Disciplina: **USP-EACH 2053 Introdução à Estatística**
Documento: **Resumo de Estatística para P2**
Livro referência: **deGroot, edição 2**
Revisão: **8** data: **2012-06-25**

Inferência Estatística
Capítulo 6

Introdução: (pag311)

Estatística: Temos dados da amostra de um evento e precisamos inferir sobre propriedades da distribuição. (indução)

Probabilidade: Temos dados sobre a distribuição e precisamos calcular a probabilidade de um evento. (dedução)

Inferência Estatística:

Notações:

θ é o parâmetro (teta) procurado
Ω é o espaço paramétrico (omega) que contém todos os valores possíveis de θ
$f(x|\theta)$ é uma função probabilidade de um evendo, dado o parâmetro θ

Parâmetros θ da distribuição pode ser μ, σ^2 em uma distribuição normal; λ em uma distribuição exponencial, etc.

Amostras aleatórias:

Amostras aleatórias são múltiplos eventos i.i.d. (independentes e identicamente distribuídos).

Notações:

$\boldsymbol{X}$ = vetor de variáveis $X_1, \ldots, X_n$ anteriores ao experimento.
$\boldsymbol{x}$ = vetor de dados de amostras $x_1, \ldots, x_n$ de um experimento.
$f_n(\boldsymbol{x}|\theta)$ = função de probabilidade conjunta de uma série de eventos, dado θ

$$f_n(\boldsymbol{x}|\theta) = f_n(x_1, .., x_n|\theta) = f(x_1|\theta) \cdot f(x_2|\theta), .., f(x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

Distribuições a priori: (pag313)

Na estatística Bayesiana podemos restringir Ω antes de coletar dados do experimento e/ou fornecer uma distribuição a priori para os valores de θ .

Notações:

$\xi(\theta)$ é uma função de probabilidade (xi) da distribuição a priori de θ

Distribuições a posteriori: (pag316)

Na estatística Bayesiana, a distribuição a posteriori é um estimador do parâmetro θ depois de ser observados os valores $x_1, .., x_n$ da amostra.

Notações:

$\xi(\theta|\boldsymbol{x})$ é a função de distribuição a posteriori de θ , dado o vetor $\boldsymbol{x}$
$\xi(X_1, .., X_n)$ ou $\xi(\boldsymbol{X})$ é um estimador, antes de ter os dados do evento
$\xi(x_1, \ldots, x_n)$ ou $\xi(\boldsymbol{x})$ é uma estimativa, depois de ter os dados do evento

Temos $f_n(\boldsymbol{x}|\theta)$ e $\xi(\theta)$ e queremos achar $\xi(\theta|\boldsymbol{x})$. Para isso podemos aplicar Bayes $P(B|A)=P(A|B)P(B)/P(A)$ calculando a função da distribuição marginal de $\boldsymbol{x}$ independente de θ :

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

A função da distribuição a posteriori de θ dado $\boldsymbol{x}$ é

$$\xi(\theta|\boldsymbol{x})=\frac{f_n(\boldsymbol{x}|\theta)\xi(\theta)}{g_n(\boldsymbol{x})}$$

Podemos chamar $g_n(\boldsymbol{x})$ de constante k pois não depende de θ

$$k=\frac{1}{\int f_n(\boldsymbol{x}|\theta)\xi(\theta)d\theta}$$

A função de distribuição de verossimilhança a posteriori de θ é

$$\xi(\theta|\boldsymbol{x})=k\cdot f_n(\boldsymbol{x}|\theta)\cdot\xi(\theta)$$

Exemplo com uma Bernoulli: (pag317)

Um processo de manufatura tem uma probabilidade p desconhecida de fabricar uma peça com defeito.

