Cómo entender el error, reducir la incertidumbre y explicar los resultados con confianza

Ing. Eduardo Lebrija
Octubre 2026

"Sin datos, solo eres otra persona más dando su opinión"
W. Edwards Deming

Agenda (120 min)

1. Características de las Pruebas Científicas
2. Errores de Prueba
3. Términos conceptuales y métricas de precisión
4. Tasas Base de Culpabilidad o Inocencia (probabilidades previas)
5. Agregación de Probabilidades Previas
6. Estadística Frecuentista y Estadística Bayesiana (Conceptos básicos)
7. Ejemplos y discusión

CONGRESO INTERNACIONAL AIPP
MÉRIDA
MÉXICO 2026

Eventos Deterministas vs Probabilísticos

Determinista: resultado exacto

- Soltar una piedra
- Colocar agua al fuego
- Soltar un globo con helio
- Agitar una botella de refresco (soda)

Probabilístico: resultado con incertidumbre

- Todas las pruebas científicas

Ejemplo (polígrafo):

•Determinista: “Mintió”

•Probabilístico: “Hay 80% de probabilidad de engaño”

Características de las Pruebas Científicas

- Objetivas
- Verificables / Empirismo
- Replicables y Metódicas
- Siguen orden lógico / procedimientos estructurados
- Falsabilidad (Provisionales)
- Revisadas por Pares
- Específicas
- Sujetas a margen de error

Las pruebas científicas no son infalibles

Una prueba cuantifica un parámetro que no se puede medir directamente, usando información proxy correlacionada con él. Por eso es probabilística: expresa la fuerza de la evidencia o el margen de incertidumbre de una conclusión.

1

De procedimiento

Mal uso del procedimiento o del instrumento

2

Aleatorio

Variación normal no controlada

3

Sistemático

Error reproducible, siempre hacia el mismo lado

Error de Procedimiento

Ejecución incorrecta de los procedimientos o uso incorrecto del instrumento.

Ejemplos

- Preguntas relevantes ambiguas o mal construidas
- Sensores mal colocados; artefactos sin marcar
- Intervalo entre preguntas menor a 20 s
- Omitir la prueba de familiarización (Estándares APA)

Cómo minimizarlos

- Estándares publicados y reglas de procedimiento
- Supervisión, control de calidad y educación continua
- Automatizar lo posible; lo sutil seguirá requiriendo criterio profesional

Error aleatorio

Variación esperada en los datos, puntajes y resultados si repitiéramos la prueba muchas veces.

Ejemplos

- Mismo examinado, dos pruebas: puntaje +5 y +3 sin cambio real
- Fatiga, ruido, movimiento, respiración irregular, temas de idoneidad
- Un puntaje cercano al corte que cambia con una repetición

Cómo minimizarlo

- Obtener más datos: ley de los grandes números y teorema del límite central ("medir dos veces, cortar una")
- Intervalos de confianza (frecuentistas) y de credibilidad (bayesianos)
- Fijar la tolerancia al error aleatorio: alfa = .05

Error sistemático

Error reproducible: falla de la teoría, los procedimientos o el aparato, que desvía los resultados siempre en la misma dirección.

Ejemplos

- Un sensor que subestima siempre la amplitud
- Técnica o regla de decisión con más falsos positivos de forma constante
- Preguntas comparativas mal discutidas que sesgan a veraces

Cómo minimizarlo

- Buscar el patrón (¿hacia dónde se desvían todos?) y ajustar, como al corregir la mira
- Estimarlo con análisis bayesiano
- Replicar en otras condiciones: el sesgo puede cambiar (otra distancia, otro blanco)

Discusión: ¿qué tipo de error es?

CONGRESO INTERNACIONAL AIPP MÉRIDA MÉXICO 2026

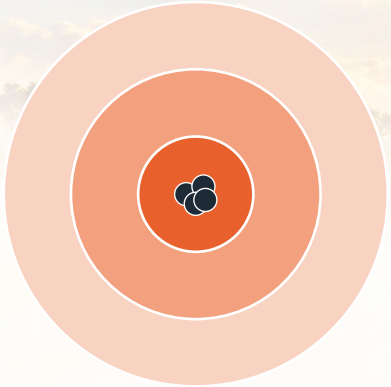
A El sensor de cardio quedó mal colocado durante toda la sesión

B El mismo examinado obtiene +7 y +3 en dos sesiones sin razón aparente

C Con un instrumento, los puntajes salen consistentemente más bajos de lo esperado

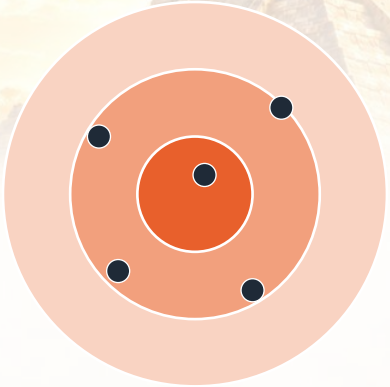
Fuera del alcance de la presentación: contramedidas e idoneidad del examinado.

Exactitud vs precisión



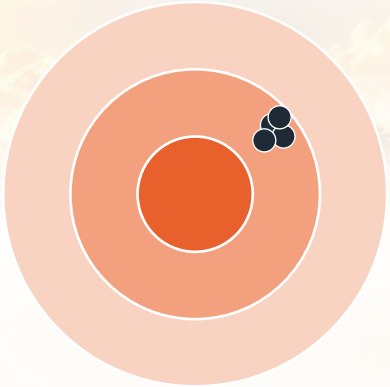
Sin error

Exacto y preciso



Error aleatorio

Exacto, poco preciso



Error sistemático

Preciso, poco exacto

Frecuentista (alfa, valor-p) se asocia con error aleatorio. Bayesiana (probabilidades previas) con error sistemático.

