



Poligrafía Práctica: Cómo escribir información de probabilidad en informes de polígrafo evidenciarios (Estándar 1.8.3)

Raymond Nelson

La Norma APA 1.8.3 establece que debe proporcionarse información probabilística junto con los resultados categóricos de las pruebas de los exámenes evidenciarios, definidos en la Sección 1.1 de los Estándares de Práctica de la APA como aquellos exámenes realizados con la intención de que los resultados se utilicen como base de información en un procedimiento legal. El suministro de información estadística se ha convertido en un estándar común en la ciencia, las pruebas científicas y la medicina forense. La razón de este requisito es educar a los consumidores en cuanto a la naturaleza inherentemente probabilística de gran parte de la información científica, y reducir las tendencias humanas a esperar la infalibilidad y a exagerar el tamaño potencial del efecto. La inclusión de información probabilística junto con resultados categóricos u opiniones profesionales, es un estándar general en la ciencia forense siguiendo a Daubert, que sostenía que la buena ciencia se caracteriza por descripciones razonables de tasas de error

conocidas o potenciales. La información de probabilidad se puede describir de varias maneras.

Referencia a estándares publicadas

La forma más sencilla de proporcionar información probabilística en apoyo del resultado de una prueba de polígrafo categórico es consultar los Estándares de Práctica de la APA que definen los requisitos de tamaño del efecto de precisión para las técnicas de polígrafo seleccionadas para diversos propósitos. Los polígrafos evidenciarios son aquellos que se llevan a cabo con el objetivo o la intención de que el resultado de la prueba se utilice como base de información en un procedimiento legal. Las pruebas por pares implican la realización de dos exámenes de polígrafo en dos testigos opuestos por parte de dos examinadores diferentes que son ciegos al resultado de la otra prueba. Los polígrafos de investigación son aquellos que se realizan sin la intención de

utilizar los resultados de la prueba en un contexto legal, como los polígrafos que se realizan en la selección de candidatos, la evaluación de riesgos de seguridad y los programas de monitoreo posterior a una condena.

La sección 1.1.7.3.1 de los Estándares de Práctica de la APA establece que las técnicas de polígrafo utilizadas para los exámenes evidenciarios deben proporcionar una tasa de precisión general de .90 o más, con una tasa de inconclusos de .20 o menos. La sección 1.1.7.3.2 establece que las técnicas de polígrafo utilizadas para las pruebas pareadas deben tener una tasa de precisión general de .86 o superior con una tasa de inconclusos de .20 o inferior (la tasa de precisión más baja está permitida porque la precisión afecta a dos resultados diferentes de los examinadores, ambos ciegos al otro resultado, y pueden combinarse como eventos de probabilidad independientes, con un tamaño de efecto resultante que supera a cualquiera de las pruebas por sí solas cuando los dos ciegos coinciden en que un testigo fue engañoso y el otro fue veraz). Los requisitos de precisión para las técnicas de polígrafo de investigación se describen en la Sección 1.1.7.3.3 de los Estándares de Práctica de la APA. Es necesario que estos proporcionen una tasa de precisión general de .80 o superior con una tasa de inconclusos de .20 o menos. El siguiente es un ejemplo de cómo informar el cumplimiento del Estándar 1.8.3 para polígrafos evidenciarios. El texto puede ser modificado para las pruebas realizadas con fines de investigación y pruebas por pares.

Los datos de prueba se registraron utilizando un formato de prueba y un método de análisis

que cumplen con los requisitos de la Sección 1.1.7.3.1 de los Estándares de Práctica de la APA, incluida una tasa de precisión general del 90% o más.

Cuando la información de precisión de la prueba se expresa en función de las métricas de precisión frecuentistas observadas en estudios publicados, no describe la solidez de la información para ningún resultado de prueba individual. En cambio, la solidez de cualquier resultado de examen individual se estima en función de los Estándares de Práctica, basados en la premisa de la realización del examen utilizando métodos que se sabe que son consistentes con las normas.

