Inferência Bayesiana: Modelo Paramétrico Lognormal

Análise de Sobrevivência

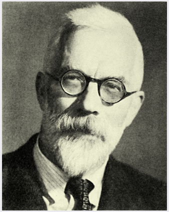
Guilherme Augusto Veloso

Prof. Dr. Enrico Antônio Colosimo - UFMG

Maio 2017

VERTENTES DA INFERÊNCIA ESTATÍSTICA

INFERÊNCIA CLÁSSICA



INFERÊNCIA BAYESIANA

Quadro Comparativo

INFERÊNCIA CLÁSSICA	INFERÊNCIA BAYESIANA
É associada à interpretação frequentista de probabilidade	A probabilidade representa uma abordagem subjetiva
Os parâmetros são considerados quantidades fixas	Os parâmetros são considerados quantidades aleatórias
A Teoria de Verossimilhança desempenha papel de destaque	Atualização da informação subjetiva via Teorema de Bayes
Intervalos de Confiança	Intervalos de Credibilidade

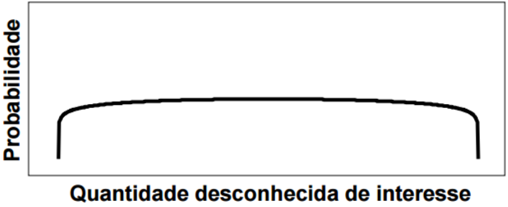
Probabilidade Subjetiva

Situação-problema:

- Qual a proporção atual de crianças com excesso de peso em Belo Horizonte?
- Vamos denominar essa proporção de θ.
- θ por ser uma proporção está entre 0 e 1. É provável que esteja próximo de 0? Próximo de 1? E de 0,25?
- Podemos construir uma distribuição de probabilidades que represente o nosso conhecimento.

Distribuição *a priori*

- A distribuição *a priori* $\pi(\theta)$ fornece a probabilidade dos diferentes valores possíveis da quantidade de interesse **ANTES** dos dados serem considerados.
- A distribuição a priori pode ser ampla, plana, uniforme (*priori* não informativa), ou pode se concentrada com um ápice se possuímos mais informação (*priori* informativa).

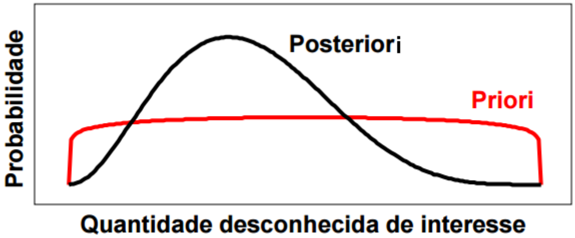


Função de Verossimilhança

- A Função de Verossimilhança $L(y|\theta)$ fornece a probabilidade de obter os dados, considerando diferentes valores possíveis da quantidade desconhecida de interesse θ.
- É calculada usando um modelo estatístico que representa o processo que produziu os dados.
- Conecta os parâmetros do modelo aos dados.

Distribuição *a posteriori*

- A distribuição *a posteriori* $\pi(\theta|y)$ fornece a probabilidade dos diferentes valores possíveis da quantidade de interesse **DEPOIS** de considerar os dados.
- Essa distribuição é combinação da *priori* (o que sabemos antes) com a Verossimilhança (o que os dados nos disseram).



Teorema de Bayes

- A informação contida na distribuição *a priori* $\pi(\theta)$ é então atualizada, através da informação dos dados contida em $L(y|\theta)$, via teorema de Bayes, levando à distribuição *a posteriori* de θ, dada por:

$$\pi(\theta|y) = \frac{L(y|\theta) \cdot \pi(\theta)}{\int\limits_{\Theta} L(y|\theta) \cdot \pi(\theta)\partial\theta}.$$

- Comumente escrevemos: $\pi(\theta|y) \propto L(y|\theta) \cdot \pi(\theta)$.

