Clasificación

Ejemplo 1: ¿Es usted un enfermo coronario?

Clasificación

Ejemplo 1: ¿Es usted un enfermo coronario?

queremos hacer un diagnostico a partir de variables clinicas

Clasificación

Ejemplo 1: ¿Es usted un enfermo coronario?

queremos hacer un diagnostico a partir de variables clinicas

Las variables registradas son:

1. $X_1 =$ **Presión Sanguínea**
2. $X_2 =$ **Sexo**: F - M
3. $X_3 =$ **Fumador**: Si - No
4. $X_4 =$**Colesterol**
5. $X_5 =$**Actividad Física**: Horas semanales de ejercicio
6. X_6 **TV**: Horas semanales de TV
7. $X_7 =$**Altura**
8. $X_8 =$**Peso**

Clasificación

Ejemplo 1: ¿Es usted un enfermo coronario?

queremos hacer un diagnostico a partir de variables clinicas

Las variables registradas son:

1. $X_1 =$ **Presión Sanguínea**
2. $X_2 =$ **Sexo**: F - M
3. $X_3 =$ **Fumador**: Si - No
4. $X_4 =$**Colesterol**
5. $X_5 =$**Actividad Física**: Horas semanales de ejercicio
6. X_6 **TV**: Horas semanales de TV
7. $X_7 =$**Altura**
8. $X_8 =$**Peso**

$$Y = \begin{cases} 1 & \text{presencia de enfermedad coronaria} \\ 0 & \text{caso contrario .} \end{cases}$$

Clasificación:

Clasificación:

Ejemplo 2: ¿Le damos un crédito?

Clasificación:

Ejemplo 2: ¿Le damos un crédito?

Podemos predecir si un cliente va a pagar?

Clasificación:

Ejemplo 2: ¿Le damos un crédito?

Podemos predecir si un cliente va a pagar?

Las variables registradas son:

1. **Veraz**
2. **Demográficas**
3. **Salario**

Clasificación:

Ejemplo 2: ¿Le damos un crédito?

Podemos predecir si un cliente va a pagar?

Las variables registradas son:

1. **Veraz**
2. **Demográficas**
3. **Salario**

$$Y = \begin{cases} 1 & \text{pagador} \\ 0 & \text{caso contrario} . \end{cases}$$

Clasificación:

Ejemplo 2: ¿Le damos un crédito?

Podemos predecir si un cliente va a pagar?

Las variables registradas son:

1. **Veraz**
2. **Demográficas**
3. **Salario**

$$Y = \begin{cases} 1 & \text{pagador} \\ 0 & \text{caso contrario} \end{cases}.$$

$$Y = \begin{cases} B & \text{Bajo riesgo crediticio} \\ M & \text{Medio riesgo crediticio} \\ A & \text{Alto riesgo crediticio} \end{cases}.$$

Pajaritos - El patito feo

Pajaritos - El patito feo

Las aves parásitas de cría ponen huevos en nidos de otras especies (hospedador), las cuales incuban los huevos y crían al pichón parásito.

Pajaritos - El patito feo

Las aves parásitas de cría ponen huevos en nidos de otras especies (hospedador), las cuales incuban los huevos y crían al pichón parásito.

En un bosque de talas de la provincia de Buenos Aires hay dos especies hospederas que son indistinguibles a simple vista.

Pajaritos - El patito feo

Las aves parásitas de cría ponen huevos en nidos de otras especies (hospedador), las cuales incuban los huevos y crían al pichón parásito.

En un bosque de talas de la provincia de Buenos Aires hay dos especies hospederas que son indistinguibles a simple vista.

Pero....una de las principales diferencias entre estas especies radica en el grado de discriminación y remoción de huevos parásitos de sus nidos.

Pajaritos - El patito feo

Una de las especies es “aceptadora“ de huevos parásitos, ya que remueve del nido sólo el 30% de los huevos parásitos.

Pajaritos - El patito feo

Una de las especies es “aceptadora“ de huevos parásitos, ya que remueve del nido sólo el 30% de los huevos parásitos.
La otra especie es “rechazadora“ ya que remueve el 80% de los huevos parásitos presentes en su nido.

Pajaritos - El patito feo

Una de las especies es “aceptadora“ de huevos parásitos, ya que remueve del nido sólo el 30% de los huevos parásitos.
La otra especie es “rechazadora“ ya que remueve el 80% de los huevos parásitos presentes en su nido.

