Universidade Federal do Rio Grande do Norte
Instituto Metrópole Digital
IMD0601 - Bioestatística

Inferência Bayesiana

Prof. Dr. Tetsu Sakamoto
Instituto Metrópole Digital - UFRN
Sala A224, ramal 182
Email: tetsu@imd.ufrn.br

Baixe a aula (e os arquivos)

- Para aqueles que não clonaram o repositório:

```
> git clone https://github.com/tetsufmbio/IMD0601.git
```

- Para aqueles que já tem o repositório local:

```
> cd /path/to/IMD0601

> git pull
```

Revisão

Distribuição multivariada

- **Multinomial**
- **Normal**
- **Dirichlet**

Inferência Bayesiana

Diferença essencial

- **Frequentista X Bayesiana** → Forma de interpretar a probabilidade

Jogada de moeda justa → P(cara) = ½;

- **Frequentista**
    - **Experimento repetido várias vezes → proporção de cara aproxima de ½;**
    - **Probabilidade é a média aritmética dos resultados do experimento.**
    - **Uma verdade, e uma medida fixa;**

Diferença essencial

- **Frequentista X Bayesiana** → Forma de interpretar a probabilidade

Jogada de moeda justa → P(cara) = ½;

- **Bayesiana**
  - **Não interpreta a probabilidade como um valor fixo ou verdadeiro;**
  - **Probabilidade como uma medida de grau pessoal de crença em relação a uma hipótese;**
  - **Não existe uma medida verdadeira, mas sim uma certa distribuição de probabilidade associada às suas crenças subjetivas;**
  - **A probabilidade de um resultado de uma jogada de uma moeda é uma variável aleatória que possui uma distribuição de probabilidade associada.**

Diferença essencial

- **Frequentista X Bayesiana** ➜ Forma de interpretar a probabilidade

Jogada de moeda justa ➜ P(cara) = **½;**

- **Bayesiana**
  - Utilizando o conhecimento subjetivo, tenta primeiro estipular qual a probabilidade de um parâmetro ➜ **distribuição a priori** ➜ **distribuição beta;**
  - Coleta de dados e baseado nos dados coletados o modelo é atualizado ➜ **distribuição a posteriori;**
  - Como escolher os parâmetros a priori? Como os dados são modelados? Como os modelos são atualizados?

Diferença essencial

- **Frequentista X Bayesiana** → Forma de interpretar a probabilidade

Jogada de moeda justa → P(cara) = ½;

- **Bayesiana**
  - **Pode escolher a prior que P(cara) = 1/10, e não ½ (Experiência com moedas não justas);**
  - **Um estatístico bayesiano pode não apresentar as mesmas probabilidades a priori;**
  - **Bayesiana utiliza as crenças subjetivas para produzir um modelo inicial de distribuição de probabilidade;**
  - **A incorporação de uma crença subjetiva é o que diferencia a bayesiana dos frequentistas.**

Por que aprender inferência bayesiana?

- **Menos teórico** → baseado em apenas um paradigma → Teorema de Bayes;
- **Bom aproveitamento do poder computacional** → pacotes em R que utilizam a abordagem bayesiana para realizar tarefas computacionalmente intensivas;
- **Flexibilidade** → Extremamente flexível em modelar distribuições de dados e pode ser aplicada em diversas situações;
- **Dados faltantes** → cobre efetivamente os dados faltantes;
- **Múltiplas variáveis**;
- **É um método de aprendizagem** → atualização constante do modelo;
- **Não requer número de amostras grandes** → Algoritmos como MCMC pode ser realizado para produzir resultados significativos mesmo com pequenas amostras.

Desvantagens sobre a inferência bayesiana

- **Críticas sobre a subjetividade do método** ➔ pode ser visto como não científico e enviesado. É a principal fonte de disputa entre os frequentistas e bayesianos;
- **Pouco entendido** ➔ Modelos bayesianos não são largamente ensinados e utilizados;
- **Complicado?** ➔ Quando apresentado matematicamente, a abordagem aparenta ser complicado;

Teorema de Bayes

Uma simples regra dentro da probabilidade.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad P(B|A) = \frac{P(B \cap A)}{P(A)}$$

$$P(A|B).P(B) = P(A \cap B) \qquad P(B|A).P(A) = P(B \cap A)$$

$$P(A \cap B) = P(B \cap A)$$

$$P(A|B).P(B) = P(B|A).P(A)$$

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

Teorema de Bayes

Teorema de Bayes

Variações do teorema de Bayes

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

$$P(A|B) \propto P(B|A).P(A)$$

P(B) constante, ∝ = “proporcional a”

$$P(H|E) = \frac{P(E|H).P(H)}{P(E)}$$

E = Evidência
H = Hipótese
P(H) = probabilidade a priori antes da evidência
P(H|E) = probabilidade a posteriori depois da evidência
P(E|H) = probabilidade da evidência dada hipótese

Aplicando o teorema de Bayes

Tabela abaixo mostra a probabilidade conjunta de dois eventos: “A” onde a proteína se liga na membrana e “B” onde a proteína possui alta proporção de resíduos hidrofóbicos.

