

¿Se podrán hacer predicciones en las ciencias sociales?

Can predictions be made in the social sciences?

Ochoa-Pachas, José Mario

Orcid: 0000-0002-0675-2196

Universidad Autónoma del Perú

jochoap@autonoma.edu.pe

Yunkor-Romero, Yurela Kosett

Orcid: 0000-0001-9902-5993

Universidad Autónoma del Perú

yurelak@autonoma.edu.pe

Recibido el: 10.11.2022

Aceptado el: 16.12.2022

RESUMEN

Durante mucho tiempo las ciencias sociales se han dedicado a realizar exclusivamente estudios cualitativos, generando metodologías diferentes a los estudios cuantitativos. Lo cierto es que estos diseños, métodos, técnicas e instrumentos no se contraponen a los estudios cuantitativos sino por el contrario se complementan. Para poder realizar estudios cuantitativos primero se deben realizar estudios cualitativos que permitan dotar de teorías, principios, leyes, axiomas que sean rigurosos, robustos y permitan el desarrollo natural de las investigaciones cuantitativas. Sin estudios cualitativos no puede haber investigaciones cuantitativas, ambas son ciencias y ambas son importantes para ella.

Una línea de investigación debe de haber pasado por el proceso natural de la ciencia: exploración, descripción, relación, explicación y con todo ese proceso llegar a la predicción. Esto permitirá darle el carácter científico al tema investigado y al mismo tiempo poder tener una visión más consistente del desarrollo de una línea de investigación.

Los estudios predictivos son poco utilizados en las ciencias sociales, al igual que las investigaciones explicativas con intervención. Esto debido a tres factores: la manipulación, el control y la aleatorización. En este aspecto el uso de la estadística es

importante, no solo porque se tenga que probar hipótesis o realizar estimación puntal, sino porque se tiene que dar consistencia al trabajo realizado. Este escrito busca dar una explicación a este campo tan fascinante como es el de la predicción sustentado en bases científicas sólidas y pertinentes, de ahí que se busca darles a los profesionales de las ciencias sociales una herramienta vital para el crecimiento y desarrollo de las mismas.

Palabras clave: investigación de nivel predictivo, multivariado, bivariado, estadística descriptiva, estadística inferencial, variables predictores, variable a predecir.

ABSTRACT

For a long time, the social sciences have dedicated themselves to carrying out exclusively qualitative studies, generating methodologies different from quantitative studies. The truth is that these designs, methods, techniques and instruments are not opposed to quantitative studies but rather complement each other. In order to carry out quantitative studies, qualitative studies must first be carried out to provide theories, principles, laws, axioms that are rigorous, robust and allow the natural development of quantitative research.

A line of investigation must have gone through the natural process of science: exploration, description, relationship, explanation and with all that process arrive at the prediction. This will give the research topic a scientific character and at the same time be able to have a more consistent vision of the development of a line of research.

Predictive studies are little used in the social sciences, as are explanatory research with intervention. This is due to three factors: manipulation, control and randomization. In this aspect, the use of statistics is important, not only because hypotheses have to be tested or point estimates made, but also because the work done has to be given consistency. This writing seeks to give an explanation to this field as fascinating as that of prediction based on solid and relevant scientific bases, hence it seeks to give social science professionals a vital tool for their growth and development.

Keywords: predictive level research, multivariate, bivariate, descriptive statistics, inferential statistics, predictor variables, variable to predict.

SUMARIO

I. Introducción; II. Los estudios predictivos en las ciencias sociales; III. Características de los estudios predictivos; IV. El marco teórico predictivo; V. Los objetivos en los estudios predictivos; VI. Las hipótesis predictivas; VII. Análisis de datos predictivos; Conclusiones; Referencias.

CONTENIDO

I. Introducción

La predicción no es casualidad, ni magia, mucho menos suerte; esta permite realizar pronósticos y prever situaciones incómodas y peligrosas, pero se debe pasar por todo un proceso indagatorio riguroso, preciso y persistente. Desde la exploración de los constructos hasta la experimentación, se requiere darle el soporte necesario y vital a este nivel de investigación.

Sin tener claro los conceptos de las variables que se van a emplear; si no se conocen las características primarias de ellas; si no se conocen con que otras variables se relacionan y que tan fuertes son las asociaciones; no se ha establecido la causalidad entre dos o más variables, es muy difícil realizar estudios predictivos que ayuden a entender cómo se genera y producen los fenómenos que se expresan en la sociedad.

