Inferência Estatística Comparada

AULA 4 - INFERÊNCIA CLÁSSICA

Introdução

- **INFERÊNCIA ESTATÍSTICA**
- Questões sobre procedimento inferencial (inferência estatística):
- (i) O que é Informação em um processo de inferência estatística e como podemos representá-la?
- (ii) Quais tipos de questões queremos responder com a informação experimental e quais procedimentos devemos/podemos adotar para responder tais questões?
- (iii) Quais critérios devemos estabelecer para avaliar o desempenho de tais procedimentos?

Introdução

- **INFERÊNCIA ESTATÍSTICA**
- Questões sobre procedimento inferencial (inferência estatística):
- (i) O que é Informação em um processo de inferência estatística e como podemos representá-la?
- (ii) Quais tipos de questões queremos responder com a informação experimental e quais procedimentos devemos/podemos adotar para responder tais questões?
- (iii) Quais critérios devemos estabelecer para avaliar o desempenho de tais procedimentos?

Introdução

- **INFERÊNCIA ESTATÍSTICA**
- Questões sobre procedimento inferencial (inferência estatística):
- (i) O que é Informação em um processo de inferência estatística e como podemos representá-la?
- (ii) Quais tipos de questões queremos responder com a informação experimental e quais procedimentos devemos/podemos adotar para responder tais questões?
- (iii) Quais critérios devemos estabelecer para avaliar o desempenho de tais procedimentos?

Introdução

- **INFERÊNCIA ESTATÍSTICA**
- Questões sobre procedimento inferencial (inferência estatística):
- (i) O que é Informação em um processo de inferência estatística e como podemos representá-la?
- (ii) Quais tipos de questões queremos responder com a informação experimental e quais procedimentos devemos/podemos adotar para responder tais questões?
- (iii) Quais critérios devemos estabelecer para avaliar o desempenho de tais procedimentos?

Introdução

- **INFERÊNCIA ESTATÍSTICA**
- Questões sobre procedimento inferencial (inferência estatística):
- (i) O que é Informação em um processo de inferência estatística e como podemos representá-la?
- (ii) Quais tipos de questões queremos responder com a informação experimental e quais procedimentos devemos/podemos adotar para responder tais questões?
- (iii) Quais critérios devemos estabelecer para avaliar o desempenho de tais procedimentos?

Introdução

- **ELEMENTOS BÁSICOS**

- θ: parâmetro
- Θ: Espaço paramétrico
- X: amostra
- $\mathcal{X}$: Espaço amostral

Introdução

- **ELEMENTOS BÁSICOS**

- θ**:** parâmetro
- Θ**:** Espaço paramétrico
- X: amostra
- $\mathcal{X}$ **:** Espaço amostral

Introdução

- **ELEMENTOS BÁSICOS**

- θ**:** parâmetro
- Θ**:** Espaço paramétrico
- X: amostra
- $\mathcal{X}$ **:** Espaço amostral

Introdução

- **ELEMENTOS BÁSICOS**

- θ**:** parâmetro
- Θ**:** Espaço paramétrico
- X: amostra
- $\mathcal{X}$ **:** Espaço amostral

Introdução

- **ELEMENTOS BÁSICOS**

- θ**:** parâmetro
- Θ**:** Espaço paramétrico
- X: amostra
- $\mathcal{X}$ **:** Espaço amostral

Introdução

- **MODELO ESTATÍSTICO (CLÁSSICO)**

- $(\mathcal{X}, \mathcal{F}, \mathcal{P})$ **:** onde
- $\mathcal{X}$ **:** Espaço amostral,
- $\mathcal{F}$ é classe de subconjuntos (σ-álgebra) de $\mathcal{X}$ e
- $\mathcal{P} = \{\mathbb{P}(\cdot|\theta) : \theta \in \Theta\}$ é família de distribuições sobre $\mathcal{F}$ indexada em θ (família paramétrica).

Introdução

- **MODELO ESTATÍSTICO (CLÁSSICO)**

- $(\mathcal{X}, \mathcal{F}, \mathcal{P})$ **:** onde
- $\mathcal{X}$ **:** Espaço amostral,
- $\mathcal{F}$ é classe de subconjuntos (σ-álgebra) de $\mathcal{X}$ e
- $\mathcal{P} = \{\mathbb{P}(\cdot|\theta) : \theta \in \Theta\}$ é família de distribuições sobre $\mathcal{F}$ indexada em θ (família paramétrica).