		Tipo de proteína	
		Liga na membrana (A)	Não liga na membrana (A^c)
Proporção de resíduos hidrofóbicos	Alta (B)	0,3	0,2
	Baixa (B^c)	0,1	0,4

Qual a probabilidade de uma proteína ter alta proporção de resíduos hidrofóbicos dado que ela é uma proteína de membrana?

Aplicando o teorema de Bayes

Qual a probabilidade de uma proteína ter alta proporção de resíduos hidrofóbicos dado que ela é uma proteína de membrana?

		Tipo de proteína	
		Liga na membrana (A)	**Não liga na membrana (A^c)**
Proporção de resíduos hidrofóbicos	**Alta (B)**	0,3	0,2
	Baixa (B^c)	0,1	0,4

$$P(B|A) = \frac{P(A|B).P(B)}{P(A)} = \frac{P(A \cap B)}{P(A)}$$

Aplicando o teorema de Bayes

		Tipo de proteína		
		Liga na membrana (A)	**Não liga na membrana (A^c)**	**marginal**
Proporção de resíduos hidrofóbicos	**Alta (B)**	0,3	0,2	0,5
	Baixa (B^c)	0,1	0,4	0,5
	marginal	0,4	0,6	

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0{,}3}{0{,}4} = 0,75$$

Aplicando o teorema de Bayes

Suponha que ter um gene X resulta na manifestação de uma doença em 50% dos casos. Suponha que a prevalência deste gene em uma população é de 1/1000 e que a prevalência da doença é de 1%. Sabendo disso, calcule a probabilidade de ter o gene X dado que você tem a doença.

Aplicando o teorema de Bayes

Suponha que ter um gene X resulta na manifestação de uma doença em 50% dos casos. Suponha que a prevalência deste gene em uma população é de 1/1000 e que a prevalência da doença é de 1%. Sabendo disso, calcule a probabilidade de ter o gene X dado que você tem a doença.

$$P(gene\ X) = 1/1000$$
$$P(doença) = 1/100$$
$$P(doença|gene\ X) = 0,5$$
$$P(gene\ X|doença) = ?$$

$$P(gene\ X|doença) = \frac{P(doença|gene\ X)P(gene\ X)}{P(doença)}$$

$$P(gene\ X|doença) = \frac{0,5.1/1000}{1/100} = 0,05$$

Estendendo o teorema de Bayes para distribuição

"Não existe uma medida verdadeira, mas sim uma certa distribuição de probabilidade associada às suas crenças subjetivas" ➜ Estatísticos bayesianos enxergam as probabilidades na forma de distribuição.

$$P(H|E) = \frac{P(E|H).P(H)}{P(E)}$$

$$P(H|E) \propto P(E|H).\, P(H)$$

Probabilidade a posteriori	Verossimilhança	Probabilidade a priori
Distribuição a posteriori	Distribuição dos dados	Distribuição a priori

Transformar as probabilidades em distribuição

Estendendo o teorema de Bayes para distribuição

$$P(H|E) \propto P(E|H).P(H)$$

Voltando ao exemplo da moeda.

P(H) → Probabilidade a priori → Crença subjetiva da proporção de caras (vamos supor que é de ¼);

Distribuição beta → modelar distribuição de probabilidades com dados de proporção;

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

Achar valores de alfa e beta onde a distribuição fique centrada em 0.25;

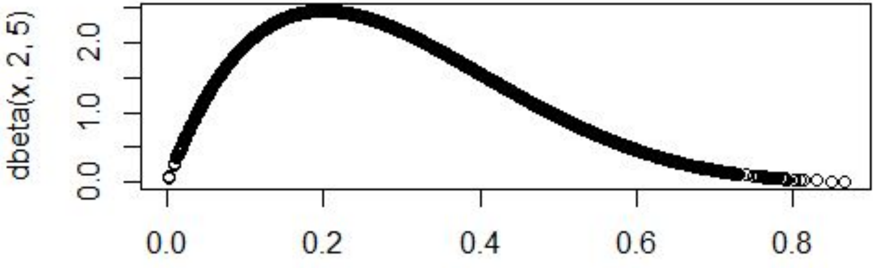
Estendendo o teorema de Bayes para distribuição

Determinar alfa e beta onde a distribuição fique centrada em 0.25;

$$\mu = \frac{\alpha}{\alpha+\beta}$$

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

Alfa = 2, Beta = 5



Estendendo o teorema de Bayes para distribuição

$$P(H|E) \propto P(E|H).P(H)$$

P(E|H) → Probabilidade da evidência dado uma hipótese;

Para obter esta probabilidade precisamos de **evidência** e evidência são os **dados**.