$\theta=p$

$\Omega=0,..,1$ o domínio de θ é o domínio de p

$\xi(\theta)=1$ é uma uniforme entre 0 e 1

x é a quantidade de sucessos até que ocorra 1 defeito

Função de probabilidaded p é uma Bernoulli

$$P(X=n)=p^n(1-p)^{(1-n)}$$

1. Encontrar a distribuição de cada evento $f(x|\theta)$; Procurando parâmetro θ como sendo a probabilidade p de um evento e x sendo a quantidade de eventos

$$f(x|\theta)=\begin{Bmatrix}\theta^x(1-\theta)^{(1-x)} se\, x=0,1\\ 0\, se\ outros\ casos\end{Bmatrix}$$

2. Calcular a distribuição conjunta $f_n(\boldsymbol{x}|\theta)$; Fazendo o produto de todos os eventos do vetor $\boldsymbol{x}=x_1,\dots,x_n$

$$f_n(\boldsymbol{x}|\theta)=\theta^y(1-\theta)^{(n-y)} \text{ quando } y=\sum_{i=1}^{n} x_i$$

3. Encontrar $\xi(\theta)$ a priori. Não temos uma distribuição a priori, portanto é uma uniforme

$$\xi(\theta)=1$$

4. Calcular a distribuição de verossimilhança a posteriori $\xi(\theta|\boldsymbol{x})=k\, f_n(\boldsymbol{x}|\theta)\xi(\theta)$

$$\xi(\theta|\boldsymbol{x})=k\,\theta^y(1-\theta)^{(n-y)} \text{ quando } y=\sum_{i=1}^{n} x_i$$

5. Encontrar $\xi(\theta|\boldsymbol{x})$ como sendo uma distribuição f.d.p a posteriori, no caso uma Beta com os parâmetros $\alpha=y+1$ e $\beta=n-y+1$

$$\xi(\theta|\boldsymbol{x})=\frac{\Gamma(n+2)}{\Gamma(y+1)\Gamma(n-y+1)}\theta^{y}(1-\theta)^{n-y} \text{ quando } y=\sum_{i=1}^{n} x_i$$

Exemplo com uma Exponencial: (pag318)

Uma lâmpada fluorescente tem uma vida útil que segue uma exponencial com parâmetro λ desconhecido. A distribuição a priori do parâmetro λ é uma Gamma com média 0.00002 e desvio padrão 0.0001.

$\alpha/\beta=0.00002$ é a média da Gamma

$\alpha/\beta^2=(0.0001)^2$ é a variância da Gamma que é o quadrado do desvio padrão

$f(x)=\lambda e^{-\lambda x}$ função da vida útil que segue uma função exponencial

1. Encontrar a distribuição de cada evento $f(x|\theta)$; foi dado do problema

$f(x)=\lambda e^{-\lambda x}$

2. Calcular a distribuição conjunta $f_n(\boldsymbol{x}|\theta)$; Fazendo o produto de todos os eventos do vetor $\boldsymbol{x}=x_1,\dots,x_n$

$$f_n(\boldsymbol{x}|\theta)=\lambda^n e^{(-\lambda y)} \text{ quando } y=\sum_{i=1}^{n} x_i$$

3. Calcular qual é a distribuição a priori $\xi(\lambda)$

$\alpha=4$

$\beta=20000$

$$\Gamma(\alpha)=\int_0^\infty x^{(\alpha-1)}e^{x}\,dx=(a-1)!$$

$$\Gamma(x|\alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^\alpha e^{\beta x} \text{ para } x>0 \text{ e } \Gamma(x|\alpha,\beta)=0 \text{ para } x\geq 0$$

$$\xi(\lambda)=\frac{(20000)^4}{3!}\lambda^3 e^{-20000\lambda}$$

$$\xi(\lambda)=k\lambda^3 e^{-20000\lambda}$$

4. Calcular a distribuição de verossimilhança a posteriori $\xi(\theta|\boldsymbol{x})=k f_n(\boldsymbol{x}|\theta)\xi(\theta)$

$$\xi(\lambda|\boldsymbol{x})=k\cdot\lambda^n e^{(-\lambda y)}\cdot\lambda^3 e^{-20000\lambda} \text{ quando } y=\sum_{i=1}^{n} x_i$$

$$\xi(\lambda|\boldsymbol{x})=k\lambda^{(n+3)}e^{-(y+20000)\lambda} \text{ quando } y=\sum_{i=1}^{n} x_i$$

5. Encontrar $\xi(\lambda|\boldsymbol{x})$ como sendo uma distribuição f.d.p a posteriori, no caso uma Gamma com os parâmetros $\alpha=n+4$ e $\beta=y+20000$

$$\xi(\lambda|\boldsymbol{x})=\frac{(y+20000)^{n+4}}{(n+3)!}\lambda^{(n+3)}e^{-(y+20000)\lambda} \text{ quando } y=\sum_{i=1}^{n} x_i$$

Observação sequencial: (pag319)

A função a priori de observações sequenciais pode ser incremental, ou pode ser absoluta (sempre se calcula a mesma apriori para toda a sequencia). Em ambos os casos, o resultado é o mesmo.

