

El Análisis de Supervivencia como Técnica para la Evaluación de la Validez Predictiva en la Psicología Jurídica

Julio Isaac Vega-Cauich

Foco Rojo: Centro de Psicología Aplicada, Mérida, México

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:
Recibido el 04 de octubre de 2017
Aceptado el 16 de mayo de 2018
Online el 27 de julio de 2018

Palabras clave:
Análisis de supervivencia
Validez predictiva
Psicología jurídica

Keywords:
Survival analysis
Predictive validity
Legal psychology

R E S U M E N

Poder determinar que un instrumento predice un comportamiento como la violencia o la reincidencia es uno de los principales objetivos de la psicología jurídica. La validez predictiva, como se conoce a ésta, permite evaluar esta característica. Este tipo de validez ha sido evaluada a través de diversas técnicas, como las tablas de contingencia, sensibilidad, especificidad y las curvas ROC. Sin embargo, estas técnicas no abordan un problema estadístico importante: la censura de los datos que ocurre en los estudios prospectivos, ya que en muchas ocasiones los sujetos no reincidirán durante parte del seguimiento realizado. El análisis de supervivencia es una técnica que permite tomar en consideración esta información y modelar el tiempo hasta que sucede un determinado evento. A través de diferentes aproximaciones, como sus técnicas descriptivas o la regresión de riesgos proporcionales de Cox, esta técnica nos permite evaluar el poder predictivo de las pruebas, así como de las intervenciones.

Survival analysis as a technique for the assessment of predictive validity in legal psychology

A B S T R A C T

Being able to determine that an instrument predicts behavior such as violence or recidivism is one of the main objectives of legal psychology. Predictive validity, as it is known, allows evaluating this characteristic. This type of validity has been evaluated through various techniques, such as contingency tables, sensitivity, specificity, and ROC curves. However, these techniques do not address a major statistical problem: data censorship that occurs in prospective studies, since in many cases the subjects will not relapse during the follow-up performed. Survival analysis is a technique that allows taking this information into consideration and modeling the time until a certain event happens. Through different approaches, such as its descriptive techniques or Cox proportional hazards models, this technique allows us to evaluate the predictive power of tests, as well as of interventions.

El concepto de validez es uno de los elementos más importantes en los instrumentos de medición de cualquier tipo, especialmente en psicología. Un instrumento que mide lo que se supone que debe medir es considerado como un instrumento válido. Esta definición clásica lleva varios años en los principales libros de metodología para ciencias sociales (p. ej., [Hernández Sampieri, Fernández Collado y Baptista Lucio, 2013](#)) y se remonta a un libro escrito por [Kelley \(1927\)](#), en uno de los primeros textos que intentaron sistematizar la evaluación educativa. Junto con la fiabilidad y la objetividad, conforma una de las características que todos los instrumentos de medición psicológica aspiran alcanzar ([Hernández Sampieri et al., 2013](#)), especialmente aquellos diseñados para la toma de decisiones.

El presente artículo retoma el análisis de supervivencia como una metodología que debería considerarse, en conjunto con las demás técnicas existentes, para evaluar la validez predictiva de una prueba o tratamiento. Se presentan los principales tipos de validez existentes, las formas en que se evalúan y las ventajas y desventajas de la técnica mencionada para evaluar la validez, haciendo énfasis en ejemplos de aplicación en el área de la psicología jurídica.

Los Tipos de Validez y su Importancia

Previo a presentar al análisis de supervivencia como una técnica para evaluar la validez predictiva, se considera necesario hacer un

breve recordatorio de los principales tipos de validez existentes ilustrándolos con algunos ejemplos de aplicación. Es bastante común que los libros de texto de metodología de la investigación en psicología consideren que la validez puede clasificarse en al menos tres tipos: 1) validez de contenido, 2) validez de constructo y 3) validez de criterio.

La validez de contenido se refiere al “grado en que un instrumento refleja un dominio específico de contenido de lo que se mide” (Hernández Sampieri et al., 2013, p. 201), es decir, qué tanto un instrumento recaba todos los posibles signos, síntomas o indicadores asociados al fenómeno que se mide. En la práctica, los profesionales esperamos definir de forma exhaustiva todos los indicadores que conforman un constructo, con la finalidad de tener una mejor validez de contenido. Por ejemplo, durante el desarrollo del *Danger Assessment* (DA), un instrumento que es utilizado para valorar el riesgo de homicidio en mujeres maltratadas, Campbell (1986) utilizó entrevistas con mujeres maltratadas, trabajadores de refugios, oficiales de policía y expertos en violencia de pareja con la finalidad de abarcar todas las posibles variables asociadas al riesgo de homicidio en mujeres maltratadas. Finalmente, los expertos fueron utilizados para validar el contenido del instrumento a partir de esta metodología. La técnica anterior es sumamente utilizada para asegurar la validez de contenido de un instrumento, ya que asegura tener múltiples fuentes de información, a la vez que ésta es valorada por jueces expertos en la materia.

Por otra parte, la validez de constructo hace referencia a que existen pruebas que “deben explicar el modelo teórico empírico que subyace a la variable de interés” (Hernández Sampieri et al., 2013, p. 203), es decir, el constructo debe ser consistente con la teoría existente. Un claro ejemplo de esto es la validación del PCL-R, un instrumento ampliamente utilizado para evaluar la presencia de psicopatía en adultos con un historial violento o delictivo (Andrés-Pueyo y Echeburúa, 2010). Durante los estudios realizados a la estructura factorial del PCL-R se encontró la existencia de dos factores, uno relacionado con características de personalidad y otro con conductas de desviación social (Harpur et al., 1988). A través de diferentes técnicas estadísticas se logró comprobar la plena existencia de estos dos factores. El primer factor mayormente relacionado con otras medidas de personalidad y narcisismo y el segundo con variables más conductuales, como un diagnóstico de trastorno antisocial de la personalidad y socialización. A su vez, estos dos factores estuvieron fuertemente correlacionados entre sí, lo que sugería una estructura de una sola dimensión.

Esto permitió concluir a los autores la teoría que el PCL mide un constructo unidimensional, el cual es matizado por dos factores que representan la parte conductual y afectiva de la psicopatía (Harpur, Hare y Hakstian, 1989). Los estudios de validación llevados a cabo por Robert D. Hare y su equipo utilizaron técnicas sencillas y elegantes para identificar la relación entre el instrumento con otras variables externas o con los mismos factores –como la correlación canónica o la correlación biserial– si bien teorías más complejas requerirán de técnicas más sofisticadas para hacerlo. Por ejemplo, recientemente el uso de modelos de ecuaciones estructurales ha demostrado que son una buena herramienta para evaluar este tipo de validez cuando se ponen a prueba teorías más complejas (Schumacker y Lomax, 2015).

Finalmente, la validez de criterio “se establece al validar un instrumento de medición al compararlo con algún criterio externo que pretende medir lo mismo” (Hernández Sampieri et al., 2013, p. 202). A su vez, ésta puede dividirse en validez concurrente y predictiva. En la primera los resultados del instrumento se relacionan con el criterio en el mismo momento. Por ejemplo, al validar el instrumento *Psychopathic Personality Inventory* (PPI), un instrumento autoadministrado que mide características de personalidad psicopática, Maltzer, Lilienfeld y Newman (2010) relacionaron las puntuaciones de éste instrumento con las obtenidas por otros instrumentos de evaluación de psicopatía como el PCL-R y el PCL-SV en tres muestras de reclusos. Aplicando en un mismo periodo de tiempo, estos tres instrumentos, identificaron que el PPI se correlaciona significativamente con las dos versiones del PCL, pero no tan fuertemente como en otras

poblaciones no penitenciarias. Esta metodología permitió a los autores determinar la validez concurrente del PPI y concluir que éste instrumento es útil para evaluar las características de personalidad de la psicopatía, especialmente en poblaciones de no delincuentes.

