes de hipóteses, são obtidas de modo similar: fixada $L : \mathcal{D} \times \Theta \to \mathbb{R}$, ao observar $x \in \mathcal{X}$, obtem-se a posteriori de θ dado $X = x$ e então toma-se $d^* \in \mathcal{D}$ solução do problema $(\mathcal{D}, \Theta.L, \mathbb{P}_x)$.

- Funções de perda comumente adotadas em estimação ($\mathcal{D} = \Theta \subset \mathbb{R}$):

- 1) $L(d, \theta) = (d - \theta)^2$ **(perda quadrática)**: se para todo $x \in \mathcal{X}$, $\mathbb{E}(\theta|X = x) < \infty$ e $\mathbb{E}(\theta|X = x) \in \Theta$, temos

  $\delta^*(x) = \mathbb{E}(\theta|X = x)$ é estimativa de Bayes com relação à perda quadrática. Nesse caso, o estimador é dado por

Estimação

- Sob certas condições, decisões ótimas (de Bayes), tais como estimativas pontuais ou intervalares, decisões em testes de hipóteses, são obtidas de modo similar: fixada $L : \mathcal{D} \times \Theta \to \mathbb{R}$, ao observar $x \in \mathcal{X}$, obtem-se a posteriori de θ dado $X = x$ e então toma-se $d^* \in \mathcal{D}$ solução do problema $(\mathcal{D}, \Theta . L, \mathbb{P}_x)$.

- Funções de perda comumente adotadas em estimação $(\mathcal{D} = \Theta \subset \mathbb{R})$:

- 1) $L(d, \theta) = (d - \theta)^2$ **(perda quadrática)**: se para todo $x \in \mathcal{X}$, $\mathbb{E}(\theta | X = x) < \infty$ e $\mathbb{E}(\theta | X = x) \in \Theta$, temos

  $\delta^*(x) = \mathbb{E}(\theta | X = x)$ é estimativa de Bayes com relação à perda quadrática. Nesse caso, o estimador é dado por

Estimação

- $\delta^*(X) = \mathbb{E}(\theta|X)$, cujo risco é
- $\rho(\delta^*, \bar{\mathbb{P}}) \;=\; \int_{\mathcal{X}} [\, \int_{\Theta} L(\delta^*(x), \theta)\; f(\theta|x)\; d\theta \,]\; f(x)dx \;=$

$$= \int_{\mathcal{X}} [\, \int_{\Theta} (\delta^*(x) - \theta)^2\; f(\theta|x)\; d\theta \,]\; f(x)dx \;=$$

$$= \int_{\mathcal{X}} [\, \int_{\Theta} (\mathbb{E}(\theta|X = x) - \theta)^2\; f(\theta|x)\; d\theta \,]\; f(x)dx \;=$$

$$= \int_{\mathcal{X}} VAR(\theta|X = x)\; f(x)dx \;=\; \mathbb{E}(VAR(\theta|X)) \;\Rightarrow$$

$$\Rightarrow\; \rho(\delta^*, \bar{\mathbb{P}}) \;=\; VAR(\theta) \;-\; VAR(\mathbb{E}(\theta|X))$$

- Note que $\rho(\delta^*, \bar{\mathbb{P}}) \;=\; \mathbb{E}((\delta^*(X) - \theta)^2)$ **NÃO** depende de θ!!!

Estimação

- $\delta^*(X) = \mathbb{E}(\theta|X)$, cujo risco é
- $\rho(\delta^*, \bar{\mathbb{P}}) \; = \; \int_{\mathcal{X}} [\; \int_{\Theta} L(\delta^*(x), \theta) \; f(\theta|x) \; d\theta \;] \; f(x)dx \; =$

$$= \; \int_{\mathcal{X}} [\; \int_{\Theta} (\delta^*(x) - \theta)^2 \; f(\theta|x) \; d\theta \;] \; f(x)dx \; =$$

$$= \; \int_{\mathcal{X}} [\; \int_{\Theta} (\mathbb{E}(\theta|X = x) - \theta)^2 \; f(\theta|x) \; d\theta \;] \; f(x)dx \; =$$

$$= \; \int_{\mathcal{X}} VAR(\theta|X = x) \; f(x)dx \; = \; \mathbb{E}(VAR(\theta|X)) \; \Rightarrow$$

$$\Rightarrow \; \rho(\delta^*, \bar{\mathbb{P}}) \; = \; VAR(\theta) \; - \; VAR(\mathbb{E}(\theta|X))$$

- Note que $\rho(\delta^*, \bar{\mathbb{P}}) \; = \; \mathbb{E}((\delta^*(X) - \theta)^2)$ **NÃO** depende de θ!!!