Introdução

- **MODELO ESTATÍSTICO (CLÁSSICO)**

- $(\mathcal{X}, \mathcal{F}, \mathcal{P})$ **:** onde
- $\mathcal{X}$ **:** Espaço amostral,
- $\mathcal{F}$ é classe de subconjuntos (σ-álgebra) de $\mathcal{X}$ e
- $\mathcal{P} = \{\mathbb{P}(\cdot|\theta) : \theta \in \Theta\}$ é família de distribuições sobre $\mathcal{F}$ indexada em θ (família paramétrica).

Introdução

- **MODELO ESTATÍSTICO (CLÁSSICO)**

- $(\mathcal{X}, \mathcal{F}, \mathcal{P})$ **:** onde
- $\mathcal{X}$ **:** Espaço amostral,
- $\mathcal{F}$ é classe de subconjuntos (σ-álgebra) de $\mathcal{X}$ e
- $\mathcal{P} = \{\mathbb{P}(\cdot|\theta) : \theta \in \Theta\}$ é família de distribuições sobre $\mathcal{F}$ indexada em θ (família paramétrica).

Introdução

- **MODELO ESTATÍSTICO (CLÁSSICO)**

- $(\mathcal{X}, \mathcal{F}, \mathcal{P})$ **:** onde
- $\mathcal{X}$ **:** Espaço amostral,
- $\mathcal{F}$ é classe de subconjuntos (σ-álgebra) de $\mathcal{X}$ e
- $\mathcal{P} = \{\mathbb{P}(\cdot|\theta) : \theta \in \Theta\}$ é família de distribuições sobre $\mathcal{F}$ indexada em θ (família paramétrica).

Introdução

- No exemplo 1 (segunda aula), $\Theta = \{1, 2, 3, ...\}$ e $X|\theta \sim U(\{1, ..., \theta\})$.
- No exemplo 2 (segunda aula), $\Theta = [0, 1]$ e $X = (X_1, ..., X_n)$, dado θ, c.i.i.d. $Ber(\theta)$
- (i) Informação sobre θ contida (revelada) na amostra (observaçâo amostral) $x \in \mathcal{X}$ expressa por

$$\mathbb{P}(X = x|\theta)$$

Introdução

- No exemplo 1 (segunda aula), $\Theta = \{1, 2, 3, ...\}$ e $X|\theta \sim U(\{1, ..., \theta\})$.
- No exemplo 2 (segunda aula), $\Theta = [0, 1]$ e $X = (X_1, ..., X_n)$, dado θ, c.i.i.d. $Ber(\theta)$
- (i) Informação sobre θ contida (revelada) na amostra (observaçâo amostral) $x \in \mathcal{X}$ expressa por

$$\mathbb{P}(X = x|\theta)$$

Introdução

- No exemplo 1 (segunda aula), $\Theta = \{1, 2, 3, ...\}$ e $X|\theta \sim U(\{1, ..., \theta\})$.
- No exemplo 2 (segunda aula), $\Theta = [0, 1]$ e $X = (X_1, ..., X_n)$, dado θ, c.i.i.d. $Ber(\theta)$
- (i) Informação sobre θ contida (revelada) na amostra (observaçâo amostral) $x \in \mathcal{X}$ expressa por

$$\mathbb{P}(X = x|\theta)$$

Introdução

- No exemplo 1 (segunda aula), $\Theta = \{1, 2, 3, ...\}$ e $X|\theta \sim U(\{1, ..., \theta\})$.
- No exemplo 2 (segunda aula), $\Theta = [0, 1]$ e $X = (X_1, ..., X_n)$, dado θ, c.i.i.d. $Ber(\theta)$
- (i) Informação sobre θ contida (revelada) na amostra (observaçâo amostral) $x \in \mathcal{X}$ expressa por

$$\mathbb{P}(X = x|\theta)$$

Introdução

- $\mathbb{P}(X = x|\theta)$

- Fixado $\theta \in \Theta$, $\mathbb{P}(X = \cdot|\theta)$ é distribuição de probabilidade para a amostra;

- Fixado $x \in \mathcal{X}$, $\mathbb{P}(X = x|\cdot)$ é função não-negativa definida em Θ: **função de verossimilhança para θ gerada por** $x \in \mathcal{X}$.

Introdução

- $\mathbb{P}(X = x|\theta)$

- Fixado $\theta \in \Theta$, $\mathbb{P}(X = \cdot|\theta)$ é distribuição de probabilidade para a amostra;

- Fixado $x \in \mathcal{X}$, $\mathbb{P}(X = x|\cdot)$ é função não-negativa definida em Θ: **função de verossimilhança para θ gerada por $x \in \mathcal{X}$**.