La validez predictiva por su parte busca predecir una situación. Por ejemplo, recientemente Grieger y Hosser (2013) utilizaron la técnica para determinar si los ocho factores de riesgo de la conducta criminal evaluados por el *Level of Service Inventory-Revised* (LSI-R) son útiles para predecir la reincidencia en delincuentes juveniles. Para ello, los investigadores evaluaron diversas variables en una muestra de 589 agresores juveniles de Alemania para obtener un diagnóstico de riesgo brindado por el LSI-R. Posteriormente, se realizó un seguimiento de seis años y medio a los jóvenes para identificar aquellos que tuvieran reincidencia. Este estudio es particularmente importante para evaluar la validez predictiva del LSI-R, pues demuestra empíricamente que los sujetos reincidentes con un diagnóstico de reincidencia bajo fueron proporcionalmente menos que aquellos con un diagnóstico de reincidencia medio y éstos a su vez que aquellos con un diagnóstico de reincidencia alto.

Formas de Evaluación de la Validez Predictiva

La validez predictiva es sin duda uno de los tipos de validez más deseados en cualquier instrumento de medición, especialmente en el campo de la psicología jurídica, razón por la cual en la actualidad existen diversas formas de evaluarla. A continuación, se describen las principales técnicas para hacerlo.

Una de las formas más utilizadas de evaluar la validez predictiva de una prueba es a través del análisis de tablas de contingencia, donde se contrasta lo predicho por la prueba con el criterio a predecir. Por ejemplo, si un instrumento dice que un agresor reincidirá y éste luego de un tiempo reincide, hablamos de lo que se conoce como un verdadero positivo, pero si no reincide, hablamos de un falso positivo. De igual forma, si el mismo instrumento dice que un interno no reincidirá y no lo hace, hablamos de un verdadero negativo, pero si reincide, se trata de un falso negativo. Estos análisis se realizan a partir de estudios prospectivos y longitudinales, en donde a los sujetos se les aplica el instrumento en cuestión y se realiza un seguimiento para identificar en cuáles de éstos se presentó el evento que se quería predecir. Una tabla de contingencia resume la frecuencia de casos en cada uno de estos posibles resultados.

Si bien las tablas de contingencia son una técnica general de clasificación binaria, existen varios estadísticos que pueden utilizarse para evaluar el poder predictivo (y por lo tanto la validez predictiva) de una prueba a través de los resultados obtenidos con estas tablas. Entre ellas destacan la tasa de falsos positivos y falsos negativos, el poder predictivo positivo y negativo, el poder diagnóstico general, la prevalencia y la sensibilidad y especificidad (Baldessarini, Finklestein y Arana, 1983). Estos dos últimos conceptos son ampliamente utilizados y se consideran como una medición de la probabilidad condicional de que un instrumento brinde un diagnóstico positivo, cuando el sujeto realmente tiene el criterio a predecir (sensibilidad), o de que el instrumento dé un diagnóstico negativo, cuando el sujeto no tiene éste criterio –especificidad– (Agresti, 2007). Un instrumento con una buena validez predictiva será aquel que tenga una menor tasa de falsos negativos y positivos o bien una alta sensibilidad y especificidad.

Si bien los estadísticos abordados anteriormente están dirigidos a profesionales del área de la salud, lo cierto es que han sido ampliamente utilizados en otras áreas de la ciencia (Farrington y Loeber, 2000). No obstante, Rice y Harris (1995) sugieren que estos estadísticos tienen ciertas deficiencias en su uso, puesto que son altamente dependientes de la prevalencia con la que se presenta un evento. Esto es especialmente problemático en criminología, debido a que la violencia es un evento atípico y su prevalencia es muy pequeña cuando se compara a nivel poblacional. Es por ello que Rice y Harris (1995)

proponen para evaluar la validez predictiva la curva ROC (acrónimo de *Receiver Operating Characteristic* - característica operativa del receptor), ya que es independiente de la prevalencia o del punto del corte para clasificar un evento.

Una forma usual de interpretar la curva ROC es analizando la proporción del área de la curva que se obtiene a partir de ésta. A mayor área, mejor predicción, ya que indica más aciertos y menos predicciones falsas (Rice y Harris, 1995). Sin embargo, una de sus principales limitaciones es que se requiere cubrir el supuesto de normalidad de las puntuaciones de una prueba. Si bien en algunos casos (como en investigaciones psiquiátricas) es factible que este supuesto se cumpla, Rice y Harris (1995) mencionan que en criminología la violencia rara vez se distribuye de forma normal. Es por ello por lo que se requiere de otras metodologías que consideren la no normalidad de la violencia o de las variables a predecir.

Sin embargo, es importante hacer notar que, además de los problemas mencionados con anterioridad, existe otro problema con mayores implicaciones metodológicas: la censura de los datos, el cual está relacionado con diseños que requieren el seguimiento de los sujetos de investigación. En muchas ocasiones el tiempo de seguimiento es inferior al tiempo que se necesita para que un evento suceda, por lo que muchas veces los eventos de interés no suceden. Esto es especialmente problemático debido a que una proporción de los participantes no presentan el evento durante el tiempo de estudio. Sin embargo, esto no quiere decir que los participantes nunca lo experimentarán en un periodo superior al seguimiento. Esto se conoce en estadística como datos con censura, o datos censurados (en específico, como datos con censura derecha, donde el evento no ha ocurrido, en oposición a la censura izquierda, donde el evento ocurrió antes del comienzo del periodo). El presente documento ahonda en el análisis de supervivencia como una técnica que puede solventar estos problemas al brindar un marco metodológico ideal para evaluar la validez predictiva en el campo de la psicología jurídica y áreas afines. Para ello, se hace una revisión de la técnica, de sus requisitos y supuestos, así como de sus formas de aplicación e interpretación, de forma práctica, a través de un ejemplo con datos reales.

El Análisis de Supervivencia: Principales Características

Definición de la Técnica

Si la validez predictiva es tan valiosa para identificar si un instrumento es útil en predecir que cierto evento suceda, es necesario contar con herramientas necesarias para evaluarlo. Ya se han mencionado algunas técnicas estadísticas que son especialmente útiles para evaluarla, pero que presentan deficiencias en su utilización (como la no normalidad de las variables o la presencia de datos censurados durante el seguimiento). El análisis de supervivencia es útil para solventar estas carencias (Kaplan y Meier, 1958).

El análisis de supervivencia hace uso de los datos de supervivencia, los cuales son “mediciones del tiempo transcurrido entre el inicio del seguimiento de un sujeto hasta la ocurrencia de un suceso o la finalización del seguimiento” (Daniel y Cross, 2013, p. 751). Supervivencia en este contexto significa únicamente que no se ha producido el suceso. Es por ello, que muchas veces ocurren lo que se conoce como datos censurados, es decir “medidas para las cuales tenemos cierta información sobre el tiempo de supervivencia, pero el tiempo de supervivencia exacto se desconoce” (Daniel y Cross, 2013, p. 752). Bien sea porque el tiempo de seguimiento es insuficiente para que se presente el evento en todos los participantes –lo que se conoce como censura administrativa– o bien porque se pierda el contacto con algunos participantes (sea por muerte, mudanzas o rechazo a continuar en el estudio). El análisis de supervivencia toma en consideración la censura de los datos, situación que la mayoría de las técnicas estadísticas no tienen en cuenta (Clark, Bradburn, Love y Altman, 2003a; Palmer Pol y Montaña Moreno, 2002; Singer y Willett, 2003). Por

ejemplo, si se tuviera el tiempo hasta que sucede el evento de todos y cada uno de los participantes, bien podría recurrirse a una regresión lineal o una regresión logística; sin embargo, esta información no está disponible en la mayoría de las ocasiones en las que se realiza un estudio prospectivo (Landau y Everitt, 2004).

Adicionalmente esta técnica toma en consideración la no normalidad del tiempo que transcurre hasta que sucede el evento, siendo el tiempo una variable que por su naturaleza no se distribuye de forma normal –solo avanza en una dirección, nunca retrocede (Gómez y Cobo, 2004). Esta consideración se realiza utilizando técnicas no paramétricas o bien técnicas paramétricas que consideran distribuciones sesgadas ideales para modelar el tiempo hasta que el evento ocurre –p. ej., la distribución exponencial o la de Weibull (Kleinbaum y Klein, 2006). Finalmente, también permite incorporar variables que sean dependientes del tiempo (como la edad, por ejemplo) y que al ser utilizadas en otras técnicas se violaría el supuesto de independencia que muchas pruebas estadísticas requieren.