Estimação

- $\delta^*(X) = \mathbb{E}(\theta|X)$, cujo risco é
- $\rho(\delta^*, \bar{\mathbb{P}}) \;=\; \int_{\mathcal{X}} [\; \int_{\Theta} L(\delta^*(x), \theta)\; f(\theta|x)\; d\theta \;]\; f(x)dx \;=$

$$= \int_{\mathcal{X}} [\; \int_{\Theta} (\delta^*(x) - \theta)^2\; f(\theta|x)\; d\theta \;]\; f(x)dx \;=$$

$$= \int_{\mathcal{X}} [\; \int_{\Theta} (\mathbb{E}(\theta|X=x) - \theta)^2\; f(\theta|x)\; d\theta \;]\; f(x)dx \;=$$

$$= \int_{\mathcal{X}} VAR(\theta|X=x)\; f(x)dx \;=\; \mathbb{E}(VAR(\theta|X)) \;\Rightarrow$$

$$\Rightarrow\; \rho(\delta^*, \bar{\mathbb{P}}) \;=\; VAR(\theta) \;-\; VAR(\mathbb{E}(\theta|X))$$

- Note que $\rho(\delta^*, \bar{\mathbb{P}}) \;=\; \mathbb{E}((\delta^*(X) - \theta)^2)$ **NÃO** depende de θ!!!

Estimação

- Exemplo: Seja $X = (X_1, ..., X_n)$ AAS Bernoulli (θ) e $\theta \sim Beta(a, b)$. Para $x \in \mathcal{X}$, sabemos que

  $\theta | X = x = (x_1, ..., x_n) \sim Beta(a + \sum_{i=1}^n x_i, b + n - \sum_{i=1}^n x_i)$

  Logo, $\delta^*(X) = \mathbb{E}(\theta|X) = \frac{a+\sum_{i=1}^n X_i}{a+b+n}$. Assim, temos

  $\rho(\delta^*, \bar{\mathbb{P}}) = VAR(\theta) - VAR(\mathbb{E}(\theta|X)) =$

  $= VAR(\theta) - VAR(\frac{a+\sum_{i=1}^n X_i}{a+b+n}) = VAR(\theta) - \frac{VAR(\sum_{i=1}^n X_i)}{(a+b+n)^2} =$

  $= VAR(\theta) - \frac{\mathbb{E}(VAR(\sum_{i=1}^n X_i|\theta)) + VAR(\mathbb{E}(\sum_{i=1}^n X_i|\theta))}{(a+b+n)^2} =$

Estimação

- Exemplo: Seja $X = (X_1, ..., X_n)$ AAS Bernoulli (θ) e $\theta \sim Beta(a, b)$. Para $x \in \mathcal{X}$, sabemos que

$$\theta|X = x = (x_1, ..., x_n) \sim Beta(a + \sum_{i=1}^n x_i, b + n - \sum_{i=1}^n x_i)$$

Logo, $\delta^*(X) = \mathbb{E}(\theta|X) = \frac{a+\sum_{i=1}^n X_i}{a+b+n}$. Assim, temos

$$\rho(\delta^*, \bar{\mathbb{P}}) = VAR(\theta) - VAR(\mathbb{E}(\theta|X)) =$$

$$= VAR(\theta) - VAR(\frac{a+\sum_{i=1}^n X_i}{a+b+n}) = VAR(\theta) - \frac{VAR(\sum_{i=1}^n X_i)}{(a+b+n)^2} =$$

$$= VAR(\theta) - \frac{\mathbb{E}(VAR(\sum_{i=1}^n X_i|\theta)) + VAR(\mathbb{E}(\sum_{i=1}^n X_i|\theta))}{(a+b+n)^2} =$$

Estimação

- $= VAR(\theta) - \frac{1}{(a+b+n)^2}\{\mathbb{E}(n\theta(1-\theta)) + VAR(n\theta)\} \Rightarrow$

  $\Rightarrow \rho(\delta^*, \bar{\mathbb{P}}) = \{1 - (\frac{n}{a+b+n})^2\}VAR(\theta) - \frac{n}{(a+b+n)^2}\mathbb{E}(\theta(1-\theta))$

