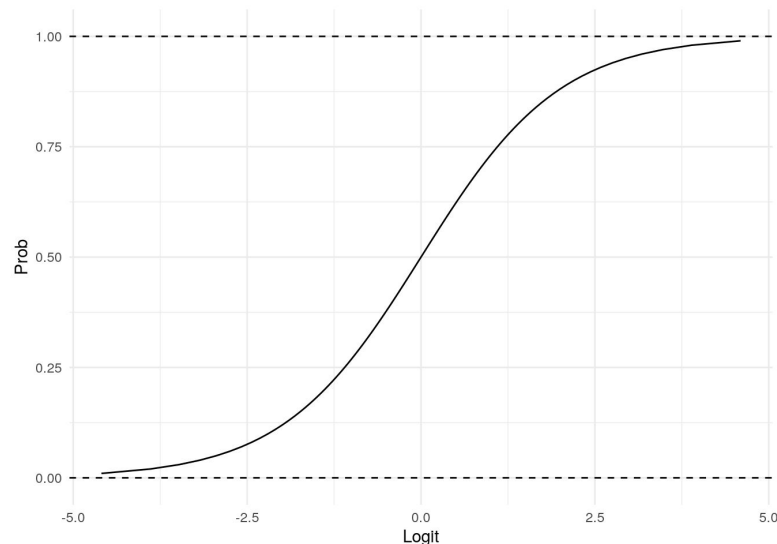


Métodos de análisis cuantitativos multivariados

Clase 5. Métricas de evaluación para regresión logística

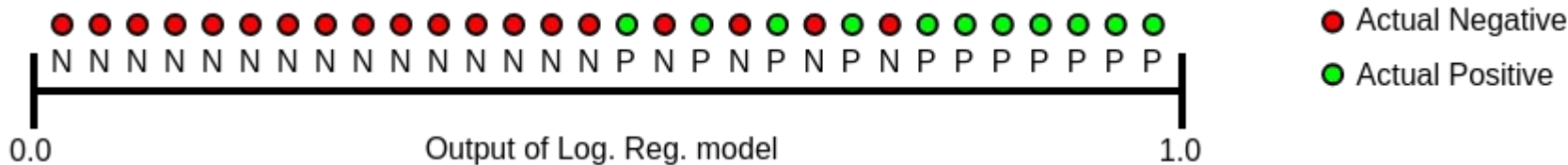
Introducción

- Siempre hemos venido trabajando con probabilidades => output del modelo de regresión logística
- Pero en general queremos predecir categorías (Sí-No; Obrero-No obrero; Éxito o Fracaso, Ocupado-No Ocupado)
- ¿Cómo pasamos de uno a otro? De una probabilidad a una categoría



De una probabilidad a una categoría

- Podemos representar en la línea las probabilidades predichas y con el color la clase observada.



De una probabilidad a una categoría

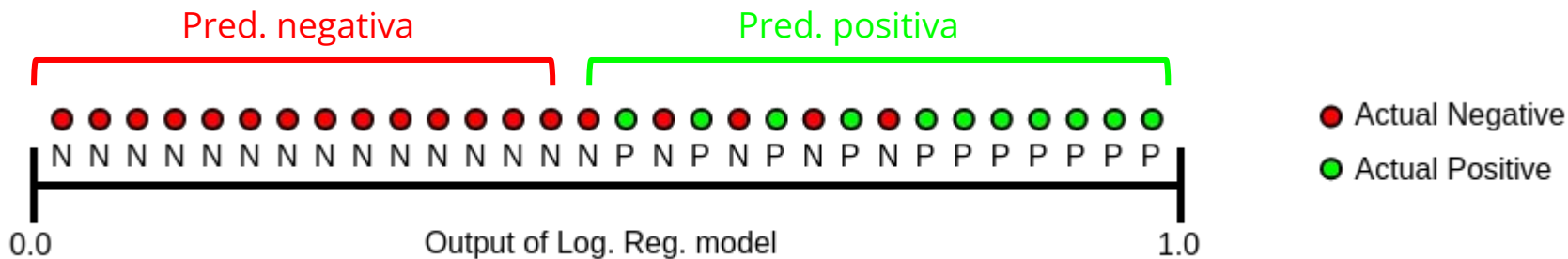
- Regresión logística (pero vale para otros)
- La predicción “nativa” de una regresión logística es una probabilidad
- Se aplica una regla de decisión para asignarle un valor positivo o negativo