Una cosa es establecer la causalidad, para lo cual previamente se ha establecido la relación, sentido y fuerza de asociación entre variables, y otra cosa es realizar predicciones claras y precisas, obviamente con su margen de error correspondiente. Si uno tiene una relación de variables fuerte, entonces puede establecer que exista causalidad, pero ello no es suficiente por lo cual se tiene que diseñar un experimento, se tiene que hacer uso de la repetibilidad y reproducibilidad de los ensayos que se han diseñado y recién ahí se puede señalar que en el fenómeno estudiado existe causalidad. Esto mismo pasa con los estudios predictivos, es decir, no puedo realizar predicciones solo si existe causalidad; se parte de ella, para mediante procesos estadísticos se puedan conformar modelos que ayuden a realizar las predicciones. Cada nivel o fase o alcance de las indagaciones científicas tienen sus características y sus límites y ello permitirá efectivizar un estudio

claro y contundente. Una cosa es la teoría y otra la práctica, el conocimiento es importante y de igual forma la experticia que proviene de realizar los trabajos en el campo; en ambos aspectos debe de haber el equilibrio adecuado para obtener el mejor provecho posible de los resultados que se han conseguido.

II. Los estudios predictivos en las ciencias sociales

Realizar una predicción implica establecer con claridad y precisión que, bajo ciertas singularidades, se va a producir un hecho. Se puede expresar bajo el siguiente silogismo: Si X es cierto, entonces Y también será cierto. En ese contexto, se recuerda que la ciencia, sea natural o social, tiene una función predictiva, ello significa que puede establecer que puede ocurrir a corto, mediano o largo plazo, por lo que, se hace necesario que se establezca el vínculo y la causalidad entre las variables que se analizan. Para realizar predicciones, se tiene que tener información y cuanto más precisa, mejor. Se puede ejercitarla, cuando, antes de leer un libro y teniendo información previa, uno puede responder las preguntas ¿de qué trata el texto? ¿qué podría pasar? ¿qué se podría imaginar?; o ante un hecho social, como una huelga y la información respectiva, uno se puede preguntar ¿la huelga se produce por exigencia de incremento salarial? ¿la huelga se va a prolongar?

Los estudios predictivos se encuentran dentro de las investigaciones cuantitativas porque mide las variables estadísticamente (Arbaiza, 2013), de tipo básico, las cuales permiten el incremento del conocimiento en una determinada área del conocimiento humano. Se pueden establecer dos tipos de predicciones, la primera se denomina categórica que se encuentra vinculada a los hechos que se pueden producir o no en un determinado contexto, como por ejemplo, mañana va hacer calor o la temperatura disminuirá en el transcurso de la semana; la segunda se llama probabilística, debido a que el suceso que se intenta predecir, se sustenta en las probabilidades, verbi gracia, se tiene un 90 % de probabilidades que se reduzca la temperatura ambiente o hay un 75 % de probabilidad de que llueva el día de mañana. En el ámbito de las ciencias sociales, se puede realizar una predicción categórica, mañana liberarán al prisionero; o una predicción probabilística, las prisiones preventivas se reducirán en un 85 %.

La predicción tiene su base en el análisis estadístico, partiendo de información histórica o nueva que son la base de las predicciones de ciertos patrones de conducta, que se pueden emplear, ya sea en el pasado, presente o futuro. En el campo de la predicción, las variables se pueden clasificar como escalables, cuando son numéricas y se caracterizan por ser continuas o discretas; y las no escalables, cuando son categóricas y su singularidad es que pueden ser nominales u ordinales. También, dentro del nivel de investigación predictivo, se pueden establecer las variables predictoras, que provienen de las variables independientes, que son los sucesos que van a generar que se produzca el evento, y las variables a predecir, ello significa que el hecho va a producirse. Los estudios predictivos manipulan y controlan las variables, por lo que son investigaciones de intervención, de acuerdo a la tabla 1.

Tabla 1

Tipo, nivel, subniveles e intervención estudios predictivos

TIPO DE INVESTIGACIÓN	NIVEL DE INVESTIGACIÓN	SUBNIVEL	INTERVENCIÓN
Básica o Pura	Predictivo	Prevención	Si
		Pronóstico	Si
		Predicción	Si

Nota. Tipo básico, con y sin intervención.