- 2) $L(d,\theta) = a_1(d-\theta)\mathbb{I}_{[\theta,\infty)}(d) + a_2(\theta-d)\mathbb{I}_{(-\infty,\theta)}(d)$, $a_1, a_2 > 0$:

  se para todo $x \in \mathcal{X}$, $\mathbb{E}(\theta|X=x) < \infty$, temos

  $\delta^*(x) = F^{-1}_{\theta|X=x}(\frac{a_1}{a_1+a_2})$, o percentil de ordem $\frac{a_1}{a_1+a_2}$ da distribuição a posteriori de θ dado $X = x$, isto é, $\delta^*(x)$ satisfaz

  $\mathbb{P}(\theta \leq \delta^*(x)|X=x) \geq \frac{a_1}{a_1+a_2}$ e $\mathbb{P}(\theta \geq \delta^*(x)|X=x) \geq \frac{a_2}{a_1+a_2}$

Estimação

- $= VAR(\theta) - \frac{1}{(a+b+n)^2}\{\mathbb{E}(n\theta(1-\theta)) + VAR(n\theta)\} \Rightarrow$

  $\Rightarrow \rho(\delta^*, \bar{\mathbb{P}}) = \{1 - (\frac{n}{a+b+n})^2\} VAR(\theta) - \frac{n}{(a+b+n)^2}\mathbb{E}(\theta(1-\theta))$

- 2) $L(d, \theta) = a_1(d-\theta)\mathbb{I}_{[\theta,\infty)}(d) + a_2(\theta - d)\mathbb{I}_{(-\infty,\theta)}(d)$, $a_1, a_2 > 0$:

  se para todo $x \in \mathcal{X}$, $\mathbb{E}(\theta|X = x) < \infty$, temos

  $\delta^*(x) = F^{-1}_{\theta|X=x}(\frac{a_1}{a_1+a_2})$, o percentil de ordem $\frac{a_1}{a_1+a_2}$ da distribuição a posteriori de θ dado $X = x$, isto é, $\delta^*(x)$ satisfaz

  $\mathbb{P}(\theta \leq \delta^*(x)|X = x) \geq \frac{a_1}{a_1+a_2}$ e $\mathbb{P}(\theta \geq \delta^*(x)|X = x) \geq \frac{a_2}{a_1+a_2}$

Estimação

- $= VAR(\theta) - \frac{1}{(a+b+n)^2}\{\mathbb{E}(n\theta(1-\theta)) + VAR(n\theta)\} \Rightarrow$

  $\Rightarrow \rho(\delta^*, \bar{\mathbb{P}}) = \{1 - (\frac{n}{a+b+n})^2\}VAR(\theta) - \frac{n}{(a+b+n)^2}\mathbb{E}(\theta(1-\theta))$

- 2) $L(d,\theta) = a_1(d-\theta)\mathbb{I}_{[\theta,\infty)}(d) + a_2(\theta - d)\mathbb{I}_{(-\infty,\theta)}(d)$, $a_1, a_2 > 0$:

  se para todo $x \in \mathcal{X}$, $\mathbb{E}(\theta|X=x) < \infty$, temos

  $\delta^*(x) = F^{-1}_{\theta|X=x}(\frac{a_1}{a_1+a_2})$, o percentil de ordem $\frac{a_1}{a_1+a_2}$ da distribuição a posteriori de θ dado $X = x$, isto é, $\delta^*(x)$ satisfaz

  $\mathbb{P}(\theta \leq \delta^*(x)|X=x) \geq \frac{a_1}{a_1+a_2}$ e $\mathbb{P}(\theta \geq \delta^*(x)|X=x) \geq \frac{a_2}{a_1+a_2}$

Estimação

- Se $a_1 = a_2$, temos $L(d, \theta) = |d - \theta|$ **(perda absoluta).**

  Nesse caso, $\delta^*(x) = F^{-1}_{\theta|X=x}(\frac{1}{2})$: mediana da distribuição a posteriori de θ dado $X = x$.

- 3) $L(d, \theta) = 1 - \mathbb{I}_{(\theta-\epsilon,\theta+\epsilon)}(d)$, $\epsilon > 0$ (θ contínuo)

  Nesse caso, $\delta^*(x)$ corresponde ao ponto central do intervalo "modal" de comprimento 2ϵ.
  Quando Θ é enumerável, podemos, alternativamente, considerar $L(d, \theta) = \mathbb{I}_{\{\theta\}}(d)$ **(função de perda 0-1).**

  Nesse caso, $\delta^*(x)$ é uma moda da distribuição a posteriori de θ dado $X = x$, isto é, $\delta^*(x)$ satisfaz $\mathbb{P}(\theta = \delta^*(x)|X = x) \geq \mathbb{P}(\theta = j|X = x), \forall j \in \Theta$.