El análisis de supervivencia también ofrece la posibilidad de comparar el tiempo de supervivencia de dos o más poblaciones diferentes (Clark et al., 2003a). Esto es especialmente útil cuando quiere contrastarse, por ejemplo, una población que recibió un diagnóstico de riesgo de reincidencia bajo, con otra de riesgo de reincidencia alto. De forma intuitiva podríamos pensar que aquellos que presentaron riesgo alto reincidirán en un menor tiempo que los participantes de riesgo bajo. Por medio de las técnicas brindadas por el análisis de supervivencia podría probarse esta hipótesis al contrastar los tiempos de supervivencia de ambos grupos, lo que brinda una herramienta de gran valor para analizar la validez predictiva dentro del campo de la psicología jurídica, en la que, por ejemplo, muchas veces quiere predecirse la probabilidad de reincidencia de agresores de diferentes delitos (Schmidt y Witte, 2012).

Estas ventajas que presenta el análisis de supervivencia son posibles gracias a que ésta analiza los datos mediante la llamada función de supervivencia. Esta función puede estimarse en cualquier momento y brinda la probabilidad de que un individuo no presente el evento en un determinado momento t del tiempo (Clark et al., 2003a; Gómez y Cobo, 2004). La estimación de estas probabilidades es lo que hace de esta técnica una herramienta inestimable para la valoración de la validez predictiva, ya que permite determinar la probabilidad de que un grupo de individuos con cierto diagnóstico o tratamiento presente el evento en un determinado periodo tiempo, en comparación con otros grupos de individuos con diferente diagnóstico o intervención, todo ello considerando la existencia de datos censurados y la asimetría de la variable tiempo.

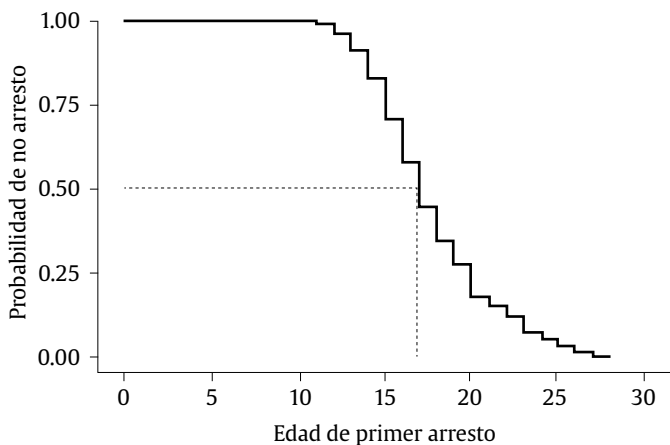
Principales Técnicas en el Análisis de Supervivencia

Las técnicas disponibles para el análisis de supervivencia pueden agruparse en dos grandes apartados (Clark et al., 2003a; Hosmer, Lemeshow y May, 2008; Singer y Willett, 2003): las técnicas descriptivas e inferenciales univariadas (como la técnica de Kaplan-Meier y las comparaciones de LogRank) y las técnicas inferenciales multivariadas de riesgos proporcionales que ajustan diversos modelos estadísticos. A su vez, estas últimas pueden dividirse en métodos semiparamétricos (que consideran algunos supuestos en los datos) y paramétricos (que ajustan una distribución de probabilidad en los datos). Las últimas dos técnicas permiten tomar en consideración covariables, que hacen posible ver la influencia de otros factores en el tiempo hasta que sucede un evento, tanto independientes del tiempo (p. ej., el consumo de drogas, tipos de delito, etc.) como dependientes del mismo (p. ej., edad, tiempo de sentencia, etc.).

A continuación se resumen brevemente cada uno ilustrándolo a través de un ejemplo con datos reales obtenidos de la base de datos de la Encuesta de Cohesión Social para la Prevención de la Violencia y la Delincuencia (ECOPRED) 2014 realizada en México por el Instituto Nacional de Estadística y Geografía (2014). Esta encuesta, realizada a

una muestra representativa de 40,366 jóvenes de entre 12 a 29 años, explora diferentes factores de riesgo asociados a la violencia y a la delincuencia (estos factores incluyen el contexto individual de los jóvenes, el desarrollo y relaciones dentro de sus familias, la influencia e interacción con amistades y compañeros y las relaciones entre los miembros de una comunidad, así como el contexto social en general). Cabe recalcar que la ECOPRED cuenta con representatividad a nivel nacional, al haber encuestado a 97,754 viviendas de 47 ciudades de México. Para este trabajo se toma en consideración como variable dependiente la edad del primer arresto o detención (en el sentido amplio de haber experimentado o no conflictos con la policía y no necesariamente un antecedente judicial), utilizando una submuestra de 1,387 jóvenes que mencionaron haber tenido un arresto o detención. El propósito de este ejemplo es explorar la validez predictiva de las variables sexo, violencia familiar, consumo de sustancias, trabajo y estudio para predecir la edad del primer arresto o detención en los jóvenes participantes.

A) Función de supervivencia



B) Función de riesgo

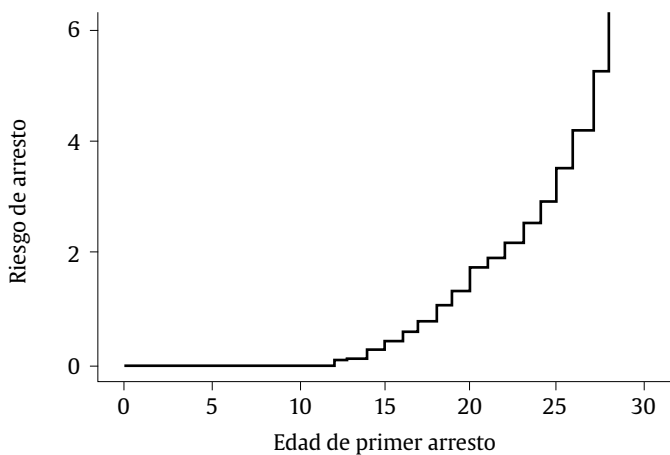


Figura 1. Curva A) de supervivencia y B) de riesgo del tiempo que transcurre hasta que se experimenta un arresto.

Técnicas descriptivas no paramétricas. Estos modelos permiten estimar la función de supervivencia sin tomar en cuenta la distribución de los datos, únicamente considerando la información prevista por la muestra. Una de las técnicas más conocidas y utilizadas para ello es el estadístico de Kaplan-Meier, también conocido como estimador producto-límite, que fue introducido por Edward Kaplan y

Paul Meier en 1958. Para ello se toma en consideración el número de eventos que suceden en un determinado tiempo t_i y el número de sujetos en riesgo n_i antes del tiempo t_i , tomando en consideración la existencia de censura en los datos. Dado que se asume que los eventos ocurren de forma independientes el uno del otro, las probabilidades de supervivencia de un tiempo t_i hasta otro tiempo t_{i+1} son multiplicados siguiendo el principio multiplicativo de la probabilidad y de esta forma se obtiene la función de supervivencia (Clark et al., 2003a), que no es más que la probabilidad de que ocurra un evento en un determinado tiempo. La función estimada a partir del método de Kaplan-Meier puede ser representada gráficamente para obtener un resultado similar a la Figura 1A, en donde el eje X representa el tiempo desde que inicia el estudio hasta que finaliza el seguimiento y en el eje Y se representa la probabilidad de que un individuo no presente el evento (en este caso, ser arrestado).

A la Figura 1A se le conoce como gráfico de supervivencia o curva de supervivencia y su interpretación es bastante sencilla. En él se puede observar como al inicio (en la niñez) la probabilidad de no ser arrestado o tener conflictos con la policía es alta, pero conforme pasa el tiempo la probabilidad disminuye y donde se puede apreciar que a los 17 años aproximadamente el 50% de los jóvenes tuvo alguna vez alguna detención o arresto.

Así mismo, también se puede trabajar con la función de riesgo [hazard function], la cual representa la probabilidad condicional de que un participante experimente el evento de interés en un determinado periodo de tiempo t dado que el evento no se había presentado en los periodos pasados (Gómez y Cobo, 2004). Esta función puede apreciarse en la Figura 1B. Nótese que esta última, a diferencia de la función de supervivencia, se enfoca en la ocurrencia del evento y no en la ausencia de este (Clark et al., 2003a). En el ejemplo se puede apreciar que a medida que aumenta la edad aumenta el riesgo de experimentar alguna detención o arresto.