Definimos,

$$Y = \begin{cases} R & \text{rechazador} \\ A & \text{aceptador}\ . \end{cases}$$

Pajaritos - El patito feo

Una de las especies es “aceptadora“ de huevos parásitos, ya que remueve del nido sólo el 30% de los huevos parásitos.
La otra especie es “rechazadora“ ya que remueve el 80% de los huevos parásitos presentes en su nido.

Definimos,

$$Y = \begin{cases} R & \text{rechazador} \\ A & \text{aceptador}\ . \end{cases}$$

Además, se sabe que el 90% de los nidos del bosque corresponden a la especie “aceptadora“ , mientras que apenas el 10% restante son nidos de la especie “rechazadora“.

Pajaritos - El patito feo

Una de las especies es “aceptadora“ de huevos parásitos, ya que remueve del nido sólo el 30% de los huevos parásitos.
La otra especie es “rechazadora“ ya que remueve el 80% de los huevos parásitos presentes en su nido.

Definimos,

$$Y = \begin{cases} R & \text{rechazador} \\ A & \text{aceptador} . \end{cases}$$

Además, se sabe que el 90% de los nidos del bosque corresponden a la especie “aceptadora“ , mientras que apenas el 10% restante son nidos de la especie “rechazadora“.

$$P(Y = A) = 0.9$$
$$P(Y = R) = 0.1$$

Pajaritos - El patito feo

El objetivo es decidir mirando un nido y según cuantos huevos sean removidos si la especie es rechazadora o aceptadora.

Pajaritos - El patito feo

El objetivo es decidir mirando un nido y según cuantos huevos sean removidos si la especie es rechazadora o aceptadora.
Es decir, conociendo el número de huevos removidos queremos predecir si la especie es rechazadora o aceptadora.

Pajaritos - El patito feo

El objetivo es decidir mirando un nido y según cuantos huevos sean removidos si la especie es rechazadora o aceptadora.
Es decir, conociendo el número de huevos removidos queremos predecir si la especie es rechazadora o aceptadora.

Supongamos que en un nido se colocan $n = 8$ huevos parasitarios.

Sea $X =$ número de huevos removidos.

Pajaritos - El patito feo

El objetivo es decidir mirando un nido y según cuantos huevos sean removidos si la especie es rechazadora o aceptadora.
Es decir, conociendo el número de huevos removidos queremos predecir si la especie es rechazadora o aceptadora.

Supongamos que en un nido se colocan $n = 8$ huevos parasitarios.

Sea $X =$ número de huevos removidos.

Distribuciones condicionales

Si sabemos que la especie es aceptadora, cuál es la distribución de X?

Distribuciones condicionales

Si sabemos que la especie es aceptadora, cuál es la distribución de X?

$X|_{Y=A} \sim Bi(8, 0.3)$, es decir $p_{X|_{Y=A}}(x) = \binom{8}{x} 0.3^x 0.7^{8-x}$

0	1	2	3	4	5	6	7	8
0.057	0.197	0.296	0.254	0.136	0.046	0.010	0.001	0.000

Distribuciones condicionales

Si sabemos que la especie es aceptadora, cuál es la distribución de X?

$X|_{Y=A} \sim Bi(8, 0.3)$, es decir $p_{X|_{Y=A}}(x) = \binom{8}{x} 0.3^x 0.7^{8-x}$

0	1	2	3	4	5	6	7	8
0.057	0.197	0.296	0.254	0.136	0.046	0.010	0.001	0.000

$X|_{Y=R} \sim Bi(8, 0.8)$, es decir $p_{X|_{Y=R}}(x) = \binom{8}{x} 0.8^x 0.2^{8-x}$

0	1	2	3	4	5	6	7	8
0.000	0.000	0.001	0.009	0.046	0.147	0.294	0.335	0.168

Distribuciones condicionales

Si sabemos que la especie es aceptadora, cuál es la distribución de X?

