Inferência Estatística Comparada

AULA 10 - PRINCÍPIOS DE INFERÊNCIA

Resumo

- Até o momento, avaliamos, comparativamente, algumas abordagens à Inferência Esatística (frequencista, bayesiana e um pouco da verossimilhancista) com relação aos seguintes aspectos:

- 1) Interpretação de probabilidade (Frequencista x Subjetiva)

- 2) Modelo Estatístico ($(\mathcal{X}, \mathcal{F}, \mathcal{P})$ versus $(\Theta \times \mathcal{X}, \mathcal{A}, \mathbb{P})$)

- 3) Natureza dos elementos básicos em inferência:
  - parâmetro
  - distribuição a priori (marginal) para o parâmetro
  - IID X Permutabilidade

Resumo

- Até o momento, avaliamos, comparativamente, algumas abordagens à Inferência Esatística (frequencista, bayesiana e um pouco da verossimilhancista) com relação aos seguintes aspectos:

- 1) Interpretação de probabilidade
(Frequencista x Subjetiva)

- 2) Modelo Estatístico
($(\mathcal{X}, \mathcal{F}, \mathcal{P})$ versus $(\Theta \times \mathcal{X}, \mathcal{A}, \mathbb{P})$)

- 3) Natureza dos elementos básicos em inferência:
  - parâmetro
  - distribuição a priori (marginal) para o parâmetro
  - IID X Permutabilidade

Resumo

- Até o momento, avaliamos, comparativamente, algumas abordagens à Inferência Esatística (frequencista, bayesiana e um pouco da verossimilhancista) com relação aos seguintes aspectos:

- 1) Interpretação de probabilidade (Frequencista x Subjetiva)

- 2) Modelo Estatístico ($(\mathcal{X}, \mathcal{F}, \mathcal{P})$ versus $(\Theta \times \mathcal{X}, \mathcal{A}, \mathbb{P})$)

- 3) Natureza dos elementos básicos em inferência:
  - parâmetro
  - distribuição a priori (marginal) para o parâmetro
  - IID X Permutabilidade

Resumo

- Até o momento, avaliamos, comparativamente, algumas abordagens à Inferência Esatística (frequencista, bayesiana e um pouco da verossimilhancista) com relação aos seguintes aspectos:

- 1) Interpretação de probabilidade (Frequencista x Subjetiva)

- 2) Modelo Estatístico ($(\mathcal{X}, \mathcal{F}, \mathcal{P})$ versus $(\Theta \times \mathcal{X}, \mathcal{A}, \mathbb{P})$)

- 3) Natureza dos elementos básicos em inferência:
  - parâmetro
  - distribuição a priori (marginal) para o parâmetro
  - IID X Permutabilidade

Resumo

- Até o momento, avaliamos, comparativamente, algumas abordagens à Inferência Esatística (frequencista, bayesiana e um pouco da verossimilhancista) com relação aos seguintes aspectos:
- 1) Interpretação de probabilidade (Frequencista x Subjetiva)
- 2) Modelo Estatístico ($(\mathcal{X}, \mathcal{F}, \mathcal{P})$ versus $(\Theta \times \mathcal{X}, \mathcal{A}, \mathbb{P})$)
- 3) Natureza dos elementos básicos em inferência:
  - parâmetro
  - distribuição a priori (marginal) para o parâmetro
  - IID X Permutabilidade

Resumo

- 4) Construção de procedimentos inferencias ótimos (estimação e testes de hipóteses):

  Otimalidade pré-experimentação (abordagem frequencista)

  X

  Otimalidade via condicionamento (Teoria da Decisão Bayesiana e abordagem verossimilhancista)

- Em geral, procedimentos frequencistas são baseados em integração sobre o espaço amostral; procedimentos bayesianos são determinados via condicionamento ao ponto amostral observado.

Resumo

- 4) Construção de procedimentos inferencias ótimos (estimação e testes de hipóteses):

  Otimalidade pré-experimentação (abordagem frequencista)

  X

  Otimalidade via condicionamento (Teoria da Decisão Bayesiana e abordagem verossimilhancista)

- Em geral, procedimentos frequencistas são baseados em integração sobre o espaço amostral; procedimentos bayesianos são determinados via condicionamento ao ponto amostral observado.

Resumo

- 4) Construção de procedimentos inferencias ótimos (estimação e testes de hipóteses):

  Otimalidade pré-experimentação (abordagem frequencista)

  X

  Otimalidade via condicionamento (Teoria da Decisão Bayesiana e abordagem verossimilhancista)

- Em geral, procedimentos frequencistas são baseados em integração sobre o espaço amostral; procedimentos bayesianos são determinados via condicionamento ao ponto amostral observado.

Resumo

- i) PROBLEMA DE ESTIMACÃO:

a) Estimador não-viesado para θ:

$\int_{\mathcal{X}} \delta(x) f(x|\theta)\, dx \;=\; \theta \,,\; \forall\theta \in \Theta$, e

$\int_{\mathcal{X}} (\delta^*(x) - \theta)^2 f(x|\theta)\, dx \;\leq\; \int_{\mathcal{X}} (\delta(x) - \theta)^2 f(x|\theta)\, dx$

b) Estimador de Bayes para θ:

$\delta^*_B(x)$ é obtido minimizando (em $\mathcal{D}$) $\mathbb{E}[\, L(d, \theta) \mid X = x \,]$.

Exemplo: $\delta(x) \;=\; \mathbb{E}(\, \theta \mid X = x \,)$

Resumo

- i) PROBLEMA DE ESTIMACÃO:

a) Estimador não-viesado para θ:

$\int_{\mathcal{X}} \delta(x) f(x|\theta)\, dx \;=\; \theta \;,\; \forall\theta \in \Theta$, e

$\int_{\mathcal{X}} (\delta^*(x) - \theta)^2 f(x|\theta)\, dx \;\leq\; \int_{\mathcal{X}} (\delta(x) - \theta)^2 f(x|\theta)\, dx$

b) Estimador de Bayes para θ:

$\delta^*_B(x)$ é obtido minimizando (em $\mathcal{D}$) $\mathbb{E}[\; L(d,\theta) \mid X = x \;]$.

Exemplo: $\delta(x) \;=\; \mathbb{E}(\; \theta \mid X = x \;)$

Resumo

- ii) PROBLEMA DE ESTIMACÃO POR INTERVALO:

  a) Intervalo de confiança para θ:

  $\mathbb{P}(\ A(X)\ \leq\ \theta\ \leq\ B(X)\ |\ \theta)\ \geq\ \gamma\ ,\ \forall\theta\in\Theta\ $, ou

  $\int_{\mathcal{X}} \mathbb{I}_{[A(x),B(x)]}(\theta)\ f(x|\theta)\ dx\ \geq\ \gamma\ ,\ \forall\theta\in\Theta$

  b) Intervalo de probabilidade (credibilidade) para θ:

  $\mathbb{P}(\ A(x)\ \leq\ \theta\ \leq\ B(x)\ |\ X=x\ )\ \geq\ \gamma\ $, ou

  $\int_{[A(x),B(x)]} f(\theta\ |\ X=x\ )\ d\theta\ \geq\ \gamma\ ,\ [A(x),B(x)]\ \subset\ \Theta$.

Resumo

- ii) PROBLEMA DE ESTIMACÃO POR INTERVALO:

  a) Intervalo de confiança para θ:

  $\mathbb{P}(\ A(X)\ \leq\ \theta\ \leq\ B(X)\mid\theta)\ \geq\ \gamma$, $\forall\theta\in\Theta$, ou

  $\int_{\mathcal{X}} \mathbb{I}_{[A(x),B(x)]}(\theta)\ f(x|\theta)\ dx\ \geq\ \gamma$, $\forall\theta\in\Theta$

  b) Intervalo de probabilidade (credibilidade) para θ:

  $\mathbb{P}(\ A(x)\ \leq\ \theta\ \leq\ B(x)\mid X=x\ )\ \geq\ \gamma$, ou

  $\int_{[A(x),B(x)]} f(\theta\mid X=x\ )\ d\theta\ \geq\ \gamma$, $[A(x),B(x)]\ \subset\ \Theta$.

Resumo

- iii) PROBLEMA DE TESTE DE HIPÓTESES:

a) Teste UMP para as hipóteses $H_0 : \theta \in \Theta_0$ versus $\theta \in \Theta_1$:

$\int_{\mathcal{X}} \mathbb{I}_{\varphi^{*-1}\{1\}}(x)\ f(x|\theta)\ dx \ \leq \ \alpha$, $\forall\theta \in \Theta_0$, e , $\forall\theta \in \Theta_1$,

$\int_{\mathcal{X}} \mathbb{I}_{\varphi^{*-1}\{1\}}(x)\ f(x|\theta)\ dx \ \geq \ \int_{\mathcal{X}} \mathbb{I}_{\varphi^{-1}\{1\}}(x)\ f(x|\theta)\ dx$, .

b) Teste de Bayes com relação à perda 0 - 1 - c para H_0 versus H_1:

$\varphi^*_B(x) \ = \ 1 \ \Leftrightarrow \ \mathbb{P}(\ \theta \in \Theta_0 \ |\ X = x\ ) \ < \ \frac{c}{c+1}$, ou

$\int_{\Theta_0} \ f(\theta \ |\ X = x\ )\ d\theta \ < \ \frac{c}{c+1}$.

