



UNIVERSIDAD
PANAMERICANA

Regresión Logística

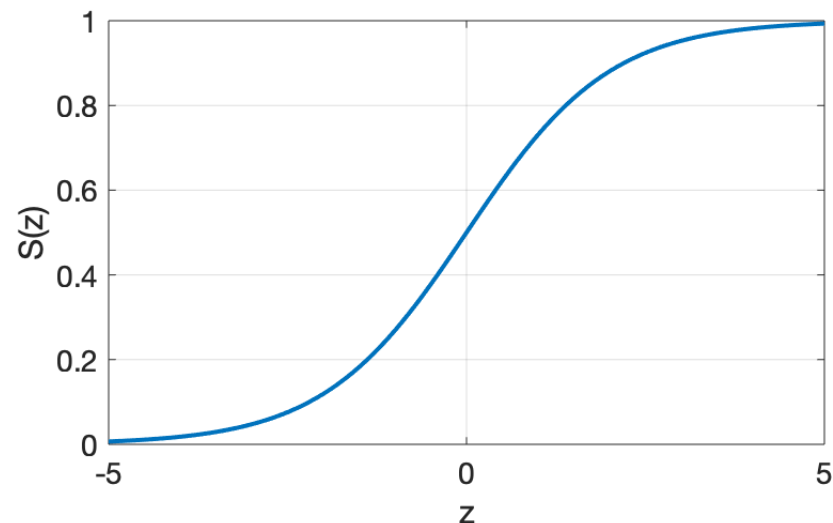
Regresión Logística

La **regresión logística** es un modelo de aprendizaje supervisado que permite hacer clasificación basada en la transformación de valores numéricos a categorías (o *clases*) utilizando la función sigmoideal.

Tipos de regresión logística:

- Binaria
- Multi-clase

$$S(z) = \frac{1}{1 + e^{-z}}$$



Se puede interpretar como la estimación de la probabilidad

Regresión Logística

La **regresión logística** es un modelo de aprendizaje supervisado que permite hacer clasificación basada en la transformación de valores numéricos a categorías (o *clases*) utilizando la función sigmoideal.

Regresión
Lineal

Función
Sigmoideal

Frontera de
Decisión

(logit)