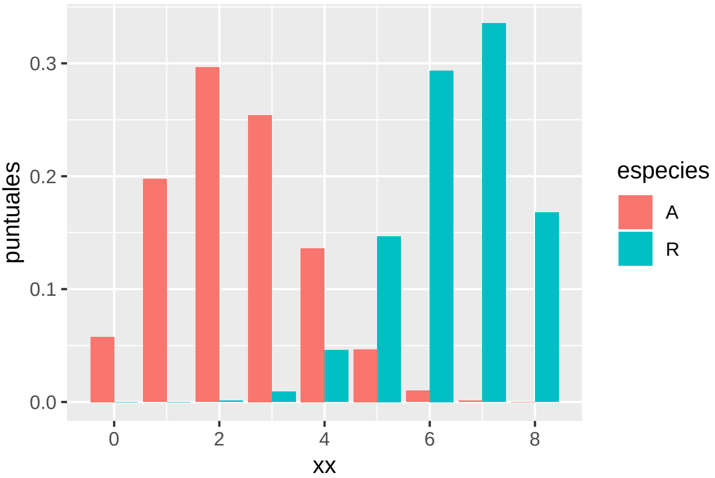
$X|_{Y=A} \sim Bi(8, 0.3)$, es decir $p_{X|_{Y=A}}(x) = \binom{8}{x} 0.3^x 0.7^{8-x}$

0	1	2	3	4	5	6	7	8
0.057	0.197	0.296	0.254	0.136	0.046	0.010	0.001	0.000

$X|_{Y=R} \sim Bi(8, 0.8)$, es decir $p_{X|_{Y=R}}(x) = \binom{8}{x} 0.8^x 0.2^{8-x}$

0	1	2	3	4	5	6	7	8
0.000	0.000	0.001	0.009	0.046	0.147	0.294	0.335	0.168

Distribuciones condicionales



Distribuciones condicionales

Cuál es la distribución de X??

Distribuciones condicionales

Cuál es la distribución de X?? $p_X(x)$?

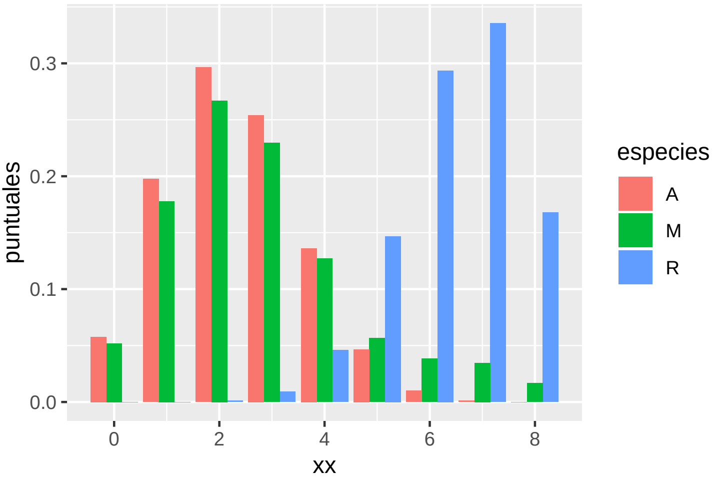
Distribuciones condicionales

Cuál es la distribución de X?? $p_X(x)$?

$$\begin{aligned} p_X(x) &= p_{X|Y=R}(x)P(Y=R) + p_{X|Y=A}(x)P(Y=A) \\ &= 0.1\binom{8}{x}0.8^x0.2^{8-x} + 0.9\binom{8}{x}0.3^x0.7^{8-x} \end{aligned}$$

0	1	2	3	4	5	6	7	8
0.052	0.178	0.267	0.230	0.127	0.057	0.038	0.035	0.017

Distribuciones condicionales y puntual



Pajaritos - El patito feo

Cómo hacemos para decidir mirando un nido con 8 huevos parasitarios y dependiendo cuantos huevos sean removidos si la especie es rechazadora o aceptadora.

Pajaritos - El patito feo

Cómo hacemos para decidir mirando un nido con 8 huevos parasitarios y dependiendo cuantos huevos sean removidos si la especie es rechazadora o aceptadora.

Si se remueven 5 huevos; es decir si $X = 5$

Pajaritos - El patito feo

Cómo hacemos para decidir mirando un nido con 8 huevos parasitarios y dependiendo cuantos huevos sean removidos si la especie es rechazadora o aceptadora.

> Si se remueven 5 huevos; es decir si $X = 5$ ¿de qué clase de nido diría que se trata?