Una de las ventajas que tiene el método de Kaplan-Meier es que permite la estimación de las funciones de supervivencia de diferentes grupos (Hosmer et al., 2008). De esta forma, se pueden utilizar medidas estadísticas para determinar si las funciones de supervivencia de un grupo son iguales o diferentes a las de otro. Por ejemplo, podrían ser utilizadas para validar si el consumo de drogas permite diferenciar adecuadamente aquellos participantes con un menor tiempo de supervivencia que el de otros (ver Figura 2), lo que representa una herramienta valiosa para determinar, por ejemplo, la validez predictiva de un instrumento de gestión del riesgo.

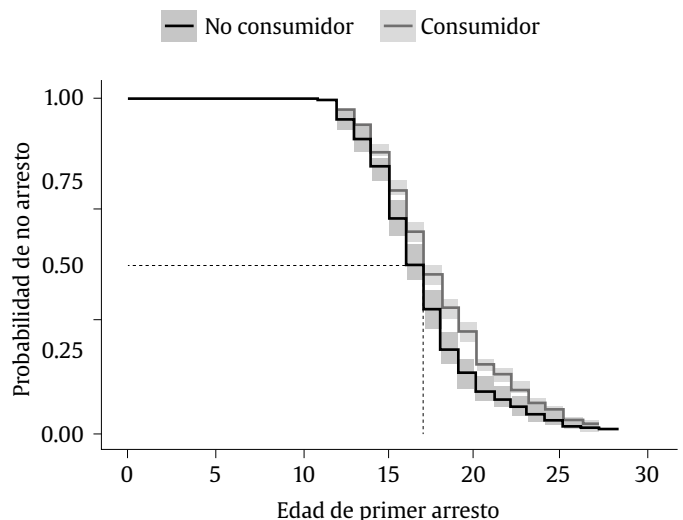


Figura 2. Comparación de la probabilidad de no ser arrestado (supervivencia) según el consumo de drogas de los participantes.

Lo anterior es posible gracias a la adaptación de estadísticos no paramétricos al análisis de supervivencia. Por ejemplo, el estadístico LogRank es una adaptación del estadístico de Mantel-Haenszel, que toma en cuenta la existencia de datos con censura. De igual forma, el estadístico de Breslow es una versión generalizada de la suma de rangos de Wilcoxon, que permite la utilización de datos censurados (Daniel y Cross, 2013; Hollander, Wolfe y Chicken, 2014; Hosmer et al., 2008). Estos estadísticos permiten probar la hipótesis nula de que la función de supervivencia de los grupos es igual frente a la hipótesis alternativa de que alguno de ellos difiere.

En el ejemplo anterior, si bien la mediana de supervivencia es prácticamente la misma para el grupo de consumidores y no consumidores –17 años–, las pruebas de hipótesis a través del estadístico LogRank ($\chi^2 = 16.55$, $gl = 1$, $p < .001$) y de Breslow ($\chi^2 = 17.69$, $gl = 1$, $p < .001$) señalan que sí existen diferencias estadísticamente significativas en el tiempo que transcurre hasta que se experimenta el primer arresto en consumidores y no consumidores de drogas. Los descriptivos señalan que los consumidores tienen un promedio de supervivencia de 16.95 años, mientras que para los no consumidores la media es 17.88 años, es decir, prácticamente una diferencia de un año, siendo los no consumidores los que tardan más tiempo en tener conflictos con la policía.

Finalmente, existen otras formas de estimar la función de supervivencia de una muestra. Si bien el método de Kaplan-Meier es con mucho el más utilizado, también existe el estadístico de Nelson-Aalen, que permite estimar esta función a partir de la función de riesgo de la muestra (Hosmer et al., 2008). A nivel práctico, ambos presentan una eficiencia relativa similar (Colosimo, Ferreira, Oliveira y Sousa, 2002; Hosmer et al., 2008). A nivel estadístico, la propuesta de Nelson y Aalen es el precedente de uno de los métodos de análisis de supervivencia más utilizados que se aborda a continuación.

Técnicas semiparamétricas que incluyen covariables. Las técnicas anteriormente explicadas si bien permiten identificar la función de supervivencia de uno o varios grupos, e incluso compararlos, no nos permiten conocer la magnitud de las diferencias de éstos, o considerar otras covariables para conocer su efecto en el tiempo hasta que sucede un evento. Una técnica que permite considerar esto es utilizar modelos de regresión que tomen en cuenta la existencia de datos censurados. Uno de los modelos más extendidos en su uso es el modelo de regresión de Cox de riesgos proporcionales (Bradburn, Clark, Love y Altman, 2003; Cox, 1972), una técnica semiparamétrica que permite estimar la función de supervivencia mediante un modelo de regresión cuya interpretación es similar a la que se hace en una regresión logística.

El modelo de riesgos proporcionales de Cox hace uso de la estimación no paramétrica propuesta por la técnica de Nelson-Aalen y permite la inclusión de covariables (Hosmer et al., 2008). Recibe el adjetivo de semiparamétrico debido a lo anterior, además de que requiere el cumplimiento de un supuesto importante: riesgos proporcionales. Este supuesto significa que la función de riesgo de diferentes valores de una covariable es proporcional a lo largo del tiempo, es decir, el riesgo es independiente del tiempo (Bradburn et al., 2003; Hess, 1995). La verificación de este supuesto resulta esencial para poder considerar válidos los resultados y predicciones brindados por el modelo resultante.

Mediante el modelo de riesgos proporcionales de Cox podría estimarse la función de supervivencia de una muestra utilizando el diagnóstico brindado por una prueba como una covariable para controlar su efecto en el tiempo hasta que ocurre el evento. Esto permite

obtener la llamada “razón de riesgo” [*risk ratio*], cuya interpretación es similar a la razón de ventajas [*odds ratio*] de una regresión logística, es decir, si la razón de riesgo es cercana a 1 significa que no existen diferencias entre un diagnóstico u otro, pero si es significativamente diferente de uno se puede concluir que un valor del diagnóstico aumenta el riesgo de presentar un evento en un determinado tiempo.

En el ejemplo utilizado podríamos explorar si el consumo de drogas sigue siendo significativo para la edad del primer arresto ante la presencia de otras covariables, como el sexo, la violencia familiar, el estudio y el trabajo. En la **Tabla 1** se observan los resultados de un modelo de regresión de Cox utilizando las variables mencionadas, donde se observa claramente que el consumo de drogas sigue siendo significativo ante la presencia de las demás variables. La razón de riesgo (1.34) nos indica que los consumidores tienen 1.34 veces más riesgo de ser arrestados en comparación con los no consumidores o, lo que es lo mismo, consumir drogas aumenta un 34% la probabilidad de ser arrestado. El intervalo de confianza del 95% nos indica que este riesgo puede ser tan bajo como un 18% o tan alto como un 52%. Además, también se puede observar que el sexo y la violencia familiar no fueron significativos, pero si lo fueron el estudiar (aumentó el riesgo en un 30%) y el trabajar (disminuyó el riesgo en un 39%).

Lo anterior es importante en términos de modelamiento estadístico, ya que al incluir covariables en el modelo de riesgos proporcionales de Cox nos permite controlar la presencia de variables significativas. Esto último se ve claramente reflejado en la **Figura 3**, donde se grafica nuevamente la variable de consumo de sustancias y su relación con la edad del primer arresto, pero esta vez controlando las demás covariables, es decir, se hizo una gráfica en función del modelo de la **Tabla 1**. En ella se observa ahora claramente la diferencia en las medianas de supervivencia (situación no tan clara en la **Figura 2**), en donde ahora los consumidores tienen una mediana de 16 años (IC 95% = 15-18), mientras que los no consumidores tienden a tener un mayor tiempo de supervivencia, con una mediana de 17 años (IC 95% = 16-19).