Resumo

- iii) PROBLEMA DE TESTE DE HIPÓTESES:

  a) Teste UMP para as hipóteses $H_0 : \theta \in \Theta_0$ versus $\theta \in \Theta_1$:

  $$\int_{\mathcal{X}} \mathbb{I}_{\varphi^{*-1}\{1\}}(x)\ f(x|\theta)\ dx \ \leq\ \alpha\ ,\ \forall\theta \in \Theta_0\ ,\ \text{e}\ ,\ \forall\theta \in \Theta_1\ ,$$

  $$\int_{\mathcal{X}} \mathbb{I}_{\varphi^{*-1}\{1\}}(x)\ f(x|\theta)\ dx \ \geq\ \int_{\mathcal{X}} \mathbb{I}_{\varphi^{-1}\{1\}}(x)\ f(x|\theta)\ dx\ ,\ .$$

  b) Teste de Bayes com relação à perda 0 - 1 - c para H_0 versus H_1:

  $$\varphi_B^*(x)\ =\ 1\ \Leftrightarrow\ \mathbb{P}(\ \theta \in \Theta_0\ |\ X = x\ )\ <\ \frac{c}{c+1}\ ,\ \text{ou}$$

  $$\int_{\Theta_0} f(\theta\ |\ X = x\ )\ d\theta\ <\ \frac{c}{c+1}.$$

Introdução

- Vamos examinar a partir de agora a concordância dos procedimentos inferenciais estudados com alguns princípios de inferência, princípios de como lidar (tratar) com a informação revelada pelos dados para fins inferenciais.

- Consideremos os seguintes exemplos:

Introdução

- Vamos examinar a partir de agora a concordância dos procedimentos inferenciais estudados com alguns princípios de inferência, princípios de como lidar (tratar) com a informação revelada pelos dados para fins inferenciais.

- Consideremos os seguintes exemplos:

Introdução

- Vamos examinar a partir de agora a concordância dos procedimentos inferenciais estudados com alguns princípios de inferência, princípios de como lidar (tratar) com a informação revelada pelos dados para fins inferenciais.

- Consideremos os seguintes exemplos:

Introdução

- 1) Seja $\Theta = (0, 1)$. Para inferir sobre θ, considere os seguintes experimentos supondo as observações (variáveis), dado θ, condicionalmente IID Bernoulli (θ):

  a) Observa-se o número de "sucessos" em 10 observações, X, e

  b) Registra-se o número de observações até a ocorrência do segundo "sucesso" (inclusive), Y.

  Consideremos a seguinte realização (tanto para experimento a) quanto para experimento b)): o primeiro "sucesso" ocorre ao observarmos a sexta variável de Bernoulli e o segundo "sucesso" ocorre na décima observação. Note que tal realização corresponde a $X = 2$ e $Y = 10$ nos experimentos a) e b), respectivamente.

Introdução

- 1) Seja $\Theta = (0, 1)$. Para inferir sobre θ, considere os seguintes experimentos supondo as observações (variáveis), dado θ, condicionalmente IID Bernoulli (θ):

  a) Observa-se o número de "sucessos" em 10 observações, X, e

  b) Registra-se o número de observações até a ocorrência do segundo "sucesso" (inclusive), Y.

  Consideremos a seguinte realização (tanto para experimento a) quanto para experimento b)): o primeiro "sucesso" ocorre ao observarmos a sexta variável de Bernoulli e o segundo "sucesso" ocorre na décima observação. Note que tal realização corresponde a $X = 2$ e $Y = 10$ nos experimentos a) e b), respectivamente.

Introdução

- Quais conclusões sobre θ podemos obter a partir dos experimentos a) e b) ? Quais inferenciais podemos fazer sobre θ a partir desses experimentos? Observe que

  Experimento a): $X|\theta \sim Binomial(10, \theta)$ e

  Experimento b): $Y|\theta \sim BN(2, \theta)$. Assim:

  $V_2^{a)}(\theta) \;=\; \mathbb{P}(X = 2|\theta) \;=\; \binom{10}{2}\theta^2(1-\theta)^8$ e

  $V_{10}^{b)}(\theta) \;=\; \mathbb{P}(Y = 10|\theta) \;=\; \binom{9}{1}\theta^2(1-\theta)^8$

  Note que, **para todo** $\theta \in [0, 1]$,

  $$V_2^{a)}(\theta) \;=\; \mathbf{5} \cdot V_{10}^{b)}(\theta)$$

Introdução

- Quais conclusões sobre θ podemos obter a partir dos experimentos a) e b) ? Quais inferenciais podemos fazer sobre θ a partir desses experimentos? Observe que

  Experimento a): $X|\theta \sim Binomial(10, \theta)$ e

  Experimento b): $Y|\theta \sim BN(2, \theta)$. Assim:

  $$V_2^{a)}(\theta) = \mathbb{P}(X = 2|\theta) = \binom{10}{2}\theta^2(1-\theta)^8 \text{ e}$$

  $$V_{10}^{b)}(\theta) = \mathbb{P}(Y = 10|\theta) = \binom{9}{1}\theta^2(1-\theta)^8$$

  Note que, **para todo** $\theta \in [0, 1]$,

  $$V_2^{a)}(\theta) = \mathbf{5} \cdot V_{10}^{b)}(\theta)$$

Introdução

- Inferências:

  1) Estimativas de máxima verossimilhança coincidem:

  $\delta_{MV}^{a)}(2) = \delta_{MV}^{b)}(10) \; = \; \frac{2}{10} \;\; (\delta_{MV}^{a)}(X) = \frac{X}{10}$ e $\delta_{MV}^{b)}(Y) = \frac{2}{Y})$

  ("likelihood intervals" também coincidem)

  2) Operação bayesiana - distribuições a posteriori coincidem:

  $\theta \mid X = 2 \quad \sim \quad \theta \mid Y = 10 \;\;,$

  contra qualquer distribuição a priori (comum sob os dois experimentos) para θ. Procedimentos derivados da posteriori, consequentemente, também coincidem.

Introdução

- Inferências:

  1) Estimativas de máxima verossimilhança coincidem:

  $\delta^{a)}_{MV}(2) = \delta^{b)}_{MV}(10) \ = \ \frac{2}{10}$ $(\delta^{a)}_{MV}(X) = \frac{X}{10}$ e $\delta^{b)}_{MV}(Y) = \frac{2}{Y})$

  ("likelihood intervals" também coincidem)

  2) Operação bayesiana - distribuições a posteriori coincidem:

  $\theta \mid X = 2 \ \sim \ \theta \mid Y = 10$,

  contra qualquer distribuição a priori (comum sob os dois experimentos) para θ. Procedimentos derivados da posteriori, consequentemente, também coincidem.

Introdução

- Inferências:

3) Estimativa baseada em estimador não-viesado:

$\delta_{NV}^{a)}(2) \; = \; \frac{2}{10}$ e $\delta_{NV}^{b)}(10) \; = \; \frac{2-1}{10-1} \; = \; \frac{1}{9}$

Lembrete:

$$Y \; \sim \; BN(k,\theta) \; , \; k > 1 \;\; \Rightarrow \;\; \mathbb{E}(\tfrac{k-1}{Y-1} \mid \theta) \; = \; \theta \;\; , \; \forall \, \theta \in (0,1)$$

Diferentemente do que ocorre com EMV e operação bayesiana, as estimativas baseadas em estimadores não-viesados são diferentes nesse caso (de proporcionalidade entre as funções de verossimilhança geradas por $X = 2$ e $Y = 10$).

Introdução

- Inferências:

  3) Estimativa baseada em estimador não-viesado:

  $\delta_{NV}^{a)}(2) = \frac{2}{10}$ e $\delta_{NV}^{b)}(10) = \frac{2-1}{10-1} = \frac{1}{9}$

  Lembrete:

  $Y \sim BN(k, \theta)\ ,\ k > 1 \ \Rightarrow\ \mathbb{E}(\frac{k-1}{Y-1} \mid \theta) = \theta\ ,\ \forall\, \theta \in (0, 1)$

  Diferentemente do que ocorre com EMV e operação bayesiana, as estimativas baseadas em estimadores não-viesados são diferentes nesse caso (de proporcionalidade entre as funções de verossimilhança geradas por $X = 2$ e $Y = 10$).

Introdução

- 2) Seja $\Theta = \{\theta_0, \theta_1\}$.

Experimento a) - Observação de X ($\mathcal{X} = \{1, 2, 3, 4, 5\}$):

x	1	2	3	4	5
$P(X = x \mid \theta_0)$	0,05	0,20	0,15	0,40	0,20
$P(X = x \mid \theta_1)$	0,20	0,40	0,15	0,20	0,05

Experimento b) - Observação de Y ($\mathcal{Y} = \{1, 2, 3, 4, 5\}$):

y	1	2	3	4	5
$P(Y = y \mid \theta_0)$	0,04	0,01	0,77	0,02	0,16
$P(Y = y \mid \theta_1)$	0,16	0,02	0,77	0,01	0,04

$V_i^{a)}(\theta) \;=\; k(i) \cdot V_i^{b)}(\theta)$, **para todo** $\theta \in \Theta$, $i \in \mathcal{X}$.

