Análise de Dados Categóricos
Função de Verossimilhança

Enrico A. Colosimo/UFMG

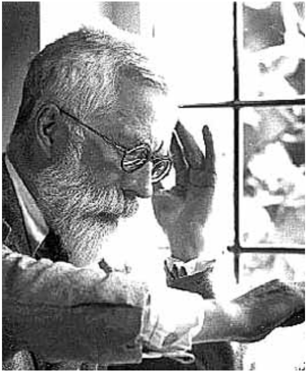
Depto. Estatística - ICEx - UFMG

Introdução

- Inferência Estatística: (Modelo + Método) + Dados
- Método: Máxima Verossimilhança
- Função de Verossimilhança: medida da informação fornecida pelos dados para um ou mais parâmetros de um modelo probabilístico.
- Ideia: valores do parâmetro que tornam os dados amostrais mais prováveis devem ser preferidos àqueles que os fazem menos provável.

Função de Verossimilhança

Sir. Ronald Fisher



Função de Verossimilhança para o Exemplo em Ecologia

- Modelo

$$X \sim bin(n = 20, \pi)$$

- Dados: $X = 12$ (valor observado)
- Função de Verossimilhança

$$L(\pi) = \prod_{i=1}^{20} \pi^{y_i}(1-\pi)^{1-y_i} = \pi^{\sum_{i=1}^{20} y_i}(1-\pi)^{20-\sum_{i=1}^{20} y_i} = \pi^{x}(1-\pi)^{(20-x)}$$

- Método: Máxima Verossimilhança

Propriedades da Função de Verossimilhança

- EMV
- Consistência e Normalidade assintótica
- Invariância
- Estatísticas
  - Estatística de Wald
  - Estatística da Razão de Verossimilhanças
  - Estatística Escore
  - Gradiente

Função de Verossimilhança para amostra Binomial

- Valores observados: $y_1, y_2, ..., y_{20}$
- $X = \sum_{i=1}^{20} y_i \sim Binomial(20, \pi)$
- $L(\pi) = \pi^{\sum y_i}(1-\pi)^{20-\sum y_i} \Rightarrow \pi^{12}(1-\pi)^8$

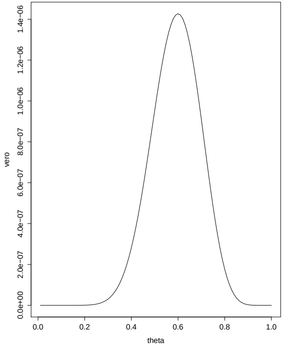
Exemplo:

$\pi = 0,1 \Rightarrow (0,1)^{12}(0,9)^8 = 4,3x10^{-13}$
$\pi = 0,5 \Rightarrow (0,5)^{12}(0,5)^8 = 9,5x10^{-7}$

A probabilidade de observar os dados amostrais é maior para $\pi = 0,5$ do que para $\pi = 0,1$

Conclusão: 0,5 é mais provável que 0,1

Função de Verossimilhança para amostra Binomial



- O estimador de máxima verossimilhança é $\widehat{\pi} = 0,6$
- $L(0,6)$ é o maior valor que $L(\pi)$ pode assumir.

$$L(0,6) = 1,43x10^{-6}.$$

Função de Verossimilhança para amostra Binomial

1. Razão de Verossimilhanças: mede o nível que a verossimilhança fornece para os valores de π

$$\lambda(\pi) = \frac{L(\pi)}{L(0,6)}$$

$0 \leq \lambda(\pi) \leq 1$

RV: $-2\log\lambda(\pi)$: útil para testar hipóteses e construir IC.

$IC(\pi, 95\%) = (\pi | -2\log\left\{\frac{L(\pi)}{L(0,6)}\right\} \leq 3,84)$

2. Log-Verossimilhança

$\log L(\pi) = (\sum y_i)\log\pi + (n - \sum y_i)\log(1-\pi)$

Propriedades da $L(\pi)$

1. $\widehat{\pi} = arg \max_{0 \leq \pi \leq 1} L(\pi) = \max_{0 \leq \pi \leq 1} \log L(\pi)$

2. Teste Wald (assintótico)

   $\widehat{\pi} \sim N(\pi, Var(\widehat{\pi}))$; $Var(\widehat{\pi}) = \mathfrak{F}^{-1}(\pi)$, em que $\mathfrak{F}^{-1}(\pi)$ é informação de Fisher.