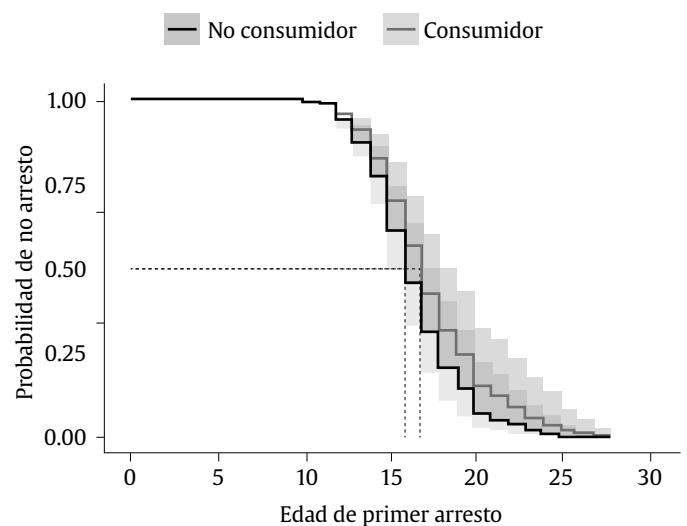


Figura 3. Comparación de la probabilidad de no ser arrestado según el consumo de drogas controlando por diversas covariables.

Tabla 1. Variables asociadas a la edad del primer arresto

Covariables	Beta	Wald	Sig.	Razón de riesgo	IC 95%
Sexo ¹	-.110	2.149	.143	0.896	[0.77, 1.03]
Violencia Familiar	.216	1.133	.287	1.241	[0.83, 1.84]
Consumo de sustancias	.296	20.502	$p < .001$	1.345	[1.18, 1.52]
Estudio	.269	19.472	$p < .001$	1.308	[1.61, 1.47]
Trabaja	-.483	59.911	$p < .001$	0.617	[0.54, 0.69]

¹Ser hombre.

Las aplicaciones de lo anterior son fácilmente identificables. Por ejemplo, podría brindar información suficiente para analizar la validez predictiva de una prueba de valoración de riesgo, ya que confirmaría la eficacia del diagnóstico para predecir si un recluso reincidirá pronto o no. Adicionalmente, el modelo de Cox nos permite incluir tantas covariables como sean de interés en nuestro estudio, por ejemplo, considerar si el recluso participó en un tratamiento, el tipo de delito, sexo, uso de drogas, etc., por lo que la posibilidad de incluir variables de control es infinita.

Técnicas paramétricas que incluyen covariables. Finalmente, existen los llamados modelos paramétricos, los cuales toman en consideración ajustar un modelo con una distribución conocida. Normalmente se consideran distribuciones como la exponencial, la Weibull o la Gompertz (Bradburn et al., 2003; Schmidt y Witte, 2012). Al igual que el modelo de Cox, también requieren del cumplimiento del supuesto de riesgos proporcionales. Los resultados obtenidos por estos modelos son interpretables de forma similar a los producidos por el modelo de Cox. Si bien su uso no es tan extendido como el modelo semiparamétricos de riesgos proporcionales, se ha podido determinar que si la función de supervivencia se ajusta adecuadamente a la distribución elegida, brinda resultados más exactos y con menores errores estándar que el modelo de Cox (Bradburn et al., 2003).

Algunos Casos del Uso del Análisis de Supervivencia en Psicología Jurídica

El uso del análisis de supervivencia para probar la validez predictiva de variables e instrumentos cada vez es mayor en esta área. Por ejemplo, en un clásico estudio realizado por Hart, Kropp y Hare (1988)

para la evaluación de la validez predictiva del *Psychopathy Checklist* (PCL) utilizaron esta técnica para determinar que el PCL es eficaz para hacer predicciones del comportamiento futuro de los criminales. De igual forma, Holsinger, Lowenkamp y Latessa (2006) utilizaron el análisis de supervivencia para poner a prueba la validez predictiva del *Level of Service Inventory-Revised* (LSI-R) en una muestra de delinquentes nativos americanos, un instrumento ampliamente utilizado para gestionar el riesgo y el tratamiento de prisioneros, demostrando que el instrumento tiene una validez predictiva modesta. Por su parte, Stansfield y Williams (2014) utilizaron el análisis de supervivencia para evaluar la validez predictiva del *Domestic Violence Screening Instrument* para predecir la violencia doméstica, con lo cual lograron identificar ítems críticos que permiten mejorar la predicción brindada por el instrumento.

Por otro lado, algunos de los sistemas de predicción de riesgo más importantes en los centros penitenciarios de diversos países han demostrado la validez predictiva de sus sistemas por medio del análisis de supervivencia utilizando la técnica propuesta por Cox. Tal es el caso del *Correctional Offender Management Profiling for Alternative Sanctions* (COMPAS) en los Estados Unidos (Brennan, Dieterich y Ehret, 2009), el *Offender Assessment System* (OASys) en Inglaterra (Howard y Dixon, 2013) o el protocolo RisCanvi en Cataluña, España (Andrés-Pueyo, 2016).

En el ejemplo presentado previamente del estudio de Grieger y Hosser con delinquentes juveniles de Alemania y el poder predictivo del LSI-R también buscaban identificar si el diagnóstico brindado por los 4 grandes factores, los 4 moderadores, y en general 8 predictores (4 moderadores + 4 grandes factores) propuestos por Andrews y Bonta (2010) eran igualmente buenos. Grieger y Hosser obtuvieron que si bien los cuatro grandes factores pueden ayudar a predecir la reincidencia general y violenta de una forma estadísticamente significativa,

Tabla 2. Principales funciones del análisis de supervivencia incluidas en SPSS y R

Funciones para realizar análisis de supervivencia	SPSS	R
	Incluido	Incluido
Descriptivos		
Creación de tablas de supervivencia	Si	Si
Resumen de eventos ocurridos y censurados	Si	Si
Descriptivos de la función de supervivencia (media y mediana)	Si	Si
Gráficas		
Gráfica de la función de supervivencia	Si	Si
Gráfica de la función de riesgo	Si	Si
Gráfica con intervalos de confianza del 95%	No	Si
Estimaciones no paramétricas		
Estimación no paramétrica de Kaplan-Meier	Si	Si
Estimación de la función de riesgo de Nelson-Aalen	Si	Si
Estimación de intervalos de confianza de las funciones de supervivencia y de riesgo	No	Si
Comparación de curvas de supervivencia por grupos		
LogRank (Mantel-Cox)	Si	Si
Breslow	Si	Si
Tarone-Ware	Si	Si ¹
Estimaciones semiparamétricas		
Modelo de regresión de Cox	Si	Si
Análisis <i>Stepwise</i>	Si	Si
Comprobación del supuesto de riesgos proporcionales		
Mediante gráfica de residuos de Schoenfeld	Si ²	Si
Mediante gráfica de log-log de la función de riesgo o supervivencia	Si	Si
Mediante pruebas estadísticas (para covariables cuantitativas)	Si ³	Si
Otros análisis específicos		
Inclusión de variables dependientes del tiempo	Si	Si
Estimación de modelos paramétricos (p. ej., Weibull, exponencial, etc.)	No	Si

¹No viene con el paquete survival por default, sino que requiere el paquete coin.

²De forma indirecta al guardar los residuos y realizar un gráfico de dispersión de residuos vs. el tiempo.

³De forma indirecta al realizar una correlación de rangos entre la variable cuantitativa y el tiempo.

la incorporación de los cuatro moderadores mejora significativamente la predicción de la reincidencia, lo que señala la importancia de considerar de forma global los ocho factores evaluados por el LSI-R. De esta forma, demostraron la validez predictiva del modelo de Riesgo-Necesidad-Responsividad (RNR) propuesto por [Andrews y Bonta \(2010\)](#).

Finalmente, la utilidad del análisis de supervivencia puede ir más allá en su uso con instrumentos de evaluación, ya que representa una forma eficaz de evaluar la efectividad de intervenciones y tratamientos. Por ejemplo, en un estudio llevado a cabo por [Rossi, Berk y Lenihan \(1980\)](#), citado en [Allison, 2010](#)) mediante un diseño completamente experimental se evalúa la utilidad de las ayudas económicas a prisioneros que se dejaron en libertad y cómo esto impacta en la reincidencia delictiva. Después de ajustar un modelo de regresión de riesgos proporcionales de Cox controlando covariables como el estado marital, libertad condicional, nivel educativo, experiencia laboral y número de condenas previas, la ayuda económica resultó ser un factor estadísticamente significativo para retrasar el tiempo de reincidencia de los exreclusos. El estudio anterior y sus datos han sido ampliamente utilizados para la enseñanza de la técnica de análisis de supervivencia como un ejemplo de su uso en la evaluación de la efectividad de los tratamientos penitenciarios. Pero sin duda, la técnica puede extrapolarse a cualquier otro tipo de evaluación de efectividad cuya finalidad sea evaluar el tiempo hasta que sucede un evento. Por ejemplo, se ha utilizado para evaluar y comparar el efecto de tratamientos residenciales y no residenciales para delincuentes sentenciados por delitos relacionados con sustancias, demostrando la eficacia de los tratamientos residenciales frente a los no residenciales ([Krebs, Strom, Koetse y Lattimore, 2009](#)).