(por exemplo, $k(1) = 1,25$; $k(2) = 20$, etc.)

Introdução

- 2) Seja $\Theta = \{\theta_0, \theta_1\}$.

Experimento a) - Observação de X ($\mathcal{X} = \{1, 2, 3, 4, 5\}$):

x	1	2	3	4	5
$P(X = x \mid \theta_0)$	0,05	0,20	0,15	0,40	0,20
$P(X = x \mid \theta_1)$	0,20	0,40	0,15	0,20	0,05

Experimento b) - Observação de Y ($\mathcal{Y} = \{1, 2, 3, 4, 5\}$):

y	1	2	3	4	5
$P(Y = y \mid \theta_0)$	0,04	0,01	0,77	0,02	0,16
$P(Y = y \mid \theta_1)$	0,16	0,02	0,77	0,01	0,04

$V_i^{a)}(\theta) \;=\; k(i) \cdot V_i^{b)}(\theta)$, **para todo** $\theta \in \Theta$, $i \in \mathcal{X}$.

(por exemplo, $k(1) = 1,25$; $k(2) = 20$, etc.)

Introdução

- Inferências:

  1) Estimadores de máxima verossimilhança: como, para todo $i \in \mathcal{X}$,

  $$V_i^{a)}(\theta) \;=\; k(i) \cdot V_i^{b)}(\theta) \;,\; \textbf{para todo } \theta \in \{\theta_0, \theta_1\} \;,$$

  EMV (não apenas uma ou outra estimativa) para θ coincidem.

  2) Operação bayesiana - (todas) distribuições a posteriori coincidem, isto é

  $$\theta \mid X = i \;\sim\; \theta \mid Y = i \;,\; \textbf{para cada } i \in \mathcal{X},$$

  contra qualquer distribuição a priori (comum sob os dois experimentos) para θ.

Introdução

- Inferências:

  1) Estimadores de máxima verossimilhança: como, para todo $i \in \mathcal{X}$,

  $$V_i^{a)}(\theta) \;=\; k(i) \cdot V_i^{b)}(\theta) \;, \; \textbf{para todo } \theta \in \{\theta_0, \theta_1\} \;,$$

  EMV (não apenas uma ou outra estimativa) para θ coincidem.

  2) Operação bayesiana - (todas) distribuições a posteriori coincidem, isto é

  $$\theta \mid X = i \;\sim\; \theta \mid Y = i \;, \; \textbf{para cada } i \in \mathcal{X},$$

  contra qualquer distribuição a priori (comum sob os dois experimentos) para θ.

Introdução

- Além disso, quaisquer procedimentos derivados da posteriori (estimativas, decisões em testes de hipóteses, etc.) também coincidem.

3) Testes MP de nível $0,05$ para $H_0 : \theta = \theta_0 \times H_1 : \theta = \theta_1$:

No experimento a): $\varphi^{a)}_{MP}(x) = \begin{cases} 1, & \text{se } x = 1 \\ 0, & \text{se } x \neq 1 \end{cases}$

No experimento b): $\varphi^{b)}_{MP}(y) = \begin{cases} 1, & \text{se } y \in \{1,2\} \\ 0, & \text{se } y \in \{3,4,5\} \end{cases}$

Note que $\varphi^{a)}_{MP}(2) \neq \varphi^{b)}_{MP}(2)$, ainda que

$V^{a)}_2(\theta) = \mathbf{20} \cdot V^{b)}_2(\theta)$, **para todo** $\theta \in \Theta$.

Introdução

- Além disso, quaisquer procedimentos derivados da posteriori (estimativas, decisões em testes de hipóteses, etc.) também coincidem.

3) Testes MP de nível $0,05$ para $H_0 : \theta = \theta_0 \times H_1 : \theta = \theta_1$:

No experimento a): $\varphi^{a)}_{MP}(x) = \begin{cases} 1, & \text{se } x = 1 \\ 0, & \text{se } x \neq 1 \end{cases}$

No experimento b): $\varphi^{b)}_{MP}(y) = \begin{cases} 1, & \text{se } y \in \{1,2\} \\ 0, & \text{se } y \in \{3,4,5\} \end{cases}$

Note que $\varphi^{a)}_{MP}(2) \neq \varphi^{b)}_{MP}(2)$, ainda que

$V_2^{a)}(\theta) = \mathbf{20} \cdot V_2^{b)}(\theta)$, **para todo** $\theta \in \Theta$ **.**

Introdução

- **Pergunta:** Nesse caso, é razoável esperar que as conclusões sobre H_0 versus H_1 sejam disitintas a partir de $X = 2$ e $Y = 2$ se

$$V_2^{a)}(\theta) \;=\; \mathbf{20} \cdot V_2^{b)}(\theta) \;, \; \textbf{para todo } \theta \in \Theta \;, \; \text{isto é,}$$

visto que as funções de verossimilhança geradas por $X = 2$ e $Y = 2$ nos experimentos a) e b), respectivamente, são proporcionais (e, consequentemente, encerram essencialmente "a mesma informação" sobre θ) ?

Discutiremos essa questão (e outras de natureza similar - e.g. : Exemplo 1) a partir de alguns príncipios de inferência. Tais princípios estabelecem como tratar a informação revelada pelos dados para fins inferenciais sobre o parâmetro de interesse; o que esperar de inferências (conclusões) feitas sobre θ.

Introdução

- **Pergunta:** Nesse caso, é razoável esperar que as conclusões sobre H_0 versus H_1 sejam disitintas a partir de $X = 2$ e $Y = 2$ se

$$V_2^{a)}(\theta) \quad = \quad \mathbf{20} \; \cdot \; V_2^{b)}(\theta) \quad , \ \textbf{para todo}\ \theta \in \Theta \quad , \ \text{isto é,}$$

visto que as funções de verossimilhança geradas por $X = 2$ e $Y = 2$ nos experimentos a) e b), respectivamente, são proporcionais (e, consequentemente, encerram essencialmente "a mesma informação" sobre θ) ?

Discutiremos essa questão (e outras de natureza similar - e.g. : Exemplo 1) a partir de alguns príncipios de inferência. Tais princípios estabelecem como tratar a informação revelada pelos dados para fins inferenciais sobre o parâmetro de interesse; o que esperar de inferências (conclusões) feitas sobre θ.

Princípios de Inferência

- Vimos que a função de verossimilhança gerada pelo ponto amostral $x \in \mathcal{X}$, $V_x(\cdot)$, contém toda informação amostral (do ponto amostral x) sobre o paramêtro θ.
- Pré-experimentação: $\mathcal{V} = \{V_x : x \in \mathcal{X}\}$
- Pós-experimentação: após a observação de $x \in \mathcal{X}$, precisamos de (toda) $\mathcal{V}$ para fins inferenciais? Ou basta V_x? Ou é necessário outro objeto?
- **Príncipio da Verossimilhança (PV)**: Toda (qualquer) inferência sobre θ (parâmetro de interesse) deve depender de $x \in \mathcal{X}$ (e do experimento que o gerou) apenas através de V_x.
- Vamos examinar esse príncipio (e alguns outros) a partir da formulação de A. Birnbaum (1962).

Princípios de Inferência

- Vimos que a função de verossimilhança gerada pelo ponto amostral $x \in \mathcal{X}$, $V_x(\cdot)$, contém toda informação amostral (do ponto amostral x) sobre o paramêtro θ.
- Pré-experimentação: $\mathcal{V} = \{V_x : x \in \mathcal{X}\}$
- Pós-experimentação: após a observação de $x \in \mathcal{X}$, precisamos de (toda) $\mathcal{V}$ para fins inferenciais? Ou basta V_x? Ou é necessário outro objeto?
- **Príncipio da Verossimilhança (PV)**: Toda (qualquer) inferência sobre θ (parâmetro de interesse) deve depender de $x \in \mathcal{X}$ (e do experimento que o gerou) apenas através de V_x.
- Vamos examinar esse príncipio (e alguns outros) a partir da formulação de A. Birnbaum (1962).

Princípios de Inferência

- Vimos que a função de verossimilhança gerada pelo ponto amostral $x \in \mathcal{X}$, $V_x(\cdot)$, contém toda informação amostral (do ponto amostral x) sobre o paramêtro θ.
- Pré-experimentação: $\mathcal{V} = \{V_x : x \in \mathcal{X}\}$
- Pós-experimentação: após a observação de $x \in \mathcal{X}$, precisamos de (toda) $\mathcal{V}$ para fins inferenciais? Ou basta V_x? Ou é necessário outro objeto?
- **Príncipio da Verossimilhança (PV)**: Toda (qualquer) inferência sobre θ (parâmetro de interesse) deve depender de $x \in \mathcal{X}$ (e do experimento que o gerou) apenas através de V_x.
- Vamos examinar esse príncipio (e alguns outros) a partir da formulação de A. Birnbaum (1962).