Si $p(y=1 | X) > 0.5 \Rightarrow$ el caso es positivo

Si $p(y=1 | X) \leq 0.5 \Rightarrow$ el caso es negativo

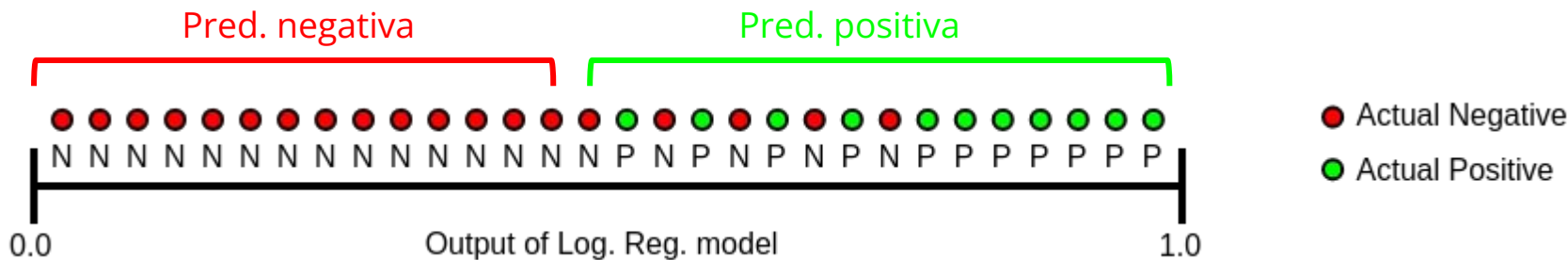
De una probabilidad a una categoría

- Todo caso que tenga una probabilidad predicha mayor a 0.5 va a ser considerado un caso “positivo” (en nuestra predicción).



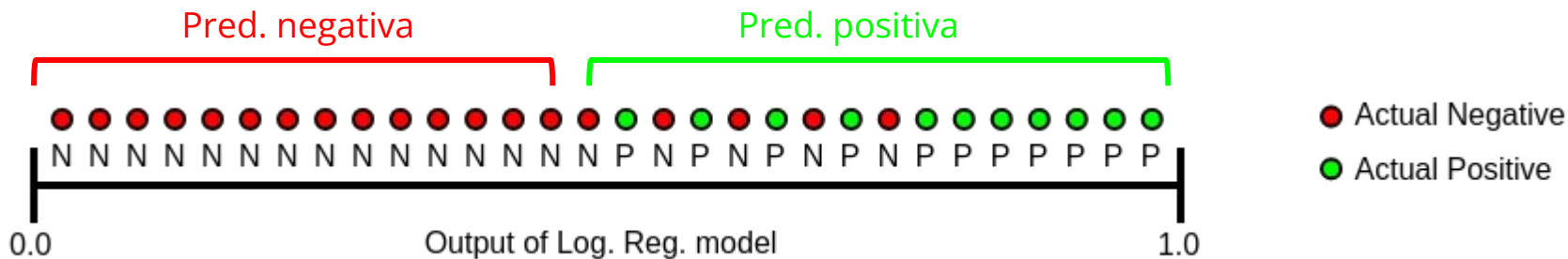
De una probabilidad a una categoría

- Esto nos va a permitir definir una herramienta clave: la matriz de confusión.



De una probabilidad a una categoría

- Esto nos va a permitir definir una herramienta clave: la matriz de confusión.



	Positivo (pred)	Negativo (pred)	Total
Positivo (obs)	11	0	11
Negativo (obs)	5	14	19
Total	16	14	30