Estimação

- Se $a_1 = a_2$, temos $L(d, \theta) = |d - \theta|$ **(perda absoluta).**

  Nesse caso, $\delta^*(x) = F^{-1}_{\theta|X=x}(\frac{1}{2})$: mediana da distribuição a posteriori de θ dado $X = x$.

- 3) $L(d, \theta) = 1 - \mathbb{I}_{(\theta-\epsilon,\theta+\epsilon)}(d)$, $\epsilon > 0$ (θ contínuo)

  Nesse caso, $\delta^*(x)$ corresponde ao ponto central do intervalo "modal" de comprimento 2ϵ.
  Quando Θ é enumerável, podemos, alternativamente, considerar $L(d, \theta) = \mathbb{I}_{\{\theta\}}(d)$ **(função de perda 0-1).**

  Nesse caso, $\delta^*(x)$ é uma moda da distribuição a posteriori de θ dado $X = x$, isto é, $\delta^*(x)$ satisfaz
  $\mathbb{P}(\theta = \delta^*(x)|X = x) \geq \mathbb{P}(\theta = j|X = x), \forall j \in \Theta$.

Estimação

- Se $a_1 = a_2$, temos $L(d, \theta) = |d - \theta|$ **(perda absoluta).**

  Nesse caso, $\delta^*(x) = F^{-1}_{\theta|X=x}(\frac{1}{2})$: mediana da distribuição a posteriori de θ dado $X = x$.

- 3) $L(d, \theta) = 1 - \mathbb{I}_{(\theta-\epsilon,\theta+\epsilon)}(d)$, $\epsilon > 0$ (θ contínuo)

  Nesse caso, $\delta^*(x)$ corresponde ao ponto central do intervalo "modal" de comprimento 2ϵ.
  Quando Θ é enumerável, podemos, alternativamente, considerar $L(d, \theta) = \mathbb{I}_{\{\theta\}}(d)$ **(função de perda 0-1).**

  Nesse caso, $\delta^*(x)$ é uma moda da distribuição a posteriori de θ dado $X = x$, isto é, $\delta^*(x)$ satisfaz
  $\mathbb{P}(\theta = \delta^*(x)|X = x) \geq \mathbb{P}(\theta = j|X = x)$, $\forall j \in \Theta$.

Estimação

- Na sequência, examinamos o problema de estimação por intervalo.

- **Problema de estimação por intervalo:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$

- $\Delta = \{I : \mathcal{X} \to \mathcal{I}\}$: conjunto de estimadores por intervalo para θ , onde $\mathcal{I} = \{[a, b] : a, b \in \mathbb{R}, a < b\}$.

- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)

- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(I, (\theta, x)) = L(I(x), \theta)$, onde $L : \mathcal{I} \times \Theta \to \mathbb{R}$ associa ao par $(I, \theta) \in \mathcal{I} \times \Theta$ a penalidade $L(I, \theta)$ para a estimativa por intervalo I quando o estado da natureza assume valor θ.

- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, examinamos o problema de estimação por intervalo.
- **Problema de estimação por intervalo:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$
- $\Delta = \{I : \mathcal{X} \to \mathcal{I}\}$: conjunto de estimadores por intervalo para θ , onde $\mathcal{I} = \{[a, b] : a, b \in \mathbb{R}, a < b\}$.
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(I, (\theta, x)) = L(I(x), \theta)$, onde $L : \mathcal{I} \times \Theta \to \mathbb{R}$ associa ao par $(I, \theta) \in \mathcal{I} \times \Theta$ a penalidade $L(I, \theta)$ para a estimativa por intervalo I quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, examinamos o problema de estimação por intervalo.
- **Problema de estimação por intervalo:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$
- $\Delta = \{I : \mathcal{X} \to \mathcal{I}\}$: conjunto de estimadores por intervalo para θ , onde $\mathcal{I} = \{[a, b] : a, b \in \mathbb{R}, a < b\}$.
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(I, (\theta, x)) = L(I(x), \theta)$, onde $L : \mathcal{I} \times \Theta \to \mathbb{R}$ associa ao par $(I, \theta) \in \mathcal{I} \times \Theta$ a penalidade $L(I, \theta)$ para a estimativa por intervalo I quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, examinamos o problema de estimação por intervalo.
- **Problema de estimação por intervalo:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$
- $\Delta = \{I : \mathcal{X} \to \mathcal{I}\}$: conjunto de estimadores por intervalo para θ , onde $\mathcal{I} = \{[a, b] : a, b \in \mathbb{R}, a < b\}$.
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(I, (\theta, x)) = L(I(x), \theta)$, onde $L : \mathcal{I} \times \Theta \to \mathbb{R}$ associa ao par $(I, \theta) \in \mathcal{I} \times \Theta$ a penalidade $L(I, \theta)$ para a estimativa por intervalo I quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, examinamos o problema de estimação por intervalo.
- **Problema de estimação por intervalo:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$
- $\Delta = \{I : \mathcal{X} \to \mathcal{I}\}$: conjunto de estimadores por intervalo para θ , onde $\mathcal{I} = \{[a, b] : a, b \in \mathbb{R}, a < b\}$.
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(I, (\theta, x)) = L(I(x), \theta)$, onde $L : \mathcal{I} \times \Theta \to \mathbb{R}$ associa ao par $(I, \theta) \in \mathcal{I} \times \Theta$ a penalidade $L(I, \theta)$ para a estimativa por intervalo I quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, examinamos o problema de estimação por intervalo.
- **Problema de estimação por intervalo:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$
- $\Delta = \{I : \mathcal{X} \to \mathcal{I}\}$: conjunto de estimadores por intervalo para θ , onde $\mathcal{I} = \{[a, b] : a, b \in \mathbb{R}, a < b\}$.
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \to \mathbb{R}$ dada por $\bar{L}(I, (\theta, x)) = L(I(x), \theta)$, onde $L : \mathcal{I} \times \Theta \to \mathbb{R}$ associa ao par $(I, \theta) \in \mathcal{I} \times \Theta$ a penalidade $L(I, \theta)$ para a estimativa por intervalo I quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Na sequência, examinamos o problema de estimação por intervalo.
- **Problema de estimação por intervalo:** $(\Delta, \Theta \times \mathcal{X}, \bar{L}, \bar{\mathbb{P}})$
- $\Delta = \{I : \mathcal{X} \rightarrow \mathcal{I}\}$: conjunto de estimadores por intervalo para θ , onde $\mathcal{I} = \{[a, b] : a, b \in \mathbb{R}, a < b\}$.
- $\Theta \times \mathcal{X}$ desconhecido (pré-experimentação)
- $\bar{L} : \Delta \times (\Theta \times \mathcal{X}) \rightarrow \mathbb{R}$ dada por $\bar{L}(I, (\theta, x)) = L(I(x), \theta)$, onde $L : \mathcal{I} \times \Theta \rightarrow \mathbb{R}$ associa ao par $(I, \theta) \in \mathcal{I} \times \Theta$ a penalidade $L(I, \theta)$ para a estimativa por intervalo I quando o estado da natureza assume valor θ.
- $\bar{\mathbb{P}}$ representa incerteza sobre (θ, X).

Estimação

- Analogamente ao problema de estimação (pontual), sob certas condições, fixada $L : \mathcal{I} \times \Theta \to \mathbb{R}$, ao observar $x \in \mathcal{X}$, obtem-se a posteriori de θ dado $X = x$ e então toma-se $I^* \in \mathcal{I}$ solução do problema $(\mathcal{I}, \Theta . L, \mathbb{P}_x)$.
- Nesse caso, a função de perda comumente leva em consideração a pertinência de θ a I e o comprimento de I. Por exemplo, para $I = [a, b]$, considere, com $c_1, c_2 > 1$,
- $L([a,b],\theta) = \begin{cases} b - a + c_1(a - \theta), & \text{se } \theta < a \\ b - a, & \text{se } a \leq \theta \leq b \\ b - a + c_2(\theta - b), & \text{se } \theta > b \end{cases}$

Estimação

- Analogamente ao problema de estimação (pontual), sob certas condições, fixada $L : \mathcal{I} \times \Theta \to \mathbb{R}$, ao observar $x \in \mathcal{X}$, obtem-se a posteriori de θ dado $X = x$ e então toma-se $I^* \in \mathcal{I}$ solução do problema $(\mathcal{I}, \Theta.L, \mathbb{P}_x)$.
- Nesse caso, a função de perda comumente leva em consideração a pertinência de θ a I e o comprimento de I. Por exemplo, para $I = [a, b]$, considere, com $c_1, c_2 > 1$,
- $L([a,b],\theta) = \begin{cases} b - a + c_1(a - \theta), & \text{se } \theta < a \\ b - a, & \text{se } a \leq \theta \leq b \\ b - a + c_2(\theta - b), & \text{se } \theta > b \end{cases}$

Estimação

- Analogamente ao problema de estimação (pontual), sob certas condições, fixada $L : \mathcal{I} \times \Theta \to \mathbb{R}$, ao observar $x \in \mathcal{X}$, obtem-se a posteriori de θ dado $X = x$ e então toma-se $I^* \in \mathcal{I}$ solução do problema $(\mathcal{I}, \Theta.L, \mathbb{P}_x)$.