Princípios de Inferência

- Vimos que a função de verossimilhança gerada pelo ponto amostral $x \in \mathcal{X}$, $V_x(\cdot)$, contém toda informação amostral (do ponto amostral x) sobre o paramêtro θ.
- Pré-experimentação: $\mathcal{V} = \{V_x : x \in \mathcal{X}\}$
- Pós-experimentação: após a observação de $x \in \mathcal{X}$, precisamos de (toda) $\mathcal{V}$ para fins inferenciais? Ou basta V_x? Ou é necessário outro objeto?
- **Príncipio da Verossimilhança (PV)**: Toda (qualquer) inferência sobre θ (parâmetro de interesse) deve depender de $x \in \mathcal{X}$ (e do experimento que o gerou) apenas através de V_x.
- Vamos examinar esse príncipio (e alguns outros) a partir da formulação de A. Birnbaum (1962).

Princípios de Inferência

- Vimos que a função de verossimilhança gerada pelo ponto amostral $x \in \mathcal{X}$, $V_x(\cdot)$, contém toda informação amostral (do ponto amostral x) sobre o paramêtro θ.
- Pré-experimentação: $\mathcal{V} = \{V_x : x \in \mathcal{X}\}$
- Pós-experimentação: após a observação de $x \in \mathcal{X}$, precisamos de (toda) $\mathcal{V}$ para fins inferenciais? Ou basta V_x? Ou é necessário outro objeto?
- **Príncipio da Verossimilhança (PV)**: Toda (qualquer) inferência sobre θ (parâmetro de interesse) deve depender de $x \in \mathcal{X}$ (e do experimento que o gerou) apenas através de V_x.
- Vamos examinar esse príncipio (e alguns outros) a partir da formulação de A. Birnbaum (1962).

Princípios de Inferência

- Vimos que a função de verossimilhança gerada pelo ponto amostral $x \in \mathcal{X}$, $V_x(\cdot)$, contém toda informação amostral (do ponto amostral x) sobre o paramêtro θ.
- Pré-experimentação: $\mathcal{V} \;=\; \{V_x : x \in \mathcal{X}\}$
- Pós-experimentação: após a observação de $x \in \mathcal{X}$, precisamos de (toda) $\mathcal{V}$ para fins inferenciais? Ou basta V_x? Ou é necessário outro objeto?
- **Príncipio da Verossimilhança (PV)**: Toda (qualquer) inferência sobre θ (parâmetro de interesse) deve depender de $x \in \mathcal{X}$ (e do experimento que o gerou) apenas através de V_x.
- Vamos examinar esse príncipio (e alguns outros) a partir da formulação de A. Birnbaum (1962).

Princípios de Inferência

- "Likelihood" - breve resumo (Edwards (1972,1974)):
- Fisher (1921)
- Bartlett (1936,1952) , Barnard (1947, 1949) , Fisher (1956): primeiras menções ao PV.
- Birnbaum (1962): desenvolvimento formal do PV a partir de conceitos e príncipios mais "simples" ("intuitivos") como suficiência e condicionalidade.
- Barnard (1962), Pratt (1965), Birnbaum (1972), Cox e Hinkley (1974), Basu (1975), Dawid (1981), Barnett (1982), Berger e Wolpert (1988) **(*)**; recentemente Wechsler, Pereira e Marques (2008), Mayo (2014), entre outros.
- Teses: Lurdes Y. T. Inoue (1995) ; V. Fossaluza (2008) ; Pedro F. C. Silva (UFMG - 2017); Jaime E. L. Curivil (2019)

Princípios de Inferência

- "Likelihood" - breve resumo (Edwards (1972,1974)):
- Fisher (1921)
- Bartlett (1936,1952) , Barnard (1947, 1949) , Fisher (1956): primeiras menções ao PV.
- Birnbaum (1962): desenvolvimento formal do PV a partir de conceitos e príncipios mais "simples" ("intuitivos") como suficiência e condicionalidade.
- Barnard (1962), Pratt (1965), Birnbaum (1972), Cox e Hinkley (1974), Basu (1975), Dawid (1981), Barnett (1982), Berger e Wolpert (1988) **(*)**; recentemente Wechsler, Pereira e Marques (2008), Mayo (2014), entre outros.
- Teses: Lurdes Y. T. Inoue (1995) ; V. Fossaluza (2008) ; Pedro F. C. Silva (UFMG - 2017); Jaime E. L. Curivil (2019)

Princípios de Inferência

- "Likelihood" - breve resumo (Edwards (1972,1974)):
- Fisher (1921)
- Bartlett (1936,1952) , Barnard (1947, 1949) , Fisher (1956): primeiras menções ao PV.
- Birnbaum (1962): desenvolvimento formal do PV a partir de conceitos e príncipios mais "simples" ("intuitivos") como suficiência e condicionalidade.
- Barnard (1962), Pratt (1965), Birnbaum (1972), Cox e Hinkley (1974), Basu (1975), Dawid (1981), Barnett (1982), Berger e Wolpert (1988) **(*)**; recentemente Wechsler, Pereira e Marques (2008), Mayo (2014), entre outros.
- Teses: Lurdes Y. T. Inoue (1995) ; V. Fossaluza (2008) ; Pedro F. C. Silva (UFMG - 2017); Jaime E. L. Curivil (2019)

Princípios de Inferência

- "Likelihood" - breve resumo (Edwards (1972,1974)):
- Fisher (1921)
- Bartlett (1936,1952) , Barnard (1947, 1949) , Fisher (1956): primeiras menções ao PV.
- Birnbaum (1962): desenvolvimento formal do PV a partir de conceitos e príncipios mais "simples" ("intuitivos") como suficiência e condicionalidade.
- Barnard (1962), Pratt (1965), Birnbaum (1972), Cox e Hinkley (1974), Basu (1975), Dawid (1981), Barnett (1982), Berger e Wolpert (1988) **(*)**; recentemente Wechsler, Pereira e Marques (2008), Mayo (2014), entre outros.
- Teses: Lurdes Y. T. Inoue (1995) ; V. Fossaluza (2008) ; Pedro F. C. Silva (UFMG - 2017); Jaime E. L. Curivil (2019)

Princípios de Inferência

- "Likelihood" - breve resumo (Edwards (1972,1974)):
- Fisher (1921)
- Bartlett (1936,1952) , Barnard (1947, 1949) , Fisher (1956): primeiras menções ao PV.
- Birnbaum (1962): desenvolvimento formal do PV a partir de conceitos e príncipios mais "simples" ("intuitivos") como suficiência e condicionalidade.
- Barnard (1962), Pratt (1965), Birnbaum (1972), Cox e Hinkley (1974), Basu (1975), Dawid (1981), Barnett (1982), Berger e Wolpert (1988) **(*)**; recentemente Wechsler, Pereira e Marques (2008), Mayo (2014), entre outros.
- Teses: Lurdes Y. T. Inoue (1995) ; V. Fossaluza (2008) ; Pedro F. C. Silva (UFMG - 2017); Jaime E. L. Curivil (2019)

Princípios de Inferência

- "Likelihood" - breve resumo (Edwards (1972,1974)):
- Fisher (1921)
- Bartlett (1936,1952) , Barnard (1947, 1949) , Fisher (1956): primeiras menções ao PV.
- Birnbaum (1962): desenvolvimento formal do PV a partir de conceitos e príncipios mais "simples" ("intuitivos") como suficiência e condicionalidade.
- Barnard (1962), Pratt (1965), Birnbaum (1972), Cox e Hinkley (1974), Basu (1975), Dawid (1981), Barnett (1982), Berger e Wolpert (1988) **(*)**; recentemente Wechsler, Pereira e Marques (2008), Mayo (2014), entre outros.
- Teses: Lurdes Y. T. Inoue (1995) ; V. Fossaluza (2008) ; Pedro F. C. Silva (UFMG - 2017); Jaime E. L. Curivil (2019)

Princípios de Inferência

- "Likelihood" - breve resumo (Edwards (1972,1974)):
- Fisher (1921)
- Bartlett (1936,1952) , Barnard (1947, 1949) , Fisher (1956): primeiras menções ao PV.
- Birnbaum (1962): desenvolvimento formal do PV a partir de conceitos e príncipios mais "simples" ("intuitivos") como suficiência e condicionalidade.
- Barnard (1962), Pratt (1965), Birnbaum (1972), Cox e Hinkley (1974), Basu (1975), Dawid (1981), Barnett (1982), Berger e Wolpert (1988) **(*)**; recentemente Wechsler, Pereira e Marques (2008), Mayo (2014), entre outros.
- Teses: Lurdes Y. T. Inoue (1995) ; V. Fossaluza (2008) ; Pedro F. C. Silva (UFMG - 2017); Jaime E. L. Curivil (2019)

Princípios de Inferência

- Sejam θ, X, Θ, $\mathcal{X}$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ como definidos anteriormente.

- **DEFINICÃO:** Um experimento E para θ é um trio $(X, \theta, \mathcal{P})$, onde X é a quantidade observada segundo $\mathbb{P}_\theta$, $\theta \in \Theta$.

  Notação: $E = (X, \theta, \mathcal{P})$ (alternativamente, $(\mathcal{X}, \Theta, \mathcal{P})$).

