Introdução a Inferência Bayesiana

Pedro Luciano de Oliveira Gomes

Centro de Ciências Exatas e da Terra
Universidade Federal do Rio Grande do Norte - UFRN
Departamento de Estatística
Seminário de Tópicos Especias - STE

13 de maio de 2015

Sumário

Sumário

Introdução

A informação que se tem sobre uma quantidade de interesse θ é fundamental na Estatística. O verdadeiro valor de θ é desconhecido e a idéia é tentar reduzir esse desconhecimento. Além disso, a intensidade da incerteza a respeito de θ pode assumir diferentes graus. Do ponto vista Bayesiano, estes diferentes graus de incerteza são representados através de modelos probabilísticos para θ.

Sumário

Teorema de Bayes

Considere uma quantidade de interesse desconhecida θ. A informação de que dispomos sobre θ, resumida probabilisticamente através de $p(\theta)$, pode ser aumentada observando uma quantidade aleatória X relacionada com θ e ,portanto, $p(x|\theta)$ define esta relação.

Teorema de Bayes

O teorema de Bayes é a regra de atualização utilizada para quantificar este aumento na iformação,

$$p(\theta|x) = \frac{p(\theta, x)}{p(x)} = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(\theta, x)d\theta}$$

Note que $\frac{1}{p(x)}$ não depende de θ e vai funcionar como uma constante normalizadora de $p(\theta|x)$.

Sumário

Distribuições a priori e a posteriori

Para um valor fixo de x, a função $Lp(\theta; x) = p(x|\theta)$ fornece a verossimilhança de cada um dos possíveis valores de θ enquanto $p(\theta)$ é chamada de distribuição a *priori* de θ.
Estas duas informações combinadas nos levam à distribuição a posteriori de θ, $p(\theta|x)$.

Distribuições a priori e a posteriori

Feito isso, temos a forma usual do teorema de Bayes:

$$p(\theta|x) \propto L(\theta; x)p(\theta).$$

Distribuição a posteriori $\propto$ verossimilhança x distribuição a priori.

Distribuições a priori e a posteriori

A constante normalizadora da posteriori pode ser facilmente recuperada pois $p(\theta|x) = kp(x|\theta)p(\theta)$ onde

$$K^{-1} = \int p(x|\theta)p(\theta)d\theta = E_\theta[p(X|\theta)] = p(x)$$

é chamada de distribuição preditiva. Esta é a distribuição esperada para a observação x dado θ. Assim,

- Antes de observar X podemos checar a adequação da priori fazendo predições via $p(x)$.
- Se X observado recebia pouca probabilidade preditiva então o modelo deve ser questionado.

Distribuições a priori e a posteriori

Suponha que estamos interessados na previsão de uma quantidade Y, também relacionado com θ, visto que já observamos uma quantidade X. E esta é descrita probabilisticamente por $p(y|\theta)$ então

$$
\begin{aligned}
p(y|\theta) = \int p(y, \theta|x) d\theta &= \int p(y|\theta, x) p(\theta|x) d\theta \\
&= \int p(y|\theta) p(\theta|x) d\theta
\end{aligned}
$$

onde a última igualdade se deve a independência entre X e Y condicionado em θ.

Distribuições a priori e a posteriori

Os conceitos de priori e posteriori são relativos àquela observação que está sendo considerada no momento. Assim, $p(\theta|x)$ é a posteriori de θ em relação a X (que já foi observado) mas é a priori de θ em relação a Y (que não foi observado ainda). Após oberserva $Y = y$ uma nova posteriori (relativa a $X = x$ e $Y = y$) é obtida aplicando-se novamente o teorema de Bayes.

Distribuições a priori e a posteriori

No entanto a posteriori final não depende da ordem em que as observação x e y foram processadas. Prova:

$$p(\theta|x_1) \propto L_1(\theta; x_1)p(\theta)$$

$$p(\theta|x_2, x_1) \propto L_2(\theta; x_2)p(\theta|x_1)$$

$$\vdots$$

$$p(\theta|x_n, x_{n-1}, ..., x_1) \propto L_n(\theta; x_n)p(\theta|x_{n-1}, ..., x_1)$$

Sumário

Exemplo 1.1

(Gamerman e Migon, 1993) Um médico, ao examinar uma pessoa, "desconfia" que ela possa ter uma certa doença. Baseado em sua experiência e nas informações dada pelo paciente ele assume que a probabilidade deste ter a doença é de 0.7. A quantidade de interesse desconheçida é o indicador da doença

$$\theta = \begin{cases} 1, & \text{se o paciente tem a doença,} \\ 0, & \text{se o paciente não tem a doença.} \end{cases}$$

Exemplo 1.1

Para aumentar sua quantidade de informação sobre a doença o médico aplica um teste X relacionado a θ através da distribuição

$$P(X = 1|\theta = 0) = 0.40 \quad \text{e} \quad P(X = 1|\theta = 1) = 0.95$$

O resultado do teste foi positivo $(X = 1)$.