Precisión frecuentista y tasas de error para el formato de prueba

Otra forma de proporcionar información de probabilidad para un examen probatorio es describir las métricas de precisión de las pruebas frecuentistas. Esto incluye métricas de precisión simples, como tasas de verdadero positivo, verdadero negativo, falso positivo y falso negativo, y también puede incluir métricas agregadas de valor predictivo positivo y valor predictivo negativo, y métricas relacionadas, como el índice de falsos positivos y el índice de falsos negativos. La descripción del porcentaje total de aciertos y el porcentaje de errores es otro ejemplo. Estas métricas se pueden estudiar para técnicas de preguntas de comparación (CQT) en su conjunto, o para formulaciones únicas de técnicas de preguntas de comparación. A continuación se muestra un ejemplo de cómo informar las métricas de precisión para un



ARTÍCULOS REGULARES

formato de prueba evidenciario de cuatro preguntas. Este texto puede ser modificado para otros formatos de prueba.

Los datos de la prueba se registraron mediante una técnica de polígrafo de 4 preguntas (técnica de Raskin). Cuando se analizó utilizando el método de análisis de datos ESS-M, los tamaños del efecto para esta técnica se han resumido en un apéndice (APA Editorial Staff, 2020) del análisis meta-analítico de 2011 sobre técnicas de polígrafo validadas (APA, 2011). La precisión general de este método se reportó en .944 (IC del 95%: .897 a .987) con una tasa de inconclusos de .031 (.01 a .092). La sensibilidad de la prueba fue de .923 y la especificidad fue de .908. La tasa de error de falsos positivos reportada fue de .046 y la tasa de falsos negativos fue de .062.

Es importante destacar que estas estadísticas no deben informarse a partir de un solo estudio, sino que deben agregarse a partir de un conjunto de datos de investigación disponibles, como se hace comúnmente en las revisiones de la literatura de investigación y la investigación meta-analítica. Otra consideración importante es que la investigación puede llevarse a cabo con grupos de estudio equilibrados, o con grupos de desequilibrio, y las estadísticas frecuentistas agregadas pueden no ser fácilmente aplicables a entornos de campo en los que la tasa de incidencia o la probabilidad previa están desequilibradas. Una tercera consideración importante es que los profesionales individuales generalmente no superan los límites o las capacidades de la ciencia subyacente, a pesar del hecho de que

a veces es socialmente gratificante hacer afirmaciones de tasas de precisión personal que son extremadamente altas. (Esto es similar a la forma en que la mayoría de los inversores, con el tiempo, no superan a los mercados financieros).

Cuando no se conocen las causas exactas de los errores de prueba, los errores de prueba se consideran análogos a los errores de medición aleatorios. Por supuesto, cuando se conoce la causa exacta de un error, se puede corregir o evitar, y así aumenta la precisión de la prueba. Pero es importante recordar siempre que no se espera que las pruebas científicas, de cualquier tipo, sean infalibles. Esto se debe a que el propósito de cualquier prueba científica es cuantificar un fenómeno de interés que no puede ser objeto de medición física (aún sujeto a error de medición aleatorio) o perfección determinista (que sería inmune tanto al error de medición aleatorio como al comportamiento humano). Las pruebas científicas hacen uso de la relación estadística entre los datos proxy disponibles y los fenómenos de interés. Debido a que los datos de la prueba y los resultados de la prueba no son mediciones físicas directas, solo se espera que cuantifiquen el margen de incertidumbre que rodea una conclusión o el resultado de la prueba.

Las métricas de precisión frecuentistas de las pruebas caracterizan todos los errores de prueba como eventos aleatorios, suponiendo que las pruebas se realicen correctamente y que los errores no sean atribuibles a un error de procedimiento. La probabilidad de un error para cualquier caso individual se estima en función de la tasa de error aleatorio. De



manera complementaria, la probabilidad de un resultado correcto de la prueba para cualquier caso individual también se estima en función de los tamaños de efecto conocidos para una técnica de polígrafo. Una limitación de este método es que algunas métricas de precisión frecuentistas de interés pueden no ser resistentes a la influencia de probabilidades previas desequilibradas. Una limitación más práctica de este método es que no proporciona información intuitivamente útil sobre las fortalezas o limitaciones de los datos para un caso individual.

Clasificadores estadísticos

Un clasificador estadístico es un valor estadístico reproducible para una prueba de polígrafo individual y se utiliza para hacer una clasificación categórica del resultado de la prueba en función de la fuerza del valor estadístico. Son posibles varios tipos diferentes de clasificadores estadísticos. Un clasificador simple común es el valor p , que se puede calcular utilizando distribuciones de referencia empíricas o teóricas. También se utilizan comúnmente clasificadores estadísticos más complejos, incluidas las funciones logísticas, los clasificadores Naïve-Bayes, las máquinas de soporte vectorial, los algoritmos de k -vecino más cercano, y otros. A continuación se muestra un ejemplo de cómo informar un clasificador estadístico (valor p) para el resultado de una prueba de polígrafo evidenciario. Este texto se puede modificar para otros formatos de prueba y otros resultados de prueba.