Distribuição conjugada: (pag321)

Amostragem de uma Bernoulli:

função de probabilidade para um evento

$f(x|\theta)=\theta^{x_i}(1-\theta)^{1-x_i}$

função de probabilidade conjunta (para um vetor de eventos)

$f_n(\boldsymbol{x}|\theta)=\theta^y(1-\theta)^{n-y}; y=\sum_{i=1}^{n} x_i$

para um θ desconhecido $0<\theta<1$

a função a priori é uma Beta com parâmetros $\alpha_1>0$ e $\beta_2>0$

a função a posteriori é uma Beta com parâmetros

$\alpha_2=\alpha_1+\sum_{i=0}^{n} x_i$ e $\beta_2=\beta_1+n-\sum_{i=0}^{n} x_i$

Amostragem de uma Uniforme:

função de probabilidade para um evento

$f(x|\theta)=\frac{1}{\theta}; 0\leq x\leq\theta$

função de probabilidade conjunta (para um vetor de eventos)

$f_n(\boldsymbol{x}|\theta)=\frac{1}{\theta^n}; 0\leq x_i\leq\theta$

Amostragem de uma Poisson:

função de probabilidade para um evento

$f(x|\theta)=\frac{e^{-\theta}\theta^x}{x!}$

função de probabilidade conjunta (para um vetor de eventos)

$f_n(\boldsymbol{x}|\theta)=(\prod_{i=1}^{n}\frac{1}{x_i!})e^{-n\theta}\theta^y; y=\sum_{i=1}^{n} x_i$

para um θ desconhecido $0<\theta$

a função a priori é uma Gamma com parâmetros $\alpha_1>0$ e $\beta_1>0$

a função a posteriori é uma Beta com parâmetros

$\alpha_2=\alpha_1+\sum_{i=0}^{n} x_i$ e $\beta_2=\beta_1+n$

Amostragem de uma Exponencial:

função de probabilidade para um evento

$f(x|\theta)=\theta e^{-\theta x}$

função de probabilidade conjunta (para um vetor de eventos)

$f_n(\boldsymbol{x}|\theta)=\theta^n e^{-\theta y}; y=\sum_{i=1}^{n} x_i$

para um θ desconhecido $0<\theta$ e σ_0^2 conhecido $\sigma_0^2>0$

a função a priori é uma Gamma com parâmetros $\alpha_1>0$ e $\beta_1>0$

a função a posteriori é uma Gamma com parâmetros

$\alpha_2=\alpha_1+n$ e $\beta_2=\beta_1+\sum_{i=0}^{n} x_i$

Amostragem de uma Normal:

função de probabilidade para um evento

$$f(x|\mu)=\frac{1}{\sigma\sqrt{2\pi}}e^{-[\frac{(x-\mu)^2}{2\sigma^2}]}$$

função de probabilidade conjunta (para um vetor de eventos)

$$f_n(\boldsymbol{x}|\mu)=\frac{1}{(\sigma 2\pi)^{n/2}}e^{-[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}]}$$

para uma média θ desconhecida $-\infty<\theta<\infty$ e σ_0^2 conhecido $\sigma_0^2>0$

a função a priori é uma Normal com parâmetros μ_1 e σ_1^2

a função a posteriori é uma Normal com parâmetros

$$\mu_2=\frac{\sigma_0^2\mu_1+\sigma_1^2 n\bar{x}_n}{\sigma_0^2+n\sigma_1^2} \text{ e } \sigma_2^2=\frac{\sigma_0^2\sigma_1^2}{\sigma_0^2+n\sigma_1^2}$$

Amostragem de uma Gamma:

função de probabilidade para um evento, esperança e variância

$$f(x|\alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}; x>0$$

$$E(X)=\frac{\alpha}{\beta}$$

$$VAR(X)=\frac{\alpha}{\beta^2}$$

função de probabilidade conjunta (para um vetor de eventos)

$$f_n(\boldsymbol{x}|\theta)=\frac{\beta^{n\alpha}}{\Gamma^n(\alpha)}(\prod_{i=1}^{n}x_i)^{\alpha-1}e^{(-\beta y)}; y=\sum_{i=1}^{n}x_i$$