- Nesse caso, a função de perda comumente leva em consideração a pertinência de θ a I e o comprimento de I. Por exemplo, para $I = [a, b]$, considere, com $c_1, c_2 > 1$,

- $$L([a,b],\theta) = \begin{cases} b - a + c_1(a - \theta), & \text{se } \theta < a \\ b - a, & \text{se } a \leq \theta \leq b \\ b - a + c_2(\theta - b), & \text{se } \theta > b \end{cases}$$

Estimação

- Analogamente ao problema de estimação (pontual), sob certas condições, fixada $L : \mathcal{I} \times \Theta \to \mathbb{R}$, ao observar $x \in \mathcal{X}$, obtem-se a posteriori de θ dado $X = x$ e então toma-se $I^* \in \mathcal{I}$ solução do problema $(\mathcal{I}, \Theta.L, \mathbb{P}_x)$.

- Nesse caso, a função de perda comumente leva em consideração a pertinência de θ a I e o comprimento de I. Por exemplo, para $I = [a, b]$, considere, com $c_1, c_2 > 1$,

- $L([a,b],\theta) = \begin{cases} b - a + c_1(a-\theta), & \text{se } \theta < a \\ b - a, & \text{se } a \leq \theta \leq b \\ b - a + c_2(\theta - b), & \text{se } \theta > b \end{cases}$

Estimação

- Nesse caso, para cada $x \in \mathcal{X}$, $I^*(x)$ é o intervalo que vai do percentil de ordem $\frac{1}{c_1}$ ao percentil de ordem $1 - \frac{1}{c_2}$ da distribuição a posteriori de θ dado $X = x$. Observe que se $c_1 = c_2 = \frac{2}{\alpha}$, $\alpha \in (0,1)$, $I^*(x)$ é o intervalo "central" de probabilidade a posteriori $1 - \alpha$. Note que
- $L([a,b],\theta) \ =$

$$= \ c_1(a-\theta)\mathbb{I}_{(-\infty,a)}(\theta) + c_2(\theta-b)\mathbb{I}_{(b,\infty)}(\theta) + b - \theta + \theta - a$$

$= \ L_1(a,\theta) \ + \ L_2(b,\theta)$, onde

$L_1(a,\theta) \ = \ (c_1 - 1)(a-\theta)\mathbb{I}_{(-\infty,a)}(\theta) + (\theta - a)\mathbb{I}_{[a,\infty)}(\theta)$ e

$L_2(b,\theta) \ = \ (b-\theta)\mathbb{I}_{(-\infty,b)}(\theta) + (c_2 - 1)\mathbb{I}_{[b,\infty)}(\theta)$.

Basta então escolher a e b que minimizam L_1 e L_2, respectivamente.

Estimação

- Nesse caso, para cada $x \in \mathcal{X}$, $I^*(x)$ é o intervalo que vai do percentil de ordem $\frac{1}{c_1}$ ao percentil de ordem $1 - \frac{1}{c_2}$ da distribuição a posteriori de θ dado $X = x$. Observe que se $c_1 = c_2 = \frac{2}{\alpha}$, $\alpha \in (0,1)$, $I^*(x)$ é o intervalo "central" de probabilidade a posteriori $1 - \alpha$. Note que

- $L([a,b],\theta) =$

$$= c_1(a-\theta)\mathbb{I}_{(-\infty,a)}(\theta) + c_2(\theta-b)\mathbb{I}_{(b,\infty)}(\theta) + b - \theta + \theta - a$$

$$= L_1(a,\theta) + L_2(b,\theta)$$, onde

$$L_1(a,\theta) = (c_1-1)(a-\theta)\mathbb{I}_{(-\infty,a)}(\theta) + (\theta-a)\mathbb{I}_{[a,\infty)}(\theta)$$ e

$$L_2(b,\theta) = (b-\theta)\mathbb{I}_{(-\infty,b)}(\theta) + (c_2-1)\mathbb{I}_{[b,\infty)}(\theta).$$

Basta então escolher a e b que minimizam L_1 e L_2, respectivamente.