Introdução

- $\mathbb{P}(X = x|\theta)$
- Fixado $\theta \in \Theta$, $\mathbb{P}(X = \cdot|\theta)$ é distribuição de probabilidade para a amostra;
- Fixado $x \in \mathcal{X}$, $\mathbb{P}(X = x|\cdot)$ é função não-negativa definida em Θ: **função de verossimilhança para θ gerada por $x \in \mathcal{X}$.**

Introdução

- $\mathbb{P}(X = x|\theta)$

- Fixado $\theta \in \Theta$, $\mathbb{P}(X = \cdot|\theta)$ é distribuição de probabilidade para a amostra;

- Fixado $x \in \mathcal{X}$, $\mathbb{P}(X = x|\cdot)$ é função não-negativa definida em Θ: **função de verossimilhança para θ gerada por** $x \in \mathcal{X}$.

Introdução

- Uma estatística $T : \mathcal{X} \rightarrow \mathcal{T}$ é suficiente para θ se preserva toda informação relevante da amostra sobre θ.Nesse caso, podemos esperar que

- $\mathbb{P}(X = x|\theta) \quad = \quad u(x).v(T(x), \theta) \;\; , \;\; \theta \in \Theta, \;\; x \in \mathcal{X}$

- **(Definição Clássica)** Uma estatística $T : \mathcal{X} \rightarrow \mathcal{T}$ é suficiente para θ se para todo valor possível t da estatística T, a distribuição condicional de X dado $T(X) = t$ e θ, não depende de θ.

- Em geral, é (parece ser) desejável que a inferência sobre θ dependa de x através (apenas) de $T(x)$.

Introdução

- Uma estatística $T : \mathcal{X} \to \mathcal{T}$ é suficiente para θ se preserva toda informação relevante da amostra sobre θ.Nesse caso, podemos esperar que
- $\mathbb{P}(X = x|\theta) \quad = \quad u(x).v(T(x), \theta) \;, \quad \theta \in \Theta, \quad x \in \mathcal{X}$
- **(Definição Clássica)** Uma estatística $T : \mathcal{X} \to \mathcal{T}$ é suficiente para θ se para todo valor possível t da estatística T, a distribuição condicional de X dado $T(X) = t$ e θ, não depende de θ.
- Em geral, é (parece ser) desejável que a inferência sobre θ dependa de x através (apenas) de $T(x)$.

Introdução

- Uma estatística $T : \mathcal{X} \to \mathcal{T}$ é suficiente para θ se preserva toda informação relevante da amostra sobre θ.Nesse caso, podemos esperar que

- $\mathbb{P}(X = x|\theta) \quad = \quad u(x).v(T(x), \theta) \ , \ \ \theta \in \Theta, \ \ x \in \mathcal{X}$

- **(Definição Clássica)** Uma estatística $T : \mathcal{X} \to \mathcal{T}$ é suficiente para θ se para todo valor possível t da estatística T, a distribuição condicional de X dado $T(X) = t$ e θ, não depende de θ.

- Em geral, é (parece ser) desejável que a inferência sobre θ dependa de x através (apenas) de $T(x)$.

Introdução

- Uma estatística $T : \mathcal{X} \to \mathcal{T}$ é suficiente para θ se preserva toda informação relevante da amostra sobre θ.Nesse caso, podemos esperar que

- $\mathbb{P}(X = x|\theta) \quad = \quad u(x).v(T(x), \theta) \ , \ \ \theta \in \Theta, \ \ x \in \mathcal{X}$

- **(Definição Clássica)** Uma estatística $T : \mathcal{X} \to \mathcal{T}$ é suficiente para θ se para todo valor possível t da estatística T, a distribuição condicional de X dado $T(X) = t$ e θ, não depende de θ.

- Em geral, é (parece ser) desejável que a inferência sobre θ dependa de x através (apenas) de $T(x)$.

Introdução

- Uma estatística $T : \mathcal{X} \rightarrow \mathcal{T}$ é suficiente para θ se preserva toda informação relevante da amostra sobre θ.Nesse caso, podemos esperar que
- $\mathbb{P}(X = x|\theta) \quad = \quad u(x).v(T(x), \theta) \ , \ \ \theta \in \Theta, \ \ x \in \mathcal{X}$
- **(Definição Clássica)** Uma estatística $T : \mathcal{X} \rightarrow \mathcal{T}$ é suficiente para θ se para todo valor possível t da estatística T, a distribuição condicional de X dado $T(X) = t$ e θ, não depende de θ.
- Em geral, é (parece ser) desejável que a inferência sobre θ dependa de x através (apenas) de $T(x)$.

Abordagem Clássica

- (ii) Questões (problemas) comumente abordados sob a perspectiva clássica (frequentista): Estimação (pontual e intervalar) e teste de hipóteses.
- Para o problema de estimação, fazemos uso de estatísticas "particulares": estimadores.
- Estimador: $\delta : \mathcal{X} \rightarrow \Theta$
- Procedimento de estimação baseado em estimadores não-viesados.