Amostragem de uma Beta:

função de probabilidade para um evento, esperança e variância

$$f(x|\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}; 0<x<1$$

$$E(X)=\frac{\alpha}{\alpha+\beta}$$

$$VAR(X)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

função de probabilidade conjunta (para um vetor de eventos)

$$f_n(\boldsymbol{x}|\theta)=\frac{\beta^{n\alpha}}{\Gamma^n(\alpha)}(\prod_{i=1}^{n}x_i)^{\alpha-1}e^{(-\beta y)}; y=\sum_{i=1}^{n}x_i$$

Estimadores de Bayes: (pag330)

Para aplicar a teoria dos estimadores Baysianos, é necessário:

1. especificar uma função de perda
2. determinar uma função a priori para o parâmetro

Um bom estimador $\delta(\boldsymbol{X})$ do parâmetro θ é aquele que faz o valor do erro definido por $\delta(\boldsymbol{X})-\theta$ tender a zero. Sendo $a=\delta(\boldsymbol{x})$ e $a\in\Omega$ uma estimativa de θ . Então $L(\theta,a)$ é a função que mede a diferença entre uma estimativa a e o seu respectivo parâmetro θ .

Diferentes funções de perta:

$$L(\theta, a) = (\theta - a)^2$$
$$L(\theta, a) = |\theta - a|$$

Sendo a esperança $E(x) = \int_{\Omega} x \cdot f(x) dx$ portanto

$$E[L(\theta, a)] = \int_{\Omega} L(\theta, a) \xi(\theta) d\theta$$

A função L para o qual o valor $E[L(\theta, a)]$ será minimizada é chamada de função de perda.

Sendo $\boldsymbol{x}$ um vetor de dados observados para estimar θ

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

A função $\delta*(\boldsymbol{X})$ é chamado de estimador de Bayes quando, para cada $\boldsymbol{x}$ de $\boldsymbol{X}$, faz com que $E[L(\theta, \delta*(\boldsymbol{x}))|\boldsymbol{x}] = min_{a \in \theta} E[L(\theta, a)|\boldsymbol{x}]$. Ou seja é o estimador que fornece a menor função de perda.

Diferentes funções de perda:

para $L(\theta, a) = (\theta - a)^2$ então $\delta*(\boldsymbol{X}) = E(\theta|\boldsymbol{X})$ será a média

para $L(\theta, a) = |\theta - a|$ então $\delta*(\boldsymbol{X}) = E(|\theta - a||\boldsymbol{X})$ será a mediana

para $L(\theta, a) = |\theta - a|^k$ e $k \in N$ ou $k > 2$ então $L(\theta, a) = \lambda(\theta)|\theta - a|^2$

O estimador de Bayes para a função de perda quardrática é:

$$\delta*(\boldsymbol{X}) = E(\theta|\boldsymbol{X})$$

O estimador de Bayes genérico é:

$$\delta*(\boldsymbol{X}) = E(L(\theta, a)|\boldsymbol{X}) = \int_{\Omega} L(\theta, a) \xi(\theta|\boldsymbol{x}) d\theta$$

Estimador de Bayes com função de perda quadrática para Bernouli:

Sendo a posteriori $\xi(\theta|\boldsymbol{x})$ uma Beta com parâmetros

$\alpha_2 = \alpha_1 + y$ e $\beta_2 = \beta_1 + n - y$ sendo $y = \sum_{i=0}^{n} x_i$

Sendo a esperança para a Beta

$$E(X) = \frac{\alpha}{\alpha + \beta}$$

O estimador de Bayes com função de perda quadrática para uma posteriori Beta é

$$\delta*(\boldsymbol{X}) = \frac{\alpha_1 + y}{\alpha_1 + \beta_1 + n} \quad \text{sendo} \quad y = \sum_{i=0}^{n} x_i$$

Estimador de Bayes com função de perda quadrática para Normal:

Sendo a posteriori $\xi(\theta|\boldsymbol{x})$ uma Normal $N(\theta, \sigma^2)$ e uma distribuição a priori $\xi(\theta)$ conhecida $N(\mu, v^2)$.