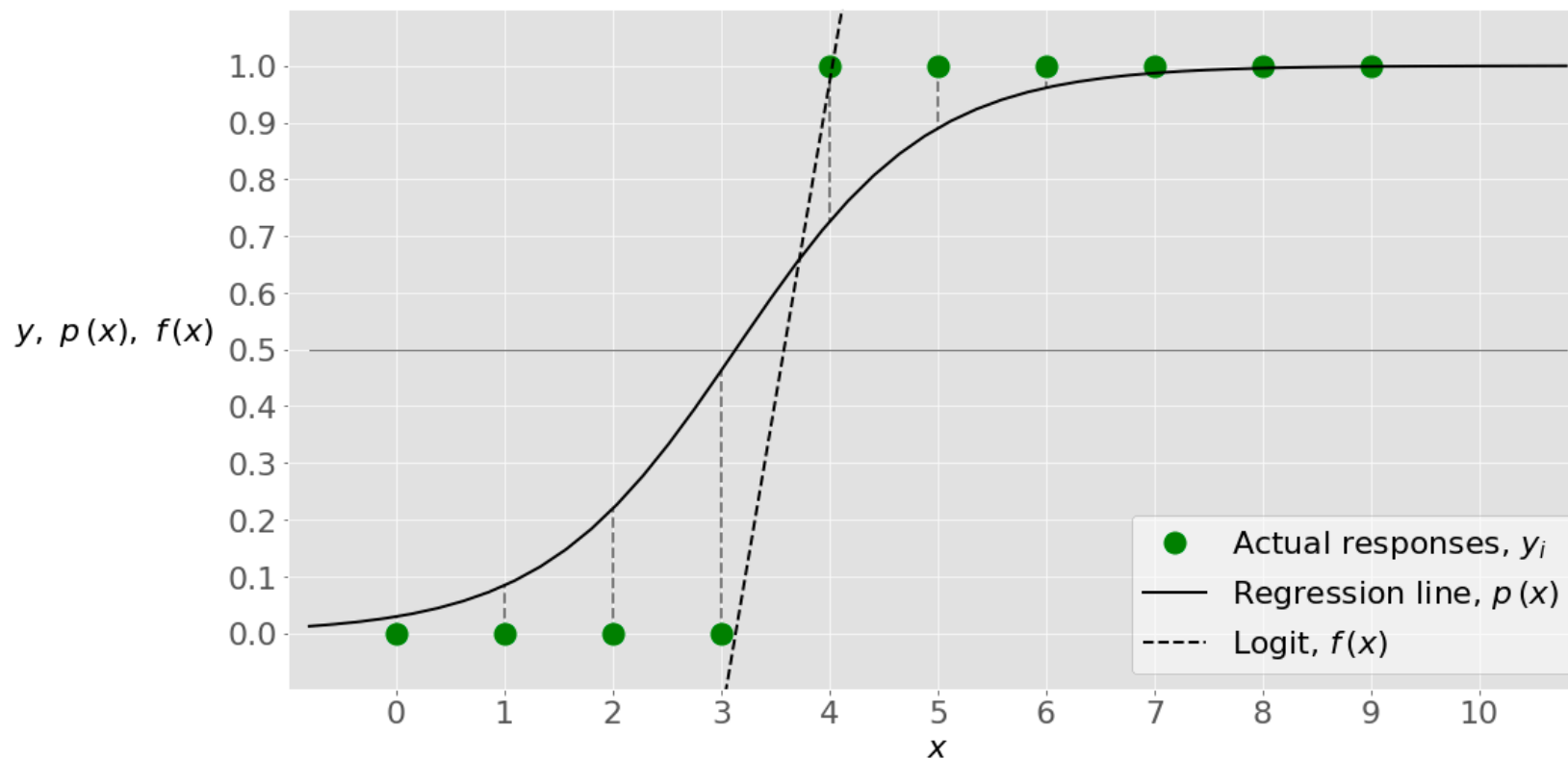
$$z = w_0 + w_1x_1 + \cdots + w_rx_r$$

$$S(z) = \frac{1}{1 + e^{-z}}$$

$$f(x) = \begin{cases} class = 1 & S(z) \geq 0.5 \\ class = 0 & S(z) < 0.5 \end{cases}$$

Regresión Logística

La **regresión logística** es un modelo de aprendizaje supervisado que permite hacer clasificación basada en la transformación de valores numéricos a categorías (o *clases*) utilizando la función sigmoidea.



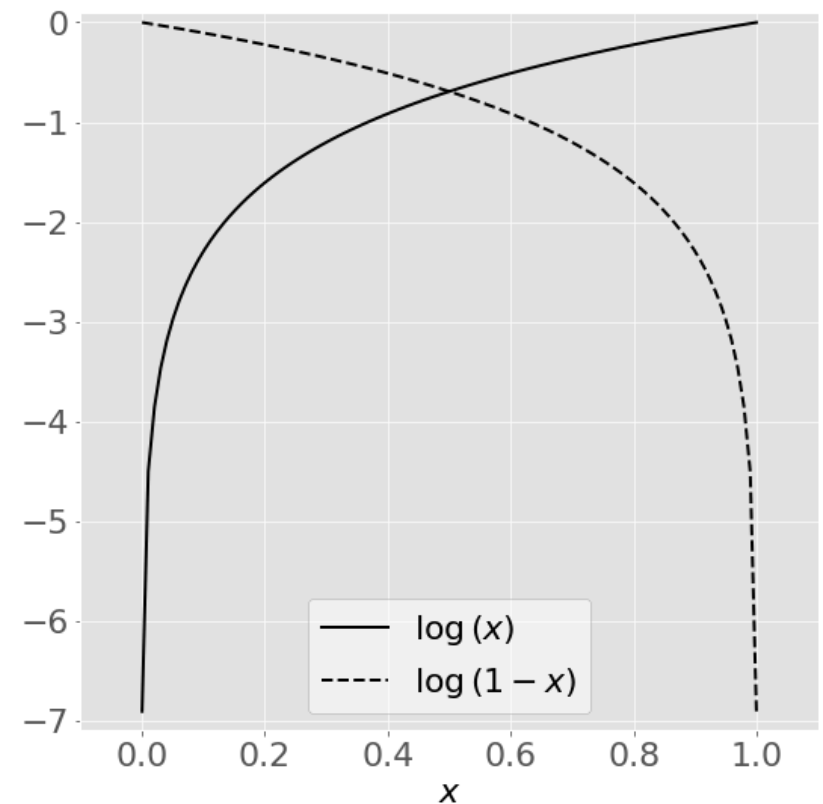
Stojiljkovic, M., "Logistic regression in Python", Real Python, 2020.

Regresión Logística

El entrenamiento de una regresión logística utiliza la función de error denominada **pérdida logística** (*cross-entropy*):

$$J = -\frac{1}{N} \sum_{i=1}^N [y_i \ln(z(x_i)) + (1 - y_i) \ln(1 - z(x_i))]$$

Penaliza las predicciones erróneas más que recompensar las predicciones correctas.



Stojiljkovic, M., "Logistic regression in Python", Real Python, 2020.

La **regresión logística multi-clase** permite hacer una multi-clasificación y se basa en la regresión logística binaria sobre cada clase.

Codificación *One-Hot*:

Cada clase en la salida se interpreta como un vector de K términos de 0's y 1's:

Valor de Salida	1	2	3	4	5	6
6	0	0	0	0	0	1
2	0	1	0	0	0	0
1	1	0	0	0	0	0
4	0	0	0	1	0	0

Se obtiene la máxima probabilidad de la estimación de la clase usando la **función exponencial normalizada** (o ***softmax***):

$$\sigma(x_i) = \frac{e^{x_i}}{\sum_{j=1}^K e^{x_{ij}}}$$