Inferência Estatística

Introdução

E.F.T[1]

[1]EACH-USP
Universidade de São Paulo

ACH2053

Outline

Outline

Natureza da Inferência Estatística

Qual a principal utilidade

Um problema de inferência estatística é um problema no qual dados que tem sido gerados de acordo com alguma distribuição de probabilidade desconhecida devem ser analisados e algum tipo de inferência sobre a distribuição desconhecida deve ser feita.

- Inferência sobre a `distribuição`.
- Inferência sobre os parâmetros da distribuição (se for conhecida a distribuição).

Parâmetros

Definição de parâmetro

Em um problema de Inferência Estatística, qualquer característica da distribuição que gerou os dados experimentais que tem um valor desconhecido (como a média μ ou desvio padrão σ é chamado de parâmetro da distribuição. O conjunto Ω de todos os possíveis valores de um parâmetro θ ou de vetor de parâmetros $(\theta_1, ..., \theta_k)$ é chamado de espaço paramétrico. Notações . . .

- Parâmetros
  - θ
  - $(\theta_1, ..., \theta_n)$ vetor de parâmetros
- Espaço paramétrico:
  - Ω.

Parâmetros

Definição de parâmetro

Em um problema de Inferência Estatística, qualquer característica da distribuição que gerou os dados experimentais que tem um valor desconhecido (como a média μ ou desvio padrão σ é chamado de parâmetro da distribuição. O conjunto Ω de todos os possíveis valores de um parâmetro θ ou de vetor de parâmetros $(\theta_1, ..., \theta_k)$ é chamado de espaço paramétrico. Notações . . .

- Parâmetros
    - θ
    - $(\theta_1, ..., \theta_n)$ vetor de parâmetros
- Espaço paramétrico:
    - Ω.

Parâmetros

Definição de parâmetro

Em um problema de Inferência Estatística, qualquer característica da distribuição que gerou os dados experimentais que tem um valor desconhecido (como a média μ ou desvio padrão σ é chamado de parâmetro da distribuição. O conjunto Ω de todos os possíveis valores de um parâmetro θ ou de vetor de parâmetros $(\theta_1, ..., \theta_k)$ é chamado de espaço paramétrico. Notações ...

- Parâmetros
  - θ
  - $(\theta_1, ..., \theta_n)$ vetor de parâmetros
- Espaço paramétrico:
  - Ω.

Outline

Distribuições a priori

Suponha um problema de inferência estatística em que as observações são tomadas de uma distribuição $f(x|\theta)$, onde θ é um parâmetro com valor desconhecido. Assumimos que o valor de θ deve pertencer a um espaço paramétrico especificado Ω. O problema de I.E. consiste em determinar na base das observações de $f(x|\theta)$ em que lugar de Ω se encontra atualmente θ.

Distribuições a priori

Em muitos problemas, antes da $f(x|\theta)$ gerar as observações, o analista será capaz de sumarizar (experiência anterior) sua informação prévia sobre em que lugar de Ω se encontra θ com maior probabilidade, sendo possível construir uma distribuição de probabilidade para θ no conjunto Ω. Esta distribuição é chamada de distribuição a priori de θ.

Distrib. a priori discretas e continuas

Se o parâmetro θ tomar um número de valores finito ou no máximo uma sequência contável infinita, então $\xi(\theta)$ será uma f.p. a priori. Se θ tomar valores num intervalo da reta, então $\xi(\theta)$ será uma f.d.p a priori de θ.

Exemplo: Moeda honesta ou de duas caras

Seja θ a probabilidade de obter cara quando certa moeda é lançada. Suponha que a moeda pode ser honesta ou ter duas caras. Portanto, os únicos valores possíveis para θ são $\theta = 1/2$ ou $\theta = 1$. Se a probabilidade a priori que a moeda é honesta é *p*, então a f.p a priori deθ é $\xi(1/2) = p$ e $\xi(1) = 1 - p$.