Errores Tipo I y II

Resultado positivo: la persona es categorizada como engañosa

Resultado negativo: la persona es categorizada como veraz.

Falso positivo: Inocente marcado como engañoso. (Tipo I)

Falso negativo: Engañoso que pasa la prueba como persona veraz. (Tipo II)

Sensibilidad y Especificidad

Sensibilidad: es la capacidad de la prueba para detectar lo que estamos buscando (detectar mentirosos)

- Capacidad de detectar los verdaderos positivos; i.e. los verdaderos engañosos
- Ej. 95% de sensibilidad: detectará 95 de cada 100 mentirosos, pero a 5 los dará falsos negativos

Especificidad: es la capacidad para discriminar lo que no estamos buscando (detectar veraces)

- Capacidad de “no detectar” los verdaderos negativos; i.e. los verdaderos inocentes
- Ej. 90% de especificidad: identificará correctamente a 90 de 100 veraces, pero a 10 los dará falsos positivos

CONGRESO INTERNACIONAL AIPP
MÉRIDA
MÉXICO 2026

Valores Predictivos

VPP: Valor predictivo positivo.

- probabilidad de culpabilidad dado resultado positivo (falló la prueba)

VPN: Valor predictivo negativo.

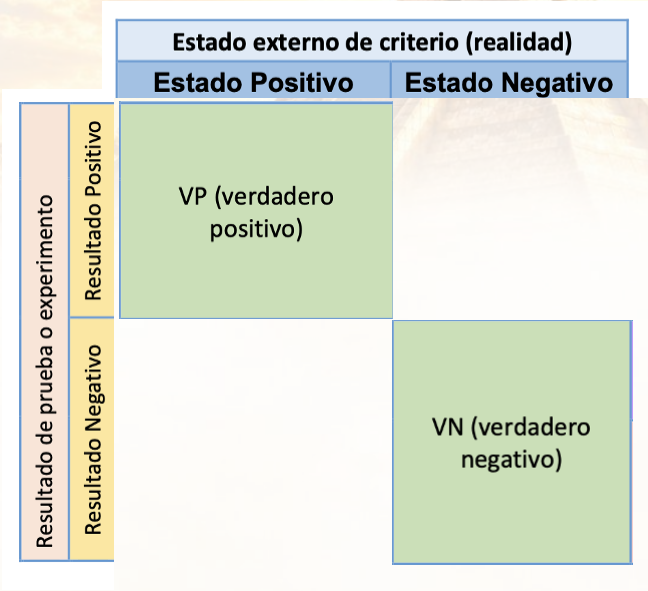
- probabilidad de inocencia dado resultado negativo (pasó la prueba)

Métricas de Precisión

		Estado externo de criterio (realidad)	
		Estado Positivo	Estado Negativo
Resultado de prueba o experimento	Resultado Positivo		
	Resultado Negativo		

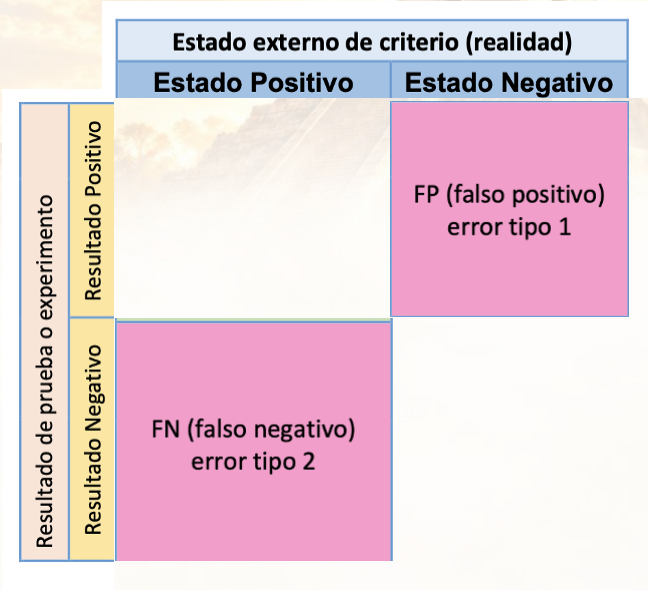
Lección de Cinco Minutos de Ciencia: Métricas de Precisión de las Pruebas
Raymond Nelson, APA Magazine 2020, 53(3)

Métricas de Precisión



Lección de Cinco Minutos de Ciencia: Métricas de Precisión de las Pruebas
Raymond Nelson, APA Magazine 2020, 53(3)

