Análisis de Datos

Clasificación Bayesiana

Teoría de decisión Bayesiana

Dr. Wilfrido Gómez Flores

Teoría de decisión Bayesiana

Teoría de probabilidad: provee un marco teórico para cuantificar y manipular la incertidumbre.

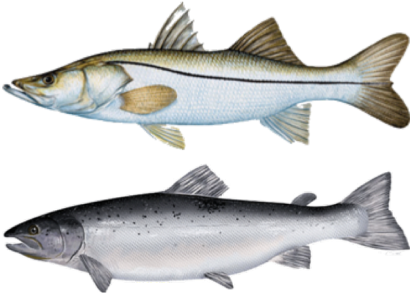
Teoría de decisión Bayesiana:

- Enfoque fundamental para la clasificación de patrones.
- Cuantifica el compromiso entre varias decisiones de clasificación usando probabilidades y sus costos asociados.
- Asume que el problema de decisión está en términos probabilísticos, y que todas las probabilidades relevantes son conocidas.

Teoría de decisión Bayesiana

Una planta empacadora de pescados quiere automatizar el proceso de separación de pescados en una banda de producción:

- Tipos de pescado: róbalo y salmón.
- Predecir cuál es el siguiente pescado en salir.
- Comportamiento aleatorio.



Dos estados de naturaleza ω_1 y ω_2.

Teoría de decisión Bayesiana

- Predicción basada en probabilidades *a priori* (o previas) mediante la **regla de decisión**:

$$\text{Decidir} \begin{cases} \omega_1 & \text{si } \; p(\omega_1) > p(\omega_2) \\ \omega_2 & \text{otro caso} \end{cases} \tag{1}$$

donde

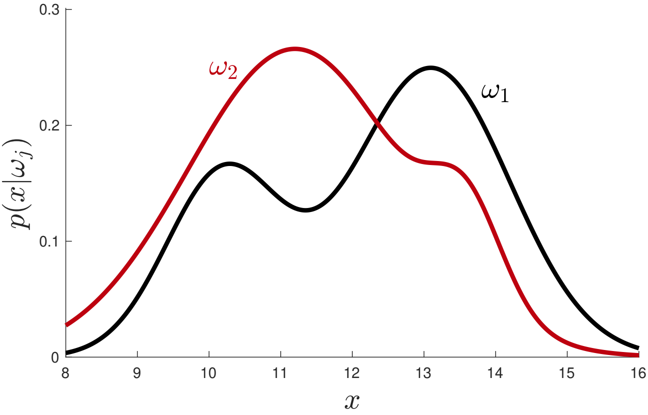
$$p(\omega_1) = \frac{n_1}{n_1 + n_2} \quad \text{y} \quad p(\omega_2) = \frac{n_2}{n_1 + n_2} \tag{2}$$

tal que $p(\omega_1) + p(\omega_2) = 1$ (exhaustividad y exclusividad).

- ¿Qué inconvenientes presenta esta regla de decisión? ¿Cuál es la probabilidad de error?

Para mejorar la regla decisión, se debe incluir información adicional:

- Una medida de luminosidad x a partir de un sensor.
- Se considera x como una variable aleatoria que depende del estado de naturaleza ω.
- Se modela en términos de una **función de densidad de probabilidad clase-condicional** $p(x|\omega)$.



Si x representa la luminosidad de un pescado, las dos funciones de densidad describen las diferencias de luminosidad de las poblaciones de dos tipos de pescado: ω_1 para el róbalo y ω_2 para el salmón.

Teoría de decisión Bayesiana

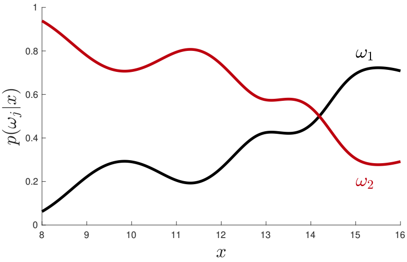
- Si se conocen $p(\omega_j)$ y $p(x|\omega_j)$, entonces el teorema de Bayes calcula la probabilidad *a posteriori* de ser el estado de naturaleza ω_j dado un valor x:

$$p(\omega_j|x) = \frac{p(x|\omega_j)p(\omega_j)}{\sum_{l=1}^{2} p(x|\omega_l)p(\omega_l)}, \quad \text{para} \;\; j = 1, 2 \tag{3}$$

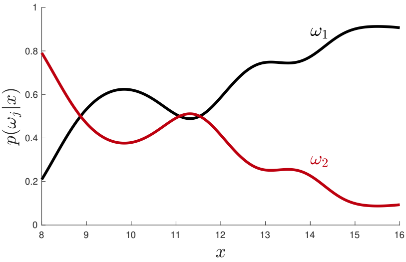
- **Regla de decisión Bayesiana** con probabilidades posteriores:

$$\text{Decidir} \begin{cases} \omega_1 & \text{si} \;\; p(\omega_1|x) > p(\omega_2|x) \\ \omega_2 & \text{otro caso} \end{cases} \tag{4}$$

Teoría de decisión Bayesiana



(a)



(b)

Probabilidades posteriores para los casos (a) $p(\omega_1) = 1/3$ y $p(\omega_2) = 2/3$, y (b) $p(\omega_1) = 2/3$ y $p(\omega_2) = 1/3$.