O estimador de Bayes com função de perda quadrática para uma posteriori Normal é

$$\delta*(\boldsymbol{X}) = \frac{\sigma^2 \mu + n v^2 \bar{X}_n}{\sigma^2 + n v^2}$$

Estimadores de Bayes para grandes amostras: (pag335)

Para grandes amostras $n \to \infty$, mesmo partindo de funções a priori diferentes, a estimativa converge para o valor de θ .

Estimador de máxima verossimilhança (EMV): (pag338)

Estimador de máxima verossimilhança é o estimador $\hat{\theta}$ cuja estimativa θ aplicada à função $f_n(\boldsymbol{x}|\theta)$ tenha a maior probabilidade de gerar o vetor de dados $\boldsymbol{x}$.

Notação:
$\hat{\theta}$ (teta chapeu) é o estimador de máxima verossimilhança
$L(\theta) = \log f_n(\boldsymbol{x}|\theta)$ é o log da função de probabilidade conjugada

O estimador de máxima verossimilhança (MVE) é o ponto máximo de $L(\theta)$ ou seja
MVE é $\dfrac{d\,L(\theta)}{d\theta} = 0$ ou $\dfrac{d}{d\,\theta}(\log f_n(\boldsymbol{x}|\theta)) = 0$

O EMV pode ser único, pode ser múltiplo ou pode ser inexistente.

Exemplo de EMV para uma Bernoulli

$$f_n(\boldsymbol{x}|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^y(1-\theta)^{n-y};\ y = \sum_{i=1}^{n} x_i$$

Aplicando um log.

$$L(\theta) = \log f_n(\boldsymbol{x}|\theta) = y\log\theta + (n-y)\log(1-\theta);\ y = \sum_{i=1}^{n} x_i$$

O ponto de mínima de $L(\theta)$ é $dL(\theta)/d\,\theta = 0$, encontramos
$\hat{\theta} = \bar{x}_n$

Mortanto o EMV de θ para uma Bernoulli é
$\hat{\theta} = \bar{X}_n$

Exemplo de EMV para uma Normal com média e variância desconhecidas.

$$\hat{\mu} = \bar{X}_n \text{ e } \hat{\sigma}^2 = \frac{1}{n}\sum_{i=0}^{n}(x_i - \bar{x}_n)^2$$

Propriedades do estimador de máxima verossimilhança: (pag338)

Estatística suficiente: (pag356)

Estatística suficiente é quando, dado observações de uma amostra $x_1, x_2, \ldots, x_n$, uma estatística $T = r(x_1, x_2, \ldots, x_n)$ e sendo função r o EMV, T é a melhor estimativa da observação, podendo substituir os dados da amostra sem nenhuma perda no resultado final.

Para uma Bernoulli $f_n(\boldsymbol{x}|\theta) = \theta^y(1-\theta)^{n-y};\ y = \sum_{i=1}^{n} x_i$, são estatística suficiente n $y = \sum_{i=1}^{n} x_i$ pois eles substituem totalmente a amostra $x_1, x_2, \ldots, x_n$.

Estatísticas suficientes mais comuns

$\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i$ é a média amostral

$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2$ é a variância amostral

$x_{(1)} = min(\boldsymbol{x}) = min(x_1, .., x_n)$ o menor valor da amostra

$x_{(n)} = max(\boldsymbol{x}) = max(x_1, .., x_n)$ o maior valor da amostra

$x_{(i)}$ a i-ésima maior observação da amostra

Limites de estatísticas suficientes.

Estimador rubusto é um estimador suficiente para ampla variedade de possíveis f.d.p's.

Estimadores não Viesados:

Um estimador não viesado $\delta(\boldsymbol{X})$ tem sua esperança $E[\delta(\boldsymbol{X})]$ igual ao valor desconhecido verdadeiro de θ .

Exemplo: se a média é a esperança da média amostral $\mu = E(\bar{\boldsymbol{X}}_{\boldsymbol{n}})$, um estimador não viesado da média $\delta(\mu) = \bar{\boldsymbol{X}}_{\boldsymbol{n}}$ é a média amostral.

Estimadores não viesados (ENV):

$\delta(\mu) = \bar{\boldsymbol{x}}_{\boldsymbol{n}} = \frac{1}{n} \sum_{i=1}^{n} x_i$ um ENV da média é a média amostral

$\delta(\sigma^2) = S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2$ um ENV da variância é a variância amostral

Teste de Hipóteses
Capítulo 8

Problemas de teste de hipótese: (pag 437)

Em um problema estatístico envolvendo o parâmetro θ , cujo valor desconhecido pertence ao espaço Ω , alguns valores de θ podem ser satisfatórios para uma determinada hipótese H_0 .