Abordagem Clássica

- (ii) Questões (problemas) comumente abordados sob a perspectiva clássica (frequentista): Estimação (pontual e intervalar) e teste de hipóteses.
- Para o problema de estimação, fazemos uso de estatísticas "particulares": estimadores.
- Estimador: $\delta : \mathcal{X} \rightarrow \Theta$
- Procedimento de estimação baseado em estimadores não-viesados.

Abordagem Clássica

- (ii) Questões (problemas) comumente abordados sob a perspectiva clássica (frequentista): Estimação (pontual e intervalar) e teste de hipóteses.
- Para o problema de estimação, fazemos uso de estatísticas "particulares": estimadores.
- Estimador: $\delta : \mathcal{X} \to \Theta$
- Procedimento de estimação baseado em estimadores não-viesados.

Abordagem Clássica

- (ii) Questões (problemas) comumente abordados sob a perspectiva clássica (frequentista): Estimação (pontual e intervalar) e teste de hipóteses.
- Para o problema de estimação, fazemos uso de estatísticas "particulares": estimadores.
- Estimador: $\delta : \mathcal{X} \rightarrow \Theta$
- Procedimento de estimação baseado em estimadores não-viesados.

Abordagem Clássica

- (ii) Questões (problemas) comumente abordados sob a perspectiva clássica (frequentista): Estimação (pontual e intervalar) e teste de hipóteses.
- Para o problema de estimação, fazemos uso de estatísticas "particulares": estimadores.
- Estimador: $\delta : \mathcal{X} \to \Theta$
- Procedimento de estimação baseado em estimadores não-viesados.

Estimação

- $\delta : \mathcal{X} \to \Theta$ é não-viesado para θ $(g(\theta))$ se, $\forall\, \theta \in \Theta$, $\mathbb{E}(\delta(X)|\theta) \ = \ \theta \ (g(\theta))$, isto é,

- $\mathbb{E}(\delta(X)|\theta) \ = \ \sum_{x\in\mathcal{X}}\delta(x)\mathbb{P}(X = x|\theta) \ = \ \theta$, $\forall\, \theta \in \Theta$. (caso discreto)

- $\mathbb{E}(\delta(X)|\theta) \ = \ \int_{\mathcal{X}}\delta(x)f(x|\theta)d\theta$ = θ , $\forall\, \theta \in \Theta$. (caso contínuo)

- No exemplo 2, considere, $n > 2$. Para qualquer $i = 1, ...n$, $\delta_i(X) = X_i$ é não-viesado para θ, do mesmo modo que $\delta'(X) = \frac{X_1+X_2}{2}$ e $\delta^* = \frac{X_1+...+X_n}{n} = \bar{X}$. Nesse caso, δ^* é melhor que qualquer outro (não-viesado).

Estimação

- $\delta : \mathcal{X} \to \Theta$ é não-viesado para θ $(g(\theta))$ se, $\forall\ \theta \in \Theta$, $\mathbb{E}(\delta(X)|\theta) \ = \ \theta \ (g(\theta))$, isto é,
- $\mathbb{E}(\delta(X)|\theta) \ = \ \sum_{x \in \mathcal{X}} \delta(x)\mathbb{P}(X = x|\theta) \ = \ \theta$, $\forall\ \theta \in \Theta$. (caso discreto)
- $\mathbb{E}(\delta(X)|\theta) \ = \ \int_{\mathcal{X}} \delta(x) f(x|\theta) d\theta$ = θ , $\forall\ \theta \in \Theta$. (caso contínuo)
- No exemplo 2, considere, $n > 2$. Para qualquer $i = 1, ...n$, $\delta_i(X) = X_i$ é não-viesado para θ, do mesmo modo que $\delta'(X) = \frac{X_1+X_2}{2}$ e $\delta^* = \frac{X_1+...+X_n}{n} = \bar{X}$. Nesse caso, δ^* é melhor que qualquer outro (não-viesado).