Exemplo: Exemplo: Proporção de itens defeituosos

Suponha que a proporção θ de itens defeituosos em um lote grande é desconhecido, e que a fdp a priori atribuida a θ é uma distribuição uniforme em (0,1). Então a fdp a priori de θ é :

$$\xi(\theta) = \begin{cases} 1 & \text{para} 0 < \theta < 1. \\ 0 & c.c. \end{cases}$$

Outline

1. Introdução à Inferência Estatística
   - Conceitos Gerais de Inferência Estatística
   - Distribuições a priori
   - Distribuições a posteriori
   - A função de Verossimilhança
   - Observações Sequenciais

2. Distribuições Conjugadas
   - Distribuições Conjugadas

Distribuições a posteriori

Suponha que n v.a. $X_1, X_2, ..., X_n$ formam uma a.a. de uma distribuição com $f(x|\theta)$. Suponha também que o valor do parâmetro θ é desconhecido e que a distribuição a priori de θ é $\xi(\theta)$. Assumimos que o espaço paramétrico é Ω.
Como as v.a. $X_1, X_2, ..., X_n$ formam uma a.a. da distribuição com $f(x|\theta)$, então a distribuição conjunta será dada por:

$$f_n(x_1, x_2, ..., x_n|\theta) = f(x_1|\theta)...f(x_n|\theta) \tag{1}$$

Distribuições a posteriori

Se denotarmos $\mathbf{x} = (x_1, x_2, ..., x_n)$, a equação 1 pode ser escrita como $f_n(\mathbf{x}|\theta)$.
O parâmetro $\theta \sim \xi(\theta)$, e a função de probabilidade(densidade) $f_n(\mathbf{x}|\theta)$, dever então ser a distribuição condicional de $X_1, X_2, ..., X_n$ para um dado valor θ. Multiplicando $f_n(\mathbf{x}|\theta)$ e $\xi(\theta)$ obtemos a conjunta $n+1$-dimensional de $X_1, X_2, ..., X_n$ e θ, na forma $f_n(\mathbf{x}|\theta)\xi(\theta)$. Assim, a marginal de $X_1, X_2, ..., X_n$ pode ser obtido integrando a conjunta sobre todos os valores de θ:

$$g_n(\mathbf{x}) = \int_{\Omega} f_n(\mathbf{x}|\theta)\xi(\theta)d(\theta) \quad (2)$$

Distribuições a posteriori

Também, a distribuição condicional de θ dado que $X_1 = x_1, ..., X_n = x_n$ que denotaremos por $\xi(\theta|\mathbf{x})$, deve ser igual à conjunta de $X_1, X_2, ..., X_n$ e θ dividida pela marginal de $X_1, X_2, ..., X_n$. Assim, temos:

$$\xi(\theta|\mathbf{x}) = \frac{f_n(\mathbf{x}|\theta)\xi(\theta)}{g_n(\mathbf{x})}, \quad \text{para} \quad \theta \in \Omega \tag{3}$$

A equação 3 é chamada a distribuição a posteriori de θ pois é a distribuição de θ após os valores de $X_1, X_2, ..., X_n$ terem sido observados.

Distribuições a posteriori

Podemos dizer que a priori $\xi(\theta)$ representa a verossimilhança relativa que o valor verdadeiro de θ está em alguma região de Ω antes de que os valores de $X_1, X_2, ..., X_n$ sejam observados, e que a posteriori $\xi(\theta|\mathbf{x})$ representa esta verossimilhança relativa após os valores $X_1 = x_1, ..., X_n = x_n$ terem sido observados.

Outline

A função de Verossimilhança

O denominador do lado direito da equação 3, é simplesmente a integral do numerador sobre todos os possíveis valores de θ. Este valor, depende então dos valores observados $x_1, ..., x_n$ e não de θ e pode ser tratado como uma constante quando o lado direito da equação 3 é uma distribuição em θ. Temos portanto a relação:

$$\xi(\theta|\mathbf{x})\alpha f_n(\mathbf{x}|\theta)\xi(\theta) \tag{4}$$

Quando a conjunta $f_n(\mathbf{x}|\theta)$ das observações numa a.a. é considerada como uma função de θ para valores dados de $x_1, ..., x_n$ é chamado de *função de verossimilhança*. Então o produto na equação 4 estabelece que a posteriori é proporcional ao produto da função de verossimilhança e a priori de θ.

Exemplo: proporção de itens defeituosos

Suponha que a proporção θ de itens defeituosos de um lote (grande) de peças é desconhecida e que a priori de θ é uma uniforme no intervalo $(0, 1)$. Suponha também que uma a.a. de n itens é tomado do lote, e para $i = 1, ..., n$ seja $X_i = 1$ se o i-ésimo item selecionado é defeituso e $X_i = 0$ caso contrário. Então, $X_1, X_2, ..., X_n$ forma n ensaios de Bernoulli com parâmetro θ. Determine a posteriori de θ.
Podemos ver que a distribuição para cada observação X_i é:

$$f(x|\theta) = \begin{cases} \theta^x(1-\theta)^{1-x} & para \quad x = 0, 1 \\ 0 & c.c \end{cases} \tag{5}$$