- Exemplo 1: Seja $n \in \mathbb{N}^*$. Consideremos $\mathcal{X} = \{0, 1, ..., n\}$, $\Theta = (0, 1)$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in (0, 1)\}$ tal que

  $$\mathbb{P}_\theta(\{x\}) = \mathbb{P}(X = x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}\, \mathbb{I}_{\{0,1,\ldots,n\}}(x)$$

  $E = (\mathcal{X}, \Theta, \mathcal{P})$ é um experimento para θ.

Princípios de Inferência

- Sejam θ, X, Θ, $\mathcal{X}$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ como definidos anteriormente.

- **DEFINICÃO:** Um experimento E para θ é um trio $(X, \theta, \mathcal{P})$, onde X é a quantidade observada segundo $\mathbb{P}_\theta$, $\theta \in \Theta$.

  Notação: $E = (X, \theta, \mathcal{P})$ (alternativamente, $(\mathcal{X}, \Theta, \mathcal{P})$).

- Exemplo 1: Seja $n \in \mathbb{N}^*$. Consideremos $\mathcal{X} = \{0, 1, ..., n\}$, $\Theta = (0, 1)$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in (0, 1)\}$ tal que

  $$\mathbb{P}_\theta(\{x\}) = \mathbb{P}(X = x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}\,\mathbb{I}_{\{0,1,\dots,n\}}(x)$$

  $E = (\mathcal{X}, \Theta, \mathcal{P})$ é um experimento para θ.

Princípios de Inferência

- Sejam θ, X, Θ, $\mathcal{X}$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ como definidos anteriormente.

- **DEFINICÃO:** Um experimento E para θ é um trio $(X, \theta, \mathcal{P})$, onde X é a quantidade observada segundo $\mathbb{P}_\theta$, $\theta \in \Theta$.

  Notação: $E = (X, \theta, \mathcal{P})$ (alternativamente, $(\mathcal{X}, \Theta, \mathcal{P})$).

- Exemplo 1: Seja $n \in \mathbb{N}^*$. Consideremos $\mathcal{X} = \{0, 1, ..., n\}$, $\Theta = (0, 1)$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in (0, 1)\}$ tal que

  $$\mathbb{P}_\theta(\{x\}) = \mathbb{P}(X = x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}\, \mathbb{I}_{\{0,1,\ldots,n\}}(x)$$

  $E = (\mathcal{X}, \Theta, \mathcal{P})$ é um experimento para θ.

Princípios de Inferência

- Sejam θ, X, Θ, $\mathcal{X}$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ como definidos anteriormente.

- **DEFINICÃO:** Um experimento E para θ é um trio $(X, \theta, \mathcal{P})$, onde X é a quantidade observada segundo $\mathbb{P}_\theta$, $\theta \in \Theta$.

  Notação: $E = (X, \theta, \mathcal{P})$ (alternativamente, $(\mathcal{X}, \Theta, \mathcal{P})$).

- Exemplo 1: Seja $n \in \mathbb{N}^*$. Consideremos $\mathcal{X} = \{0, 1, ..., n\}$, $\Theta = (0, 1)$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in (0, 1)\}$ tal que

  $$\mathbb{P}_\theta(\{x\}) \;=\; \mathbb{P}(X = x|\theta) \;=\; \binom{n}{x}\theta^x(1-\theta)^{n-x}\,\mathbb{I}_{\{0,1,...,n\}}(x)$$

  $E \;=\; (\mathcal{X}, \Theta, \mathcal{P})$ é um experimento para θ.

Princípios de Inferência

- Exemplo 2: Seja $k \in \mathbb{N}^*$. Consideremos os conjuntos $\mathcal{Y} = \{k, k+1, k+2, ...\}$, $\Theta = (0,1)$ e $\mathcal{P}' = \{\mathbb{P}'_\theta : \theta \in (0,1)\}$ tal que

  $$\mathbb{P}'_\theta(\{y\}) \;=\; \mathbb{P}'(Y = y|\theta) \;=\; \binom{y-1}{k-1}\theta^k(1-\theta)^{y-k}\,\mathbb{I}_{\{k,k+1,...\}}(y)$$

  $E' \;=\; (\mathcal{Y}, \Theta, \mathcal{P}')$ é um experimento para θ.

- O conjunto de experimentos para (o mesmo) θ é denotado por $\mathcal{E}_\theta$ (ou $\mathcal{E}_\Theta$).

- Dos exemplos 1 e 2, temos que $E, E' \in \mathcal{E}_\theta$ (admitindo que tais experimentos dizem respeito a mesma quantidade de interesse θ).

Princípios de Inferência

- Exemplo 2: Seja $k \in \mathbb{N}^*$. Consideremos os conjuntos $\mathcal{Y} = \{k, k+1, k+2, ...\}$, $\Theta = (0,1)$ e $\mathcal{P}' = \{\mathbb{P}'_\theta : \theta \in (0,1)\}$ tal que

  $$\mathbb{P}'_\theta(\{y\}) \;=\; \mathbb{P}'(Y = y|\theta) \;=\; \binom{y-1}{k-1}\theta^k(1-\theta)^{y-k}\,\mathbb{I}_{\{k,k+1,...\}}(y)$$

  $E' \;=\; (\mathcal{Y}, \Theta, \mathcal{P}')$ é um experimento para θ.

- O conjunto de experimentos para (o mesmo) θ é denotado por $\mathcal{E}_\theta$ (ou $\mathcal{E}_\Theta$).

- Dos exemplos 1 e 2, temos que $E, E' \in \mathcal{E}_\theta$ (admitindo que tais experimentos dizem respeito a mesma quantidade de interesse θ).

Princípios de Inferência

- Exemplo 2: Seja $k \in \mathbb{N}^*$. Consideremos os conjuntos $\mathcal{Y} = \{k, k+1, k+2, ...\}$, $\Theta = (0,1)$ e $\mathcal{P}' = \{\mathbb{P}'_\theta : \theta \in (0,1)\}$ tal que

$$\mathbb{P}'_\theta(\{y\}) \;=\; \mathbb{P}'(Y = y|\theta) \;=\; \binom{y-1}{k-1}\theta^k(1-\theta)^{y-k}\,\mathbb{I}_{\{k,k+1,...\}}(y)$$

$E' \;=\; (\mathcal{Y}, \Theta, \mathcal{P}')$ é um experimento para θ.

- O conjunto de experimentos para (o mesmo) θ é denotado por $\mathcal{E}_\theta$ (ou $\mathcal{E}_\Theta$).

- Dos exemplos 1 e 2, temos que $E, E' \in \mathcal{E}_\theta$ (admitindo que tais experimentos dizem respeito a mesma quantidade de interesse θ).

Princípios de Inferência

- Exemplo 2: Seja $k \in \mathbb{N}^*$. Consideremos os conjuntos $\mathcal{Y} = \{k, k+1, k+2, ...\}$, $\Theta = (0,1)$ e $\mathcal{P}' = \{\mathbb{P}'_\theta : \theta \in (0,1)\}$ tal que

$$\mathbb{P}'_\theta(\{y\}) \;=\; \mathbb{P}'(Y = y|\theta) \;=\; \binom{y-1}{k-1}\theta^k(1-\theta)^{y-k}\, \mathbb{I}_{\{k,k+1,...\}}(y)$$

$E' \;=\; (\mathcal{Y}, \Theta, \mathcal{P}')$ é um experimento para θ.

- O conjunto de experimentos para (o mesmo) θ é denotado por $\mathcal{E}_\theta$ (ou $\mathcal{E}_\Theta$).

- Dos exemplos 1 e 2, temos que $E, E' \in \mathcal{E}_\theta$ (admitindo que tais experimentos dizem respeito a mesma quantidade de interesse θ).

Princípios de Inferência

- Exemplo 3: Consideremos os conjuntos $\mathcal{X} = \mathbb{N}^n$, $\Theta = \mathbb{R}_+$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta > 0\}$ tal que

$$\begin{aligned} \mathbb{P}_\theta(\{(x_1, ..., x_n)\}) &= \mathbb{P}(X = (x_1, ..., x_n)|\theta) = \\ &= \mathbb{P}(X_1 = x_1, ..., X_n = x_n|\theta) = \\ &= \frac{e^{-n\theta}\ \theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \prod_{i=1}^n \mathbb{I}_{\mathbb{N}}(x_i) \end{aligned}$$

$E = (\mathcal{X}, \Theta, \mathcal{P})$ é um experimento para θ.

- Finalmente, apresentamos a seguir o conceito de experimento misto (aleatorizado).

Princípios de Inferência

- Exemplo 3: Consideremos os conjuntos $\mathcal{X} = \mathbb{N}^n$, $\Theta = \mathbb{R}_+$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta > 0\}$ tal que

$$\mathbb{P}_\theta(\{(x_1, ..., x_n)\}) \;=\; \mathbb{P}(X = (x_1, ..., x_n)|\theta) \;=$$

$$= \;\mathbb{P}(X_1 = x_1, ..., X_n = x_n|\theta) \;=$$

$$= \;\frac{e^{-n\theta}\;\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}\;\prod_{i=1}^n \mathbb{I}_{\mathbb{N}}(x_i)$$

$E \;=\; (\mathcal{X}, \Theta, \mathcal{P})$ é um experimento para θ.