Pajaritos - El patito feo

Cómo hacemos para decidir mirando un nido con 8 huevos parasitarios y dependiendo cuantos huevos sean removidos si la especie es rechazadora o aceptadora.

- Si se remueven 5 huevos; es decir si $X = 5$ ¿de qué clase de nido diría que se trata?
- Si se remueven 3 huevos; (X=3) ¿de qué clase de nido diría que se trata?

Queremos construir un Clasificador

Queremos construir un Clasificador

Una Regla (de clasificación) asigna a $x \in \{0, 1, \ldots, 8\}$ un tipo de hospedador: $\{A, R\}$

Queremos construir un Clasificador

Una Regla (de clasificación) asigna a $x \in \{0, 1, \ldots, 8\}$ un tipo de hospedador: $\{A, R\}$ es decir, buscamos

$$h : \{0, 1, \ldots, 8\} \to \{A, R\}$$

Queremos construir un Clasificador

Una Regla (de clasificación) asigna a $x \in \{0, 1, \ldots, 8\}$ un tipo de hospedador: $\{A, R\}$ es decir, buscamos

$$h : \{0, 1, \ldots, 8\} \to \{A, R\}$$

Por ejemplo,

x	0	1	2	3	4	5	6	7	8

Queremos construir un Clasificador

Una Regla (de clasificación) asigna a $x \in \{0, 1, \ldots, 8\}$ un tipo de hospedador: $\{A, R\}$ es decir, buscamos

$$h : \{0, 1, \ldots, 8\} \to \{A, R\}$$

Por ejemplo,

x	0	1	2	3	4	5	6	7	8
clasificador h_1	A	A	A	A	R	R	R	R	R

Queremos construir un Clasificador

Una Regla (de clasificación) asigna a $x \in \{0, 1, \ldots, 8\}$ un tipo de hospedador: $\{A, R\}$ es decir, buscamos

$$h : \{0, 1, \ldots, 8\} \to \{A, R\}$$

Por ejemplo,

x	0	1	2	3	4	5	6	7	8
clasificador h_1	A	A	A	A	R	R	R	R	R

Otro clasificador

x	0	1	2	3	4	5	6	7	8
clasificador h_2	A	R	A	R	R	R	R	R	R

Queremos construir un Clasificador

Queremos construir un Clasificador

Otra manera de escribirlo

x	0	1	2	3	4	5	6	7	8
clasificador h_1	A	A	A	A	R	R	R	R	R

$$h_1(x) = \begin{cases} A & x \in \{0,1,2,3\} \\ R & x \in \{4,5,6,7,8\} \end{cases}$$

Queremos construir un Clasificador

Otra manera de escribirlo

x	0	1	2	3	4	5	6	7	8
clasificador h_1	A	A	A	A	R	R	R	R	R

$$h_1(x) = \begin{cases} A & x \in \{0, 1, 2, 3\} \\ R & x \in \{4, 5, 6, 7, 8\} \end{cases}$$

x	0	1	2	3	4	5	6	7	8
clasificador h_2	A	R	A	R	R	R	R	R	R

$$h_2(x) = \begin{cases} A & x \in \{0, 2, 3\} \\ R & x \in \{1, 4, 5, 6, 7, 8\} \end{cases}$$

Queremos construir un Clasificador

Otra manera de escribirlo

x	0	1	2	3	4	5	6	7	8
clasificador h_1	A	A	A	A	R	R	R	R	R

$$h_1(x) = \begin{cases} A & x \in \{0,1,2,3\} \\ R & x \in \{4,5,6,7,8\} \end{cases}$$

x	0	1	2	3	4	5	6	7	8
clasificador h_2	A	R	A	R	R	R	R	R	R

$$h_2(x) = \begin{cases} A & x \in \{0,2,3\} \\ R & x \in \{1,4,5,6,7,8\} \end{cases}$$

Cuál es más razonable??

Clasificación: Marco Teórico

- $X \in \mathcal{X}$, $Y \in \mathcal{Y}$. (en nuestro caso $\mathcal{X} = \{0, 1, \ldots, 8\}$ y $\mathcal{Y} = \{A, R\}$)
- (X, Y) vector aleatorio, con puntual p_{XY}.
- Clasificador: Regla (de clasificación) que asigna a cada $x \in \mathcal{X}$ un elemento $y \in \mathcal{Y}$

$$\text{Clasificador} \quad h : \mathcal{X} \to \mathcal{Y}$$

- Error de Clasificación Medio del clasificador h

$$L(h) = \mathbb{P}(h(X) \neq Y)$$

- Objetivo (teórico): Encontrar h que minimice el error medio de clasificación.