Métricas de Precisión



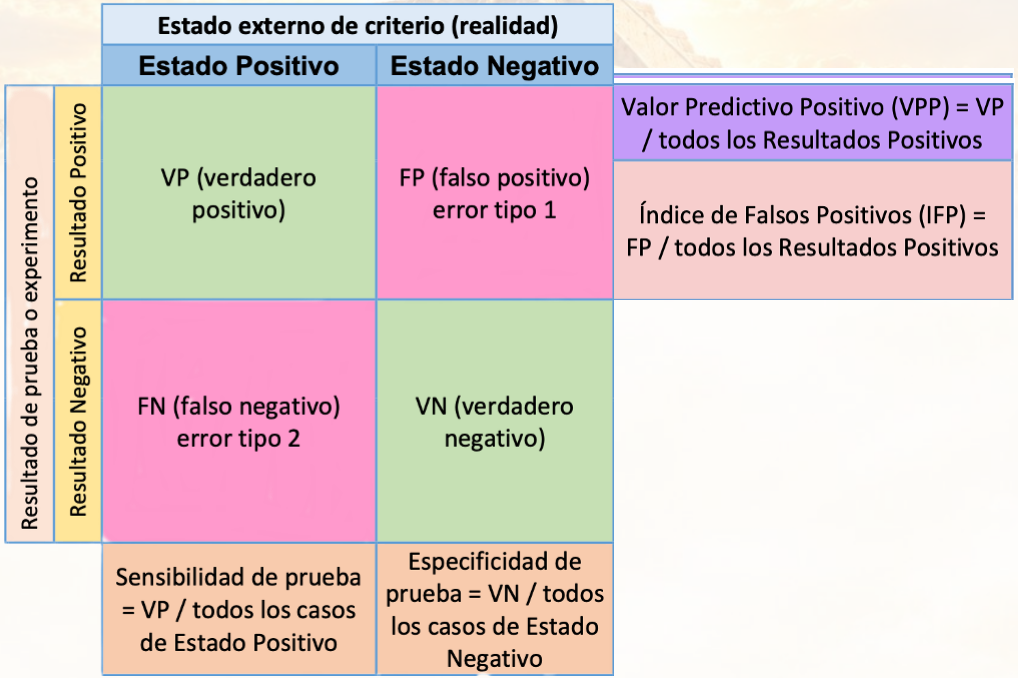
Lección de Cinco Minutos de Ciencia: Métricas de Precisión de las Pruebas
Raymond Nelson, APA Magazine 2020, 53(3)

Métricas de 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
	Resultado Negativo	FN (falso negativo) error tipo 2	VN (verdadero negativo)
		Sensibilidad de prueba = VP / todos los casos de Estado Positivo	Especificidad de prueba = VN / todos los casos de Estado Negativo

Lección de Cinco Minutos de Ciencia: Métricas de Precisión de las Pruebas
Raymond Nelson, APA Magazine 2020, 53(3)

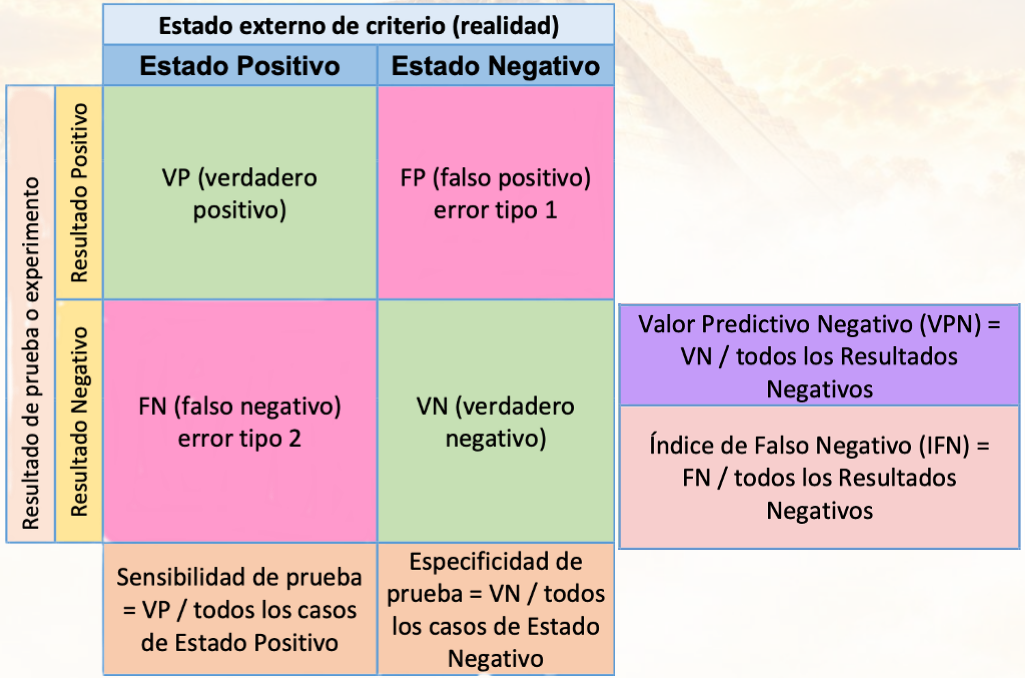
Métricas de Precisión



Resultado de prueba o experimento		Estado externo de criterio (realidad): Estado Positivo	Estado externo de criterio (realidad): Estado Negativo	
	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)	
		Sensibilidad de prueba = VP / todos los casos de Estado Positivo	Especificidad de prueba = VN / todos los casos de Estado Negativo	

Lección de Cinco Minutos de Ciencia: Métricas de Precisión de las Pruebas
Raymond Nelson, APA Magazine 2020, 53(3)