Modelo Paramétrico Exponencial

Função de Verossimilhança:

Suponha que os tempos de sobrevivência $\boldsymbol{t} = (t_1, t_2, \cdots, t_n)'$ possuem distribuição $Exp(\lambda)$, de modo que:

$$f(t_i|\lambda) = \lambda exp(-\lambda t_i), \text{ com } i = 1, 2, \cdots, n.$$

- A Função de Sobrevivência é: $S(t_i|\lambda) = exp(-\lambda t_i)$.
- Seja $D = (n, \mathbf{t}, \delta)$. A Função de Verossimilhança é:

$$L(D|\lambda) = \prod_{i=1}^{n} f(t_i|\lambda)^{\delta_i} S(t_i|\lambda)^{(1-\delta_i)} = \lambda^{\sum_{i=1}^{n} \delta_i} \exp\left(-\lambda \sum_{i=1}^{n} t_i\right).$$

Modelo Paramétrico Exponencial

Distribuições *a priori* e *a posteriori*:

- A distribuição *a priori* de λ é uma *Gama*(α, β):

$$\pi(\lambda|\alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}\exp(-\beta\lambda).$$

- Desse modo, temos que a distribuição *a posteriori* de λ é:

$$\begin{aligned}
\pi(\lambda|D) &\propto L(D|\lambda)\cdot\pi(\lambda|\alpha,\beta) \\
&\propto \left[\lambda^{\sum_{i=1}^{n}\delta_i}\exp\left(-\lambda\sum_{i=1}^{n}t_i\right)\right]\cdot\left[\lambda^{\alpha-1}\exp(-\beta\lambda)\right] \\
&= \lambda^{\alpha+\sum_{i=1}^{n}\delta_i-1}exp\left[-\left(\beta+\sum_{i=1}^{n}t_i\right)\lambda\right].
\end{aligned}$$

Modelo Paramétrico Exponencial

Distribuição *a posteriori* - Interpretação:

- Desse modo, $\pi(\lambda|D) \sim Gama\left(\alpha + \sum_{i=1}^{n} \delta_i; \beta + \sum_{i=1}^{n} t_i\right)$
- $E(\lambda|D) = \frac{\alpha+\sum_{i=1}^{n} \delta_i}{\beta+\sum_{i=1}^{n} t_i}$
- $V(\lambda|D) = \frac{\alpha+\sum_{i=1}^{n} \delta_i}{\left(\beta+\sum_{i=1}^{n} t_i\right)^2}$

Exemplo: Lâmpadas de Projetores LCD

- A falha mais comum desses projetores é na lâmpada.
- Um fabricante alega que os usuários podem esperar cerca de 1500 horas de projeção em cada lâmpada usada em condições normais de operação.

- Podemos usar essa informação para construir a distribuição *a priori* para λ?

Exemplo: Lâmpadas de Projetores LCD

Construção da Distribuição *a priori*:

- Seja$(\lambda|\alpha,\beta) \sim$ *Gama*(α,β).
- Se $\kappa = \frac{1}{\lambda}$, entao $(\kappa|\alpha,\beta) \sim$ *Gama Inversa*(α,β).
- Considerando as expressões para a média e a variância da distribuição Gama Inversa, tem-se que:

$$\begin{cases} \frac{\beta}{\alpha-1} = 1500 \\ \sqrt{\frac{\beta^2}{(\alpha-1)^2(\alpha-2)}} = 2000 \end{cases} \Rightarrow \alpha = 2,5 \text{ e } \beta = 2350.$$

- Desse modo, $\pi(\lambda|\alpha,\beta) \sim$ *Gama*$(2,5; 2350)$.

Exemplo: Lâmpadas de Projetores LCD

- Foram registradas as horas de projeção até a falha de lâmpadas idênticas que estavam em 31 projetores LCD.

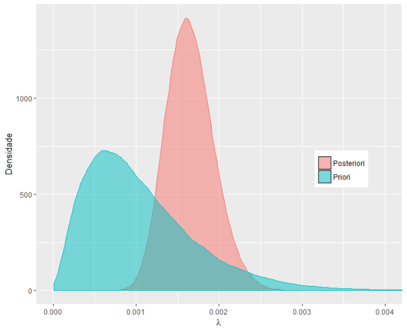
Construção da Distribuição *a posteriori*:

- $\sum_{i=1}^{n} t_i = 17907$ e $\sum_{i=1}^{n} \delta_i = n = 31$
- $\pi(\lambda|D) \sim Gama\left(\alpha + \sum_{i=1}^{n} \delta_i; \beta + \sum_{i=1}^{n} t_i\right)$

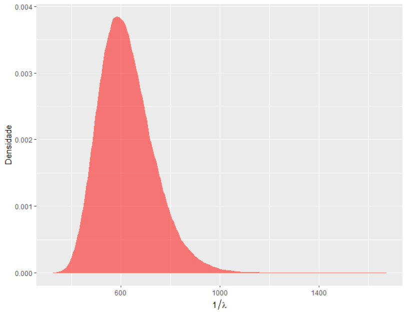
$$\sim Gama(2,5 + 31; 2350 + 17907)$$
$$\sim Gama(33,5; 20257)$$

Exemplo: Lâmpadas de Projetores LCD

Distribuições *a priori* e *a posteriori* de λ



Distribuição *a posteriori* de $\frac{1}{\lambda}$



Exemplo: Lâmpadas de Projetores LCD

Estimativas *a posteriori* para quantidades de interesse:

Parâmetro λ:

- $\hat{\lambda}_B = E(\lambda|D) = \frac{33.5}{20257} = 0.001653$
- $\hat{\lambda}_{MV} = \frac{1}{\bar{t}} = \frac{1}{577.6452} = 0.001731$
- Intervalo de Credibilidade de 90% para λ: $[0,00121; 0,00214]$

Tempo médio até a falha $\left(\frac{1}{\lambda} = \kappa\right)$:

- $\hat{\kappa} = 623,44$
- Intervalo de Credibilidade de 90% para κ : $[465,17; 824,76]$

Modelo Paramétrico Lognormal

Função de Verossimilhança:

Suponha que os tempos de sobrevivência $\boldsymbol{t} = (t_1, t_2, \cdots, t_n)'$ possuem distribuição $LN(\mu; \sigma^2)$, de modo que:

$$f(t_i|\mu,\sigma^2) = (2\pi)^{-1/2}\,(t_i\sigma)^{-1}\,exp\left[-\frac{1}{2\sigma^2}\,(log(t_i)-\mu)^2\right].$$

- A Função de Sobrevivência é: $S(t_i|\mu,\sigma^2) = 1 - \Phi\left(\frac{log(t_i)-\mu}{\sigma}\right)$.
- Seja $D = (n, \mathbf{t}, \delta)$. A Função de Verossimilhança é:

$$L(D|\mu,\sigma^2) = \prod_{i=1}^{n} f(t_i|\mu,\sigma^2)^{\delta_i} S(t_i|\mu,\sigma^2)^{(1-\delta_i)}$$

$$= \left(2\pi\sigma^2\right)^{-\frac{\sum_{i=1}^{n}\delta_i}{2}} exp\left[-\tfrac{1}{2\sigma^2}\textstyle\sum_{i=1}^{n}\delta_i\,(log(t_i)-\mu)^2\right] \times \textstyle\prod_{i=1}^{n} t_i^{-\delta_i}\left[1-\Phi\left(\frac{log(t_i)-\mu}{\sigma}\right)\right]^{1-\delta_i}.$$

Modelo Paramétrico Lognormal

Distribuições *a priori*:

- Seja $\tau = 1/\sigma^2$.
- Não existe *priori* conjugada para (μ, τ) quando ambos são quantidades desconhecidas.
- Seja $\pi(\mu, \tau) = \pi(\mu|\tau) \cdot \pi(\tau)$.
- Nesse caso, uma típica escolha de *priori* conjunta é:
  - $\pi(\mu|\tau) \sim N\left(\mu_0, \frac{1}{\tau\tau_0}\right)$;
  - $\pi(\tau) \sim Gama\left(\frac{\alpha}{2}, \frac{\beta}{2}\right)$.