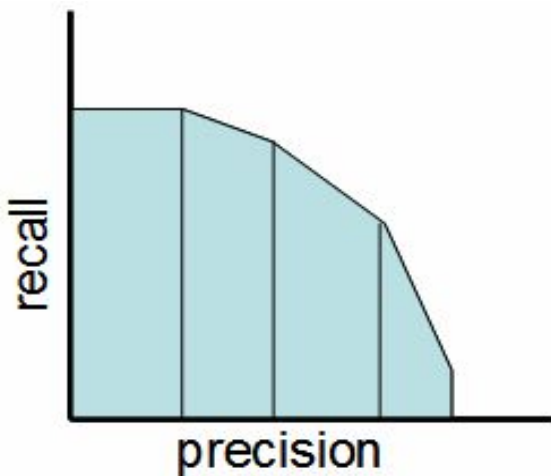
Matriz de confusión

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

También llamada **recall**

Evaluando regresión logística

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$



- Es la media armónica entre precision y recall
- Balancea ambas métricas
- Si precision es baja => F1 es baja
- Si recall es baja => F1 es baja

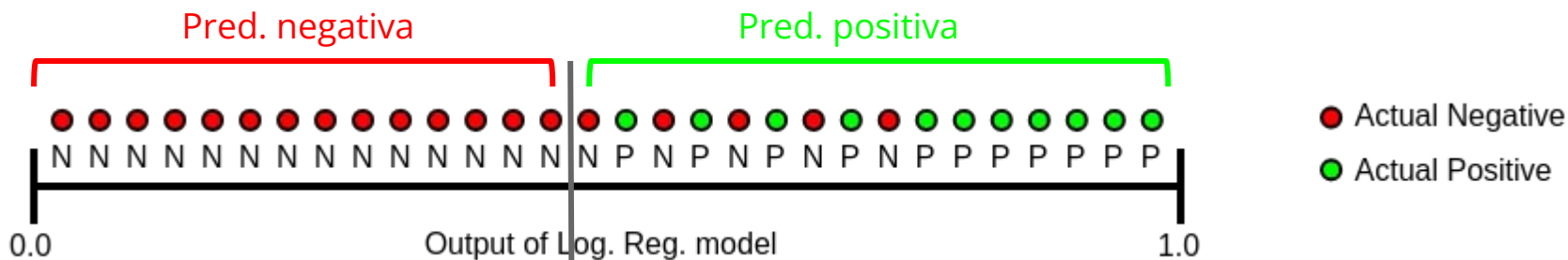
Evaluando regresión logística

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}.$$

- F1 da igual peso a precision y recall
- Fb permite balancear ese peso, mediante el parámetro beta
- Beta mide cuánta importancia le damos a precision en relación al recall

De una probabilidad a una categoría

- Variando el umbral (0.5), puedo hacer hacer variar las métricas basadas en la matriz de confusión

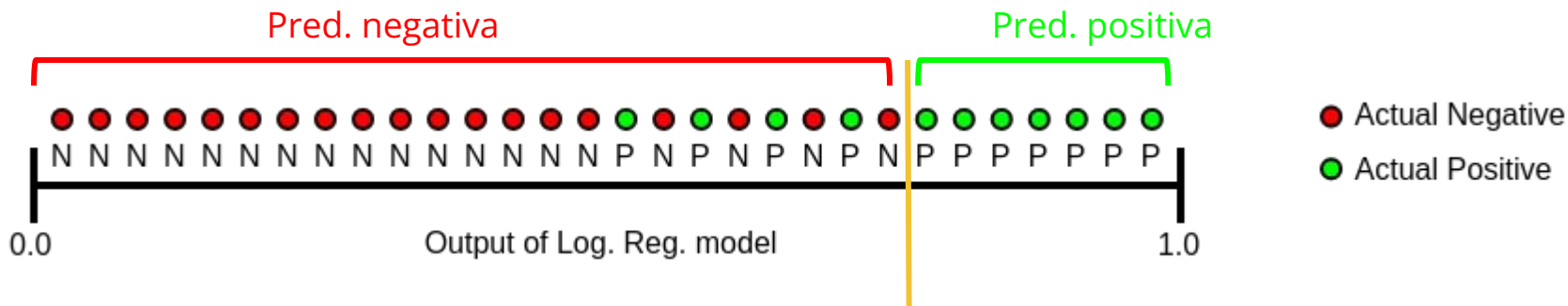


Umbral = 0.5
Accuracy = 0.83
Precision = 0.69
Recall = 1.0

	Positivo (pred)	Negativo (pred)	Total
Positivo (obs)	11	0	11
Negativo (obs)	5	14	19
Total	16	14	30