Métricas de Precisión



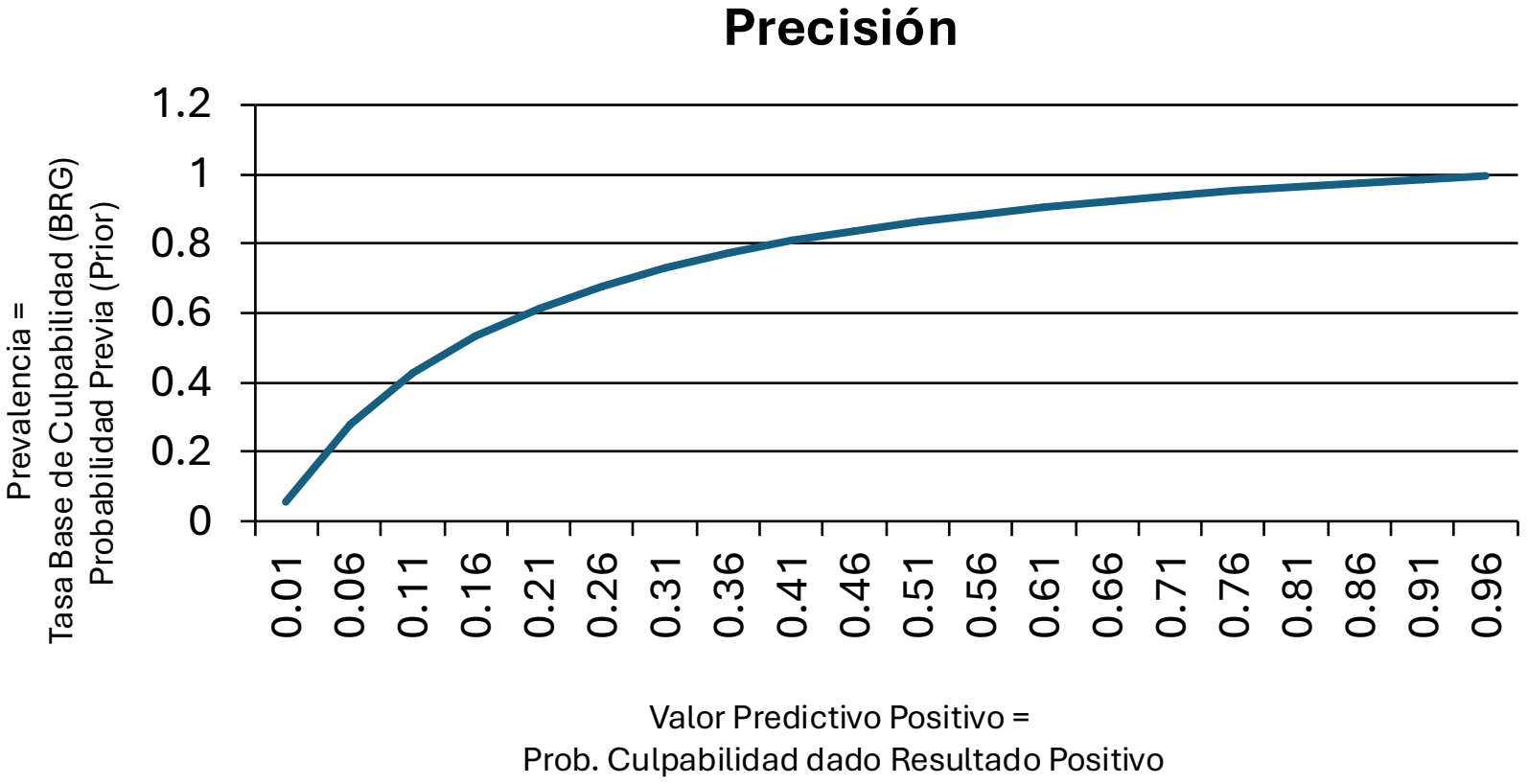
Lección de Cinco Minutos de Ciencia: Métricas de Precisión de las Pruebas
Raymond Nelson, APA Magazine 2020, 53(3)

Métricas de 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	

Lección de Cinco Minutos de Ciencia: Métricas de Precisión de las Pruebas
Raymond Nelson, APA Magazine 2020, 53(3)

VPP vs Prevalencia (Sens=90%, Esp=85%)



Estimación de las tasas base a partir de resultados observados

Se usan las proporciones V y NV observadas, junto con Sens, Spec, FN y FP de la técnica.

Ecuaciones iniciales

$$NV = (Sens \times BRG) + (FP \times BRI)$$
$$V = (Spec \times BRI) + (FN \times BRG)$$

Relaciones entre variables

$BRG = 1 - BRI$

FP y FN se introducen como tasas independientes

... desarrollo algebraico omitido ...

Relación entre las tasas base

$$BRG = 1 - BRI$$

Fórmula de la tasa base de inocentes

$$BRI = (V - FN) / (Spec - FN)$$

Definiciones

Sens	Sensibilidad	**Spec**	Especificidad	**FN**	Tasa de falsos negativos	**FP**	Tasa de falsos positivos
BRG	Tasa base de culpables	**BRI**	Tasa base de inocentes	**NV**	Proporción de resultados No Veraz	**V**	Proporción de resultados Veraz

V y NV son proporciones observadas; FN y FP son tasas de error. Puede haber resultados inconclusos.

Basado en Grubin, Handler, Nelson y Raskin (2016), APA Magazine, 49(5), pp. 39–47.