Estimação

- $\delta : \mathcal{X} \to \Theta$ é não-viesado para θ $(g(\theta))$ se, $\forall\ \theta \in \Theta$, $\mathbb{E}(\delta(X)|\theta) \ = \ \theta\ (g(\theta))$, isto é,

- $\mathbb{E}(\delta(X)|\theta) \ = \ \sum_{x\in\mathcal{X}}\delta(x)\mathbb{P}(X = x|\theta) \ = \ \theta$, $\forall\ \theta \in \Theta$. (caso discreto)

- $\mathbb{E}(\delta(X)|\theta) \ = \ \int_{\mathcal{X}}\delta(x)f(x|\theta)d\theta \ = \ \theta$, $\forall\ \theta \in \Theta$. (caso contínuo)

- No exemplo 2, considere, $n > 2$. Para qualquer $i = 1, ...n$, $\delta_i(X) = X_i$ é não-viesado para θ, do mesmo modo que $\delta'(X) = \frac{X_1+X_2}{2}$ e $\delta^* = \frac{X_1+...+X_n}{n} = \bar{X}$. Nesse caso, δ^* é melhor que qualquer outro (não-viesado).

Estimação

- $\delta : \mathcal{X} \to \Theta$ é não-viesado para θ $(g(\theta))$ se, $\forall\ \theta \in \Theta$, $\mathbb{E}(\delta(X)|\theta) = \theta\ (g(\theta))$, isto é,
- $\mathbb{E}(\delta(X)|\theta) = \sum_{x \in \mathcal{X}} \delta(x) \mathbb{P}(X = x|\theta) = \theta$, $\forall\ \theta \in \Theta$. (caso discreto)
- $\mathbb{E}(\delta(X)|\theta) = \int_{\mathcal{X}} \delta(x) f(x|\theta) d\theta$ = θ , $\forall\ \theta \in \Theta$. (caso contínuo)
- No exemplo 2, considere, $n > 2$. Para qualquer $i = 1, ...n$, $\delta_i(X) = X_i$ é não-viesado para θ, do mesmo modo que $\delta'(X) = \frac{X_1+X_2}{2}$ e $\delta^* = \frac{X_1+...+X_n}{n} = \bar{X}$. Nesse caso, δ^* é melhor que qualquer outro (não-viesado).

Estimação

- $\delta : \mathcal{X} \to \Theta$ é não-viesado para θ $(g(\theta))$ se, $\forall\ \theta \in \Theta$, $\mathbb{E}(\delta(X)|\theta) = \theta\ (g(\theta))$, isto é,

- $\mathbb{E}(\delta(X)|\theta) = \sum_{x \in \mathcal{X}} \delta(x)\mathbb{P}(X = x|\theta) = \theta$, $\forall\ \theta \in \Theta$. (caso discreto)

- $\mathbb{E}(\delta(X)|\theta) = \int_{\mathcal{X}} \delta(x) f(x|\theta) d\theta$ = θ , $\forall\ \theta \in \Theta$. (caso contínuo)

- No exemplo 2, considere, $n > 2$. Para qualquer $i = 1, ...n$, $\delta_i(X) = X_i$ é não-viesado para θ, do mesmo modo que $\delta'(X) = \frac{X_1+X_2}{2}$ e $\delta^* = \frac{X_1+...+X_n}{n} = \bar{X}$. Nesse caso, δ^* é melhor que qualquer outro (não-viesado).

Estimação

- Para $g(\theta) = \theta^2$, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para $g(\theta)$, do mesmo modo que $\delta^{(i,j)}(X) = X_i.X_j$, $i \neq j$. Qual é o melhor nesse caso?

- Pode-se mostrar que $\delta^{**}(X) = \frac{\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} X_i - 1\right)}{n(n-1)}$, que é função da estatística suficiente $\sum_{i=1}^{n} X_i$, é o "melhor" estimador não-viesado para θ^2.

- Sob certas condições, pode-se obter o melhor estimador não-viesado para $g(\theta)$ tomando-se um estimador não-viesado qualquer para $g(\theta)$ e condicionando tal estimador a um estatística suficiente (e completa) para θ (RAO-BLACKWELLIZACÃO).

Estimação

- Para $g(\theta) = \theta^2$, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para $g(\theta)$, do mesmo modo que $\delta^{(i,j)}(X) = X_i.X_j$, $i \neq j$. Qual é o melhor nesse caso?
- Pode-se mostrar que $\delta^{**}(X) = \frac{\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} X_i - 1\right)}{n(n-1)}$, que é função da estatística suficiente $\sum_{i=1}^{n} X_i$, é o "melhor" estimador não-viesado para θ^2.
- Sob certas condições, pode-se obter o melhor estimador não-viesado para $g(\theta)$ tomando-se um estimador não-viesado qualquer para $g(\theta)$ e condicionando tal estimador a um estatística suficiente (e completa) para θ (RAO-BLACKWELLIZACÃO).

Estimação

- Para $g(\theta) = \theta^2$, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para $g(\theta)$, do mesmo modo que $\delta^{(i,j)}(X) = X_i.X_j$, $i \neq j$. Qual é o melhor nesse caso?