De una probabilidad a una categoría

- Variando el umbral (0.5), puedo hacer hacer variar las métricas basadas en la matriz de confusión

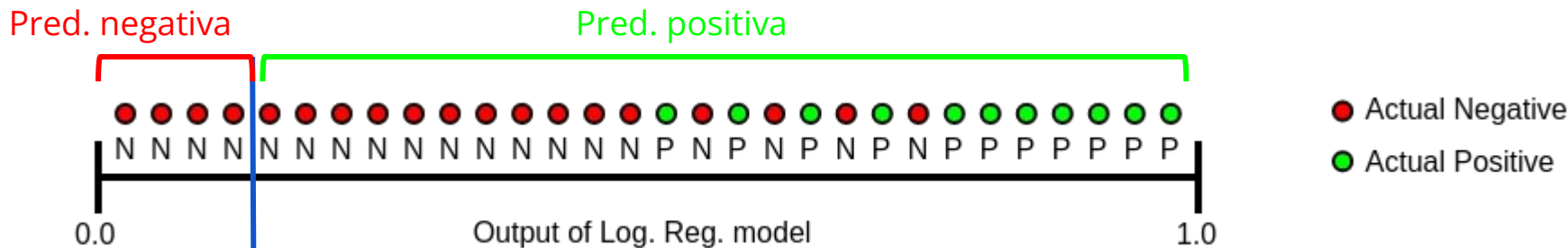


	Positivo (pred)	Negativo (pred)	Total
Positivo (obs)	7	4	11
Negativo (obs)	0	19	19
Total	7	23	30

Umbral = 0.76
Accuracy = 0.87
Precision = 1.0
Recall = 0.64

De una probabilidad a una categoría

- Variando el umbral (0.5), puedo hacer hacer variar las métricas basadas en la matriz de confusión



Umbral = 0.13

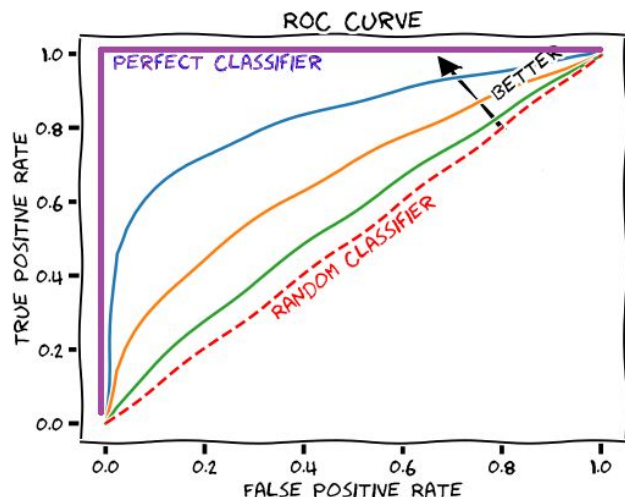
Accuracy = ?

Precision = ?

Recall = ?

	Positivo (pred)	Negativo (pred)	Total
Positivo (obs)	11	0	11
Negativo (obs)	15	4	19
Total	26	4	30