Estimación del error de clasificación de h

Como calcular el error de Clasificación Medio ??

$$L(h) = \mathbb{P}(h(X) \neq Y)$$

Por ejemplo con el clasificador h_1,

$$h_1(x) = \begin{cases} A & x \in \{0, 1, 2, 3\} \\ R & x \in \{4, 5, 6, 7, 8\} \end{cases}$$

$P\left(h_1(X) \neq Y\right)$

Estimación del error de clasificación de h

Como calcular el error de Clasificación Medio ??

$$L(h) = \mathbb{P}(h(X) \neq Y)$$

Por ejemplo con el clasificador h_1,

$$h_1(x) = \begin{cases} A & x \in \{0,1,2,3\} \\ R & x \in \{4,5,6,7,8\} \end{cases}$$

$$P\left(h_1(X) \neq Y\right) = P(h_1(X) \neq Y | Y = R) P(Y = R) + P(h_1(X) \neq Y | Y = A) P(Y = A)$$

Estimación del error de clasificación de h

Como calcular el error de Clasificación Medio ??

$$L(h) = \mathbb{P}(h(X) \neq Y)$$

Por ejemplo con el clasificador h_1,

$$h_1(x) = \begin{cases} A & x \in \{0, 1, 2, 3\} \\ R & x \in \{4, 5, 6, 7, 8\} \end{cases}$$

$$\begin{aligned} P\left(h_1(X) \neq Y\right) &= P(h_1(X) \neq Y | Y = R)P(Y = R) + P(h_1(X) \neq Y | Y = A)P(Y = A) \\ &= P(X \in \{0, 1, 2, 3\} | Y = R)P(Y = R) + P(X \in \{4, 5, 6, 7, 8\} | Y = A)P(Y = A) \end{aligned}$$

Estimación del error de clasificación de h

Como calcular el error de Clasificación Medio ??

$$L(h) = \mathbb{P}(h(X) \neq Y)$$

Por ejemplo con el clasificador h_1,

$$h_1(x) = \begin{cases} A & x \in \{0,1,2,3\} \\ R & x \in \{4,5,6,7,8\} \end{cases}$$

$$\begin{aligned}
&P\left(h_1(X) \neq Y\right) = P(h_1(X) \neq Y | Y = R) P(Y = R) + P(h_1(X) \neq Y | Y = A) P(Y = A) \\
&= P(X \in \{0,1,2,3\} | Y = R) P(Y = R) + P(X \in \{4,5,6,7,8\} | Y = A) P(Y = A) \\
&= \sum_{x=0}^{3} \binom{8}{x} 0.8^x 0.2^{8-x} 0.1 + \sum_{x=4}^{8} \binom{8}{x} 0.3^x 0.7^{8-x} 0.9 = 0.1757
\end{aligned}$$

Estimación del error de clasificación de h

Como calcular el error de Clasificación Medio ??

$$L(h) = \mathbb{P}(h(X) \neq Y)$$

Por ejemplo con el clasificador h_1,

$$h_1(x) = \begin{cases} A & x \in \{0, 1, 2, 3\} \\ R & x \in \{4, 5, 6, 7, 8\} \end{cases}$$

$$\begin{aligned} &P\left(h_1(X) \neq Y\right) = P(h_1(X) \neq Y | Y = R) P(Y = R) + P(h_1(X) \neq Y | Y = A) P(Y = A) \\ &= P(X \in \{0, 1, 2, 3\} | Y = R) P(Y = R) + P(X \in \{4, 5, 6, 7, 8\} | Y = A) P(Y = A) \\ &= \sum_{x=0}^{3} \binom{8}{x} 0.8^x 0.2^{8-x} 0.1 + \sum_{x=4}^{8} \binom{8}{x} 0.3^x 0.7^{8-x} 0.9 = 0.1757 \end{aligned}$$

$$h_1(x) = \begin{cases} A & x \in \{0, 2, 3\} \\ R & x \in \{1, 4, 5, 6, 7, 8\} \end{cases}$$