En el campo de las ciencias sociales, si no existen, estudios experimentales, menos realizan investigaciones predictivas, ya que, para poder realizarlos, primero se debe conocer a profundidad los constructos, las teorías y los principios en los cuales se rige la temática abordada; luego se debe conocer si existe o no relación entre las variables involucradas; después se debe comprobar experimentalmente la causalidad entre las variables. Solo en ese caso se puede realizar la respectiva predicción, siguiendo la respectiva línea de investigación. A esto se puede agregar, que estos estudios son longitudinales, y pueden ser prospectivos o retrospectivos.

Por otra parte, si bien es cierto que las investigaciones predictivas son de carácter básico, también se enmarcan dentro de los estudios cuantitativos. El enfoque es uno solo, científico, y los tipos de investigaciones pueden ser básicas o puras cuando buscan ampliar el conocimiento, formulan teorías, presentan principios, y las investigaciones aplicadas, las cuales buscan mejorar, resolver un problema o diseñar una herramienta. Pero también

pueden ser de tipo cualitativa, cuando buscan definir un término, plantear una nueva teoría, interpretar un hecho o resolver un problema desde el punto de vista teórico y no emplea estadística, sino métrica, es decir, datos recolectados y procesados por terceros; y el tipo cuantitativo, que emplea la estadística para recolectar y procesar sus propios datos, de acuerdo a la tabla 2.

Tabla 2

Tipo, nivel, subnivel de los estudios predictivos

TIPO DE INVESTIGACIÓN	NIVEL DE INVESTIGACIÓN	SUBNIVEL	INTERVENCIÓN
Cuantitativo	Predictivo	Prever	SI
		Pronosticar	SI
		Predecir	SI

Nota: Tipo cuantitativo y predictivo.

Para vincular las clases de indagación por el aumento de los saberes y el empleo de la estadística, estableciéndose en función de la tabla 3, el estudio predictivo con sus subniveles de investigación (Supo, 2015)

Tabla 3

Tipos y niveles de Investigación

TIPO DE INVESTIGACIÓN (I)	NIVEL Y SUBNIVEL DE INVESTIGACIÓN	TIPO DE INVESTIGACIÓN (II)
Básico	Predictivo:	
	- Prevención	Cuantitativo
	- Pronóstico	
	- Predicción	

Nota: Básico, cuantitativo predictivo.

III. Características del nivel predictivo

Los estudios predictivos son multivariados, pero frecuentemente se utilizan dos variables (bivariado), dos elementos que cambian, uno denominado predictor y otro llamado a predecir, la primera proviene de la variable independiente y otra se deriva de la variable dependiente. Las variables predictoras son las que se manipulan y controlan para saber los efectos que tienen sobre la variable a predecir. Al igual que los demás estudios básicos,

se pasa por un proceso, que no necesariamente es lineal, pero por razones didácticas sigue un orden, tal como un proceso de observación, descripción, relación, explicación y la respectiva predicción de los hechos que se estudian. Todo ello se debe sustentar en una teoría (Shmueli & Koppius, 2011), ya que el principal método que se utiliza, es el deductivo (Bunge, 2019), entonces se parte de la teoría para predecir el suceso que se va a producir. Por ello, es importante emplear una teoría robusta y que permita dar una predicción certera del suceso analizado.

En ese contexto, los problemas se deben plantear de manera adecuada y pertinente, y ello requiere de un trabajo epistemológico y ello implica que el indagador debe empezar la búsqueda, a partir de su curiosidad, pasión e interés con el propósito de formularlo y ubicarlo dentro de un ámbito del conocimiento científico (Lavado, 2010)

Como se ha indicado, la investigación predictiva es deductiva, es la teoría científica la que va a predecir el fenómeno que se va a producir. Es imprescindible tener un profundo conocimiento de la teoría que se va a emplear, para darle el entendimiento adecuado y dar la explicación correcta, reduciendo los sesgos en las medidas e interpretaciones (Bunge, 1997). En ese orden de ideas, si se va a desarrollar un estudio predictivo sobre la prisión preventiva, se debe de sustentar en una teoría madura que pueda explicar y predecir el hecho analizado. Asimismo, se tienen dos modelos predictivos, el primero se denomina de clasificación, debido a que permite predecir la pertenencia de una clase, ello implica que se taxonomiza a los grupos que se estudian y determinar qué elementos son más propensos a una y otra decisión, siendo los resultados de carácter binario. El otro modelo es de regresión, debido a que permite la predicción de un valor de un determinado segmento, de un beneficio o un prejuicio que puede tener una persona que ha perdido su libertad por una acción ilegal que cometió.