Modelo Paramétrico Lognormal

Distribuição *a posteriori*:

$$
\begin{aligned}
\pi(\mu,\tau|D) &\propto L(D|\mu,\tau)\cdot\pi(\mu,\tau|\mu_0,\tau_0,\alpha,\beta)\\
&\propto L(D|\mu,\tau)\cdot\pi(\mu|\tau,\mu_0,\tau_0)\cdot\pi(\tau|\alpha,\beta)\\
&\propto \tau^{\frac{\alpha_0+\sum_{i=1}^{n}\delta_i+1}{2}-1} exp\left\{-\frac{\tau}{2}\left[\sum_{i=1}^{n}\delta_i\left(log(t_i)-\mu\right)^2+\tau_0\left(\mu-\mu_0\right)^2+\beta\right]\right\}\\
&\text{x}\quad \prod_{i=1}^{n} t_i^{-\delta_i}\left[1-\Phi\left(\tau^{\frac{1}{2}}\left(log(t_i)-\mu\right)\right)\right]^{1-\delta_i}
\end{aligned}
$$

Amostrador de Gibbs

Algoritmo:

Seja (μ, τ)os parâmetros de interesse e $\pi(\mu, \tau|D)$ a distribuição *a posteriori*. Segue o esquema básico do Amostrador de Gibbs:

PASSO 0: Escolha valores iniciais arbitrários para $\left(\mu^{(0)}, \tau^{(0)}\right)$.

PASSO 1: Gere $\left(\mu^{(i+1)}, \tau^{(i+1)}\right)$ da seguinte maneira:

- Gere $\mu^{(i+1)} \sim \pi(\mu|\tau^{(i)}, D)$;
- Gere $\tau^{(i+1)} \sim \pi(\tau|\mu^{(i+1)}, D)$.

PASSO 2: Determine $i = i + 1$ e volte ao Passo 1 até que $i = M$, com M (Número de Iterações) pré-especificado.

Exemplo: Dados de Câncer de Bexiga

Distribuições *a priori (Não-Informativas)*:

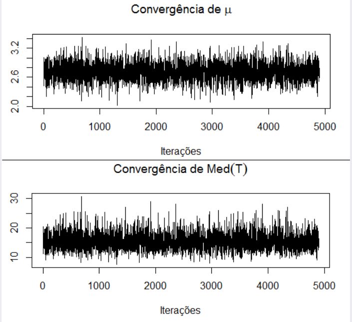
- $\pi(\mu|\tau, \mu_0, \tau_0) \sim N\left(\mu_0, \frac{1}{\tau\tau_0}\right) = N\left(0, \frac{1}{\tau 10^{-4}}\right);$
- $\pi(\tau|\alpha, \beta) \sim Gamma\left(\frac{\alpha}{2}, \frac{\beta}{2}\right) = Gamma\left(\frac{2}{2}, \frac{0{,}002}{2}\right).$

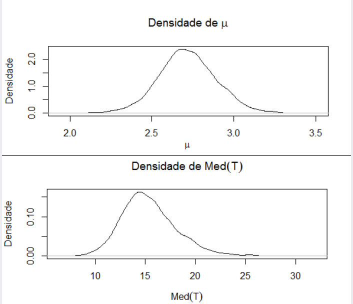
Quantidades de Interesse:

- μ;
- $\sigma = \sqrt{\frac{1}{\tau}}$;
- Tempo Mediano de Falha: $t_{0.5} = exp(\mu)$;
- Tempo Médio de Falha: $E(T) = exp\left(\mu + \frac{1}{2\tau}\right)$.

Exemplo: Dados de Câncer de Bexiga

Convergência das Cadeias e Densidade Marginal *a posteriori*





Exemplo: Dados de Câncer de Bexiga

Estimativas Pontuais:

	μ	σ	$E(T)$	$t_{0,5}$
Inferência Clássica	2,72	0,76	20,3	15,2
Inferência Bayesiana	2,72	0,75	20,8	15,4

Estimativas Intervalares 95%:

	$E(T)$		$t_{0,5}$	
	LI	LS	LI	LS
Inferência Clássica	11,9	28,9	10,0	20,4
Inferência Bayesiana	13,5	31,0	10,5	21,1

Referências

- COLOSIMO, E. e GIOLO, S., Análise de Sobrevivência Aplicada, ABE-Projeto Fisher, 2006.
- HAMADA, M. et.al., Bayesian Reliability, Springer, 2008.
- IBRAHIM, J. et.al., Bayesian Survival Analysis, Springer, 2001.