Hipótese nula e alternativa: (pag 437)

O conjunto dos valores satisfatórios é chamado de Ω_0 e o conjunto dos valores não satisfatórios é chamado de Ω_1

Notação:

H_0 Hipótese nula $\theta \in \Omega_0$, ou seja, hipótese de interesse

H_1 Hipótese alternativa $\theta \in \Omega_1$

θ Parâmetro que está sendo testado

Ω_0 Espaço paramétrico da hipótese nula (teta zero), ou seja, de interesse.

Ω_1 Espaço paramétrico da hipótese alternativa (teta um)

Ω Espaço paramétrico total (teta) onde $\Omega_0 \cup \Omega_1 = \Omega$

$\theta = \theta_0$ Parâmetro dentro do espaço paramêtrico da hipótese nula $\theta \in \Omega_0$

$\theta = \theta_1$ Parâmetro dentro do espaço paramêtrico da hipótese alternativa $\theta \in \Omega_1$

Região Crítica: (pag 438)

Região do espaço amostral S_1 que rejeita a hipótese nula, ou seja, que está associada ao parâmetro dentro da hipótese alternativa $\theta\in\Omega_1$.

Notação:

S Espaço amostral

S_0 Espaço amostral que aceita a hipótese nula

S_1 ou C ou RC **Região crítica** é o espaço amostral que rejeita a hipótese nula

Função Poder: (pag 438)

É a função $\pi(\theta)$ que dá a probabilidade de rejeitar a hipótese nula H_0 . Ou seja é a probabilidade de $\theta\in\Omega_1$. A função poder, como toda probabilidade, varia no intervalo de zero a um.

Notação:

$\pi(\theta)$ = função poder (pi) é a probabilidade de $\theta\in\Omega_1$ ou seja, rejeitar H_0

$\boldsymbol{x}$ = vetor de dados de amostras $x_1, \ldots, x_n$ de um experimento

$f(x|\theta)$ = função de probabilidade de um evento, dado θ

$F(x|\theta)$ = função de probabilidade acumulada de um evento, dado θ

$f_n(\boldsymbol{x}|\theta)$ = função de probabilidade conjunta de uma série de eventos, dado θ

Probabilidade conjunta:

$$f_n(\boldsymbol{x}|\theta)=f_n(x_1, .., x_n|\theta)=f(x_1|\theta)\cdot f(x_2|\theta), .., f(x_n|\theta)=\prod_{i=1}^{n} f(x_i|\theta)$$

Função poder:

$$\pi(\theta)=f_n(\boldsymbol{x}\in S_1|\theta); \theta\in\Omega$$

Tamanho do Teste:

Tamanho do teste α é a maior probabilidade, entre todos os valores de θ , que satisfaça a hipótese nula em não fazer uma decisão errada.

Notação:

α_0 = nível de significância $0<\alpha_0<1$

α = tamanho do teste

Tamanho do teste:

$$\alpha=\sup_{\theta\in\Omega}\pi(\theta|\delta)$$

Para uma hipótese nula simples, o tamanho do teste é

$$\alpha=\pi(\theta_0|\delta)$$

Teste:

Um teste δ é a definição de uma função poder $\pi(\theta|\delta)$.

Notação:

$\pi(\theta|\delta)$ = função poder (pi) de um teste δ

$\alpha(\delta)$ = tamanho do teste δ

Roteiro:

(1) Identificar o parâmetro θ e sua distribuição

$Uniforme(0,\theta)$ ou $Uniforme(\theta,0)$

$Normal(\theta,\sigma^2)$

$Poison(\theta)$

$Exponencial(\theta)$

$Polinomial(\theta)$

(2) Identificar a hipótese de interesse H_0

$H_0 = 3 \leq \theta \leq 4$

(3) Identificar o espaço paramêtrico Ω , Ω_0 e Ω_1

Ω Espaço paramétrico total (teta) onde $\Omega_0 \cup \Omega_1 = \Omega$

Ω_0 Espaço paramétrico da hipótese nula (teta zero), ou seja, de interesse.