Utilizando el ESS, un protocolo estandarizado, normativo y basado en la evidencia para el

análisis de datos de pruebas, la puntuación total general de -5 es igual o superior a la puntuación de corte requerida de -4 para clasificaciones engañosas. El nivel de significación estadística se calcula en $p = .032$, que es igual o menor que el límite alfa requerido ($\alpha = .05$), e indica que solo se puede esperar que una pequeña proporción de personas verídicas (3.2%) produzca una puntuación igual o inferior en la prueba. Estos resultados apoyan la conclusión de que existe engaño indicado (DI) en las respuestas fisiológicas a las preguntas de estímulo de la prueba durante este examen.

Los clasificadores estadísticos son abstracciones matemáticas que proporcionan una base objetiva para la toma de decisiones a nivel de caso individual. Sin embargo, no describen por sí solas la probabilidad práctica de clasificación correcta o incorrecta a nivel de un caso individual o de un grupo de casos. No intentan proporcionar información que pueda ser fácil o intuitivamente descriptiva de la fuerza práctica de la información a nivel de caso individual. Sin embargo, los clasificadores estadísticos pueden combinarse con otros métodos, especialmente el Teorema de Bayes, para proporcionar información posterior o de resultados que tenga un mayor significado práctico e intuitivo.

Probabilidades posteriores o de resultado

Las probabilidades condicionales posteriores, también conocidas como probabilidades posteriores o distribuciones posteriores, y los niveles de confianza en los resultados son conceptos relacionados en probabilidad y estadística. Son probabilidades asignadas a



eventos o resultados específicos después de observar y analizar los datos de prueba. Las probabilidades condicionales posteriores proporcionan una base estadística para la clasificación, al tiempo que proporcionan información que describe la fuerza práctica de la información para cada caso. En el contexto del análisis bayesiano, las probabilidades posteriores se calculan utilizando el teorema de Bayes, que relaciona la probabilidad condicional de un evento A dada la evidencia B con la probabilidad previa de A y la probabilidad de B dado A. Es importante destacar que tanto la construcción como la realización de la prueba del polígrafo (incluida la entrevista, la secuencia de la pregunta de la prueba, la instrumentación, y sensores) y el análisis de los datos de la prueba se basan en la teoría básica de la prueba del polígrafo: que los mayores cambios en la actividad fisiológica se cargan en diferentes tipos de estímulos de prueba en función del engaño y la verdad en respuesta a estímulos objetivo relevantes.

Las probabilidades *a posteriori* son esenciales en la estadística bayesiana, ya que nos permiten utilizar nuevos datos o evidencias para realizar modificaciones objetivas y matemáticas a las probabilidades asociadas a diferentes resultados posibles. Las probabilidades bayesianas *a posteriori* a menudo se expresan en forma de probabilidades y conducen fácilmente al cálculo del Factor de Bayes, que describe, en forma de probabilidades, el cambio o aumento en la fuerza probabilística de la información en apoyo de una clasificación o conclusión particular. A continuación se muestra un ejemplo de cómo reportar la información posterior en apoyo de una

conclusión categórica para un examen de polígrafo evidenciario. El texto se puede modificar para otros resultados de la prueba.

Utilizando el ESS-M, un protocolo estandarizado, normativo y basado en la evidencia para el análisis de datos de pruebas, la puntuación total general de 19 igualó o superó la puntuación de corte numérica requerida (3). La posibilidad posterior de veracidad fue de 87 a 1, para lo cual la probabilidad posterior fue de 0,98. El límite inferior del intervalo de credibilidad bayesiano 1-alfa fue de 15 a 1, lo que superó las probabilidades previas (1 a 1). Esto indica una probabilidad superior al 95% de que las probabilidades posteriores de verdad superen las probabilidades previas. La información posterior para este examen se incrementó en un factor de Bayes de 87 veces. Estos resultados analíticos apoyan la conclusión de que no hubo reacciones significativas indicativas de engaño en la distribución de los cambios registrados en la actividad fisiológica en respuesta a los estímulos de prueba relevantes durante este examen.