Estimação

- Nesse caso, para cada $x \in \mathcal{X}$, $I^*(x)$ é o intervalo que vai do percentil de ordem $\frac{1}{c_1}$ ao percentil de ordem $1 - \frac{1}{c_2}$ da distribuição a posteriori de θ dado $X = x$. Observe que se $c_1 = c_2 = \frac{2}{\alpha}$, $\alpha \in (0, 1)$, $I^*(x)$ é o intervalo "central" de probabilidade a posteriori $1 - \alpha$. Note que
- $L([a, b], \theta) \; =$

$$= \; c_1(a - \theta)\mathbb{I}_{(-\infty,a)}(\theta) + c_2(\theta - b)\mathbb{I}_{(b,\infty)}(\theta) + b - \theta + \theta - a$$

$$= \; L_1(a, \theta) \; + \; L_2(b, \theta)$$, onde

$$L_1(a, \theta) \; = \; (c_1 - 1)(a - \theta)\mathbb{I}_{(-\infty,a)}(\theta) + (\theta - a)\mathbb{I}_{[a,\infty)}(\theta)$$ e

$$L_2(b, \theta) \; = \; (b - \theta)\mathbb{I}_{(-\infty,b)}(\theta) + (c_2 - 1)\mathbb{I}_{[b,\infty)}(\theta).$$

Basta então escolher a e b que minimizam L_1 e L_2, respectivamente.

Estimação

- Exemplo 1: $X = (X_1, ..., X_n)$ AAS do modelo normal de média θ e variância $v^2 > 0$. Suponhamos $\theta \sim N(\mu_0, \sigma_0^2)$, com $\mu_0 \in \mathbb{R}$ e $\sigma_0^2 > 0$. Então,

  $\theta | X_1 = x_1, ..., X_n = x_n \sim N(\mu_x, \sigma_x^2)$, onde

  $\mu_x = \frac{\sigma_0^2 n\bar{x} + v^2 \mu_o}{\sigma_0^2 n + v^2}$ e $\sigma_x^2 = \frac{\sigma_0^2 v^2}{v^2 + n\sigma_0^2}$.

  O intervalo "central" de probabilidade (a posteriori) $1-\alpha$ é

  $I^*(x) = [\mu_x - \Phi^{-1}(1-\frac{\alpha}{2})\sigma_x \, , \, \mu_x + \Phi^{-1}(1-\frac{\alpha}{2})\sigma_x]$

  O estimador por intervalo é então

  $I^*(X) = [A(X) - C(n, v^2, \sigma_0^2) \, , \, A(X) + C(n, v^2, \sigma_0^2)]$, com

  $A(X) = \frac{\sigma_0^2 n\bar{X} + v^2 \mu_o}{\sigma_0^2 n + v^2}$ e $C(n, v^2, \sigma_0^2) = \Phi^{-1}(1-\frac{\alpha}{2})\sqrt{\frac{\sigma_0^2 v^2}{v^2 + n\sigma_0^2}}$

Estimação

- Exemplo 1: $X = (X_1, ..., X_n)$ AAS do modelo normal de média θ e variância $v^2 > 0$. Suponhamos $\theta \sim N(\mu_0, \sigma_0^2)$, com $\mu_0 \in \mathbb{R}$ e $\sigma_0^2 > 0$. Então,

  $\theta | X_1 = x_1, ..., X_n = x_n \ \sim N(\mu_x, \sigma_x^2)$, onde

  $\mu_x = \frac{\sigma_0^2 n\bar{x} + v^2 \mu_o}{\sigma_0^2 n + v^2}$ e $\sigma_x^2 = \frac{\sigma_0^2 v^2}{v^2 + n\sigma_0^2}$.

  O intervalo "central" de probabilidade (a posteriori) $1 - \alpha$ é

  $$I^*(x) = [\mu_x - \Phi^{-1}(1 - \tfrac{\alpha}{2})\sigma_x \ , \ \mu_x + \Phi^{-1}(1 - \tfrac{\alpha}{2})\sigma_x]$$

  O estimador por intervalo é então

  $I^*(X) = [A(X) - C(n, v^2, \sigma_0^2) \ , \ A(X) + C(n, v^2, \sigma_0^2)]$, com

  $A(X) = \frac{\sigma_0^2 n\bar{X} + v^2 \mu_o}{\sigma_0^2 n + v^2}$ e $C(n, v^2, \sigma_0^2) = \Phi^{-1}(1 - \frac{\alpha}{2})\sqrt{\frac{\sigma_0^2 v^2}{v^2 + n\sigma_0^2}}$

Estimação

- Exemplo 2: $X = (X_1, ..., X_n)$ AAS do modelo Uniforme em $(0, \theta)$, $\theta > 0$. Suponhamos $\theta \sim Pareto(a, b)$, $a, b > 0$. Então,

  $\theta | X_1 = x_1, ..., X_n = x_n \ \sim Pareto(a + n, \max\{b, x_{(n)}\})$, com $x_{(n)} = \max\{x_1, ..., x_n\}$.