Estimación del error de clasificación de h

Como calcular el error de Clasificación Medio ??

$$L(h) = \mathbb{P}(h(X) \neq Y)$$

Por ejemplo con el clasificador h_1,

$$h_1(x) = \begin{cases} A & x \in \{0, 1, 2, 3\} \\ R & x \in \{4, 5, 6, 7, 8\} \end{cases}$$

$$\begin{aligned} P\left(h_1(X) \neq Y\right) &= P(h_1(X) \neq Y | Y = R) P(Y = R) + P(h_1(X) \neq Y | Y = A) P(Y = A) \\ &= P(X \in \{0, 1, 2, 3\} | Y = R) P(Y = R) + P(X \in \{4, 5, 6, 7, 8\} | Y = A) P(Y = A) \\ &= \sum_{x=0}^{3} \binom{8}{x} 0.8^x 0.2^{8-x} 0.1 + \sum_{x=4}^{8} \binom{8}{x} 0.3^x 0.7^{8-x} 0.9 = 0.1757 \end{aligned}$$

$$h_1(x) = \begin{cases} A & x \in \{0, 2, 3\} \\ R & x \in \{1, 4, 5, 6, 7, 8\} \end{cases}$$

$$P\left(h_2(X) \neq Y\right) = 0.3536$$

Clasificador Óptimo

Teorema: El clasificador definido como

$$h_{opt}(x) = \begin{cases} A & x \in \{z : P(Y = A|X = z) \geq P(Y = R|X = z)\} \\ R & x \in \{z : P(Y = R|X = z) > P(Y = A|X = z)\} \end{cases}$$

es óptimo en el sentido que minimiza $L(h) = \mathbb{P}(h(X) \neq Y)$.

Clasificador Óptimo

Teorema: El clasificador definido como

$$h_{opt}(x) = \begin{cases} A & x \in \{z : P(Y = A|X = z) \geq P(Y = R|X = z)\} \\ R & x \in \{z : P(Y = R|X = z) > P(Y = A|X = z)\} \end{cases}$$

es óptimo en el sentido que minimiza $L(h) = \mathbb{P}(h(X) \neq Y)$.

Primero notemos que

$$\begin{aligned} h_{opt}(x) &= \begin{cases} A & x \in \{z : p_{Y|X=z}(A) \geq p_{Y|X=z}(R)\} \\ R & x \in \{z : p_{Y|X=z}(R) > p_{Y|X=z}(A)\} \end{cases} \\ &= \begin{cases} A & x \in \{z : \frac{p_{X|Y=A}(z)p_Y(A)}{p_X(z)} \geq \frac{p_{X|Y=R}(z)p_Y(R)}{p_X(z)}\} \\ R & x \in \{z : \frac{p_{X|Y=R}(z)p_Y(R)}{p_X(z)} > \frac{p_{X|Y=A}(z)p_Y(A)}{p_X(z)}\} \end{cases} \end{aligned}$$

Clasificador Óptimo

Luego

$$h_{opt}(x) \quad = \quad \begin{cases} A & x \in \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\} \\ R & x \in \{z : p_{X|Y=R}(z)p_Y(R) > p_{X|Y=A}(z)p_Y(A)\} \end{cases}$$

Clasificador Óptimo

Luego

$$h_{opt}(x) \quad = \quad \begin{cases} A & x \in \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\} \\ R & x \in \{z : p_{X|Y=R}(z)p_Y(R) > p_{X|Y=A}(z)p_Y(A)\} \end{cases}$$

Notemos que cualquier clasificador es de la forma

$$h(x) = \begin{cases} A & x \in C_A \\ R & x \in C_R \end{cases}$$

Entonces debemos probar que si tomamos
$C_A = \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\}$ minimiza $L(h)$.