Los niveles de confianza en los resultados son otra forma de expresar el grado de certeza o confianza asociado con una predicción o decisión particular tomada por un modelo probabilístico, como una red bayesiana u otro clasificador estadístico. Estos niveles de confianza a menudo se expresan como proporciones decimales entre 0 y 1, o como un porcentaje, con valores más altos que indican una mayor confianza en el resultado o la predicción. Los niveles de confianza en los resultados suelen derivarse de las probabilidades condicionales posteriores y



pueden ser útiles para la toma de decisiones y la evaluación de riesgos.

En la práctica, tanto las probabilidades condicionales posteriores como los niveles de confianza en los resultados son valiosos por su capacidad para proporcionar descripciones prácticas e intuitivamente útiles de la fuerza de la información en apoyo de las

predicciones, y clasificaciones realizadas por los modelos probabilísticos. Se pueden utilizar para establecer cortes de decisión (por ejemplo, clasificar puntos de datos en función de un nivel mínimo de confianza) y para proporcionar información adicional a los usuarios sobre el nivel de certeza que se puede asignar razonablemente a una clasificación o conclusión.



Referencias

- Asociación Americana de Polígrafo (2011). Estudio metaanalítico de la precisión de los criterios de las técnicas de polígrafo validadas. *Polígrafo*, 40(4), 196-305.
- Asociación Americana de Polígrafo. (25 de agosto de 2023). *Estándares de práctica de la APA*. Obtenido de <https://www.polygraph.org> el 7 de noviembre de 2023.
- Redacción APA. (2020). Anexo a la encuesta metaanalítica de 2011: la prueba de cuatro preguntas de Utah ("técnica de Raskin") / ESS. *Polígrafo y Evaluación de Credibilidad Forense*, 49(2), 73-78.
- Daubert v. Merrel Dow Pharmaceuticals, 509 U.S. 579 (1993).
- Honts C. R., Thurber, S. y Handler, M. (2021) Un metaanálisis exhaustivo de la prueba de polígrafo de preguntas de comparación. *Psicología Cognitiva Aplicada* 35(2), 411–427.
- Consejo Superior de Investigaciones Científicas. (2003). *El polígrafo y la detección de mentiras*. Washington, DC: The National Academies Press.
- Consejo Superior de Investigaciones Científicas. (2011). *Manual de Referencia sobre Evidencia Científica*. Washington, DC: The National Academies Press.
- Nelson, R. (2016). Teoría científica (analítica) de las pruebas de polígrafo. *Revista APA* 2016, 49(5), 69-82.



ARTÍCULOS REGULARES

- Nelson, R. (2017). Distribuciones de referencia multinomiales para el sistema empírico de puntuación. *Polígrafo y Evaluación de Credibilidad Forense*, 46(2), 81-115.
- Nelson, R. (2018). Guía sobre cómo utilizar las tablas de referencia ESS-Multinomial en cuatro pasos. *Revista APA*, 51(2), 78-89.
- Nelson, R. (2018). Distribuciones de referencia multinomiales para puntuaciones de tres posiciones de los exámenes de polígrafo con preguntas de comparación. *Polígrafo y Evaluación de Credibilidad Forense*, 47(2), 158-175.
- Nelson, R. (2018). Lección científica de cinco minutos: Teorema de Bayes y análisis bayesiano. *Revista APA 2018*, 51 (5), 65-78.
- Nelson, R. (2018). Polígrafo práctico: un tutorial (con gráficos) sobre resultados posteriores e intervalos creíbles utilizando el clasificador bayesiano ESS-M. *Revista APA*, 51(4), 66-87.
- Nelson, R. (2019). Explicación del resumen narrativo de la ESS-M. *Revista APA 2019*, 52(6), 30-48.
- Nelson, R. (2022). Distribuciones de referencia multinomiales para puntuaciones Likert de siete posiciones de las pruebas de polígrafo de preguntas de comparación. *Polígrafo y Evaluación de Credibilidad Forense*, 51(1), 41-67.
- Nelson, R. y Handler, M. (2012). Utilización de datos normativos de referencia con exámenes diagnósticos y el Sistema de Puntuación Empírica. *Revista APA*, 45(3), 61-69.
- Nelson, R. y Handler, M. (2015). Distribuciones estadísticas de referencia para polígrafos de preguntas de comparación. *Polígrafo*, 44(1), 91-114.
- Nelson, R. y Rider, J. (2018). Polígrafo práctico: ESS-M simplificado. *Revista APA*, 51(6), 55-62.