Programas Computacionales para el Análisis de Supervivencia

En la actualidad, existen diversos programas que permiten aplicar la técnica, tanto de pago (como SPSS o Minitab) como de acceso libre (por ejemplo, R). En la [Tabla 2](#) se resumen brevemente las principales funciones del análisis de supervivencia de las que consta el SPSS (uno de los programas estadísticos más utilizados por los psicólogos) y R (un programa libre y multiplataforma muy utilizado en el análisis de datos).

En el caso de SPSS, el programa ofrece muchas funciones relacionadas con esta técnica. Tal como se aprecia en la [Tabla 2](#), únicamente la representación gráfica de los intervalos de confianza, algunos elementos de la exploración de supuestos y el ajuste de modelos paramétricos no están considerados entre las funciones que brinda este software. A través del menú “analizar - supervivencia”, se pueden obtener las siguientes opciones: tablas de supervivencia, Kaplan Meier, regresión de Cox y Cox con variable dependiente del tiempo. Las primeras dos opciones brindan los descriptivos de la función de supervivencia y de riesgo (incluyendo las comparaciones de grupos), mientras que las últimas dos permiten el ajuste de un modelo de riesgos proporcionales de Cox tanto sin variables dependientes del tiempo como con ellas.

En el caso de R, al ser un programa de acceso abierto y respaldado por una comunidad de programadores y estadísticos, el programa permite todas las opciones brindadas por SPSS así como otras no incluidas, tales como la representación gráfica con intervalos de confianza y el ajustar modelos estadísticos con una distribución en particular (p. ej., Weibull) e incluso otros métodos más sofisticados, como el análisis de supervivencia multivariado. Si bien su uso requiere de conocimientos básicos de programación, lo cierto es que existen diversas interfaces gráficas que permiten emular ambientes similares a SPSS, tales como R Commander ([Fox, 2005](#)), el cual incluye un *plug in* para realizar análisis de supervivencia ([Fox y Carvalho, 2012](#)) y que ejecuta los principales comandos del paquete *survival* ([Therneau y Grambsch, 2000](#)).

En resumen, se aprecia que existen varias alternativas para poder ejecutar el análisis de supervivencia a través de software, tanto de pago como de acceso libre. Si se desea un resumen más detallado de los pasos exactos para hacer los análisis, se recomienda a los usuarios de SPSS el libro de [Landau y Everitt \(2004\)](#), *A Handbook of Statistical Analyses using SPSS*, cuyo capítulo diez aborda en detalle esta técnica (incluyendo la exploración de los supuestos para el modelo de riesgos proporcionales de Cox). Para los usuarios de R, si bien el libro de [Kleinbaum y Klein \(2006\)](#) incluye anexos para SPSS, Stata y R, hay que destacar que en general brinda una introducción muy completa de la técnica para los principiantes, así como de la forma de ejecutar el código de análisis de supervivencia en R paso por paso.

Conclusiones

Si bien se ha podido presentar esta técnica como una herramienta más para evaluar la validez predictiva de un instrumento o variable, hay que mencionar algunas de sus limitaciones, entre las que se encuentran la censura informativa, el tiempo de seguimiento y la existencia de covariables dependientes del tiempo.

Con respecto a la censura, uno de los requerimientos de la técnica es la consideración de que la censura de los datos no es informativa, es decir, no aporta información del fenómeno en cuestión. Sin embargo, es posible que en algunos casos la censura de los datos sí tenga un significado ([Bradburn et al., 2003](#)). Por ejemplo, si se estudiara delincuentes inmigrantes, la censura podría deberse a que, posterior a la liberación, estos hayan sido deportados a su país de origen, por lo que la censura en el tiempo de reincidencia estaría asociada con la deportación. Es por ello que resulta vital evaluar si la existencia de censuras es informativa o no.

Otro aspecto es la consideración del tiempo de seguimiento. Si bien no existe un criterio que defina el tiempo de seguimiento en un estudio, tiene que someterse a consideración varios aspectos, desde metodológicos hasta financieros. Por ejemplo, el periodo de seguimiento no debe ser muy corto, puesto que se corre el riesgo de que no se presente ningún evento, y un tiempo demasiado largo implica más gasto de recursos. Una recomendación es analizar estudios similares para identificar los tiempos utilizados por otros investigadores. [Capdevila et al., \(2015\)](#) encontraron que un periodo de seguimiento de entre 2 y 3 años es suficiente para los estudios de reincidencia penitenciaria. Por otro lado, el tiempo variará en función del evento que se desea observar; por ejemplo, en el campo de la medicina los estudios oncológicos consideran distintos periodos de seguimiento dependiendo del tipo de cáncer diagnosticado.

Así mismo hay que tomar en consideración la existencia de variables dependientes del tiempo. Por ejemplo, algunos estudios han demostrado que la edad es un factor de riesgo de reincidencia y que los reclusos jóvenes tienen más probabilidad de reincidir que sus homólogos más longevos ([Uggen, 2000](#)). Dado que el riesgo depende del tiempo, esto violaría el supuesto del riesgo proporcional que requieren los modelos paramétricos y semiparamétricos, aunque existen metodologías que permiten analizar estas variables considerando su dependencia del tiempo ([Hosmer et al., 2008](#)), y la mayoría de los softwares estadísticos permiten realizarlo. No obstante, su utilización es poco común y suelen emplearse otras técnicas, como dicotomizar variables continuas dependientes del tiempo ([Clark, Bradburn, Love y Altman, 2003b](#)), técnica que ha resultado útil para poder utilizar modelos de riesgos proporcionales.

Como se ha visto, todas las desventajas anteriores han tenido formas de contrarrestarlas, ya sea en el proceso del análisis o en la planificación de la metodología del estudio. Además, entre las ventajas del método, como ya se mencionó, se encuentra la consideración de datos censurados, distintas formas de comparar tiempos de supervivencia de varios grupos, así como su inclusión en softwares comerciales bien conocidos, como SPSS, o de código abierto, como R ([Fox y](#)

Carvalho, 2012; Kleinbaum y Klein, 2006). Además, hay que destacar que el análisis de supervivencia no es excluyente para el uso de otras técnicas para evaluar la validez predictiva; muchos de los estudios citados en los ejemplos de aplicación conjugan la técnica de análisis de supervivencia con las curvas ROC. Así, lejos de ser una técnica que pretenda sustituir a las existentes, es una herramienta más para poder brindar evidencia empírica de la validez predictiva.

Finalmente, otra gran ventaja es la posible extrapolación de la técnica para evaluar la validez predictiva de diferentes instrumentos y técnicas, que no necesariamente tienen que estar circunscritos en el diagnóstico, o incluso en el campo de la psicología jurídica. Por ejemplo, ha sido utilizado para evaluar la eficacia de ciertos tratamientos con delinquentes (McNiell y Binder, 2007) o bien para evaluar otros tipos de instrumentos no criminológicos y sus variables relacionadas, como en algunos estudios de depresión (Tohen, Watermaux y Tsuang, 1990). En un ejemplo aplicado a la psicología clínica, Anderson y Lambert (2001) utilizaron la técnica de análisis de supervivencia para evaluar el número de sesiones que se necesitan para que un paciente tenga un cambio clínicamente significativo y Willett y Singer (1991) han sugerido su uso en la psicología educativa para estudiar la deserción escolar y los factores que pueden predecir cuándo ocurrirá.