Exemplo 1.1

Dado este resultado, usaremos o teorema de Bayes para quantificar esse aumento.

$$P(\theta = 1|X = 1) \propto L(\theta = 1; X = 1)p(\theta = 1) = (0.95)(0.7) = 0.665$$

$$P(\theta = 0|X = 1) \propto L(\theta = 0; X = 1)p(\theta = 0) = (0.40)(0.3) = 0.120.$$

Exemplo 1.1

A constante normalizadora é tal que

$$P(\theta = 1|X = 1) + P(\theta = 0|X = 1) = 1$$

Logo,

$$k(0.665) + k(0.120) = 1$$

e $k = \frac{1}{0.785}$. Portanto, a distribuição a posteriori de θ é

$$P(\theta = 1|X = 1) = \frac{0.665}{0.785} = 0.847$$

$$P(\theta = 0|X = 1) = \frac{0.120}{0.785} = 0.153$$

Exemplo 1.1

Agora o médico aplica outro teste Y cujo resultado está relacionado a θ através da seguinte distribuição

$$P(Y=1|\theta=0)=0.04 \quad \text{e} \quad P(Y=1|\theta=1)=0.99$$

Exemplo 1.1

Antes de observar o resultado do teste é interessante obter sua distribuição preditiva. Como θ é uma quantidade discreta segue que

$$p(y|x) = \sum_{\theta} p(y|\theta)p(\theta|x)$$

note que $p(\theta|x)$ é a priori em relação a Y. Assim,

$$\begin{aligned} P(Y=1|X=1) &= P(Y=1|\theta=0)P(\theta=0|X=1) \\ &+ P(Y=1|\theta=1)P(\theta=1|X=1) \\ &= (90.04)(0.153) + (0.99)(0.847) = 0.845 \\ P(Y=0|X=1) &= 1 - P(Y=1|X=1) = 0.155 \end{aligned}$$

Exemplo 1.1

Sabendo que o resultado do teste Y foi negativo, quantificaremos essa informação usando o teorema de Bayes,

$$
\begin{aligned}
P(\theta = 1|X = 1, Y = 0) &\propto L(\theta = 1; Y = 0)P(\theta = 1|X = 1) \\
&\propto (0.01)(0.847) = 0.0085 \\
P(\theta = 0|X = 1, Y = 0) &\propto L(\theta = 0; Y = 0)P(\theta = 0|X = 1) \\
&\propto (0.96)(0.153) = 0.1469.
\end{aligned}
$$

Exemplo 1.1

A constante normalizadora é

$$k = \frac{1}{(0.0085 + 0.1469)} = \frac{1}{0.1554}$$

e assim a distribuição a posteriori de θ é

$$\begin{aligned} P(\theta = 1 | X = 1, Y = 0) &= \frac{0.0085}{0.1554} = 0.055 \\ P(\theta = 0 | X = 1, Y = 0) &= \frac{0.1469}{0.1554} = 0.945. \end{aligned}$$

Exemplo 1.1

Vejamos como a probabilidade da doença se alterou ao longo dos exames

$$P(\theta = 1) = \begin{cases} 0.7, & \text{antes dos testes,} \\ 0.847, & \text{após o teste X,} \\ 0.055, & \text{após X e Y.} \end{cases}$$

Exemplo 1.1

Notemos que o valor observado de Y recebia pouca probabilidade preditiva. Isto pode levar o médico a repensar o modelo,

(i) Será que $P(\theta = 1) = 0.7$ é uma priori adequada?

(ii) Será que as distribuições amostrais de X e Y estão corretas? O teste X é tão inexpressivo e Y é realmente tão poderoso?

Sumário

Resultados

Tabela: Tabela de prioris

	Priori		
Posteriori	0.3	0.5	0.7
$P(\theta = 1 \mid X = 1)$	0.504	0.703	0.847
$P(\theta = 0 \mid X = 1)$	0.495	0.296	0.153
$P(\theta = 1 \mid X = 1, Y = 0)$	0.010	0.0241	0.055
$P(\theta = 0 \mid X = 1, Y = 0)$	0.989	0.9756	0.945

Sumário

Conclusão

Vejamos que é possível inferir sobre um parâmetro mesmo sem ter observado uma amostra, o contrário do que aconteçe em inferência clássica.

Sumário

Referências

EHLERS, Ricardo S..**Introdução a Inferência Bayesiana**, Paraná, v. 1, p.2-8, jun. 2003.