Ejemplo 1: DLST/ESS — estimación de las tasas base

Sens = 0,81 • Spec = 0,75 • FP = 0,15 • FN = 0,11

Resultados observados

65% pasan (NSR) • 35% no pasan según el reporte de la agencia

Cálculo

BRI = (0,65 − 0,11) / (0,75 − 0,11) = 0,84
BRG = 1 − 0,84 = 0,16

Tasas base estimadas a partir del porcentaje que pasa

BRG = 16%
≈ 16 de cada 100 casos

BRI = 84%
≈ 84 de cada 100 casos

NPV = 63 / (63 + 1,76) ≈ 97% • **PPV = 12,96 / (12,96 + 12,60) ≈ 51%**

La tabla 1 estima 9,64% inconclusos; 25,6% SR + 9,64% inconclusos ≈ 35% que no pasan.

Basado en Grubin, Handler, Nelson y Raskin (2016), APA Magazine, 49(5), pp. 39–47.

Ejemplo 2: AFMGQT/LEPET (formato MGQT)

Sens = 0,76 • Spec = 0,25 • FP = 0,11 • FN = 0,06

Resultados observados

17% pasan (NSR) • 83% no pasan según el reporte de la agencia

Cálculo

BRI = (0,17 – 0,06) / (0,25 – 0,06) = 0,58
BRG = 1 – 0,58 = 0,42

Tasas base estimadas a partir del porcentaje que pasa

BRG = 42%
≈ 42 de cada 100 casos

BRI = 58%
≈ 58 de cada 100 casos

NPV = 14,5 / (14,5 + 2,52) ≈ 85% • **PPV = 31,92 / (31,92 + 6,38) ≈ 83%**

La tabla 2 estima 44,6% inconclusos; 38,4% SR + 44,6% inconclusos = 83% que no pasan.

Basado en Grubin, Handler, Nelson y Raskin (2016), APA Magazine, 49(5), pp. 39–47.

Apéndice E-2: Puntuaciones de corte de ESS-M para exámenes asuntos múltiples

Puntuaciones ESS-M / Examen de asuntos múltiples

Probabilidades previas de engaño	Probabilidad previa	Alfa (veracidad/engaño)								
		.01/.01	.01/.05	.01/.10	.05/.01	.05/.05	.05/.10	.10/.01	.10/.05	.10/.10
9 a 1 (9 en 10)	.90	(+8) / ninguno	(+8) / ninguno	(+8) / -9	(+7) / ninguno	(+7) / ninguno	(+7) / -9	(+7) / ninguno	(+7) / ninguno	(+7) / -9
8 a 1 (8 en 9)	.89	(+7) / ninguno	(+7) / ninguno	(+7) / -6	(+7) / ninguno	(+7) / ninguno	(+7) / -6	(+7) / ninguno	(+7) / ninguno	(+7) / -6
7 a 1 (7 en 8)	.88	(+7) / ninguno	(+7) / -10	(+7) / -6	(+7) / ninguno	(+7) / -10	(+7) / -6	(+7) / ninguno	(+7) / -10	(+7) / -6
6 a 1 (6 en 7)	.86	(+7) / ninguno	(+7) / -8	(+7) / -5	(+6) / Ninguno	(+6) / -8	(+6) / -4	(+6) / Ninguno	(+6) / -8	(+6) / -5
5 a 1 (5 en 6)	.83	(+6) / -14	(+6) / -5	(+6) / -4	(+6) / -14	(+6) / -5	(+6) / -4	(+6) / -14	(+6) / -5	(+6) / -4
4 a 1 (4 en 5)	.80	(+6) / -9	(+6) / -5	(+6) / -4	(+5) / -9	(+5) / -5	(+5) / -4	(+5) / -9	(+5) / -5	(+5) / -4
3 a 1 (3 en 4)	.75	(+5) / -6	(+5) / -4	(+5) / -3	(+5) / -6	(+5) / -4	(+5) / -3	(+4) / -6	(+4) / -4	(+4) / -3
2 a 1 (2 en 3)	.67	(+4) / -5	(+4) / -3	(+4) / -3	(+3) / -5	(+3) / -3	(+3) / -3	(+3) / -5	(+3) / -3	(+3) / -3
1 a 1 (1 en 2)	**.50**	**(+2) / -4**	**(+2) / -3**	**(+2) / -2**	**(+1) / -4**	**(+1) / -3**	**(+1) / -2**	**(+1) / -4**	**(+1) / -3**	**(+1) / -2**
De 1 a 2 (1 en 3)	.33	(+1) / -6	(+1) / -5	(+1) / -4	(0) / -6	(0) / -5	(0) / -4	(0) / -6	(0) / -5	(0) / -4
De 1 a 3 (1 en 4)	.25	(0) / -7	(0) / -6	(0) / -6	(0) / -7	(0) / -6	(0) / -6	(0) / -7	(0) / -6	(0) / -6
De 1 a 4 (1 en 5)	.20	(+1) / -7	(0) / -7	(0) / -6	(0) / -7	(0) / -7	(0) / -6	(0) / -7	(0) / -7	(0) / -6
De 1 a 5 (1 en 6)	.17	(+4) / -8	(+4) / -7	(+4) / -7	(0) / -8	(0) / -7	(0) / -7	(0) / -8	(0) / -7	(0) / -7
De 1 a 6 (1 en 7)	.14	(Ninguno) / -8	(Ninguno) / -8	(Ninguno) / -7	(0) / -8	(0) / -8	(0) / -7	(0) / -8	(0) / -8	(0) / -7
De 1 a 7 (1 en 8)	.13	(Ninguno) / -9	(Ninguno) / -8	(Ninguno) / -8	(0) / -9	(0) / -8	(0) / -8	(0) / -9	(0) / -8	(0) / -8
De 1 a 8 (1 en 9)	.11	(Ninguno) / -9	(Ninguno) / -8	(Ninguno) / -8	(Ninguno) / -9	(Ninguno) / -8	(Ninguno) / -8	(0) / -9	(0) / -8	(0) / -8
De 1 a 9 (1 en 10)	.10	(Ninguno) / -9	(Ninguno) / -9	(Ninguno) / -8	(Ninguno) / -9	(Ninguno) / -9	(Ninguno) / -8	(0) / -9	(0) / -9	(0) / -8

Los paréntesis indican el uso de una corrección estadística para los efectos de multiplicidad.