En conclusión, el análisis de supervivencia representa una metodología que a partir de ventajas analíticas prevé una herramienta inestimable para determinar empíricamente la validez predictiva de un instrumento, un signo o un tratamiento, técnica que además es conjugable con otros tipos de metodologías que permiten brindar una mejor evidencia de validez. El aprendizaje, el manejo correcto y la difusión de esta técnica son esenciales para cualquier profesional que desee validar la predicción de un instrumento o la eficacia de un tratamiento. Si el lector desea comprender mejor la técnica desde un punto de vista teórico-matemático, se recomienda encarecidamente el libro de Hosmer et al., (2008), donde se desglosan los principios estadísticos que subyacen a las técnicas planteadas. Si lo que se desea es una introducción sencilla, amena y práctica, se recomienda revisar el libro de Kleinbaum y Klein (2006), donde se aborda de forma clara y completa la técnica del análisis de supervivencia, incluyendo anexos para su uso en softwares bien conocidos como Stata, SPSS o R.

Extended Summary

Validity is one of the most important elements in measuring instruments of any kind, especially in psychology. An instrument that measures what it is supposed to measure is considered as a valid test. The present article proposes the survival analysis as a methodology that should be considered, in conjunction with the other existing techniques, to evaluate the predictive validity of a test or treatment.

Predictive validity is a type of criterion validity that seeks to predict a situation. A clear example of this validity is the study by Grieger and Hosser (2013) that attempts to determine whether the eight risk factor model for criminal behavior assessed by Level of Service Inventory-Revised (LSI-R) is useful in predicting recidivism in juvenile delinquents. The study by Grieger and Hosser demonstrates empirically that recidivists subjects with a diagnosis of low recidivism were proportionately fewer than those with a diagnosis of medium and high recidivism.

Ways of Assessing Predictive Validity

Predictive validity is undoubtedly one of the most desired types of validity in any measurement instrument, especially in the field of legal psychology, which is why there are now several ways of assessing it. These techniques are briefly described below.

One of the most used ways of evaluating the predictive validity of a test is through the analysis of contingency tables where the diagnosis predicted by the test is compared with the criterion

to be predicted. This technique allows the identification of false positives and false negatives, positive and negative predictive power, general diagnostic power, prevalence and sensitivity and specificity. However, Rice and Harris (1995) suggest that these statistics have certain deficiencies in their use since they are highly dependent on the prevalence in which an event occurs.

Rice and Harris (1995) propose to evaluate predictive validity by using the ROC Curve (Receiver Operating Characteristic) since it is independent of the prevalence or cut-off point to classify an event. However, one of its main limitations is that it is necessary to cover the normality assumption of test scores. Rice and Harris mention that, in criminology, violence rarely has a normal distribution.

In addition to this, there is also the problem of monitoring the research subjects. In many cases, the tracking time is less than the time it takes for an event to happen so that many times the events of interest do not happen. This predicament is especially problematic because a proportion of participants do not present the event during the follow-up period. However, this does not mean that participants will never experience the event in a period longer than the follow-up. This phenomenon is known in statistics as censored data. Survival analysis is a statistical technique that takes these disadvantages into account and provides information on the predictive validity of a test.

Survival Analysis: Principal Features

Definition. Survival analysis makes use of survival data, which are “measurements of elapsed time between the initial enrollment in a study and the final disposition of the study subject” (Daniel & Cross, 2013, p. 751). Survival in this context means only that the event has not occurred. Survival analysis takes into account censored data, a situation that most statistical techniques do not perform (Clark, Bradburn, Love, & Altman, 2003a).

Survival analysis also offers the possibility of comparing the survival time of two or more different populations (Clark et al., 2003a). This is especially useful when you want to contrast, for example, people who received a diagnosis of low risk of recidivism, with one high risk of recidivism.

Main techniques in Survival Analysis

Non-parametric descriptive techniques. These models allow estimating the survival function without taking into account the distribution of the data, only considering the information provided by the sample. One of the best known and used techniques for this is the Kaplan-Meier statistic. The function estimated from the Kaplan-Meier method can be plotted to obtain an image similar to Figure 1A (Hosmer et al., 2008). In this plot, the x-axis represents the time from the start of the study until it ends of the follow-up, and on the y-axis represents the probability that an individual will not present the event (in this case, be arrested).

One of the advantages of the Kaplan-Meier method is that it allows the estimation of the survival functions of different groups (Hosmer et al., 2008). This is possible thanks to the adaptation of non-parametric statistics to the survival analysis. For example, the LogRank statistic, or the Breslow statistic, are adaptations of the chi-square test that takes into account the existence of censored data.

Semi-parametric techniques including covariates. One of the most popular models in its use is the Cox proportional hazards model (Bradburn, Clark, Love, & Altman, 2003), a semiparametric regression technique that allows us to estimate the survival function using a regression model that can include covariates. Also, the interpretation of its results is similar to a logistic regression.

Using the Cox proportional hazards model, the survival function of a sample could be estimated, using the diagnosis provided by a test

as a covariate to control its effect over time until the event occurs. This allows us to obtain the so-called risk ratio, whose interpretation is similar to the odds ratio of a logistic regression. That is, if the risk ratio is close to 1, it means that there is no difference between one diagnosis or another.

Parametric techniques including covariates. Finally, there are the so-called parametric models, which take into account to fit a model with a known distribution. Usually, distributions such as exponential, Weibull, or Gompertz distributions are considered (Bradburn et al., 2003; Schmidt & Witte, 2012). The results obtained by these models are interpretable in a similar way to those produced by the Cox proportional hazard model.

Utility and examples of its use. Some of the most important risk prediction systems in prisons in several countries have demonstrated the predictive validity of their systems using survival analysis, employing the technique of the proportional hazard model. Such is the case of the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) in the United States (Brennan, Dieterich, & Ehret, 2009), the Offender Assessment System (OASys) in England (Howard & Dixon, 2013), or the RisCanvi protocol in Catalonia, Spain (Andrés-Pueyo, 2016).

Conclusion

Survival analysis represents a methodology that, based on analytical advantages, provides an invaluable tool to assess, empirically, the predictive validity of an instrument or a treatment. This technique is also conjugable with other types of methodologies, which allows better evidence of its validity. The learning, proper management, and diffusion of this technique are essential for any professional who wishes to validate the prediction of an instrument or the effectiveness of a treatment.