Exemplo: proporção de itens defeituosos

Seja $y = \sum_{i=1}^{n} x_i$, então a densidade conjunta de $X_1, X_2, ..., X_n$ pode ser escrita para $x_i = 0$ ou $x_i = 1$:

$$f_n(\mathbf{x}|\theta) = \theta^y (1-\theta)^{n-y}$$

como a priori $xi(\theta)$ é dado por

$$\xi(\theta) = \begin{cases} 1 & \text{para} 0 < \theta < 1. \\ 0 & c.c. \end{cases}$$

então, para $0 < \theta < 1$:

$$f_n(\mathbf{x}|\theta) = \theta^y (1-\theta)^{n-y}$$

vemos que exceto uma constante, a posteriori tem a forma de uma distribuição Beta com parâmetros $\alpha = y + 1$ e $\beta = n - y + 1$.

Outline

Observações Sequenciais

Em muitos experimentos, as observações $X_1, X_2, ..., X_n$ que formam uma a.a. devem ser obtidos sequencialmente, isto é, um por vez. Em tais experimentos, o valor de X_1 é observado primeiro, depóis é observado X_2, e assim sucessivamente. Suponha que a priori de θ é $\xi(\theta)$. Depois que o valor de X_1 for observado, a posteriori $\xi(\theta|x_1)$ pode ser calculado segundo a relação:

$$\xi(\theta|x_1)\alpha f(x_1|\theta)\xi(\theta)$$

Observações Sequenciais

Esta posteriori servirá como priori de θ quando o valor de X_2 seja calculado. Assim, após o valor x_2 de X_2 for observado, a posteriori $\xi(\theta|x_1, x_2)$ pode ser calculado da relação:

$$\xi(\theta|x_1, x_2)\alpha f(x_2|\theta)\xi(\theta|x_1)$$

Se continuarmos neste cálculo, teremos a seguinte relação:

$$\xi(\theta|\mathbf{x})\alpha f(x_n|\theta)\xi(\theta|x_1, ...x_{n-1})$$

Outline

1. Introdução à Inferência Estatística
   - Conceitos Gerais de Inferência Estatística
   - Distribuições a priori
   - Distribuições a posteriori
   - A função de Verossimilhança
   - Observações Sequenciais

2. Distribuições Conjugadas
   - Distribuições Conjugadas

Distribuições conjugadas

Algumas distribuições a priori são particularmente convenientes de serem usadas com amostras de outras distribuições.

- Amostragem de uma distribuição de Bernoulli
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Bernoulli com
  - θ desconhecido $(0 < \theta < 1)$
  - a priori de θ é uma Beta com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Beta com parâmetros $\alpha + \sum_{i=1}^{n} x_i$ e $\beta + n - \sum_{i=1}^{n} x_i$

Distribuições conjugadas

Algumas distribuições a priori são particularmente convenientes de serem usadas com amostras de outras distribuições.

- Amostragem de uma distribuição de Bernoulli
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Bernoulli com
  - θ desconhecido $(0 < \theta < 1)$
  - a priori de θ é uma Beta com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Beta com parâmetros $\alpha + \sum_{i=1}^{n} x_i$ e $\beta + n - \sum_{i=1}^{n} x_i$

Distribuições conjugadas

Algumas distribuições a priori são particularmente convenientes de serem usadas com amostras de outras distribuições.

- Amostragem de uma distribuição de Bernoulli
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Bernoulli com
  - θ desconhecido $(0 < \theta < 1)$
  - a priori de θ é uma Beta com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Beta com parâmetros $\alpha + \sum_{i=1}^{n} x_i$ e $\beta + n - \sum_{i=1}^{n} x_i$

Distribuições conjugadas

Algumas distribuições a priori são particularmente convenientes de serem usadas com amostras de outras distribuições.

- Amostragem de uma distribuição de Bernoulli
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Bernoulli com
  - θ desconhecido $(0 < \theta < 1)$
  - a priori de θ é uma Beta com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Beta com parâmetros $\alpha + \sum_{i=1}^{n} x_i$ e $\beta + n - \sum_{i=1}^{n} x_i$

Distribuições conjugadas

- Amostragem de uma distribuição de Poisson
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Poisson com
  - θ desconhecido $(0 < \theta)$
  - a priori de θ é uma Gamma com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Beta com parâmetros $\alpha + \sum_{i=1}^{n} x_i$ e $\beta + n$

Distribuições conjugadas

- Amostragem de uma distribuição de Poisson
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Poisson com
  - θ desconhecido $(0 < \theta)$
  - a priori de θ é uma Gamma com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Beta com parâmetros $\alpha + \sum_{i=1}^{n} x_i$ e $\beta + n$

