Probabilidad y estadística para la Inteligencia Artificial, clase 7

Camilo Enrique Argoty Pulido

Especialización en Inteligencia Artificial

18 de noviembre de 2024

Inferencia bayesiana

Distribuciones a priori y a posteriori

Desde el enfoque bayesiano, la función de densidad de una variable representa una «creencia»sobre una variable Θ.

Distribuciones a priori y a posteriori

Desde el enfoque bayesiano, la función de densidad de una variable representa una «creencia»sobre una variable Θ.

De acuerdo a lo anterior, es natural que nuestra creencia sobre una variable cambie según la información que entreguen los datos disponibles.

Distribuciones a priori y a posteriori

Desde el enfoque bayesiano, la función de densidad de una variable representa una «creencia»sobre una variable Θ.

De acuerdo a lo anterior, es natural que nuestra creencia sobre una variable cambie según la información que entreguen los datos disponibles.

De aquí que en el proceso de inferencia bayesiana tenemos dos densidades distintas para Θ:

A priori: Es la creencia inicial, antes de la exposición a los datos. Se la denota por $\pi(\theta)$

Distribuciones a priori y a posteriori

Desde el enfoque bayesiano, la función de densidad de una variable representa una «creencia»sobre una variable Θ.

De acuerdo a lo anterior, es natural que nuestra creencia sobre una variable cambie según la información que entreguen los datos disponibles.

De aquí que en el proceso de inferencia bayesiana tenemos dos densidades distintas para Θ:

A priori: Es la creencia inicial, antes de la exposición a los datos. Se la denota por $\pi(\theta)$

A posteriori: Es la creencia modificada, una vez se tiene esposición a los datos. Esta es una densidad condicional, que depende de los datos. Por esto, se denota por $f_{\Theta|\overline{X}=\bar{x}}(\theta)$

Teorema de Bayes de nuevo

Recordemos que, en el teorema de Bayes, si (B_i) es una partición de Ω y $A \subseteq \Omega$, entonces dado un i fijo:

Teorema de Bayes de nuevo

Recordemos que, en el teorema de Bayes, si (B_i) es una partición de Ω y $A \subseteq \Omega$, entonces dado un i fijo:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_j P(A|B_j)P(B_j)}$$

Generalización del Teorema de Bayes

Para el caso de la inferencia bayesiana, esto se puede generalizar a lo siguiente:

Generalización del Teorema de Bayes

Para el caso de la inferencia bayesiana, esto se puede generalizar a lo siguiente:

Theorem

Dada una muestra aleatoria $\overline{X} = (X_1, \ldots, X_n)$ *con valores observados* $\bar{x} = (x_1, \ldots, x_n)$*, y dada una densidad a priori* $\pi(\theta)$*, la densidad a posteriori está dada por:*

Generalización del Teorema de Bayes

Para el caso de la inferencia bayesiana, esto se puede generalizar a lo siguiente:

Theorem

Dada una muestra aleatoria $\overline{X} = (X_1, \dots, X_n)$ *con valores observados* $\bar{x} = (x_1, \dots, x_n)$, *y dada una densidad a priori* $\pi(\theta)$, *la densidad a posteriori está dada por:*

$$f_{\Theta|\overline{X}=\bar{x}}(\theta) = \frac{f_{\overline{X}|\Theta=\theta}(\bar{x})\pi(\theta)}{\int f_{\overline{X}|\Theta=\theta}(\bar{x})\pi(\theta)d\theta}$$

Distribuciones conjugadas

El cálculo anterior puede ser demasiado engorroso. Para hacerlo más fácil, suelen usarse parejas de distribuciones **conjugadas**. Esto quiere decir, que si Θ representa un parámetro dentro de una distribución, si $\pi(\theta)$ tiene cierta forma conocida, la distribución a posteriori también tendrá una forma conocida.

Distribuciones conjugadas

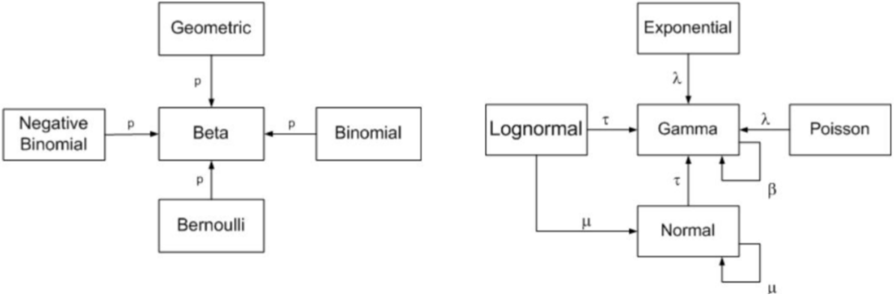
El cálculo anterior puede ser demasiado engorroso. Para hacerlo más fácil, suelen usarse parejas de distribuciones **conjugadas**. Esto quiere decir, que si Θ representa un parámetro dentro de una distribución, si $\pi(\theta)$ tiene cierta forma conocida, la distribución a posteriori también tendrá una forma conocida.

El siguiente diagrama muestra el comportamiento de las distintas distribuciones conjugadas:

Distribuciones conjugadas

El cálculo anterior puede ser demasiado engorroso. Para hacerlo más fácil, suelen usarse parejas de distribuciones **conjugadas**. Esto quiere decir, que si Θ representa un parámetro dentro de una distribución, si $\pi(\theta)$ tiene cierta forma conocida, la distribución a posteriori también tendrá una forma conocida.

El siguiente diagrama muestra el comportamiento de las distintas distribuciones conjugadas:



Ejemplo

Example

La cantidad de accidentes semanales en una planta industrial se distribuye Poisson de media μ. En una muestra de 100 semanas se observaron las siguientes frecuencias:

Cantidad de accidentes	0	1	2	3	4	5
Frecuencia (en semanas)	10	29	25	17	13	6

A priori, μ tiene una distribución exponencial de media 2. Hallar la distribución a posteriori de μ.