- Pode-se mostrar que $\delta^{**}(X) = \frac{\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} X_i - 1\right)}{n(n-1)}$, que é função da estatística suficiente $\sum_{i=1}^{n} X_i$, é o "melhor" estimador não-viesado para θ^2.

- Sob certas condições, pode-se obter o melhor estimador não-viesado para $g(\theta)$ tomando-se um estimador não-viesado qualquer para $g(\theta)$ e condicionando tal estimador a um estatística suficiente (e completa) para θ (RAO-BLACKWELLIZACÃO).

Estimação

- Para $g(\theta) = \theta^2$, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para $g(\theta)$, do mesmo modo que $\delta^{(i,j)}(X) = X_i.X_j$, $i \neq j$. Qual é o melhor nesse caso?

- Pode-se mostrar que $\delta^{**}(X) = \frac{\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} X_i - 1\right)}{n(n-1)}$, que é função da estatística suficiente $\sum_{i=1}^{n} X_i$, é o "melhor" estimador não-viesado para θ^2.

- Sob certas condições, pode-se obter o melhor estimador não-viesado para $g(\theta)$ tomando-se um estimador não-viesado qualquer para $g(\theta)$ e condicionando tal estimador a um estatística suficiente (e completa) para θ (RAO-BLACKWELLIZACÃO).

Estimação

- No exemplo acima, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para θ^2. Com efeito, $\mathbb{E}(\delta^{(1,2)}(X)|\theta) = \mathbb{E}(X_1X_2|\theta) = \mathbb{E}(X_1|\theta)\mathbb{E}(X_2|\theta) = \theta^2$, $\forall\ \theta \in \Theta$.

- Além disso, $T(X) = \sum_{i=1}^{n} X_i$ é suficiente para θ:

  $\mathbb{P}(X_1 = x_1, ..., X_n = x_n|\theta) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i|\theta) =$

  $\prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}$

- Assim, podemos Rao-Blackwellizar $\delta^{(1,2)}$, obtendo

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)})|T(X))$.

Estimação

- No exemplo acima, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para θ^2. Com efeito, $\mathbb{E}(\delta^{(1,2)}(X)|\theta) = \mathbb{E}(X_1X_2|\theta) = \mathbb{E}(X_1|\theta)\mathbb{E}(X_2|\theta) = \theta^2$, $\forall\ \theta \in \Theta$.

- Além disso, $T(X) = \sum_{i=1}^n X_i$ é suficiente para θ:
 $\mathbb{P}(X_1 = x_1, ..., X_n = x_n|\theta) = \prod_{i=1}^n \mathbb{P}(X_i = x_i|\theta) = \prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}$

- Assim, podemos Rao-Blackwellizar $\delta^{(1,2)}$, obtendo

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)})|T(X))$.

Estimação

- No exemplo acima, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para θ^2. Com efeito,
$\mathbb{E}(\delta^{(1,2)}(X)|\theta) = \mathbb{E}(X_1 X_2|\theta) = \mathbb{E}(X_1|\theta)\mathbb{E}(X_2|\theta) = \theta^2$,
$\forall\ \theta \in \Theta$.

- Além disso, $T(X) = \sum_{i=1}^{n} X_i$ é suficiente para θ:
$\mathbb{P}(X_1 = x_1, ..., X_n = x_n|\theta) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i|\theta) =$
$\prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}$

- Assim, podemos Rao-Blackwellizar $\delta^{(1,2)}$, obtendo

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)})|T(X))$.

Estimação

- No exemplo acima, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para θ^2. Com efeito, $\mathbb{E}(\delta^{(1,2)}(X)|\theta) = \mathbb{E}(X_1X_2|\theta) = \mathbb{E}(X_1|\theta)\mathbb{E}(X_2|\theta) = \theta^2$, $\forall\ \theta \in \Theta$.

- Além disso, $T(X) = \sum_{i=1}^n X_i$ é suficiente para θ:
  $\mathbb{P}(X_1 = x_1, ..., X_n = x_n|\theta) = \prod_{i=1}^n \mathbb{P}(X_i = x_i|\theta) =$
  $\prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}$

- Assim, podemos Rao-Blackwellizar $\delta^{(1,2)}$, obtendo

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)})|T(X))$.

Estimação

- No exemplo acima, $\delta^{(1,2)}(X) = X_1.X_2$ é não-viesado para θ^2. Com efeito, $\mathbb{E}(\delta^{(1,2)}(X)|\theta) = \mathbb{E}(X_1X_2|\theta) = \mathbb{E}(X_1|\theta)\mathbb{E}(X_2|\theta) = \theta^2$, $\forall\ \theta \in \Theta$.

