

Extracto de:

Five Minute Science Lesson: Test Accuracy Metrics

Raymond Nelson

APA MAGAZINE

THE MAGAZINE FOR THE POLYGRAPH PROFESSIONAL

MAY• JUN Vol. 53.3

Traductor:

rodolfo@poligrafia.com.mx

Términos conceptuales para entender la “Precisión”

		Estado externo de criterio (realidad)		
		Estado Positivo	Estado Negativo	
Resultado de prueba o experimento	Resultado Positivo	VP (verdadero positivo)	FP (falso positivo) error tipo 1	Valor Predictivo Positivo (VPP) = VP / todos los Resultados Positivos Índice de Falsos Positivos (IFP) = FP / todos los Resultados Positivos
	Resultado Negativo	FN (falso negativo) error tipo 2	VN (verdadero negativo)	Valor Predictivo Negativo (VPN) = VN / todos los Resultados Negativos Índice de Falso Negativo (IFN) = FN / todos los Resultados Negativos
		Sensibilidad de prueba = VP / todos los casos de Estado Positivo	Especificidad de prueba = VN / todos los casos de Estado Negativo	Precisión

Estado del criterio (criterio externo): es cualquier fenómeno de interés que queremos cuantificar, clasificar o predecir. Esto se conoció en algún momento como el *parámetro* o el *fenómeno desconocido* de interés.

Resultado de prueba: a menudo se refiere al resultado categórico de la prueba. Los resultados de las pruebas científicas no se discuten en términos de *aprobar* o *reprobar* - aunque en algunos contextos podría existir una tendencia hacia interpretaciones prácticas en este nivel - sino que utilizan los términos *positivo* y *negativo* para indicar si la probabilidad es suficiente para apoyar una conclusión categórica sobre la presencia o ausencia del parámetro desconocido.

Resultado positivo: es un término que significa que un resultado de prueba probabilística apoya una conclusión categórica de que el parámetro o fenómeno desconocido de interés está *presente* en un caso.

Resultado negativo: significa que un resultado probabilístico de prueba apoya una conclusión categórica de que el parámetro o fenómeno desconocido de interés está *ausente* en un caso.

Estado positivo: se refiere a si el parámetro o fenómeno desconocido de interés está *realmente presente* (en la realidad) en un caso.

Estado negativo: se refiere a si el parámetro o fenómeno desconocido de interés está *realmente ausente* (en la realidad) en un caso.

Verdadero positivo (VP): describe un *resultado positivo* que es concurrente con un *estado positivo*. La prueba ha identificado *correctamente* la *presencia* de los fenómenos desconocidos de interés de un caso.

Verdadero negativo (VN): describe un *resultado negativo* que es concurrente con un *estado negativo*. La prueba ha identificado *correctamente* la *ausencia* de los fenómenos desconocidos de interés de un caso.

Falso positivo (FP): describe un *resultado positivo* que es concurrente con un *estado negativo*. Una prueba ha identificado *incorrectamente* la *presencia* de los fenómenos desconocidos de un caso para el que el estado del criterio externo (realidad) es realmente negativo (los fenómenos desconocidos de interés están realmente *ausentes*). FP a veces se calcula como 1-especificidad. Sin embargo, el cálculo de 1-especificidad será incorrecto para las pruebas que incluyen el uso

de clasificaciones de inconclusos (cuando un resultado probabilístico no es estadísticamente significativo para la clasificación positiva o negativa).

Falso negativo (FN): describe un *resultado negativo* que es concurrente con un *estado positivo*. Una prueba ha identificado *incorrectamente* la *presencia* de los fenómenos desconocidos en un caso para el que el estado del criterio externo (realidad) es realmente positivo (los fenómenos desconocidos de interés están realmente *presentes*).

FP a veces se calcula como 1-sensibilidad. Sin embargo, el cálculo de 1-sensibilidad será incorrecto para las pruebas que incluyen el uso de clasificaciones de inconclusos (cuando un resultado de la prueba no es positivo ni negativo).

Valor predictivo positivo (VPP): se refiere a la proporción de resultados VP con respecto a los resultados (VP + FN) de un grupo de casos en los que se conoce el estado positivo o negativo real. (Los científicos de los datos, a veces se refieren a estos como *casos etiquetados*.) Puede ser útil para estimar la probabilidad de que un resultado positivo sea *correcto* en un caso desconocido. Sin embargo, el VPP no es resistente a las diferencias en la proporción previa de casos con estado positivo y estado negativo (tasa base o tasa de incidencia). Es decir, el VPP – la posibilidad que un resultado positivo sea correcto – estará en función de la proporción de casos de estado positivo dentro del grupo de casos, además de estar influenciado por la sensibilidad de la prueba y las tasas de FP.

Valor predictivo negativo (VPN): se refiere a la proporción de resultados TN con respecto a los resultados (TN + FN) de un grupo de casos etiquetados (es decir, casos para los que se conoce el estado positivo real o el estado negativo). Puede ser útil para estimar la probabilidad de que un resultado negativo sea *correcto* en un caso desconocido. El VPN tampoco es resistente a las diferencias en la proporción previa de casos de estado positivo y estado negativo (tasa base o tasa de incidencia). El VPN – la probabilidad de que un resultado negativo sea correcto – variará tanto con la tasa de especificidad de la prueba como con la proporción de casos de estado positivo (es decir, probabilidad previa, tasa base o tasa de incidencia) en un grupo de casos.

Índice de falsos positivos (IFP): se refiere a la proporción de resultados PF con respecto a todos los resultados positivos (PF + TP) para un grupo de casos etiquetados. Puede ser útil para estimar la probabilidad de que un resultado positivo sea *incorrecto* en un caso

desconocido. El FPI no es resistente al desequilibrio grupal (es decir, diferencias en la proporción de casos con estado positivo y estado negativo) y variará tanto con la tasa base previa o la tasa de incidencia como con la tasa de sensibilidad de la prueba.

Índice de falsos negativos (IFN): se refiere a la proporción de resultados FN con respecto a todos los resultados negativos (FN + TN) para un grupo de casos etiquetados. Puede ser útil para estimar la probabilidad de que un resultado negativo sea *incorrecto* en un caso desconocido. IFN no es resistente al desequilibrio grupal. El IFN variará con la tasa de especificidad de la prueba y también con las diferencias en la proporción de casos de estado negativo y estado positivo (tasa base previa o tasa de incidencia).

Sensibilidad: se refiere a la proporción de un grupo de casos con estado positivo para los que una prueba puede identificar *correctamente* la *presencia* de los fenómenos desconocidos de interés. Debido a que se calcula solo dentro del subgrupo de casos de estado positivo, la sensibilidad de la prueba es resistente a las diferencias en la tasa de incidencia previa. Es decir, la sensibilidad de la prueba será invariable con respecto al desequilibrio del grupo.

Especificidad: se refiere a la proporción de un grupo de casos de estado negativo para los que una prueba puede identificar *correctamente* la *ausencia* de los fenómenos desconocidos de interés. La especificidad de la prueba se calcula dentro del subgrupo de casos de estado negativo y es resistente a las diferencias de desequilibrio del grupo (probabilidad previa, tasa base o tasa de incidencia).