Exemplo de Inferência bayesiana

"What is Bayesian statistics and why everything else is wrong."

Michael Lavine

Exemplo

Câncer na escola Slater. Paul Brodeur in the *New Yorker* in Dec. 1992.

→ A escola Slater é uma escola elementar onde os empregados estão preocupados com a hipótese de que a alta incidência de câncer possa se dever à proximidade com a linha de transmissão de alta voltagem.

Fatos

→ 8 casos de câncer entre os 145 empregados com idade entre 40 e 44 anos.

→ baseado na taxa de câncer nacional, (aproximadamente 3/100), o número de casos esperado seria de 4.2

Suposições:

1) Os empregados desenvolveram câncer independentemente.

2) A chance de ter câncer, θ, era a mesma para todos os 145 empregados.

Assim, o número de casos de câncer segue uma distribuição binomial $X \sim \text{Bin}(145, \theta)$

Como é que cada uma das 4 teorias simplificadas explica os dados?

Teoria A: $\theta = .03$

Teoria B: $\theta = .04$

Teoria C: $\theta = .05$

Teoria D: $\theta = .06$

A verossimilhança das teorias A-D

Para comparar estas teorias devemos ver como cada uma delas explica os dados. Isto é, para cada θ, usa-se os resultados elementares da binomial para calcular:

$$\Pr(X = 8 \mid \theta) = \binom{145}{8} \theta^{8} (1-\theta)^{137}$$

teoria A: $\Pr(X = 8 \mid \theta = .03) \approx .036$

teoria B: $\Pr(X = 8 \mid \theta = .04) \approx .096$

teoria C: $\Pr(X = 8 \mid \theta = .05) \approx .134$

teoria D: $\Pr(X = 8 \mid \theta = .06) \approx .136$

Isto dá uma taxa aproximada de 1:3:3:4. Portanto , a teoria B explica os dados cerca de 3 vezes melhor que a teoria A.

A verossimilhança

$$\Pr(X \mid \theta, 145) = \binom{145}{X} \theta^{X} (1-\theta)^{145-X}$$

Inicialmente, Pr(X | θ) é função de duas variáveis: X e θ.

Uma vez que foi observados X = 8, então Pr(X | θ) descreve quão bem cada teoria, ou valor de θ explica os dados. Outros valores de X não são relevantes.

O princípio da verossimilhança diz que desde que X tenha sido observado, digamos X = 8, então nenhum outro valor de X importa e devemos tratar Pr(X | θ) simplesmente como Pr(X = 8 | θ).

O princípio da verossimilhança é fundamental no pensamento Bayesiano.

Uma análise Bayesiana

Existem outras fontes de informação sobre se o câncer pode ser induzido pela proximidade com a alta voltagem.

- Alguns cientistas mostram correlações positivas entre câncer e proximidade.

- Outros não encontram estas correlações e os físicos e biólogos acreditam que a energia nos campos magnéticos associados com alta voltagem é muito pequeno para ter efeito sobre as células.

Se julgarmos que os prós e contras são igualmente confiáveis, A teoria A (sem efeito) é tão aceitável quanto as Teorias B, C, e D juntas, e se acreditarmos que as teorias B, C, e D são igualmente prováveis.

$Pr(A) \approx .5 \approx Pr(B) + Pr(C) + Pr(D)$

e, $Pr(B) \approx Pr(C) \approx Pr(D) \approx 1/6$

Estas quantidades vão ser nossas priores.

Teorema de Bayes

Baseado na definição de probabilidade condicional sabemos que:

$$\Pr(A \mid X = 8) = \frac{\Pr(\mathrm{A\ e\ X} = 8)}{\Pr(\mathrm{X} = 8)}$$

$$\Pr(A \mid X = 8) = \frac{\Pr(\mathrm{A})\Pr(\mathrm{X} = 8 \mid \mathrm{A})}{\Pr(\mathrm{A})\Pr(\mathrm{X} = 8 \mid \mathrm{A}) + \Pr(\mathrm{B})\Pr(\mathrm{X} = 8 \mid \mathrm{B}) + \Pr(\mathrm{C})\Pr(\mathrm{X} = 8 \mid \mathrm{C}) + \Pr(\mathrm{D})\Pr(\mathrm{X} = 8 \mid \mathrm{D})}$$

$$\Pr(A \mid X = 8) = \frac{(1/2)(.036)}{(1/2)(.036) + (1/6)(.096) + (1/6)(.134) + (1/6)(.136)}$$

$$\Pr(A \mid X = 8) = 0.23$$

Igualmente,

Pr(B | X = 8) = .21

Pr(C | X = 8) = .28

Pr(D | X = 8) = .28

De acordo com estes resultados deve-se dizer que cada uma das quatro teorias é igualmente provável e, que as chances são de 3:1 de que a taxa de a incidência de câncer na escola seja maior que .03

Análise Freqüentista

Estatísticos Clássicos calculam o p-value para testar a hipótese de que Ho: $\theta = .03$ contra a hipótese alternativa. O valor de p é definido como a probabilidade de se observar um valor tão ou mais alto que o observado, dado Ho seja verdadeira.

i.e. para o problema em questão:

$$\text{p-value} = \Pr(X=8 \mid \theta = .03) + \Pr(X=9 \mid \theta = .03) + \Pr(X=10 \mid \theta = .03)$$

$$+ \ldots + \Pr(X=145 \mid \theta = .03)$$

$$\approx .07$$

Observe que utilizando a abordagem clássica, nós poderíamos rejeitar a hipótese nula de que não existe efeito das linhas de energia na escola.

A análise bayesiana revelou que a probabilidade que $\Pr(\theta > .03) \approx .77$

O que não seria suficiente para rejeitar a hipótese nula.

Críticas aos P-values

Bayesianos dizem que o p-value não pode ser usado para comparar hipóteses porque:

1) as hipóteses devem ser comparadas com base no quão bem elas explicam os dados.

2) Os p-value não consideram como a hipótese alternativa explica os dados.

3) Os valores de p-value são irrelevantes porque eles não informam quão bem cada hipóteses explica os dados.

Em resumo, os p-value não obedecem ao princípio da verossimilhança porque ele usa $Pr(X=x|\theta)$ para outros valores de x além do observado $x=8$.

O mesmo é verdade para toda a classe de testes de hipóteses e para os intervalos de confiança.

Criticas da abordagem Bayesiana

1) **Os resultados são subjetivos.** Com poucos dados as estimativas dos parâmetros podem ser sensíveis às priores escolhidas No caso da Escola poderíamos ter um grande leque de resultados dependendo da priori.

   Resposta dos Bayesianos: Os bayesianos usam “priores difusas,” e análises de sensibilidade para diminuir a influência das priores nos seus resultados.

2) **As análises bayesiana não tem uma base filosófica clara.** Os bayesianos tratam θ como uma variável aleatória enquanto para os clássicos θ é fixo. Os estatísticos clássicos acreditam que θ é ou não é igual a .03.

   Resposta dos Bayesianos: Tratar θ como aleatório não necessariamente significa que θ seja aleatório. Estamos na verdade expressando nossa incerteza sobre θ.