Apéndice E-1: Puntuaciones de corte de ESS-M para exámenes de un solo tema

Puntuaciones ESS-M / Examen de evento específico										
		Alfa (veracidad/engaño)								
Probabilidades previas de engaño	Probabilidad previa	.01/.01	.01/.05	.01/.10	.05/.01	.05/.05	.05/.10	.10/.01	.10/.05	.10/.10
9 a 1 (9 en 10)	.90	+14 / -9 (ninguno)	+14 / -6 (ninguno)	+14 / -4 (-21)	+13 / -9 (ninguno)	+13 / -6 (ninguno)	+13 / -4 (-21)	+13 / -9 (ninguno)	+13 / -6 (ninguno)	+13 / -4 (-21)
8 a 1 (8 en 9)	.89	+13 / -8 (ninguno)	+13 / -5 (ninguno)	+13 / -4 (-20)	+13 / -8 (ninguno)	+13 / -5 (ninguno)	+13 / -4 (-20)	+12 / -8 (ninguno)	+12 / -5 (ninguno)	+12 / -4 (-20)
7 a 1 (7 en 8)	.88	+13 / -7 (ninguno)	+13 / -5 (-22)	+13 / -4 (-18)	+12 / -7 (ninguno)	+12 / -5 (-22)	+13 / -4 (-18)	+12 / -7 (ninguno)	+12 / -5 (-22)	+12 / -4 (-18)
6 a 1 (6 en 7)	.86	+12 / -7 (ninguno)	+12 / -4 (-20)	+12 / -4 (-16)	+11 / -7 (ninguno)	+11 / -4 (-20)	+11 / -4 (-16)	+11 / -7 (ninguno)	+11 / -4 (-20)	+11 / -4 (-16)
5 a 1 (5 en 6)	.83	+11 / -6 (ninguno)	+11 / -4 (-17)	+11 / -3 (-15)	+10 / -6 (ninguno)	+10 / -4 (-17)	+10 / -3 (-15)	+10 / -6 (ninguno)	+10 / -4 (-17)	+10 / -3 (-15)
4 a 1 (4 en 5)	.80	+10 / -5 (-21)	+10 / -4 (-15)	+10 / -3 (-14)	+10 / -5 (-21)	+10 / -4 (-15)	+10 / -3 (-14)	+9 / -5 (-21)	+9 / -4 (-15)	+9 / -3 (-14)
3 a 1 (3 en 4)	.75	+9 / -5 (-16)	+9 / -3 (-13)	+9 / -3 (-12)	+8 / -5 (-16)	+8 / -3 (-13)	+8 / -3 (-12)	+8 / -5 (-16)	+8 / -3 (-13)	+8 / -3 (-12)
2 a 1 (2 en 3)	.67	+7 / -4 (-13)	+7 / -3 (-11)	+7 / -2 (-10)	+6 / -4 (-13)	+6 / -3 (-11)	+6 / -2 (-10	+6 / -4 (-13)	+6 / -3 (-11)	+6 / -2 (-10)
1 a 1 (1 en 2)	**.50**	**+4 / -4 (-9)**	**+4 / -3 (-7)**	**+4 / -2 (-6)**	**+3 / -4 (-9)**	**+3 / - 3 (-7)**	**+3 / -2 (-6)**	**+2 / -4 (-9)**	**+2 / -3 (-7)**	**+2 / -2 (-6)**
De 1 a 2 (1 en 3)	.33	+4 / -7 (-11)	+4 / -6 (-9)	+4 / -6 (-8)	+3 / -7 (-11)	+3 / -6 (-9)	+3 / -6 (-8)	+2 / -7 (-11)	+2 / -6 (-9)	+2 / -6 (-8)
De 1 a 3 (1 en 4)	.25	+5 / -9 (-11)	+5 / -8 (-9)	+5 / -8 (-8)	+3 / -9 (-11)	+3 / -8 (-9)	+3 / -8 (-8)	+3 / -9 (-11)	+3 / -8 (-9)	+3 / -8 (-8)
De 1 a 4 (1 en 5)	.20	+5 / -10 (-12)	+5 / -10 (-10)	+5 / -9 (-9)	+4 / -10 (-12)	+4 / -10 (-10)	+4 / -9 (-9)	+3 / -10 (-12)	+3 / -10 (-10)	+3 / -9 (-9)
De 1 a 5 (1 en 6)	.17	+6 / -11 (-12)	+6 / -10 (-10)	+6 / -10 (-10)	+4 / -11 (-12)	+4 / -10 (-10)	+4 / -10 (-10)	+3 / -11 (-12)	+3 / -10 (-10)	+3 / -10 (-10)
De 1 a 6 (1 en 7)	.14	+7 / -12 (-13)	+7 / -11 (-11)	+7 / -11 (-10)	+4 / -12 (-13)	+4 / -11 (-11)	+4 / -11 (-10)	+4 / -12 (-13)	+4 / -11 (-11)	+4 / -11 (-10)
De 1 a 7 (1 en 8)	.13	+7 / -12 (-13)	+7 / -12 (-11)	+7 / -11 (-10)	+5 / -12 (-13)	+5 / -12 (-11)	+5 / -11 (-10)	+4 / -12 (-13)	+4 / -12 (-11)	+4 / -11 (-10)
De 1 a 8 (1 en 9)	.11	+8 / -13 (-13)	+7 / -13 (-11)	+8 / -12 (-10)	+5 / -13 (-13)	+5 / -13 (-11)	+5 / -12 (-10)	+4 / -13 (-13)	+4 / -13 (-11)	+4 / -12 (-10)
De 1 a 9 (1 en 10)	.10	+9 / -14 (-13)	+9 / -13 (-12)	+9 / -13 (-11)	+6 / -14 (-13)	+6 / -13 (-12)	+6 / -13 (-11)	+4 / -14 (-13)	+4 / -13 (-12)	+4 / -13 (-11)
Los paréntesis indican el uso de una corrección estadística para los efectos de multiplicidad.										