3. Teste Escore (assintótico)

   $S(\pi) = \frac{dlogL(\pi)}{d\pi} \sim N(0, \mathfrak{F}(\pi))$

4. TRV (assintótico)

   RV$= -2\log(\lambda(\pi)) = -2\log\frac{L(\pi)}{L(\widehat{\pi})} \sim \chi^2_1$

Estimador de Máxima Verossimilhança - EMV

$X = \sum_{i=1}^{20} Y_i$

$L(\pi) = \pi^x (1-\pi)^{n-x}$

$l(\pi) = \log L(\pi) = x \log(\pi) + (n-x) \log(1-\pi)$

$$S(\pi) = \frac{x}{\pi} - \frac{n-x}{1-\pi}$$

EMV: $x(1-\pi) = (n-x)\pi \Rightarrow \widehat{\pi} = \frac{x}{n} = 0,6$

$\mathfrak{F}(\pi)$ - Informação de Fisher

$$\mathfrak{F}(\pi) = -E\left[\frac{d^2 l(\pi)}{d\pi^2}\right]$$

$$\frac{d^2 l(\pi)}{d\pi^2} = -[\frac{x}{\pi^2} + \frac{n-x}{(1-\pi)^2}]$$

$$\mathfrak{F}(\pi) = \frac{E(x)}{\pi^2} + \frac{E(n-x)}{(1-\pi)^2} = \frac{n\pi}{\pi^2} + \frac{n(1-\pi)}{(1-\pi)^2}$$

$$\mathfrak{F}(\pi) = \frac{n}{\pi(1-\pi)}$$

$$Var(\widehat{\pi}) = [\frac{\pi(1-\pi)}{n}]$$

Neste caso a variância assintótica é exata

$$Var(\widehat{\pi}) = Var(\frac{x}{n}) = \frac{1}{n^2} n\pi(1-\pi) = \frac{\pi(1-\pi)}{n}$$

1- Estatística de Wald

$$W = \frac{\widehat{\pi} - \pi}{\sqrt{\dfrac{\pi(1-\pi)}{n}}} \sim N(0,1)$$

$$W = \frac{\frac{x}{n} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{x - n\pi}{\sqrt{n\pi(1-\pi)}}$$

No exemplo

$$W = \frac{12-10}{\sqrt{5}} = \frac{2}{\sqrt{5}} \Rightarrow valor - p = 0,3711$$

Intervalo de 95% de Confiança para π:

$$\widehat{\pi} \pm 1,96\sqrt{\frac{\widehat{\pi}(1-\widehat{\pi})}{n}} \Longrightarrow (0,385; 0,815)$$

Obs. Aproximação normal apresentada anteriormente.

2 - Estatística Escore

$$S = \frac{\frac{x}{\pi} - \frac{n-x}{1-\pi}}{\sqrt{\left(\frac{n}{\pi(1-\pi)}\right)}} = \frac{x - n\pi}{\sqrt{n\pi(1-\pi)}} \sim N(0,1)$$

Exemplo: $S = \frac{2}{\sqrt{5}} = W$

Escore = Wald

Obs.: Não foi necessário utilizar o EMV na estatística S.

3 - Estatística RV

$$RV = -2log\frac{L(\pi)}{L(\widehat{\pi})} = 2(l(\widehat{\pi}) - l(\pi)) \sim \chi_1^2$$
$$RV = 2[xlog(\frac{x}{n}) + (n-x)log\frac{n-x}{n} - xlog(\pi) - (n-x)log(1-\pi)]$$

$$RV = 2[xlog(\frac{x}{n\pi}) + (n-x)log(\frac{n-x}{n(1-\pi)})]$$

Exemplo
$H_0 : \pi = \frac{1}{2}$

x=12

n=20

$$RV = 2[12log(\frac{12}{10}) + 8log(\frac{8}{10})] = 0,805 \Rightarrow valor - p = 0,369$$

$$IC(\pi, 95\%) = (\pi| - 2\log\left\{\frac{L(\pi)}{L(0,6)}\right\} \leq 3,84)$$

Exemplo

Resumo

Teste	Valor-p
Exato	0,5034
$MC \approx Boostrap$	0,5
$Qui - Quadrado = Aprox.Normal$	0,371
W=S	0,371
RV	0,369

OBS: Podemos mostrar que, neste caso, Qui-Quadrado = W = S = Ap. normal.

Função de Verossimilhança Aproximada

- Modelo
$$X \sim N(n\pi, n\pi(1-\pi))$$
- $X = 12$
- Método: Máxima Verossimilhança.