- Além disso, $T(X) = \sum_{i=1}^{n} X_i$ é suficiente para θ:
  $\mathbb{P}(X_1 = x_1, ..., X_n = x_n|\theta) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i|\theta) =$
  $\prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}$

- Assim, podemos Rao-Blackwellizar $\delta^{(1,2)}$, obtendo

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)})|T(X))$.

Estimação

- Para $t \in \{0, 1, ..., n\}$, temos:

- $\mathbb{E}(\delta^{(1,2)})(X)|T(X)) = t, \theta) = \mathbb{E}(X_1 X_2 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\mathbb{P}(X_1 X_2 = 1 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\mathbb{P}(X_1 = X_2 = 1 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\frac{\mathbb{P}(X_1 = X_2 = 1, \sum_{i=1}^{n} X_i = t | \theta)}{\mathbb{P}(\sum_{i=1}^{n} X_i = t | \theta)} = \frac{\mathbb{P}(X_1 = X_2 = 1, \sum_{i=3}^{n} X_i = t - 2 | \theta)}{\mathbb{P}(\sum_{i=1}^{n} X_i = t | \theta)} =$

  $\frac{\theta \ \theta \ C_{t-2}^{n-2} \theta^{t-2} (1-\theta)^{n-t}}{C_t^n \theta^t (1-\theta)^{n-t}} = \frac{t(t-1)}{n(n-1)}$. Assim,

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)}(X))|T(X)) = \frac{(\sum_{i=1}^{n} X_i) \ (\sum_{i=1}^{n} X_i - 1)}{n \ (n-1)}$

Estimação

- Para $t \in \{0, 1, ..., n\}$, temos:

- $\mathbb{E}(\delta^{(1,2)})(X)|T(X)) = t, \theta) = \mathbb{E}(X_1 X_2 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\mathbb{P}(X_1 X_2 = 1 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\mathbb{P}(X_1 = X_2 = 1 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\frac{\mathbb{P}(X_1 = X_2 = 1, \sum_{i=1}^{n} X_i = t | \theta)}{\mathbb{P}(\sum_{i=1}^{n} X_i = t | \theta)} = \frac{\mathbb{P}(X_1 = X_2 = 1, \sum_{i=3}^{n} X_i = t-2 | \theta)}{\mathbb{P}(\sum_{i=1}^{n} X_i = t | \theta)} =$

  $\frac{\theta \ \theta \ C_{t-2}^{n-2} \theta^{t-2} (1-\theta)^{n-t}}{C_t^n \theta^t (1-\theta)^{n-t}} = \frac{t(t-1)}{n(n-1)}$. Assim,

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)}(X))|T(X)) = \frac{(\sum_{i=1}^{n} X_i) \ (\sum_{i=1}^{n} X_i - 1)}{n \ (n-1)}$

Estimação

- Para $t \in \{0, 1, ..., n\}$, temos:

- $\mathbb{E}(\delta^{(1,2)})(X)|T(X)) = t, \theta) = \mathbb{E}(X_1 X_2 | \sum_{i=1}^{n} X_i = t, \theta) =$
  $\mathbb{P}(X_1 X_2 = 1 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\mathbb{P}(X_1 = X_2 = 1 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\frac{\mathbb{P}(X_1 = X_2 = 1, \sum_{i=1}^{n} X_i = t | \theta)}{\mathbb{P}(\sum_{i=1}^{n} X_i = t | \theta)} = \frac{\mathbb{P}(X_1 = X_2 = 1, \sum_{i=3}^{n} X_i = t-2 | \theta)}{\mathbb{P}(\sum_{i=1}^{n} X_i = t | \theta)} =$

  $\frac{\theta \; \theta \; C_{t-2}^{n-2} \theta^{t-2} (1-\theta)^{n-t}}{C_t^n \theta^t (1-\theta)^{n-t}} = \frac{t(t-1)}{n(n-1)}$. Assim,

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)}(X))|T(X)) = \frac{(\sum_{i=1}^{n} X_i) \; (\sum_{i=1}^{n} X_i - 1)}{n \; (n-1)}$

Estimação

- Para $t \in \{0, 1, ..., n\}$, temos:

- $\mathbb{E}(\delta^{(1,2)})(X)|T(X)) = t, \theta) = \mathbb{E}(X_1 X_2 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\mathbb{P}(X_1 X_2 = 1 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\mathbb{P}(X_1 = X_2 = 1 | \sum_{i=1}^{n} X_i = t, \theta) =$

  $\frac{\mathbb{P}(X_1 = X_2 = 1, \sum_{i=1}^{n} X_i = t|\theta)}{\mathbb{P}(\sum_{i=1}^{n} X_i = t|\theta)} = \frac{\mathbb{P}(X_1 = X_2 = 1, \sum_{i=3}^{n} X_i = t-2|\theta)}{\mathbb{P}(\sum_{i=1}^{n} X_i = t|\theta)} =$