Distribuições conjugadas

- Amostragem de uma distribuição de Poisson
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Poisson com
  - θ desconhecido ($0 < \theta$)
  - a priori de θ é uma Gamma com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ ($i = 1, ..., n$) será uma Beta com parâmetros $\alpha + \sum_{i=1}^{n} x_i$ e $\beta + n$

Distribuições conjugadas

- Amostragem de uma distribuição de Poisson
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Poisson com
  - θ desconhecido ($0 < \theta$)
  - a priori de θ é uma Gamma com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ ($i = 1, ..., n$) será uma Beta com parâmetros $\alpha + \sum_{i=1}^{n} x_i$ e $\beta + n$

Distribuições conjugadas

- Amostragem de uma distribuição Normal
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Normal com
  - média θ desconhecido $(-\infty < \theta < \infty)$, σ^2 conhecido $(\sigma^2 > 0)$
  - a priori de θ é uma normal com média μ e variância ν^2
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma normal com parâmetros μ_1 e variância ν_1:

$$\mu_1 = \frac{\sigma^2\mu + n\nu^2\bar{x}_n}{\sigma^2 + n\nu^2}$$

$$\nu_1^2 = \frac{\sigma^2\nu^2}{\sigma^2 + n\nu^2}$$

Distribuições conjugadas

- Amostragem de uma distribuição Normal
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Normal com
  - média θ desconhecido $(-\infty < \theta < \infty)$, σ^2 conhecido $(\sigma^2 > 0)$
  - a priori de θ é uma normal com média μ e variância ν^2
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma normal com parâmetros μ_1 e variância ν_1:

$$\mu_1 = \frac{\sigma^2\mu + n\nu^2\bar{x}_n}{\sigma^2 + n\nu^2}$$

$$\nu_1^2 = \frac{\sigma^2\nu^2}{\sigma^2 + n\nu^2}$$

Distribuições conjugadas

- Amostragem de uma distribuição Normal
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Normal com
  - média θ desconhecido $(-\infty < \theta < \infty)$, σ^2 conhecido $(\sigma^2 > 0)$
  - a priori de θ é uma normal com média μ e variância ν^2
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma normal com parâmetros μ_1 e variância ν_1:

$$\mu_1 = \frac{\sigma^2\mu + n\nu^2\bar{x}_n}{\sigma^2 + n\nu^2}$$

$$\nu_1^2 = \frac{\sigma^2\nu^2}{\sigma^2 + n\nu^2}$$

Distribuições conjugadas

- Amostragem de uma distribuição Normal
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma Normal com
  - média θ desconhecido $(-\infty < \theta < \infty)$, σ^2 conhecido $(\sigma^2 > 0)$
  - a priori de θ é uma normal com média μ e variância ν^2
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma normal com parâmetros μ_1 e variância ν_1:

$$\mu_1 = \frac{\sigma^2\mu + n\nu^2\bar{x}_n}{\sigma^2 + n\nu^2}$$

$$\nu_1^2 = \frac{\sigma^2\nu^2}{\sigma^2 + n\nu^2}$$

Distribuições conjugadas

- Amostragem de uma distribuição exponencial
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma exponencial com
  - θ desconhecido $(0 < \theta)$
  - a priori de θ é uma Gamma com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Gamma com parâmetros $\alpha + n$ e $\beta + \sum_{i=1}^{n} x_i$

Distribuições conjugadas

- Amostragem de uma distribuição exponencial
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma exponencial com
  - θ desconhecido ($0 < \theta$)
  - a priori de θ é uma Gamma com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ ($i = 1, ..., n$) será uma Gamma com parâmetros $\alpha + n$ e $\beta + \sum_{i=1}^{n} x_i$

Distribuições conjugadas

- Amostragem de uma distribuição exponencial
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma exponencial com
  - θ desconhecido $(0 < \theta)$
  - a priori de θ é uma Gamma com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Gamma com parâmetros $\alpha + n$ e $\beta + \sum_{i=1}^{n} x_i$

Distribuições conjugadas

- Amostragem de uma distribuição exponencial
  - $X_1, X_2, ..., X_n$ forma uma a.a. de uma exponencial com
  - θ desconhecido $(0 < \theta)$
  - a priori de θ é uma Gamma com parâmetros $\alpha > 0$ e $\beta > 0$
  - a posteriori de θ dado que $X_i = x_i$ $(i = 1, ..., n)$ será uma Gamma com parâmetros $\alpha + n$ e $\beta + \sum_{i=1}^{n} x_i$