- Finalmente, apresentamos a seguir o conceito de experimento misto (aleatorizado).

Princípios de Inferência

- Exemplo 3: Consideremos os conjuntos $\mathcal{X} = \mathbb{N}^n$, $\Theta = \mathbb{R}_+$ e $\mathcal{P} = \{\mathbb{P}_\theta : \theta > 0\}$ tal que

$$\begin{aligned} \mathbb{P}_\theta(\{(x_1, ..., x_n)\}) &= \mathbb{P}(X = (x_1, ..., x_n)|\theta) = \\ &= \mathbb{P}(X_1 = x_1, ..., X_n = x_n|\theta) = \\ &= \frac{e^{-n\theta} \ \theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \prod_{i=1}^n \mathbb{I}_{\mathbb{N}}(x_i) \end{aligned}$$

  $E = (\mathcal{X}, \Theta, \mathcal{P})$ é um experimento para θ.

- Finalmente, apresentamos a seguir o conceito de experimento misto (aleatorizado).

Princípios de Inferência

- **DEFINICÃO :** Sejam $E_1 = (X_1, \theta, \mathcal{P}_1)$ e $E_2 = (X_2, \theta, \mathcal{P}_2)$ dois experimentos para θ, com $X_j \in \mathcal{X}_j$ e $\mathcal{P}_j = \{\mathbb{P}_\theta^{(j)} : \theta \in \Theta\}$, $j = 1, 2$. Seja J uma variável aleatória Uniforme em $\{1, 2\}$ (independente de X_1 e X_2 e θ).

  O experimento $E^* = (X^*, \theta, \mathcal{P}^*)$, com

  $X^* = \{1\} \times \mathcal{X}_1 \ \cup \ \{2\} \times \mathcal{X}_2$ e

  $\mathcal{P}^*$ satisfazendo

  $$\mathbb{P}_\theta^*(\{(j, x_j)\}) = \mathbb{P}^*(X^* = (j, x_j)|\theta) = \frac{1}{2}\,\mathbb{P}_j(X_j = x_j|\theta)$$

  é chamado experimento misto de E_1 e E_2.

- O experimento E^* consiste na escolha, segundo a variável J, do experimento (E_1 ou E_2) a ser conduzido e, então, registrar uma observação do experimento escolhido.

Princípios de Inferência

- **DEFINICÃO :** Sejam $E_1 = (X_1, \theta, \mathcal{P}_1)$ e $E_2 = (X_2, \theta, \mathcal{P}_2)$ dois experimentos para θ, com $X_j \in \mathcal{X}_j$ e $\mathcal{P}_j = \{\mathbb{P}_\theta^{(j)} : \theta \in \Theta\}$, $j = 1, 2$. Seja J uma variável aleatória Uniforme em $\{1, 2\}$ (independente de X_1 e X_2 e θ).

  O experimento $E^* = (X^*, \theta, \mathcal{P}^*)$, com

  $X^* = \{1\} \times \mathcal{X}_1 \ \cup \ \{2\} \times \mathcal{X}_2$ e

  $\mathcal{P}^*$ satisfazendo

  $$\mathbb{P}_\theta^*(\{(j, x_j)\}) = \mathbb{P}^*(X^* = (j, x_j)|\theta) = \frac{1}{2}\,\mathbb{P}_j(X_j = x_j|\theta)$$

  é chamado experimento misto de E_1 e E_2.

- O experimento E^* consiste na escolha, segundo a variável J, do experimento (E_1 ou E_2) a ser conduzido e, então, registrar uma observação do experimento escolhido.

Princípios de Inferência

- **DEFINICÃO :** Sejam $E_1 = (X_1, \theta, \mathcal{P}_1)$ e $E_2 = (X_2, \theta, \mathcal{P}_2)$ dois experimentos para θ, com $X_j \in \mathcal{X}_j$ e $\mathcal{P}_j = \{\mathbb{P}_\theta^{(j)} : \theta \in \Theta\}$, $j = 1, 2$. Seja J uma variável aleatória Uniforme em $\{1, 2\}$ (independente de X_1 e X_2 e θ).

  O experimento $E^* = (X^*, \theta, \mathcal{P}^*)$, com

  $X^* = \{1\} \times \mathcal{X}_1 \;\cup\; \{2\} \times \mathcal{X}_2$ e

  $\mathcal{P}^*$ satisfazendo

  $$\mathbb{P}_\theta^*(\{(j, x_j)\}) = \mathbb{P}^*(X^* = (j, x_j)|\theta) = \frac{1}{2}\, \mathbb{P}_j(X_j = x_j|\theta)$$

  é chamado experimento misto de E_1 e E_2.

- O experimento E^* consiste na escolha, segundo a variável J, do experimento (E_1 ou E_2) a ser conduzido e, então, registrar uma observação do experimento escolhido.

Princípios de Inferência

- Exemplo 4: Retomando os exemplos 1) e 2), onde $\Theta = (0,1)$, consideremos o experimento misto $E^* = (X^*, \theta, \mathcal{P}^*)$, com

  $$X^* = \{1\} \times \{0,1,...,n\} \;\cup\; \{2\} \times \{k, k+1, k+2, ...\} \quad \text{e}$$

  $\mathcal{P}^*$ satisfazendo

  $$\mathbb{P}^*_\theta(\{(1,x)\}) \;=\; \mathbb{P}^*(X^* = (1,x)|\theta) \;=\; \tfrac{1}{2}\,\mathbb{P}(X = x|\theta) \quad \text{e}$$

  $$\mathbb{P}^*_\theta(\{(2,y)\}) \;=\; \mathbb{P}^*(X^* = (2,y)|\theta) \;=\; \tfrac{1}{2}\,\mathbb{P}'(Y = y|\theta)$$

  O experimento E^* consiste em escolher qual dos experimentos deve ser conduzido, E (observação de $X|\theta \sim Bin(n,\theta)$) ou E' (observação de $Y|\theta \sim BN(k,\theta)$), e então registrar o valor da variável a ser observada segundo o experimento escolhido (X ou Y).

Princípios de Inferência

- Exemplo 4: Retomando os exemplos 1) e 2), onde $\Theta = (0, 1)$, consideremos o experimento misto $E^* = (X^*, \theta, \mathcal{P}^*)$, com

  $$X^* = \{1\} \times \{0, 1, ..., n\} \;\; \cup \;\; \{2\} \times \{k, k+1, k+2, ...\} \quad \text{e}$$

  $\mathcal{P}^*$ satisfazendo

  $$\mathbb{P}^*_\theta(\{(1, x)\}) = \mathbb{P}^*(X^* = (1, x)|\theta) = \tfrac{1}{2}\,\mathbb{P}(X = x|\theta) \quad \text{e}$$

  $$\mathbb{P}^*_\theta(\{(2, y)\}) = \mathbb{P}^*(X^* = (2, y)|\theta) = \tfrac{1}{2}\,\mathbb{P}'(Y = y|\theta)$$

  O experimento E^* consiste em escolher qual dos experimentos deve ser conduzido, E (observação de $X|\theta \sim Bin(n, \theta)$) ou E' (observação de $Y|\theta \sim BN(k, \theta)$), e então registrar o valor da variável a ser observada segundo o experimento escolhido (X ou Y).

Princípios de Inferência

- A partir da condução do experimento E e da observação de x, podemos inferir, tirar conclusões sobre θ.Tal inferência, conclusão sobre θ é denominada (Birnbaum (1962)) **Evidência sobre θ a partir de E e x, $Ev(e, x)$.**
- Afinal, qual é o significado de $Ev(E, x)$?
- Em princípio, não há restrição sobre $Ev(E, x)$: estimativa, decisão sobre hipóteses, etc.. (não precisa ser restrita aos problemas tradicionais de inferência estatística). Ênfase da abordagem é sobre **como** $Ev(E, x)$ **deve depender de E e x.**
- Críticas sobre a formulação original; revisão da formulação (Birnbaum (1972), dentre outros); introdução de relações de equivalência em $\mathcal{R}_\theta = \cup_{E \in \mathcal{E}_\theta} \{E\} \times \mathcal{X}_E$ (conjunto de todas as realizações de experimentos da forma (E, x)).

Princípios de Inferência

- A partir da condução do experimento E e da observação de x, podemos inferir, tirar conclusões sobre θ.Tal inferência, conclusão sobre θ é denominada (Birnbaum (1962)) **Evidência sobre θ a partir de E e x, $Ev(e, x)$.**
- Afinal, qual é o significado de $Ev(E, x)$?
- Em princípio, não há restrição sobre $Ev(E, x)$: estimativa, decisão sobre hipóteses, etc.. (não precisa ser restrita aos problemas tradicionais de inferência estatística). Ênfase da abordagem é sobre **como** $Ev(E, x)$ **deve depender de** E **e** x.
- Críticas sobre a formulação original; revisão da formulação (Birnbaum (1972), dentre outros); introdução de relações de equivalência em $\mathcal{R}_\theta = \cup_{E \in \mathcal{E}_\theta} \{E\} \times \mathcal{X}_E$ (conjunto de todas as realizações de experimentos da forma (E, x)).