Una característica que se tiene que resaltar, son las comparaciones que se realizan en las investigaciones predictivas, en el campo social, se vinculan con que pasará en el porvenir; se puede hacer una comparación, del pronóstico de cuantas empresas se van a formalizar, pronosticar cuantas personas pueden tener prisión preventiva, pronóstico de madres gestantes que darán a luz en prisión; de acuerdo a las singularidades y particularidades

de cada persona. Dos personas que se encuentran internas en un centro penitenciario, no se van a encontrar en la misma situación jurídica, entonces tendrán diferentes pronósticos.

Se puede hacer comparaciones del tiempo de vida media, aplicando el análisis de supervivencia en dos grupos de internos acusados de homicidio calificado, que pueden ser defendidos de diversas formas; los resultados de la defensa deben estar en concordancia con las evidencias, las pruebas y el enfoque que el abogado defensor utilice para armar su caso, donde se puede aplicar una técnica estadística comparativa que se llama análisis de supervivencia log-rank2, la que es una prueba no paramétrica o también el análisis de regresión de Cox para variables categóricas.

El problema predictivo debe de formularse a partir de los elementos constitutivos: línea de investigación, propósito del estudio y población. Los problemas investigativos tienen su propia personalidad debido a que se suele confundir con los de enseñanza-aprendizaje. Las preguntas tienen que ubicarse dentro de la predicción, pronóstico y prevención, eso implica que su construcción tiene que ver con el impacto y los efectos que se presentan al manipular las variables predictoras sobre la variable a predecir. Si se tiene dificultades para la formalización de una empresa, dentro del marco legal peruano, que a la fecha se tiene un 87 % de informalidad, ¿se podría predecir si el porcentaje de informalidad puede aumentar en los siguientes dos años? Se podría formular la pregunta predictiva de la siguiente manera:

¿Se predecirá que la informalidad de las empresas peruanas se irá incrementando producto de la normativa legal vigente?

¿La informalidad de las empresas peruanas tendrán un pronóstico negativo en los siguientes dos años por la normativa legal vigente?

¿Se preverá que se reducirá la informalidad de las empresas peruanas al modificar la normativa legal vigente?

En los tres subniveles predictivos se buscó formular preguntas que puedan ubicar en estos subniveles. Como existen pocos estudios de este grado en las ciencias sociales y existe poco desarrollo de estos tipos de indagaciones.

IV. El marco teórico predictivo

Todo marco teórico presenta tres elementos esenciales: antecedentes de estudio, los cimientos teóricos y el marco conceptual, que también se suele denominar como definición de las variables y sus dimensiones. El marco teórico en el campo predictivo, debe tener en consideración que se debe empezar declarando la teoría, que “son los artefactos conceptuales de una arquitectura bien construida” (Lavado, 2018, p.187) en la cual se va a sustentar el trabajo y sobre ella se va a elaborar las bases teóricas. En ese sentido, no se debe de olvidar que “El teórico que se interesa por la verdad, debe interesarse también por la falsedad, pues descubrir que un enunciado es falso equivale a descubrir que su negación es verdadera” (Popper, 2007, 30). Como son estudios que emplean el método deductivo, entonces su base son las teorías que, “están hechas de proposiciones ordenadas de manera deductiva” (Lavado, 2020, p. 68), principios que sean consistentes, pertinentes, suficientes y robustas. Los principios, las leyes y los conceptos les dan soporte a las teorías. En el ámbito de las leyes, se pueden tener teorías como la de Kelsen, teoría pura del derecho; Reale, teoría tridimensional del derecho; Alexy, teoría de los derechos humanos.

La definición de la terminología empleada, se debe sustentar en los conceptos de las variables y de sus dimensiones en la teoría que se ha elegido, para que se pueda esclarecer mediante estos conceptos la manera de medir los constructos correspondientes. Cada concepto definido debe estar sustentado en la teoría que se seleccionó, para lo cual es necesario citar la fuente de donde se obtuvo la información.