Referencias

- Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- Allison, P. D. (2010). *Survival analysis using SAS: A practical guide* (2nd ed.). Cary, NC: SAS Publishing.
- Anderson, E. M. y Lambert, M. J. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology, 57*, 875-888. <https://doi.org/10.1002/jclp.1056>
- Andrés-Pueyo, A. (2016). ¿Es técnicamente posible anticipar la reincidencia delictiva? El protocolo RisCanvi en las prisiones de Cataluña. En *IX Jornadas de ATIP Almagro 2016* (pp. 53-78). Asociación de Técnicos Superiores de Instituciones Penitenciarias. Almagro.
- Andrés-Pueyo, A. y Echeburúa, E. (2010). Valoración del riesgo de violencia: instrumentos disponibles e indicaciones de aplicación. *Psicothema, 22*, 403-409.
- Andrews, D. A. y Bonta, J. (2010). *The psychology of criminal conduct* (5nd ed.). Cincinnati, OH: Anderson Publishing.
- Baldessarini, R. J., Finklestein, S. y Arana, G. W. (1983). The predictive power of diagnostic tests and the effect of prevalence of illness. *Archives of General Psychiatry, 40*, 569-573. <https://doi.org/10.1001/archpsyc.1983.01790050095011>
- Bradburn, M. J., Clark, T. G., Love, S. B. y Altman, D. G. (2003). Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods. *British Journal of Cancer, 89*, 431-436. <https://doi.org/10.1038/sj.bjc.6601119>
- Brennan, T., Dieterich, W. y Ehret, B. (2009). Evaluating the predictive validity of the COMPAS risk and needs assessment system. *Criminal Justice and Behavior, 36*, 21-40. <https://doi.org/10.1177/0093854808326545>
- Campbell, J. C. (1986). Nursing assessment for risk of homicide with battered women. *Advances in Nursing Science, 8*(4), 36-51.
- Capdevila, C., Blanch Serentill, M., Ferrer Puig, M., Andrés Pueyo, A., Framis Ferrer, B., Comas López, N., ... Mora Encinas, J. (2015). *Tasa de Reincidencia Penitenciaria 2014*. Barcelona, España: Centro de Estudios Jurídico y Formación Especializada.
- Clark, T. G., Bradburn, M. J., Love, S. B. y Altman, D. G. (2003a). Survival analysis part I: Basic concepts and first analyses. *British Journal of Cancer, 89*, 232-238. <https://doi.org/10.1038/sj.bjc.6601118>
- Clark, T. G., Bradburn, M. J., Love, S. B. y Altman, D. G. (2003b). Survival analysis part IV: Further concepts and methods in survival analysis. *British Journal of Cancer, 89*, 781-786. <https://doi.org/10.1038/sj.bjc.6601117>
- Colosimo, E., Ferreira, F., Oliveira, M. y Sousa, C. (2002). Empirical comparisons between Kaplan-Meier and Nelson-Aalen survival function estimators. *Journal of Statistical Computation and Simulation, 72*, 299-308. <https://doi.org/10.1080/00949650212847>
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological), 34*, 187-220.
- Daniel, W. y Cross, C. L. (2013). *Biostatistics: A foundation for analysis in the health sciences* (10th ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- Farrington, D. P. y Loeber, R. (2000). Some benefits of dichotomization in psychiatric and criminological research. *Criminal Behaviour and Mental Health, 10*, 100-122. <https://doi.org/10.1002/cbm.349>
- Fox, J. (2005). The R Commander: A basic statistics graphical user interface to R. *Journal of Statistical Software, 14*(9), 1-42. <https://doi.org/10.18637/jss.v014.i09>
- Fox, J. y Carvalho, M. S. (2012). The RcmdrPlugin.survival package: Extending the R Commander interface to survival analysis. *Journal of Statistical Software, 49*(7), 1-32. <https://doi.org/10.18637/jss.v049.i07>
- Gómez, G. y Cobo, E. (2004). Hablemos de... Análisis de Supervivencia. *Gastroenterología y Hepatología Continuada, 3*, 185-191. Recuperado de <http://www.elsevier.es/en/node/2121697>
- Grieger, L. y Hosser, D. (2013). Which risk factors are really predictive?: An analysis of Andrews and Bonta's "Central Eight" Risk Factors for recidivism in German youth correctional facility inmates. *Criminal Justice and Behavior, 41*, 613-634. <https://doi.org/10.1177/0093854813511432>
- Harpur, T. J., Hare, R. D. y Hakstian, A. R. (1989). Two-factor conceptualization of psychopathy: Construct validity and assessment implications. *Journal of Consulting and Clinical Psychology, 1*, 6-17. <https://doi.org/10.1037/1040-3590.1.1.6>
- Harpur, T. J., Ralph Hakstian, A., Hare, R. D., Newman, J., Peterson, B., Raine, A. y Serin, R. (1988). Factor structure of the psychopathy checklist. *Journal of Consulting and Clinical Psychology, 56*, 741-747. <https://doi.org/10.1037/0022-006X.56.5.741>
- Hart, S. D., Kropp, P. R. y Hare, R. D. (1988). Performance of male psychopaths following conditional release from prison. *Journal of Consulting and Clinical Psychology, 56*, 227-232. <https://doi.org/10.1037/0022-006X.56.2.227>
- Hernández Sampieri, R., Fernández Collado, C. y Baptista Lucio, M. (2013). *Metodología de la investigación* (6 ed.). México, DF: McGraw-Hill.
- Hess, K. R. (1995). Graphical methods for assessing violations of the proportional hazards assumption in cox regression. *Statistics in Medicine, 14*, 1707-1723. <https://doi.org/10.1002/sim.4780141510>
- Hollander, M., Wolfe, D. A. y Chicken, E. (2014). *Nonparametric statistical methods*. Hoboken, NJ: John Wiley & Sons, Inc.
- Holsinger, A. M., Lowenkamp, C. T. y Latessa, E. J. (2006). Exploring the validity of the Level of Service Inventory-Revised with native American offenders. *Journal of Criminal Justice, 34*, 331-337. <https://doi.org/10.1016/j.jcrimjus.2006.03.009>
- Hosmer, D. W., Lemeshow, S. y May, S. (2008). *Applied survival analysis: Regression modeling of time-to-event data* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- Howard, P. D. y Dixon, L. (2013). Identifying change in the likelihood of violent recidivism: Causal dynamic risk factors in the OASys violence predictor. *Law and Human Behavior, 37*, 163-174. <https://doi.org/10.1037/lhb0000012>
- Instituto Nacional de Estadística y Geografía. (2014). *Encuesta de Cohesión Social para la Prevención de la Violencia y la Delincuencia (ECOPRED) 2014*. México, DF: Instituto Nacional de Estadística y Geografía. Recuperado de <http://www.beta.inegi.org.mx/proyectos/enchogares/especiales/ecopred/2014/>
- Kaplan, E. L. y Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*(282), 457-481. <https://doi.org/10.1080/01621459.1958.10501452>
- Kelley, T. L. (1927). *Interpretation of educational measurements*. New York, NY: Macmillan.
- Kleinbaum, D. G. y Klein, M. (2006). *Survival analysis: A self-learning text*. New York, NY: Springer Science & Business Media. <https://doi.org/10.1007/978-1-4419-6646-9>
- Krebs, C. P., Strom, K. J., Koetse, W. H. y Lattimore, P. K. (2009). The impact of residential and nonresidential drug treatment on recidivism among drug-involved probationers: A survival analysis. *Crime & Delinquency, 55*, 442-471. <https://doi.org/10.1177/001128707307174>
- Landau, S. y Everitt, B. S. (2004). *A handbook of statistical analysis using SPSS*. Boca Raton, FL: Chapman & Hall/CRS Press LLC.
- Malterer, M. B., Lilienfeld, S. O. y Newman, J. P. (2010). Concurrent validity of the Psychopathic Personality Inventory with offender and community samples. *Assessment, 17*, 3-15. <https://doi.org/10.1177/1073191109349743>
- McNiel, D. E. y Binder, R. L. (2007). Effectiveness of a mental health court in reducing criminal recidivism and violence. *American Journal of Psychiatry, 164*, 1395-1403. <https://doi.org/10.1176/appi.ajp.2007.06101664>
- Palmer Pol, A. y Montañó Moreno, J. J. (2002). Redes neuronales artificiales aplicadas al análisis de supervivencia: un estudio comparativo con el modelo de regresión de Cox en su aspecto predictivo. *Psicothema, 14*, 630-636.
- Rice, M. E. y Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63*, 737-748. <https://doi.org/10.1037/0022-006X.63.5.737>

- Rossi, P. H., Berk, R. A. y Lenihan, K. J. (1980). *Money, work, and crime: Experimental evidence*. New York, NY: Elsevier.
- Schmidt, P. y Witte, A. (2012). *Predicting recidivism using survival models*. New York, NY: Springer Science & Business Media.
- Schumacker, R. E. y Lomax, R. G. (2015). *A beginner's guide to Structural Equation Modeling*. New York, NY: Routledge, Taylor & Francis Group.
- Singer, J. D. y Willett, J. B. (2003). Survival analysis. En J. A. Schinka y W. F. Velicer (Eds.), *Handbook of psychology. Volume 2 Research methods in psychology* (pp. 555-580). Hoboken, NJ: John Wiley & Sons, Inc. <https://doi.org/10.1002/0471264385.wei0222>
- Stansfield, R. y Williams, K. R. (2014). Predicting family violence recidivism using the DVSI-R: Integrating survival analysis and perpetrator characteristics. *Criminal Justice and Behavior*, 41, 163-180. <https://doi.org/10.1177/0093854813500776>
- Therneau, T. M. y Grambsch, P. M. (2000). *Modeling survival data: Extending the Cox Model*. New York, NY: Springer.
- Tohen, M., Watermaux, C. M. y Tsuang, M. T. (1990). Outcome in Mania: A 4-year prospective follow-up of 75 patients utilizing survival analysis. *Archives of General Psychiatry*, 47, 1106-1111. <https://doi.org/10.1001/archpsyc.1990.01810240026005>
- Uggen, C. (2000). Work as a turning point in the life course of criminals: A duration model of age, employment, and recidivism. *American Sociological Review*, 67, 529-546. <https://doi.org/10.2307/2657381>
- Willett, J. B. y Singer, J. D. (1991). From whether to when: New methods for studying student dropout and teacher attrition. *Review of Educational Research*, 61, 407-450. <https://doi.org/10.3102/00346543061004407>