Princípios de Inferência

- A partir da condução do experimento E e da observação de x, podemos inferir, tirar conclusões sobre θ.Tal inferência, conclusão sobre θ é denominada (Birnbaum (1962)) **Evidência sobre θ a partir de E e x,** $Ev(e, x)$.
- Afinal, qual é o significado de $Ev(E, x)$?
- Em princípio, não há restrição sobre $Ev(E, x)$: estimativa, decisão sobre hipóteses, etc.. (não precisa ser restrita aos problemas tradicionais de inferência estatística). Ênfase da abordagem é sobre **como** $Ev(E, x)$ **deve depender de** E **e** x.
- Críticas sobre a formulação original; revisão da formulação (Birnbaum (1972), dentre outros); introdução de relações de equivalência em $\mathcal{R}_\theta = \cup_{E \in \mathcal{E}_\theta} \{E\} \times \mathcal{X}_E$ (conjunto de todas as realizações de experimentos da forma (E, x)).

Princípios de Inferência

- A partir da condução do experimento E e da observação de x, podemos inferir, tirar conclusões sobre θ.Tal inferência, conclusão sobre θ é denominada (Birnbaum (1962)) **Evidência sobre θ a partir de E e x**, $Ev(e, x)$.

- Afinal, qual é o significado de $Ev(E, x)$?

- Em princípio, não há restrição sobre $Ev(E, x)$: estimativa, decisão sobre hipóteses, etc.. (não precisa ser restrita aos problemas tradicionais de inferência estatística). Ênfase da abordagem é sobre **como** $Ev(E, x)$ **deve depender de** E **e** x.

- Críticas sobre a formulação original; revisão da formulação (Birnbaum (1972), dentre outros); introdução de relações de equivalência em $\mathcal{R}_\theta = \cup_{E \in \mathcal{E}_\theta} \{E\} \times \mathcal{X}_E$ (conjunto de todas as realizações de experimentos da forma (E, x)).

Princípios de Inferência

- A partir da condução do experimento E e da observação de x, podemos inferir, tirar conclusões sobre θ.Tal inferência, conclusão sobre θ é denominada (Birnbaum (1962)) **Evidência sobre θ a partir de E e x**, $Ev(e, x)$.
- Afinal, qual é o significado de $Ev(E, x)$?
- Em princípio, não há restrição sobre $Ev(E, x)$: estimativa, decisão sobre hipóteses, etc.. (não precisa ser restrita aos problemas tradicionais de inferência estatística). Ênfase da abordagem é sobre **como $Ev(E, x)$ deve depender de E e x**.
- Críticas sobre a formulação original; revisão da formulação (Birnbaum (1972), dentre outros); introdução de relações de equivalência em $\mathcal{R}_\theta = \cup_{E \in \mathcal{E}_\theta} \{E\} \times \mathcal{X}_E$ (conjunto de todas as realizações de experimentos da forma (E, x)).

Princípios de Inferência

- Exemplo 1: Consideremos o exemplo da aula passada em que $\Theta = (0,1)$ e os experimentos $E_1 = (X, \theta, \mathcal{P}_1)$ e $E_2 = (Y, \theta, \mathcal{P}_2)$, com $X|\theta \sim Bin(10, \theta)$ e $Y|\theta \sim BN(2, \theta)$. Consideremos as seguintes evidências sobre θ: para $R = (E, x) \in \mathcal{R}_\theta$.

  (i) $Ev_1(E, x) = argmax\ V_x^{(E)}(\cdot)$

  Assim, $Ev_1(E_1, 4) = \frac{4}{10}$ e $Ev_1(E_2, 9) = \frac{2}{9}$

  (ii) Suponha que exista uma distribuição (a priori) contínua para θ (com densidade f). Definimos $Ev_2(E, x) = f_E(\cdot|x)$, com $f_E(\theta|x) = \frac{V_x^{(E)}(\theta)\ f(\theta)}{\int_\Theta V_x^{(E)}(\theta')\ f(\theta')d\theta'}$

  Assim, se $f(\theta) = \mathbb{I}_{(0,1)}(\theta)$, $Ev_2(E_1, 10) = 11\ \theta^{10}\ \mathbb{I}_{(0,1)}(\theta)$ e

  $Ev_2(E_2, 6) = 105\ \theta^2\ (1-\theta)^4\ \mathbb{I}_{(0,1)}(\theta)$.

Princípios de Inferência

- Exemplo 1: Consideremos o exemplo da aula passada em que $\Theta = (0,1)$ e os experimentos $E_1 = (X, \theta, \mathcal{P}_1)$ e $E_2 = (Y, \theta, \mathcal{P}_2)$, com $X|\theta \sim Bin(10, \theta)$ e $Y|\theta \sim BN(2, \theta)$. Consideremos as seguintes evidências sobre θ: para $R = (E, x) \in \mathcal{R}_\theta$.

  (i) $Ev_1(E, x) = argmax\ V_x^{(E)}(\cdot)$

  Assim, $Ev_1(E_1, 4) = \frac{4}{10}$ e $Ev_1(E_2, 9) = \frac{2}{9}$

  (ii) Suponha que exista uma distribuição (a priori) contínua para θ (com densidade f). Definimos $Ev_2(E, x) = f_E(\cdot|x)$, com $f_E(\theta|x) = \frac{V_x^{(E)}(\theta)\ f(\theta)}{\int_\Theta V_x^{(E)}(\theta')\ f(\theta') d\theta'}$

  Assim, se $f(\theta) = \mathbb{I}_{(0,1)}(\theta)$, $Ev_2(E_1, 10) = 11\ \theta^{10}\ \mathbb{I}_{(0,1)}(\theta)$ e

  $Ev_2(E_2, 6) = 105\ \theta^2\ (1-\theta)^4\ \mathbb{I}_{(0,1)}(\theta)$.

Princípios de Inferência

- (iii) $Ev_3(E,x) = \delta_{NV}^{(E)}(x)$, onde $\delta_{NV}^{(E)} : \mathcal{X}_E \to \Theta$ é estimador não-viesado ótimo sob E.

Assim, $Ev_3(E_1,6) = \frac{6}{10}$ e $Ev_1(E_2,21) = \frac{1}{20}$

Nesse exemplo, podemos definir as relações de equivalência $\cong_i \subset \mathcal{R}_\theta \times \mathcal{R}_\theta$, $i = 1,2,3$, por

$(E,x) \cong_i (E',x') \Leftrightarrow Ev_i(E,x) = Ev_i(E',x')$

Assim,

$(E_1,2) \cong_1 (E_2,10)$, $(E_1,5) \cong_1 (E_2,4)$, mas

$(E_1,2) \ncong_1 (E_1,7)$. Além disso,

$(E_1,2) \cong_2 (E_2,10)$, mas $(E_1,5) \ncong_2 (E_2,4)$. Por fim,

$(E_1,2) \ncong_3 (E_2,10)$ e mas $(E_1,1) \cong_3 (E_2,11)$

Princípios de Inferência

- (iii) $Ev_3(E,x) = \delta_{NV}^{(E)}(x)$, onde $\delta_{NV}^{(E)} : \mathcal{X}_E \to \Theta$ é estimador não-viesado ótimo sob E.

  Assim, $Ev_3(E_1, 6) = \frac{6}{10}$ e $Ev_1(E_2, 21) = \frac{1}{20}$

  Nesse exemplo, podemos definir as relações de equivalência $\cong_i \subset \mathcal{R}_\theta \times \mathcal{R}_\theta$, $i = 1, 2, 3$, por

  $(E,x) \cong_i (E',x') \Leftrightarrow Ev_i(E,x) = Ev_i(E',x')$

  Assim,

  $(E_1, 2) \cong_1 (E_2, 10)$, $(E_1, 5) \cong_1 (E_2, 4)$, mas

  $(E_1, 2) \ncong_1 (E_1, 7)$. Além disso,

  $(E_1, 2) \cong_2 (E_2, 10)$, mas $(E_1, 5) \ncong_2 (E_2, 4)$. Por fim,

  $(E_1, 2) \ncong_3 (E_2, 10)$ e mas $(E_1, 1) \cong_3 (E_2, 11)$

Princípios de Inferência

- **PRINCÍPIO DA SUFICIÊNCIA (WSP)**

  Considere o experimento $E = (X, \theta, \mathcal{P}) = (\mathcal{X}, \Theta, \mathcal{P})$ para θ. Seja $T : \mathcal{X} \to \mathcal{T}$ uma estatística suficiente para θ. Então, para $x_1, x_2 \in \mathcal{X}$,

  $$T(x_1) = T(x_2) \;\Rightarrow\; Ev(E, x_1) = Ev(E, x_2)$$

- Dois pontos amostrais que levam a um mesmo valor de uma estatística suficiente T para θ devem produzir inferências idênticas (coincidentes) sobre θ.
- "... the Weak Sufficiency Principle is, itself, a cornerstone of classical statistics" (Berger and Wolpert (1988)).