Ω_1 Espaço paramétrico da hipótese alternativa (teta um)

Traçar a linha do espaço parâmetrico

(4) Identificar o teste δ com o seu espaço amostral S_1

S_1 ou C ou RC **Região crítica** é o espaço amostral que rejeita a hipótese nula

Traçar a linha do espaço amostral

(5) identificar a função a acumulada $F(x|\theta)$ e a conjunta $f_n(\boldsymbol{x}|\theta)$

$f(x|\theta)$ para $x = k$

$F(x|\theta)$ para $k \leq x \leq m$

$f_n(\boldsymbol{x}|\theta)$ a função conjugada

(6) Achar a função poder para cada sub-espaço do espaço paramétrico

se $\pi(\theta \in \Omega_a|\delta)$ calcular $f_n(\boldsymbol{x} \in S_{1a}|\theta) + f_n(\boldsymbol{x} \in S_{1b}|\theta) + ...$

se $\pi(\theta \in \Omega_b|\delta)$ calcular $f_n(\boldsymbol{x} \in S_{1a}|\theta) + f_n(\boldsymbol{x} \in S_{1b}|\theta) + ...$

(7) Calcular o tamanho do teste

$\alpha(\delta) = \sup_{\theta \in \Omega} \pi(\theta|\delta)$

Exemplo:

Dada uma amostra aleatória $x_1, ..., x_n$ de uma distribuição uniforme no intervalo de $(0,\theta)$ sendo θ desconhecido e $\theta > 0$. A hipótese testa é $H_0 = 3 \leq \theta \leq 4$.

Estimativa suficiente para o máximo da uniforme é

$x_{(n)} = max(x_1, .., x_n)$

O teste pode ser definido como, a hipótese será expandida para $H_0 = 2.9 \leq \theta \leq 4$ portanto

$\pi(\theta|\delta) = P(x_{(n)} < 2.9|\theta) + P(x_{(n)} > 4|\theta)$

$f(x|\theta) = \frac{1}{\theta}$

$F(x < 2.9|\theta) = \frac{2.9}{\theta}$

$f_n(\boldsymbol{x_{(n)}} < 2.9|\theta) = (\frac{2.9}{\theta})^n$

$F(x > 4|\theta) = 1 - F(x < 4|\theta) = 1 - \frac{4}{\theta}$

$f_n(\boldsymbol{x_{(n)}} < 4|\theta) = 1 - (\frac{4}{\theta})^n$

Calculando a função poder

$$\pi(\theta|\delta) = \left(\frac{2.9}{\theta}\right)^n + 1 - \left(\frac{4}{\theta}\right)^n$$

Calculando o tamanho do teste

$$\alpha = \sup_{\theta \in \Omega} \pi(3) = \left(\frac{2,9}{3}\right)^n$$

Testando hipótese simples: (pag 442)

Erro tipo 1: Rejeitar hipótese H_0 quando de fato ela é verdadeira.

Erro tipo 2: Aceitar hipótese H_0 quando de fato ela é falsa.

δ Procedimento de teste

Erro tipo 1

$$\alpha(\theta) = P(\textit{rejeitar}\ H_0 | \theta = \theta_0)$$

Erro tipo 2

$$\beta(\theta) = P(\textit{aceitar}\ H_0 | \theta = \theta_1)$$

O ideal é encontrar um procedimento de teste que mantenha $\alpha(\theta)$ e $\beta(\theta)$ o menor possível.

Minimizando a combinação linear

$$a\alpha(\delta^*) + b\beta(\delta^*) \leq a\alpha(\delta) + b\beta(\delta)$$

Minimizando a combinação linear

$$a\alpha(\delta_1) + b\beta(\delta_1) = a\sum_{x \in R} f_0(\boldsymbol{x}) + b\sum_{x \in R^c} f_1(\boldsymbol{x})$$

$$= a\sum_{x \in R} f_0(\boldsymbol{x}) + b(1 - \sum_{x \in R} f_1(\boldsymbol{x}))$$

$$= b + \sum_{x \in R} [af_0(\boldsymbol{x}) - b f_1(\boldsymbol{x})]$$

Problema de multi-decisão: (pag 456)

Teste uniformemenete mais poderoso: (pag 466)

Selecionando procedimento de teste: (pag 477)

Teste T: (pag 485)

Discussão sobre metodologia de teste de hipótese: (pag 494)

Distribuição F: (pag 499)

Comparando média de duas distribuições normais: (pag 506)