Teoría de decisión Bayesiana

- Probabilidad de error de la regla de clasificación Bayesiana para una observación x:

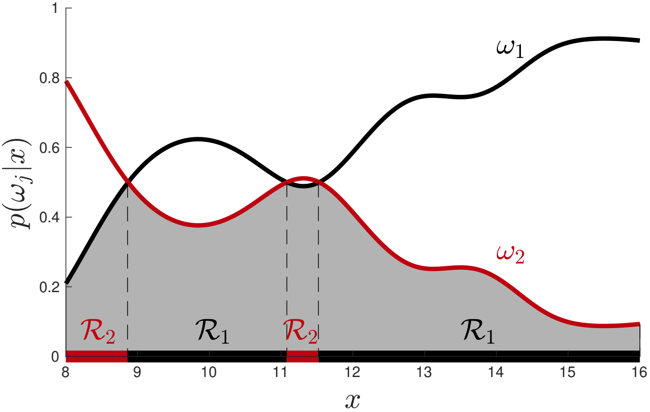
$$p(\text{error}|x) = \begin{cases} p(\omega_1|x) & \text{si se decide } \omega_2 \\ p(\omega_2|x) & \text{si se decide } \omega_1 \end{cases}$$

esto es, $p(\text{error}|x) = \text{mín}\,[p(\omega_1|x), p(\omega_2|x)]$.

- Probabilidad de error promedio:

$$p(\text{error}) = \int_{-\infty}^{\infty} p(\text{error}|x)p(x)dx \tag{5}$$

Teoría de decisión Bayesiana



Probabilidad de error promedio para el caso $p(\omega_1) = 2/3$ y $p(\omega_2) = 1/3$. El área sombreada representa $p(\mathrm{error}|x)$ para cada punto en x y el promedio de error es $p(\mathrm{error}) = 0.35$. Las líneas discontinuas separan las regiones de decisión $\mathcal{R}_j$ de acuerdo a la regla de decisión Bayesiana.

Teoría de decisión Bayesiana

Generalización de los conceptos de la teoría de decisión Bayesiana para permitir un vector de características $\mathbf{x} \in \mathbb{R}^d$ y un conjunto con c estados de naturaleza (i.e., clases) $\Omega = \{\omega_j | j = 1, \ldots, c\}$:

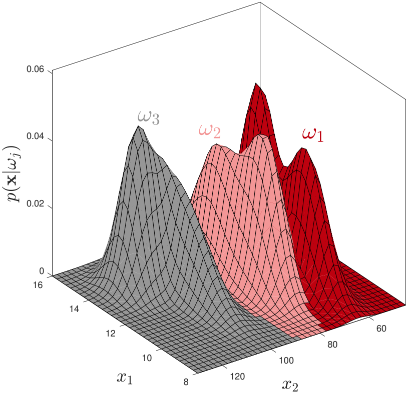
- **Probabilidades posteriores:**

$$p(\omega_j|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_j)p(\omega_j)}{\sum_{l=1}^{c} p(\mathbf{x}|\omega_l)p(\omega_l)} \tag{6}$$

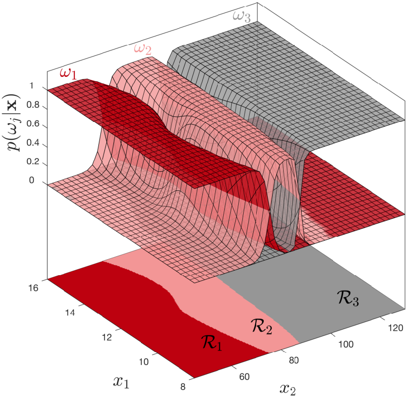
- **Regla de decisión Bayesiana:**

$$\text{Decidir } \omega_i \text{ si } p(\omega_i|\mathbf{x}) > p(\omega_j|\mathbf{x}) \text{ para todo } i \neq j \tag{7}$$

Teoría de decisión Bayesiana



(a)



(b)

Suponga que ahora se tienen una medida de luminosidad (x_1) y una medida del largo (x_2), y tres estados de naturaleza: róbalo (ω_1), salmón (ω_2), y atún (ω_3): (a) funciones de verosimilitud y (b) probabilidades posteriores; las regiones de decisión $\mathcal{R}_j$ se definieron mediante la regla de decisión Bayesiana en (7).

Teoría de decisión Bayesiana

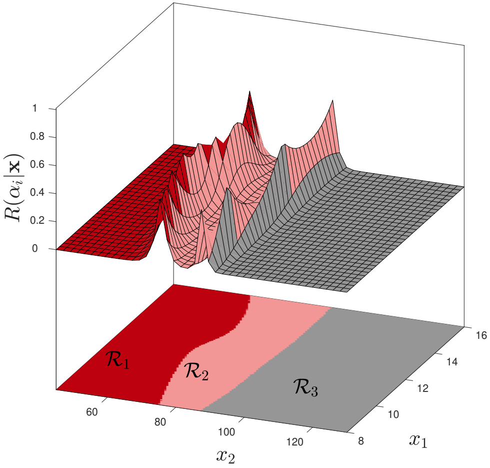
- La acción α_i es interpretada como la decisión de que el verdadero estado de naturaleza es ω_i.
- El **riesgo condicional** expresa la probabilidad de error de decidir la acción α_i dado un patrón $\mathbf{x}$:

$$R(\alpha_i|\mathbf{x}) = \sum_{j=1}^{c} \lambda(\alpha_i|\omega_j) p(\omega_j|\mathbf{x}) \tag{8}$$

donde λ es la *función de pérdida simétrica*:

$$\lambda(\alpha_i|\omega_j) = \begin{cases} 0 & i = j \\ 1 & i \neq j \end{cases} \tag{9}$$

Teoría de decisión Bayesiana



Riesgo condicional para el problema de clasificación de pescados. El error aumenta en las zonas cercanas a las fronteras entre regiones de decisión $\mathcal{R}_j$.