Evaluando regresión logística

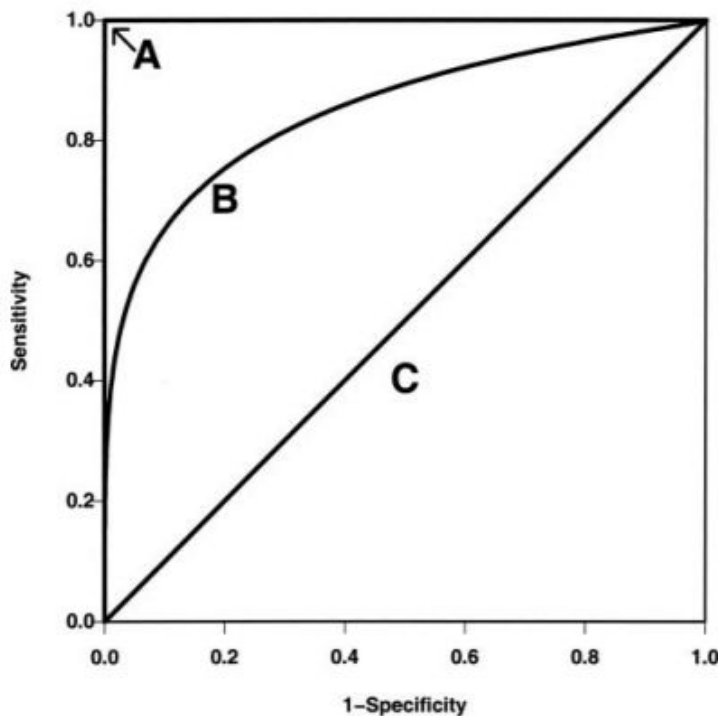


- Curva ROC y métrica AUC: cuantificar la performance agregada de diferentes clasificadores para todos los umbrales posibles
- Eje X: tasa de falsos positivos (1 - especificidad)
- Eje Y: tasa de verdaderos positivos (sensibilidad)
- Querríamos clasificadores que scoreen alto en Y y bajo en X

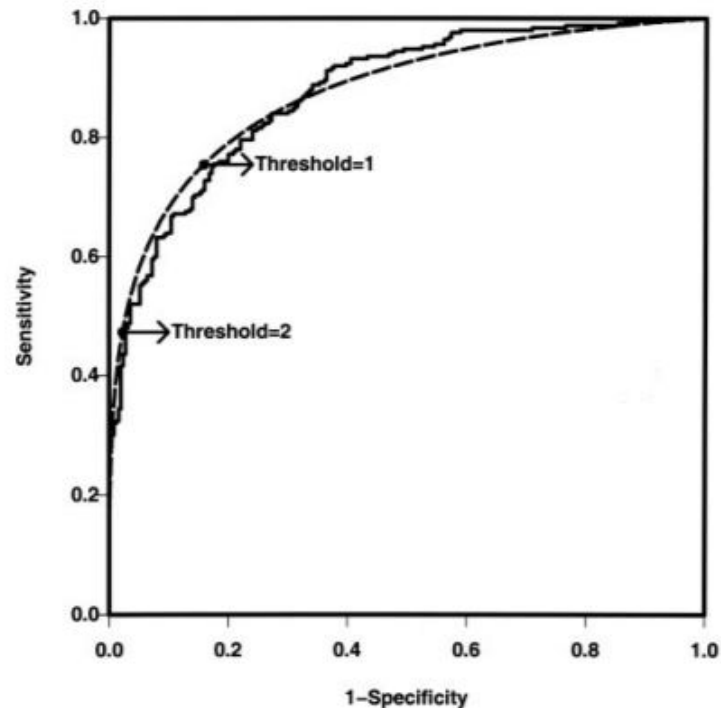
Evaluando regresión logística

True positive rate (o recall)

$$\frac{tp}{tp + fn}$$

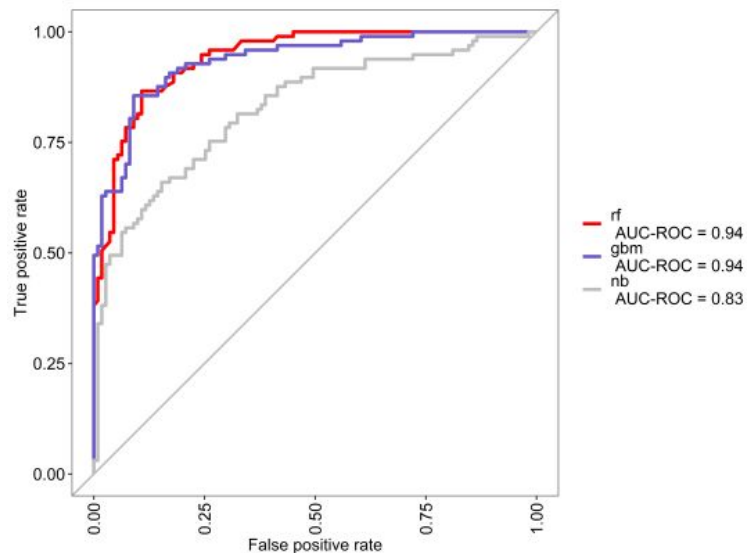


False positive rate



$$\frac{FP}{FP + TN}$$

Evaluando regresión logística



- En este ejemplo tenemos plotados las curvas ROC de tres modelos de regresión
- ¿Cuál es mejor? ¿Por qué?