Clasificador Óptimo

Luego

$$h_{opt}(x) \;=\; \begin{cases} A & x \in \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\} \\ R & x \in \{z : p_{X|Y=R}(z)p_Y(R) > p_{X|Y=A}(z)p_Y(A)\} \end{cases}$$

Notemos que cualquier clasificador es de la forma

$$h(x) = \begin{cases} A & x \in C_A \\ R & x \in C_R \end{cases}$$

Entonces debemos probar que si tomamos
$C_A = \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\}$ minimiza $L(h)$.

$$L(h) \;=\; \mathbb{P}(h(X) \neq Y) = \sum_{x \in C_A} p_{X|Y=R}(z)p_Y(R) + \sum_{x \in C_R} p_{X|Y=A}(z)p_Y(A)$$

Clasificador Óptimo

Luego

$$h_{opt}(x) \;=\; \begin{cases} A & x \in \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\} \\ R & x \in \{z : p_{X|Y=R}(z)p_Y(R) > p_{X|Y=A}(z)p_Y(A)\} \end{cases}$$

Notemos que cualquier clasificador es de la forma

$$h(x) = \begin{cases} A & x \in C_A \\ R & x \in C_R \end{cases}$$

Entonces debemos probar que si tomamos
$C_A = \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\}$ minimiza $L(h)$.

$$\begin{aligned} L(h) &= \mathbb{P}(h(X) \neq Y) = \sum_{x \in C_A} p_{X|Y=R}(z)p_Y(R) + \sum_{x \in C_R} p_{X|Y=A}(z)p_Y(A) \\ &= \sum_{x \in C_A} p_{X|Y=R}(z)p_Y(R) + \sum_{x \in C_R} p_{X|Y=A}(z)p_Y(A) \\ &+ \sum_{x \in C_A} p_{X|Y=A}(z)p_Y(A) - \sum_{x \in C_A} p_{X|Y=A}(z)p_Y(A) \end{aligned}$$

Clasificador Óptimo

$$
\begin{aligned}
L(h) &= \sum_{x \in C_A} p_{X|Y=R}(z) p_Y(R) + \sum_{x \in C_R} p_{X|Y=A}(z) p_Y(A) \\
&+ \sum_{x \in C_A} p_{X|Y=A}(z) p_Y(A) - \sum_{x \in C_A} p_{X|Y=A}(z) p_Y(A)
\end{aligned}
$$

Clasificador Óptimo

$$
\begin{aligned}
L(h) &= \sum_{x\in C_A} p_{X|Y=R}(z)p_Y(R) + \sum_{x\in C_R} p_{X|Y=A}(z)p_Y(A) \\
&+ \sum_{x\in C_A} p_{X|Y=A}(z)p_Y(A) - \sum_{x\in C_A} p_{X|Y=A}(z)p_Y(A) \\
&= \sum_{x\in C_A} (p_{X|Y=R}(z)p_Y(R) - p_{X|Y=A}(z)p_Y(A)) \\
&+ \sum_{x\in C_R} p_{X|Y=A}(z)p_Y(A) + \sum_{x\in C_A} p_{X|Y=A}(z)p_Y(A) \\
&= \sum_{x\in C_A} (p_{X|Y=R}(z)p_Y(R) - p_{X|Y=A}(z)p_Y(A)) + p_Y(A)
\end{aligned}
$$

Entonces $L(h)$ es mínimo si para todo $x \in C_A$ se tiene que $(p_{X|Y=R}(z)p_Y(R) - p_{X|Y=A}(z)p_Y(A)) < 0$. Pero justamente

$$
h_{opt}(x) = \begin{cases} A & x \in \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\} \\ R & x \in \{z : p_{X|Y=R}(z)p_Y(R) > p_{X|Y=A}(z)p_Y(A)\} \end{cases}
$$

Clasificador Óptimo: Pátito feo

$$h_{opt}(x) = \begin{cases} A & x \in \{z : p_{X|Y=A}(z)p_Y(A) \geq p_{X|Y=R}(z)p_Y(R)\} \\ R & x \in \{z : p_{X|Y=R}(z)p_Y(R) > p_{X|Y=A}(z)p_Y(A)\} \end{cases}$$

$p_{X|Y=A}(z)p_Y(A)$

x	0	1	2	3	4	5	6	7	8
	.052	0.178	0.267	0.229	0.123	0.042	0.009	0.001	0.00

$p_{X|Y=R}(z)p_Y(R)$

x	0	1	2	3	4	5	6	7	8
	0.0	0.0	0.0	0.001	0.005	0.015	0.029	0.034	0.017

$$h_{opt}(x) = \begin{cases} A & x \in \{0, 1, 2, 3, 4, 5\} \\ R & x \in \{6, 7, 8\} \end{cases}$$

$$L(h_{opt}) = \mathbb{P}(h_{opt}(X) \neq Y) = 0.0304$$