Princípios de Inferência

- **PRINCÍPIO DA SUFICIÊNCIA (WSP)**

  Considere o experimento $E = (X, \theta, \mathcal{P}) = (\mathcal{X}, \Theta, \mathcal{P})$ para θ. Seja $T : \mathcal{X} \to \mathcal{T}$ uma estatística suficiente para θ. Então, para $x_1, x_2 \in \mathcal{X}$,

$$T(x_1) \ = \ T(x_2) \ \ \Rightarrow \ \ Ev(E, x_1) \ = \ Ev(E, x_2)$$

- Dois pontos amostrais que levam a um mesmo valor de uma estatística suficiente T para θ devem produzir inferências idênticas (coincidentes) sobre θ.

- "... the Weak Sufficiency Principle is, itself, a cornerstone of classical statistics" (Berger and Wolpert (1988)).

Princípios de Inferência

- **PRINCÍPIO DA SUFICIÊNCIA (WSP)**

  Considere o experimento $E = (X, \theta, \mathcal{P}) = (\mathcal{X}, \Theta, \mathcal{P})$ para θ. Seja $T : \mathcal{X} \to \mathcal{T}$ uma estatística suficiente para θ. Então, para $x_1, x_2 \in \mathcal{X}$,

$$T(x_1) = T(x_2) \Rightarrow Ev(E, x_1) = Ev(E, x_2)$$

- Dois pontos amostrais que levam a um mesmo valor de uma estatística suficiente T para θ devem produzir inferências idênticas (coincidentes) sobre θ.

- "... the Weak Sufficiency Principle is, itself, a cornerstone of classical statistics" (Berger and Wolpert (1988)).

Princípios de Inferência

- **PRINCÍPIO DA SUFICIÊNCIA (WSP)**

  Considere o experimento $E = (X, \theta, \mathcal{P}) = (\mathcal{X}, \Theta, \mathcal{P})$ para θ. Seja $T : \mathcal{X} \to \mathcal{T}$ uma estatística suficiente para θ. Então, para $x_1, x_2 \in \mathcal{X}$,

$$T(x_1) \;=\; T(x_2) \;\;\Rightarrow\;\; Ev(E, x_1) \;=\; Ev(E, x_2)$$

- Dois pontos amostrais que levam a um mesmo valor de uma estatística suficiente T para θ devem produzir inferências idênticas (coincidentes) sobre θ.
- "... the Weak Sufficiency Principle is, itself, a cornerstone of classical statistics" (Berger and Wolpert (1988)).

Princípios de Inferência

- Exemplo: Consideremos o experimento $E = (\mathcal{X}, \Theta, \mathcal{P})$, $\mathcal{X}$ enumerável, tal que, para todo $x \in \mathcal{X}$, $V_x : \Theta \to \mathbb{R}_+$ é limitada. Seja T uma estatística suficiente para θ. Consideremos a evidência sobre θ dada por

  $Ev(E, x) = argmax\, V_x(\cdot)$.

  Sejam $x_1, x_2 \in \mathcal{X}$ tais que $T(x_1) = T(x_2)$. Assim,

  $$Ev(E, x_1) \ = \ argmax\, V_{x_1}(\cdot) \ = \ argmax\, a(x_1)\, b(T(x_1), \cdot) \ =$$

  $$= \ argmax\, b(T(x_1), \cdot) \ = \ argmax\, b(T(x_2), \cdot) \ =$$

  $$= \ argmax\, a(x_2)\, b(T(x_2), \cdot) \ = \ argmax\, V_{x_2}(\cdot) \ = \ Ev(E, x_2)$$

  Logo, $T(x_1) \ = \ T(x_2) \ \Rightarrow \ Ev(E, x_1) \ = \ Ev(E, x_2)$.

Princípios de Inferência

- Exemplo: Consideremos o experimento $E = (\mathcal{X}, \Theta, \mathcal{P})$, $\mathcal{X}$ enumerável, tal que, para todo $x \in \mathcal{X}$, $V_x : \Theta \to \mathbb{R}_+$ é limitada. Seja T uma estatística suficiente para θ. Consideremos a evidência sobre θ dada por

  $$Ev(E, x) = argmax\ V_x(\cdot).$$

  Sejam $x_1, x_2 \in \mathcal{X}$ tais que $T(x_1) = T(x_2)$. Assim,

  $$\begin{aligned} Ev(E, x_1) &= argmax\, V_{x_1}(\cdot) = argmax\, a(x_1)\, b(T(x_1), \cdot) = \\ &= argmax\, b(T(x_1), \cdot) = argmax\, b(T(x_2), \cdot) = \\ &= argmax\, a(x_2)\, b(T(x_2), \cdot) = argmax\, V_{x_2}(\cdot) = Ev(E, x_2) \end{aligned}$$

  Logo, $T(x_1) = T(x_2) \Rightarrow Ev(E, x_1) = Ev(E, x_2)$.

Princípios de Inferência

- Exemplo: Consideremos o experimento $E = (\mathcal{X}, \Theta, \mathcal{P})$, $\mathcal{X}$ enumerável. Suponhamos que exista uma distribuição de probabilidade (a priori) contínua para θ com densidade f. Seja T uma estatística suficiente para θ. Consideremos a seguinte evidência sobre θ: $Ev(E, x) = f_E(\cdot|x)$, onde

$$f_E(\theta|x) = f_E(\theta|X = x) = \frac{V_x^{(E)}(\theta)\ f(\theta)}{\int_\Theta V_x^{(E)}(\theta')\ f(\theta')d\theta'} \propto V_x^{(E)}(\theta)\ f(\theta)$$

Sejam $x_1, x_2 \in \mathcal{X}$ tais que $T(x_1) = T(x_2)$. Da definição de Ev e de suficiência (bayesiana), temos que

$$Ev(E, x_1) = f_E(\cdot|X = x_1) = f_E(\cdot|T(X) = T(x_1)) =$$

$$= f_E(\cdot|T(X) = T(x_2)) = f_E(\cdot|X = x_2) = Ev(E, x_2)$$

Logo, $T(x_1) = T(x_2) \Rightarrow Ev(E, x_1) = Ev(E, x_2)$.

Princípios de Inferência

- Exemplo: Consideremos o experimento $E = (\mathcal{X}, \Theta, \mathcal{P})$, $\mathcal{X}$ enumerável. Suponhamos que exista uma distribuição de probabilidade (a priori) contínua para θ com densidade f. Seja T uma estatística suficiente para θ. Consideremos a seguinte evidência sobre θ: $Ev(E, x) = f_E(\cdot|x)$, onde

$$f_E(\theta|x) = f_E(\theta|X = x) = \frac{V_x^{(E)}(\theta)\ f(\theta)}{\int_\Theta V_x^{(E)}(\theta')\ f(\theta')d\theta'} \propto V_x^{(E)}(\theta)\ f(\theta)$$

Sejam $x_1, x_2 \in \mathcal{X}$ tais que $T(x_1) = T(x_2)$. Da definição de Ev e de suficiência (bayesiana), temos que

$$Ev(E, x_1) = f_E(\cdot|X = x_1) = f_E(\cdot|T(X) = T(x_1)) =$$

$$= f_E(\cdot|T(X) = T(x_2)) = f_E(\cdot|X = x_2) = Ev(E, x_2)$$

Logo, $T(x_1) = T(x_2) \Rightarrow Ev(E, x_1) = Ev(E, x_2)$.

Princípios de Inferência

- Exemplo: Consideremos o experimento $E = (\mathcal{X}, \Theta, \mathcal{P})$, $\mathcal{X}$ enumerável. Seja T uma estatística suficiente e completa. Suponhamos que exista estimador não-viesado para θ, $\delta_0 : \mathcal{X} \to \Theta$. Seja $\delta^*_E(X) = \mathbb{E}_E(\delta_0(X)|T(X)) = g_E(T(X))$ o ENVVUM para θ. Consideremos a seguinte evidência sobre θ a partir de E e x:

$$Ev(E, x) = \delta^*_E(x) = g_E(T(x))$$

Sejam $x_1, x_2 \in \mathcal{X}$ tais que $T(x_1) = T(x_2)$. Então,

$$Ev(E, x_1) = \delta^*_E(x_1) = g_E(T(x_1)) = g_E(T(x_2)) =$$

$$= \delta^*_E(x_2) = Ev(E, x_2)$$

Logo, $T(x_1) = T(x_2) \Rightarrow Ev(E, x_1) = Ev(E, x_2)$.

Princípios de Inferência

- Exemplo: Consideremos o experimento $E = (\mathcal{X}, \Theta, \mathcal{P})$, $\mathcal{X}$ enumerável. Seja T uma estatística suficiente e completa. Suponhamos que exista estimador não-viesado para θ, $\delta_0 : \mathcal{X} \to \Theta$. Seja $\delta_E^*(X) = \mathbb{E}_E(\delta_0(X)|T(X)) = g_E(T(X))$ o ENVVUM para θ. Consideremos a seguinte evidência sobre θ a partir de E e x:

  $$Ev(E, x) = \delta_E^*(x) = g_E(T(x))$$

  Sejam $x_1, x_2 \in \mathcal{X}$ tais que $T(x_1) = T(x_2)$. Então,

  $$Ev(E, x_1) = \delta_E^*(x_1) = g_E(T(x_1)) = g_E(T(x_2)) =$$

  $$= \delta_E^*(x_2) = Ev(E, x_2)$$

  Logo, $T(x_1) = T(x_2) \Rightarrow Ev(E, x_1) = Ev(E, x_2)$.