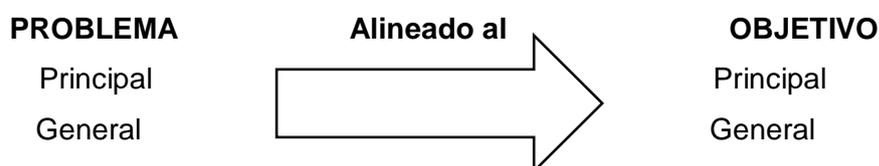
V. Los objetivos estadísticos predictivos

El objetivo general o primario o principal en las investigaciones predictivas que se deriva del propósito del estudio, se enmarca dentro de la predicción, pronóstico o prevención. Se debe tener en cuenta que estos objetivos deben estar alineados con el título y el problema general del estudio. Un alineamiento indica que, si el problema es predictivo, el objetivo también lo tiene que ser; no se puede plantear un problema de predicción y redactar un objetivo aplicativo o explicativo ya que no habría correspondencia con el propósito del estudio. Se debe recordar que el objetivo debe redactarse con

un verbo en infinitivo (ar, er, ir) y presenta los mismos componentes que el problema de investigación que se formuló como una interrogante. La base del desarrollo es el enunciado del estudio que se corresponde con el título de la investigación que debe alinearse con el problema principal, que a su vez debe de corresponder con el objetivo principal de la investigación.

Figura 1

Alineamiento problemas y objetivos



En las ciencias sociales, se le suele llamar problema general y objetivo general, sin embargo, también se le puede denominar problema principal y objetivo principal. La finalidad del estudio direcciona el alcance o nivel de indagación, se le suele denominar al problema general especificidad del estudio. Los problemas y objetivos que se derivan de los generales, se les asigna el término secundario, específico, auxiliar, intermedio o también operacional, ya que son fases previas que permiten el logro del objetivo general.

Como el nivel predictivo se encuentra entre los estudios básicos y cuantitativos, los objetivos que se utilizaran son estadísticos, para lo cual se muestra la siguiente tabla 4.

Tabla 4

Objetivos estadísticos predictivos

Nivel	Objetivos estadísticos
Predictivo	Prever
	Pronosticar
	Predecir

El objetivo estadístico predecir, indica la probabilidad de que se produzca un evento o suceso, en función de un modelo predictivo que previamente se aplicó. Esto es frecuente en la justicia peruana, debido a que, en el año 2019, el 39 % de los internos tenían mandato de prisión preventiva; ello significa que existe un 39 % de probabilidad de que al ciudadano se le pueda dictar

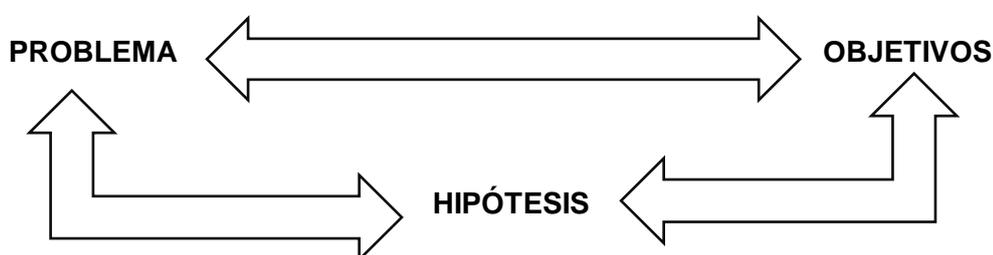
dicha medida de excepción. Por otra parte, el verbo pronosticar nos indica la probabilidad de que se produzca un evento en función del tiempo, en ese sentido, como van a confluír un conjunto de variables a predecir, se debe edificar un modelo predictivo que tenga el mínimo de errores o sesgos en las mediciones correspondientes; y el verbo prever, permite tomar las medidas necesarias para evitar cualquier siniestro que se pueda predecir y pronosticar, o tener en consideración las acciones que se deben realizar ante la probabilidad de que pase algún suceso favorable.