Verossimilhança aproximada

$X \sim bin(n, \pi)$ pode ser aproximada por

$X \sim N(n\pi, n\pi(1-\pi))$, em que, $\sigma^2 = n\pi(1-\pi)$ e $\mu = n\pi$

$$L(\pi) \propto \frac{1}{\sigma} exp[-\frac{1}{2}\frac{(x-n\pi)^2}{\sigma^2}] = \frac{1}{\sqrt{n\pi(1-\pi)}} exp[-\frac{1}{2}\frac{(x-n\pi)^2}{n\pi(1-\pi)}]$$

$$l(\pi) = logL(\pi) \propto -\frac{1}{2}log(n\pi(1-\pi)) - \frac{(x-n\pi)^2}{2n\pi(1-\pi)}$$

Como $x - n\hat{\pi} = 0$ temos $\hat{\pi} = \frac{x}{n}$

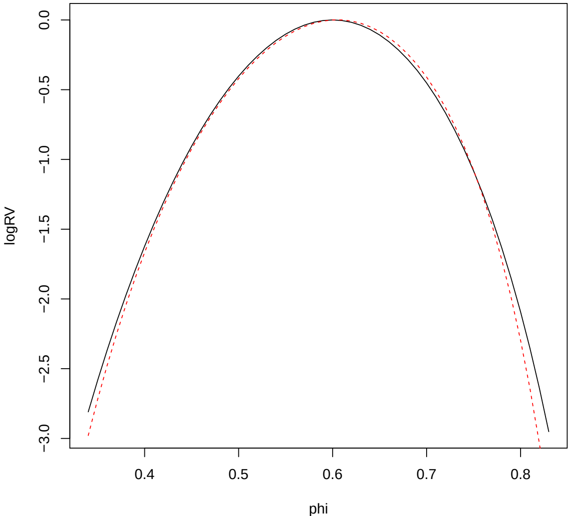
$$l(\hat{\pi}) = -\frac{1}{2}log(n\hat{\pi}(1-\hat{\pi}))$$

$$log\frac{L(\pi)}{L(\hat{\pi})} = l(\pi) - l(\hat{\pi}) = -\frac{1}{2}logn\pi(1-\pi) - \frac{(x-n\pi)^2}{2n\pi(1-\pi)} + \frac{1}{2}logn\hat{\pi}(1-\hat{\pi})$$

Quão boa é esta aproximação?

De forma conservadora x e $n - x$ devem ser ≥ 10

Vermelho: Aproximação normal e Preto: verdadeira (binomial)



Intervalo de Confiança da RV

- Intervalo de $(1-\alpha)100\%$ de confiança

$$-2log\frac{L(\pi)}{L(\hat{\pi})} = \chi^2_{1;1-\alpha} \Rightarrow log\frac{L(\pi)}{L(\hat{\pi})} = -\frac{\chi^2_{1;0,95}}{2} = -\frac{1,96^2}{2} = -1,921$$

- Exemplo: x=12 e n=20 e 95% de confiança

$-\frac{1,96^2}{2} = -1,921$ Usual:

$$\hat{\pi} \pm 1,96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0,6 \pm 1,96\sqrt{\frac{0,6\times0,4}{20}} : (0,385; 0,814)$$

Intervalo de confiança

1. Verossimilhança verdadeira

   $$logRV = xlog(\frac{n\pi}{x}) + (n-x)log(\frac{n(1-\pi)}{n-x})$$

   $$logRV = 12log(\frac{20\pi}{12}) + 8log(\frac{20(1-\pi)}{8}) = -1,921$$

   (0,383;0,793)

2. Verossimilhança aproximada

   $$logRV = \frac{1}{2}[-log20\pi(1-\pi) + \frac{(12-20\pi)^2}{20\pi(1-\pi)} + log12(0,4)]$$

   (0,386;0,778)

Resumo

Intervalo confiança (Clássico/Frequentista)

Intervalo de credibilidade

π quantidade de interesse

Método	Intervalo	Valor-p
Exato	(0.361;0.809)	0.5034
$Ap.Normal \approx W \approx S \approx Qui$	(0.385;0.814)	0.371
R. Verossimilhança	(0.386;0.78)	0.369
$MC \approx$	(0.38;0.78)	0.5
Bootstrap	(0.40;0.80)	0.5
Bayesiano	(0.39;0.79)	