Agregación de probabilidades previas

Nelson (2026) • Pruebas de selección de múltiples asuntos

Idea central

- En una prueba de múltiples asuntos se evalúan varios comportamientos posibles.
- La decisión compuesta es si ocurrió al menos una conducta relevante.
- Al combinar asuntos independientes, aumenta la probabilidad previa del evento compuesto.

Fórmula bajo el supuesto de independencia

$$P(\text{al menos uno}) = 1 - \prod_{i=1}^{k} (1 - p_i)$$

Dos asuntos: $P(A \cup B) = 1 - (1 - pA)(1 - pB)$

Supuesto y alcance

- La fórmula supone independencia estadística entre los asuntos evaluados.
- La independencia es un modelo de referencia; no demuestra que los comportamientos sean independientes en la práctica.
- La agregación estima el prior de “cualquier conducta relevante”, no la probabilidad posterior tras conocer el resultado.

Fuente: Nelson, R. (2026), Polygraph & Forensic Credibility Assessment, 55(1)

Ejemplo: combinar dos probabilidades previas

Un asunto con prior bajo y otro con prior moderado • Se asume independencia

Asunto 1

$p_1 = 0{,}10$

+

Asunto 2

$p_2 = 0{,}50$

→

Prior agregado

$P = 0{,}55$

Cálculo

$$P(\text{al menos uno}) = 1 - (1 - 0{,}10)(1 - 0{,}50)$$
$$= 1 - (0{,}90 \times 0{,}50) = 1 - 0{,}45 = 0{,}55$$

Interpretación

El prior de al menos una conducta pasa a 55%.

No significa

No significa que cada asunto tenga prior de 55%, ni que la prueba tenga 55% de probabilidad de acertar.

Fuente: Nelson, R. (2026), Polygraph & Forensic Credibility Assessment, 55(1)

Discusión y conclusiones

Qué permite concluir el modelo y qué queda pendiente de demostrar

Discusión

- El prior agregado depende del número de asuntos y de sus probabilidades individuales.
- La agregación puede mover un prior bajo hacia un nivel más moderado.
- Con sensibilidad y especificidad constantes, un prior mayor eleva el PPV y reduce la proporción de FP entre resultados positivos.
- La correlación entre conductas puede reducir o aumentar el efecto frente al cálculo por independencia.

Conclusiones

- La fórmula demuestra cómo se agrega el prior del evento compuesto “al menos uno”.
- No cambia por sí misma la sensibilidad ni la especificidad intrínsecas de la técnica.
- El artículo no evalúa empíricamente resultados operativos: no demuestra una mejora real del desempeño.
- Se requiere investigación adicional, especialmente cuando los asuntos no son independientes.

Mensaje clave

La agregación cambia el contexto probabilístico; su utilidad práctica debe verificarse empíricamente y no interpretarse como una mejora automática del instrumento.

Fuente: Nelson, R. (2026), Polygraph & Forensic Credibility Assessment, 55(1)

Frecuentista vs bayesiana

Frecuentista (ESS)

- Pregunta: ¿qué tan raros serían estos datos si el examinado fuera veraz?
- Probabilidad como frecuencia a largo plazo; sin probabilidades previas
- Resultado: valor-p, alfa = .05, intervalos de confianza

Bayesiana (ESS-M)

- Pregunta: dados los datos, ¿qué tan probable es que sea veraz o engañoso?
- Probabilidad como grado de certeza: previa × verosimilitud
- Resultado: odds o probabilidad posterior, intervalos de credibilidad

Un mismo resultado, dos lecturas

CONGRESO INTERNACIONAL AIPP
MÉRIDA
MÉXICO 2026

Frecuentista

- Puntaje total de −4 con corte de −3 (alfa = .05, prueba diagnóstica)
- Conclusión: significativo para engaño
- Error esperado controlado por alfa

Bayesiana

- Odds posteriores = odds previas × razón de verosimilitud
- Ilustrativo: previa 50% (1:1) y BF de 10 dan 10:1, es decir 91%
- Si la previa cambia, el resultado cambia

Cortes de ESS-M: ±3 en pruebas diagnósticas con alfa = .05. Los números bayesianos son ilustrativos.

Por qué importan las probabilidades previas

1,000 examinados, prevalencia de engaño 10%, sensibilidad y especificidad de 90% (ilustrativo)

90%

P(resultado positivo | engañoso)

50%

P(engañoso | positivo): 90 de 180

99%

P(veraz | negativo): 810 de 820

Con una previa de 50%, el valor predictivo positivo subiría a 90%.