  O intervalo "central" de probabilidade (a posteriori) $1 - \alpha$ é

  $$I^*(x) \ = \ [\max\{b, x_{(n)}\}(1 - \tfrac{\alpha}{2})^{-\frac{1}{a+n}} \ , \ \max\{b, x_{(n)}\}(\tfrac{\alpha}{2})^{-\frac{1}{a+n}}]$$

  O estimador por intervalo é então

  $$I^*(X) \ =$$

  $$= \ [\ \max\{b, X_{(n)}\}(1 - \tfrac{\alpha}{2})^{-\frac{1}{a+n}} \ , \ \max\{b, X_{(n)}\}(\tfrac{\alpha}{2})^{-\frac{1}{a+n}} \ ]$$

Estimação

- Exemplo 2: $X = (X_1, ..., X_n)$ AAS do modelo Uniforme em $(0, \theta)$, $\theta > 0$. Suponhamos $\theta \sim Pareto(a, b)$, $a, b > 0$. Então,

  $\theta | X_1 = x_1, ..., X_n = x_n \; \sim Pareto(a + n, \max\{b, x_{(n)}\})$, com $x_{(n)} = \max\{x_1, ..., x_n\}$.

  O intervalo "central" de probabilidade (a posteriori) $1 - \alpha$ é

  $$I^*(x) \; = \; [\max\{b, x_{(n)}\}(1 - \tfrac{\alpha}{2})^{-\frac{1}{a+n}} \, , \; \max\{b, x_{(n)}\}(\tfrac{\alpha}{2})^{-\frac{1}{a+n}}]$$

  O estimador por intervalo é então

  $$I^*(X) \; =$$

  $$= \; [\; \max\{b, X_{(n)}\}(1 - \tfrac{\alpha}{2})^{-\frac{1}{a+n}} \, , \; \max\{b, X_{(n)}\}(\tfrac{\alpha}{2})^{-\frac{1}{a+n}} \;]$$

Estimação

- No Exemplo 2, observe que o estimador $I' : \mathcal{X} \to \mathcal{I}$ que associa a cada $x \in \mathcal{X}$ o intervalo

  $$[\,\max\{b, x_{(n)}\}\ ,\ \max\{b, x_{(n)}\}(\alpha)^{-\frac{1}{a+n}}\,]$$

  também satisfaz (como I^*) que,

  $$\forall x \in \mathcal{X},\ \mathbb{P}(\theta \in I'(x)|X = x) = 1 - \alpha.$$

  No entanto, para todo $x \in \mathcal{X}$, $\lambda(I'(x)) < \lambda(I^*(x))$.

  Em cenários como do Exemplo 2, estimativas diferentes de $I^*(x)$, por exemplo $I'(x)$, podem ser mais "interessantes" como "resumo" da distribuição a posteriori de θ dado $X = x$, $x \in \mathcal{X}$. Examinaremos agora tais estimadores (e otimalidade).

Estimação

- No Exemplo 2, observe que o estimador $I' : \mathcal{X} \to \mathcal{I}$ que associa a cada $x \in \mathcal{X}$ o intervalo

  $$[\ \max\{b, x_{(n)}\}\ ,\ \max\{b, x_{(n)}\}(\alpha)^{-\frac{1}{a+n}}\ ]$$

  também satisfaz (como I^*) que,

  $$\forall x \in \mathcal{X},\ \mathbb{P}(\theta \in I'(x)|X = x) = 1 - \alpha.$$

  No entanto, para todo $x \in \mathcal{X}$, $\lambda(I'(x)) \ < \ \lambda(I^*(x))$.

  Em cenários como do Exemplo 2, estimativas diferentes de $I^*(x)$, por exemplo $I'(x)$, podem ser mais "interessantes" como "resumo" da distribuição a posteriori de θ dado $X = x$, $x \in \mathcal{X}$. Examinaremos agora tais estimadores (e otimalidade).