  $\frac{\theta \ \theta \ C_{t-2}^{n-2} \theta^{t-2} (1-\theta)^{n-t}}{C_t^n \theta^t (1-\theta)^{n-t}} = \frac{t(t-1)}{n(n-1)}$. Assim,

- $\delta^{**}(X) = \mathbb{E}(\delta^{(1,2)}(X))|T(X)) = \frac{(\sum_{i=1}^{n} X_i) \ (\sum_{i=1}^{n} X_i \ - \ 1)}{n \ (n-1)}$

Estimação

- $\delta^{**}(X)$ satisfaz: (Teorema de Rao-Blackwell)
- $\mathbb{E}(\delta^{**}(X)|\theta) = \theta \; , \; \forall\theta \in \Theta$ e
- $Var(\delta^{**}(X)|\theta) \; \leq \; Var(\delta^{(1,2)}(X)|\theta) \; , \; \forall\theta \in \Theta.$
- Esse procedimento (Rao-Blackwellização) pode ser adotado para um estimadores não-viesados sob condições relativamente gerais.
- Há outras alternativas de estimadores para θ $(g(\theta))$: por exemplo, estimador de máxima verossimilhança.

Estimação

- $\delta^{**}(X)$ satisfaz: (Teorema de Rao-Blackwell)
- $\mathbb{E}(\delta^{**}(X)|\theta) = \theta \;,\; \forall \theta \in \Theta$ e
- $Var(\delta^{**}(X)|\theta) \;\leq\; Var(\delta^{(1,2)}(X)|\theta) \;,\; \forall \theta \in \Theta.$
- Esse procedimento (Rao-Blackwellização) pode ser adotado para um estimadores não-viesados sob condições relativamente gerais.
- Há outras alternativas de estimadores para θ $(g(\theta))$: por exemplo, estimador de máxima verossimilhança.

Estimação

- $\delta^{**}(X)$ satisfaz: (Teorema de Rao-Blackwell)
- $\mathbb{E}(\delta^{**}(X)|\theta) = \theta \ , \ \forall\theta \in \Theta$ e
- $Var(\delta^{**}(X)|\theta) \ \leq \ Var(\delta^{(1,2)}(X)|\theta) \ , \ \forall\theta \in \Theta.$
- Esse procedimento (Rao-Blackwellização) pode ser adotado para um estimadores não-viesados sob condições relativamente gerais.
- Há outras alternativas de estimadores para θ $(g(\theta))$: por exemplo, estimador de máxima verossimilhança.

Estimação

- $\delta^{**}(X)$ satisfaz: (Teorema de Rao-Blackwell)
- $\mathbb{E}(\delta^{**}(X)|\theta) = \theta \ , \ \forall\theta \in \Theta$ e
- $Var(\delta^{**}(X)|\theta) \ \leq \ Var(\delta^{(1,2)}(X)|\theta) \ , \ \forall\theta \in \Theta.$
- Esse procedimento (Rao-Blackwellização) pode ser adotado para um estimadores não-viesados sob condições relativamente gerais.
- Há outras alternativas de estimadores para θ $(g(\theta))$: por exemplo, estimador de máxima verossimilhança.

Estimação

- $\delta^{**}(X)$ satisfaz: (Teorema de Rao-Blackwell)

- $\mathbb{E}(\delta^{**}(X)|\theta) = \theta \ , \ \forall\theta \in \Theta$ e

- $Var(\delta^{**}(X)|\theta) \ \leq \ Var(\delta^{(1,2)}(X)|\theta) \ , \ \forall\theta \in \Theta.$

- Esse procedimento (Rao-Blackwellização) pode ser adotado para um estimadores não-viesados sob condições relativamente gerais.

- Há outras alternativas de estimadores para θ $(g(\theta))$: por exemplo, estimador de máxima verossimilhança.

Estimação

- $\delta^{**}(X)$ satisfaz: (Teorema de Rao-Blackwell)
- $\mathbb{E}(\delta^{**}(X)|\theta) = \theta \ , \ \forall\theta \in \Theta$ e
- $Var(\delta^{**}(X)|\theta) \ \leq \ Var(\delta^{(1,2)}(X)|\theta) \ , \ \forall\theta \in \Theta.$
- Esse procedimento (Rao-Blackwellização) pode ser adotado para um estimadores não-viesados sob condições relativamente gerais.
- Há outras alternativas de estimadores para θ $(g(\theta))$: por exemplo, estimador de máxima verossimilhança.