Probabilidad Bayesiana

- Integra información previa y evidencia nueva
- Teorema de Bayes

$$P(A|B)=\frac{P(A)*P(B|A)}{P(B)}$$

A y B son eventos, y además: P(B) ≠ 0.
P(A|B): es la probabilidad de que ocurra A, dado que ha ocurrido B.
P(B|A): es la probabilidad de que ocurra B, dado que ha ocurrido A.
P(A): es la probabilidad de que ocurra A.
P(B): es la probabilidad de que ocurra B.

Ejemplo de Aplicación

Supongamos que para el día de mañana, tenemos los siguientes datos del meteorólogo:

Probabilidad de Lluvia P(A) = 30%
Probabilidad de Nublado P(B) = 50%
Sabemos además, que la probabilidad de que haya nubes dado que está lloviendo, expresado como P(B|A) = 100%, es decir, siempre hay nubes si está lloviendo.

Ahora, nos levantamos al día siguiente, y resulta ser que está nublado. El teorema de Bayes nos permite calcular una nueva probabilidad de que llueva, dado que está nublado P(A|B)

$$P(A|B) = \frac{P(A) * P(B|A)}{P(B)} = \frac{30\% * 100\%}{50\%} = 60\%$$

Nuestra nueva probabilidad de que llueva; dado que SI hay nubes, es ahora de 60%. Hemos actualizado nuestra probabilidad de lluvia.

T. de Bayes en Términos de una Hipótesis

$$Posterior(H|D) = \frac{Prior(H) * P(D|H)}{Prior(D)}$$

Se lee de la siguiente manera: la probabilidad posterior de que la hipótesis (o conclusión) sea correcta, debido a que observamos el dato (D) es igual al PRODUCTO DE la probabilidad previa de que la hipótesis sea correcta POR la probabilidad de observar el dato (D) si la hipótesis es correcta, DIVIDIDO ENTRE la probabilidad previa de observar el dato (D)

Interpretación simple:

• Probabilidad posterior = evidencia × creencia previa

Cómo reportar resultados estadísticos (ESS-M)

- **Para la toma de registros, se utilizó una técnica científica validada para casos exploratorios de asunto múltiple conocida como MGQT. Para analizar los resultados, se utilizó el sistema de puntuación empírica multinomial (ESS-M)**, un protocolo estandarizado basado en evidencia que utiliza Estadística Bayesiana con una distribución de referencia multinomial, para determinar la probabilidad posterior al examen de que la evaluada respondió o no con veracidad. Los resultados se calcularon utilizando una probabilidad previa de .50 lo que significa que las posibilidades de engaño antes de practicar la prueba eran de 1 a 1. Se calculó un intervalo de credibilidad con un alfa unilateral de .05 para reportar resultados significativos(*).
- En este examen se utilizó una regla de decisión por sub-totales, que asume que las preguntas realizadas con el polígrafo son de carácter independiente.
- La mínima puntuación por pregunta obtenida en esta prueba es de +3, que igualó o excedió el puntaje de corte numérico requerido de +1 para ser clasificado como veraz.
- Estos resultados son significativos(*) con una probabilidad del 95%.
- El factor de Bayes asociado a esta calificación es de 19; por lo que **las posibilidades de veracidad posteriores a la prueba son de 19 a 1**; lo que puede expresarse también como una **probabilidad de veracidad posterior a la prueba de 95%.**
- Estos resultados analíticos apoyan la conclusión categórica de que la evaluada **no presentó reacciones significativas de engaño en esta prueba.**

Cómo reportar resultados estadísticos (ESS-M)

* En el sistema ESS-M, un resultado de prueba se considera estadísticamente significativo cuando el límite inferior del intervalo de credibilidad de las posibilidades posteriores excede el valor de las posibilidades previas, o el de la tasa de corte mínima requerida.

Ideas clave

- Toda prueba tiene error: de procedimiento, aleatorio y sistemático.
- Exactitud y precisión muestran qué tipo de error hay.
- La inferencia estadística debe expresar la incertidumbre, no ocultarla.
- Los valores predictivos dependen de la eficiencia de la prueba y de la probabilidad previa.
- Estadística frecuentista y bayesiana responden preguntas distintas.
- Hay otros conceptos estadísticos importantes que quedaron fuera de nuestra discusión el día de hoy.

Referencias principales

- Nelson, R. (2018). A Brief Discussion of Some Causes of Testing Errors. APA Magazine, 51(6), 45–46.
- Nelson, R. (2015). The Polygraph Test as a Single-Subject Scientific Experiment. APA Magazine, 48(5).
- Nelson, R. (2016). Five Minute Science Lesson: Prior Probabilities and the Algebra to Calculate 'Em. APA Magazine, 49(5).
- Nelson, R. (2020). Conceptual Terms for Understanding Test Accuracy Metrics. APA Magazine, 53(3).
- Nelson, R. (2020). Cómo reportar resultados estadísticos / Reporting Statistical Results. APA Magazine, 53(3).
- Kircher, J. C., & Raskin, D. C. (1987). Polygraph countermeasures: Current status and research. / Statistical precision reference used in the prior Journal Club work.
- Raskin, D. C. (1987). Methodological issues in estimating polygraph accuracy in field applications. Canadian Journal of Behavioural Science, 19(4), 389–404.

Fuente oficial consultada para Nelson (2018): polygraph.org/docs/516_preview_2.pdf