

IntML2019 - Teoria bayesiana - Discriminantes y Errores

April 30, 2019

Copyright (c) 2019 Georgina Flesia

<http://www.famaf.proed.unc.edu.ar/course/view.php?id=470>

1 IntML2019 - Teoria Bayesiana - Discriminantes y Errores

2 Ejercicio 1.

En el caso de dos categorías, la regla de decisión de Bayes el error condicional esta dada por la ecuacion (7)

$$P(error|x) = \min[P(\omega_1|x), P(\omega_2|x)]$$

. Incluso si las densidades a posteriori son continuas, esta forma de la condicional de errores casi siempre conduce a un integrando discontinuo en el calculo del error total en la ecuacion (5)

$$P(error) = \int P(error|x)p(x)dx$$

.

- (a) Demostrar que para densidades arbitrarias, podemos obtener una cota superior para el error total usando en la ecuacion anterior

$$P(error|x) = 2P(\omega_1|x)P(\omega_2|x)$$

.

- (b) Demostrar que si en la ecuacion (5), utilizamos $P(error|x) = \alpha P(\omega_1|x)P(\omega_2|x)$, con $\alpha < 2$, entonces no podemos garantizar que la integral de una cota superior para el error.
- (c) Analogamente, demostrar que podemos utilizar

$$P(error|x) = P(\omega_1|x)P(\omega_2|x)$$

y obtener una cota inferior para el error total.

- (d) Demostrar que si $P(error|x) = \beta P(\omega_1|x)P(\omega_2|x)$ con $\beta > 1$, entonces no podemos garantizar que la integral de una cota inferior para el error.

3 Ejercicio 2.

Suponga que se tienen dos variables independiente con igual funcion densidad

$$p(x|\omega_i) \propto e^{(-|x-a_i|/b_i)}$$

para $i = 1, 2$ and $0 < b_i$.

- (a) Escriba una expresion analitica para cada densidad, es decir normalize cada funcion para parametros a_i, b_i arbitrarios, b_i positivo.
- (b) Calcule el radio de verosimilitud como funcion de los parametros.
- (c) Graficar el radio $p(x|\omega_1)/p(x|\omega_2)$ para el caso $a_1 = 0, b_1 = 1, a_2 = 1$ y $b_2 = 2$.

4 Ejercicio 3.

Si las distribuciones condicionales para las dos categorias en el problema unidimensional son distribuciones de Cauchy

$$p(x|\omega_i) = \frac{1}{\pi b} \cdot \frac{1}{1 + (\frac{x-a_i}{b})^2}$$

- (a) Suponiendo que $P(\omega_1) = P(\omega_2)$, muestre que $P(\omega_1|x) = P(\omega_2|x)$ si $x = \frac{a_1+a_2}{2}$, es decir, el minimo para las cotas de la probabilidad de error en la decision es el punto medio entre los picos de las dos distribuciones, independientemente de la b .
- (b) Graficar $P(\omega_1|x)$ para el caso de $a_1 = 3, a_2 = 5$ y $b = 1$.
- (c) ¿Como se comportan $P(\omega_1|x)$ y $P(\omega_2|x)$ cuando $x \rightarrow \infty$ y cuando $x \rightarrow -\infty$? Explicar.

5 Ejercicio 4.

Si las distribuciones condicionales para las dos categorias en el problema unidimensional son distribuciones de Cauchy, y asuma la igualdad entre las probabilidades a priori para las categorias.

- (a) Demostrar que la probabilidad de error minimo esta dado por

$$P(error) = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \left| \frac{a_2 - a_1}{2} \right|$$

- (b) Graficar esto como una funcion de $\frac{a_2 - a_1}{2}$.
- (c) ¿Cual es el maximo de $P(error)$ y en que condiciones ocurre esto? Explicar.

6 Ejercicio 5.

Considere la siguiente regla de decisión para el problema de dos categorías unidimensional: Decido por ω_1 si $x > \theta$; en otro caso decida por ω_2 .

- (a) Demuestre que la probabilidad de error para esta regla esta dada por

$$P(\text{error}) = \int P(\text{error}|x)p(x)dx = P(\omega_1) \int_{-\infty}^{\theta} p(x|\omega_1)dx + P(\omega_2) \int_{\theta}^{\infty} p(x|\omega_2)dx$$

- (b) Derivando, demuestre que una condicion necesaria para minimizar el error es

$$p(\theta|\omega_1)p(\omega_1) = p(\theta|\omega_2)p(\omega_2)$$

- (c) Define esta ecuación un θ unico?
- (d) De un ejemplo de un valor de θ en el que el error se maximice.

7 Ejercicio 6.

Supongamos que se sustituye la funcion de decision deterministica $\alpha(x)$ por la regla aleatoria dada por la probabilidad $P(\alpha_i|x)$ de tomar la decisi'on α_i dado que se observe x .

- (a) Mostrar que el riesgo resultante viene dado por

$$R = \int \left[\sum_{i=1}^a R(\alpha_i|x)P(\alpha_i|x) \right] p(x)dx$$

- (b) Ademas, demostrar que R se minimiza para $P(\alpha_i|x) = 1$ para la accion α_i asociada con el riesgo condicional minimo $R(\alpha_i|x)$, lo que demuestra que no obtenemos ningun beneficio haciendo aleatoria la regla de decision.
- (c) ¿Podemos beneficiarnos aleatorizando una regla suboptima? Explicar.