Para obter dados, precisamos realizar **experimento**: Jogadas de 10 moedas:

- cada jogada é um Bernoulli {cara, coroa}.
- Combinação das 10 jogadas → distribuição binomial.

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Estendendo o teorema de Bayes para distribuição

Para obter dados, precisamos realizar **experimento**: Jogadas de 10 moedas:

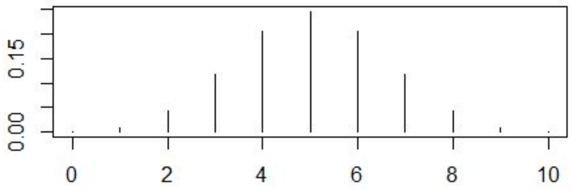
- cada jogada é um Bernoulli {cara, coroa}.
- Combinação das 10 jogadas ➜ distribuição binomial.

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Suponha que das 10 jogadas, 5 deram cara.

- Frequentista ➜ P(cara) = p = proporção de caras = 0.5

$$P(k) = \binom{10}{k} 0.5^k (1-0.5)^{10-k}$$



Estendendo o teorema de Bayes para distribuição

Suponha que das 10 jogadas, 5 deram cara.

- Bayesiana ➜ P(cara)=p é uma variável aleatória ➜ infinitos modelos plausíveis, um para cada valor de p.

$$P(5\ caras\ em\ 10|p) = \binom{10}{5} p^5 (1-p)^5$$

Probabilidade de um resultado dado um p ➜ **Verossimilhança(p)**

Estendendo o teorema de Bayes para distribuição

$$P(H|E) \propto P(E|H).\,P(H)$$

$$P(H) = \frac{1}{\beta(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

$$P(H) = \frac{1}{\beta(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

$$P(E|H) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$P(H|E) \propto \binom{n}{k} p^k (1-p)^{n-k}.\, \frac{1}{\beta(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

Estendendo o teorema de Bayes para distribuição

$$P(H|E) \propto P(E|H).P(H)$$

$$P(H|E) \propto \binom{n}{k} p^k (1-p)^{n-k} . \frac{1}{\beta(\alpha,\beta)} p^{\alpha-1} (1-p)^{\beta-1}$$

$a^m.a^n = a^{m+n}$

$$P(H|E) \propto \binom{n}{k} \frac{1}{\beta(\alpha,\beta)} p^{k+\alpha-1} (1-p)^{n-k+\beta-1}$$

Considerar constante como c

$$P(H|E) \propto c p^{k+\alpha-1} (1-p)^{n-k+\beta-1}$$

Estendendo o teorema de Bayes para distribuição

$$P(H|E) \propto cp^{k+\alpha-1}(1-p)^{n-k+\beta-1}$$

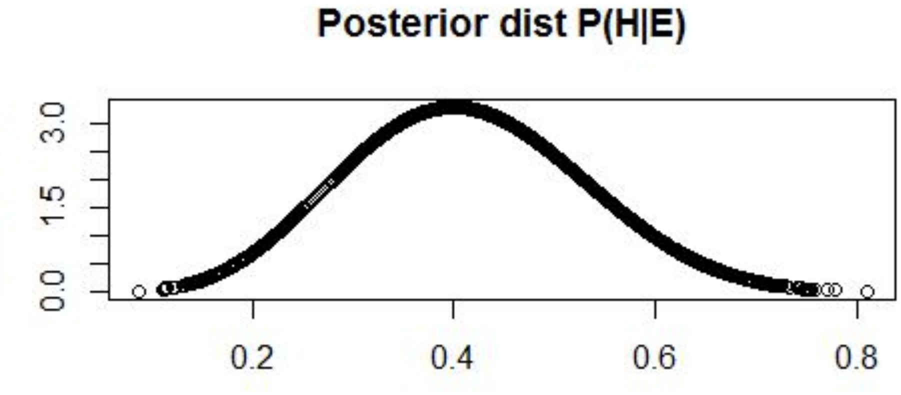
Segue uma distribuição beta:

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

onde:

- **alfa2 = k + alfa** ➔ 5 + 2 = 7
- **beta2 = n - k + beta** ➔ 10 - 5 + 5 = 10

Distribuição posteriori = beta(7, 10)



Estendendo o teorema de Bayes para distribuição

$$P(H|E) \propto cp^{k+\alpha-1}(1-p)^{n-k+\beta-1}$$

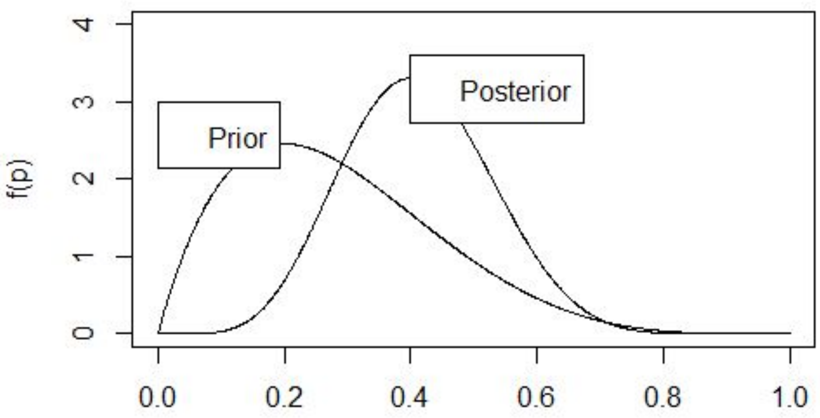
Segue uma distribuição beta:

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

onde:

- **alfa2** = **k** + **alfa** ➔ 5 + 2 = 7
- **beta2** = **n** - **k** + **beta** ➔ 10 - 5 + 5 = 10

Distribuição posteriori = beta(7, 10)



Atualizando o modelo com novos dados

Exercício:

Vamos supor que você tenha feito um novo experimento com 10 jogadas de moedas e chegou nos mesmos resultados. Utilize o modelo posteriori calculado anteriormente como modelo priori e determine um novo modelo posteriori.

Atualizando o modelo com novos dados

Exercício:

Vamos supor que você tenha feito um novo experimento com 10 jogadas de moedas e chegou nos mesmos resultados (5 caras em 10 jogadas). Utilize o modelo posteriori calculado anteriormente como modelo priori e determine um novo modelo posteriori.

$$P(H|E) \propto cp^{k+\alpha-1}(1-p)^{n-k+\beta-1}$$

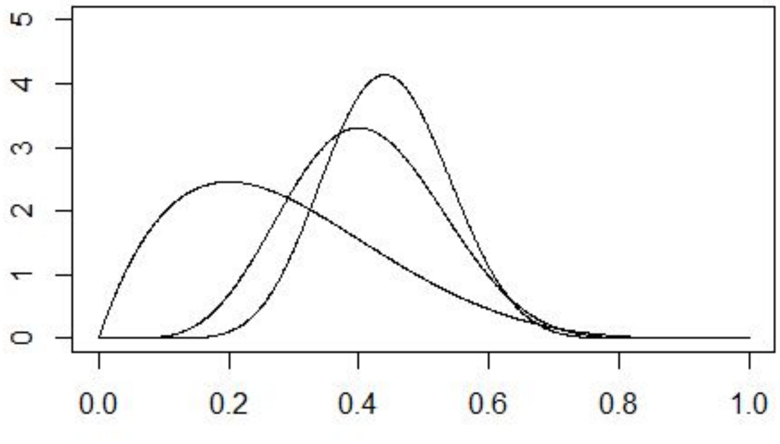
- **alfa2 = k + alfa** ➔ 5 + 7 = 12
- **beta2 = n - k + beta** ➔ 10 - 5 + 10 = 15

Atualizando o modelo com novos dados

Exercício:

Vamos supor que você tenha feito um novo experimento com 10 jogadas de moedas e chegou nos mesmos resultados (5 caras em 10 jogadas). Utilize o modelo posteriori calculado anteriormente como modelo priori e determine um novo modelo posteriori.

$$P(H|E) \propto cp^{k+\alpha-1}$$



- **alfa2** = **k** + **alfa** → 5 + 7 = 12
- **beta2** = **n** - **k** + **beta** → 10 - 5 + 10 = 15

Essência do algoritmo bayesiano

$$P(Modelo|dados) \propto P(Dados|modelo)P(Modelo)$$

$$Posteriori \propto Verossimilhança.Priori$$

"Modelo subjetivo a priori é atualizado por um modelo de dados (verossimilhança) para produzir uma distribuição posteriori, combinando a subjetividade do conhecimento do pesquisador com a objetividade dos dados observados."