VI. Las hipótesis predictivas

Toda hipótesis es “una proposición o un sistema –conjunto ordenado de proposiciones que puede ser puesto a prueba, corroborado sólo indirectamente, o sea, a través de sus consecuencias” (Rodríguez, 2007, pp. 130-131) Una conjetura predictiva tiene que cumplir los siguientes criterios: las variables deben estar relacionadas fuertemente, tienen que haber causalidad entre las variables que implica que la variable independiente es la causa de la variable dependiente teniendo como sustento la experimentación con la correspondiente repetibilidad y reproducibilidad. Sobre este sustento y teniendo una teoría robusta, entonces se plantean los problemas y los objetivos para formular las hipótesis de carácter predictivo.

Figura 2

Vínculo problema, objetivos, hipótesis



Por otra parte, se pueden clasificar las hipótesis: a) de primer orden, las cuales muestran generalizaciones de carácter empírico, que las personas las observan de manera directa y que por ello se les debe someter a prueba para que se les acepte en el ámbito de la ciencia; b) de segundo orden, que expresan vínculos entre uniformidades, que señalan causalidad; c) de tercer

orden, que indican asociación entre sistemas de variables (Tafur & Izaguirre, 2015), las cuales se encuentran en el campo predictivo.

Sobre esta base, se deben de emplear dos tipos de deducciones, la primera que se sustenta en una sola premisa y que tiene como sostén la teoría que se ha seleccionado; y la deducción indirecta, que se sostiene en la teoría y presenta varias premisas, siendo esta más compleja. En ese sentido, lo importante de la deducción se manifiesta en: a) se fundamenta en premisas, hipótesis, teorías que se contrastan y se argumentan de forma indiscutible, prediciendo hechos buscando su aplicación en hechos concretos y singulares teniendo en cuenta la teoría elegida; b) cuando se confrontan las premisas, se deducen nuevas premisas que se desconocían lo que enriquece el grupo de las mismas y produciendo nuevos temas por investigar.

VII. Análisis de datos predictivos

Los análisis estadísticos que se realizan en las investigaciones predictivas llevan a un modelo matemático (Biecek, 2018) que permitirá realizar la predicción. Estos análisis son los siguientes:

- a. Predicción con regresión logística binaria, que se emplea cuando las variables predictoras y la a predecir son categóricas nominales y dicotómicas para estudios transversales. El resultado es un modelo que permitirá realizar la predicción.
- b. Predicción lineal, que se utiliza cuando las variables predictoras son numéricas u ordinales y la variable a predecir es numérica, para estudios transversales, también se le denomina método de los mínimos cuadrados. El modelo obtenido asocia las dos variables estudiadas.
- c. Pronóstico de tiempo de vida media, que permite calcular la probabilidad de ocurrencia y el pronóstico del tiempo media hasta que se produzca el evento.
- d. Análisis de supervivencia, donde se analizan los procesos aleatorios relacionados con el fallo de los sistemas mecánicos y la muerte de los organismos vivos. Se suele emplear el método actuarial, el método de Kaplan-Meier o el long-rank.

CONCLUSIONES

Primera: los estudios predictivos son cuantitativos, básicos, deductivos y se sustentan en una teoría.

Segunda: las investigaciones predictivas presentan variables predictoras y variables a predecir.

Tercera: los estudios predictivos desarrollan modelos que permiten realizar las predicciones, los pronósticos y prever cualquier contingencia.

Cuarta: En las ciencias sociales en general y en el derecho en particular se pueden realizar estudios predictivos.

REFERENCIAS

- Arbaiza, L. (2013). *Cómo elaborar una tesis de grado*. Lima, Perú: ESAN.
- Biecek, P. (2018). Dalex: Explaniers for Complex Predictive Models in R. *Journal of Machine Learning Research*(19), 1-5. Obtenido de <https://www.jmlr.org/papers/volume19/18-416/18-416.pdf>
- Bunge, M. (s.f.).
- Bunge, M. (1997). *La Causalidad. El principio de la causalidad en la ciencia moderna*. Buenos Aires, Argentina: Sudamericana.
- Bunge, M. (2019). *Investigación científica: enfoque, método y evaluación*. Lima, Perú: Grijley.
- Lavado, L. (2010). *Ciencia y Tecnología. Proceso de investigación*. Universidad Inca Garcilaso de la Vega, Fondo Editorial.
- Lavado, L. (2018). *Métodos de Investigación en ciencias sociales*. Lima, Perú: Grijley.
- Lavado, L. (2020). *Epistemología e investigación*. Fondo Editorial Universidad Nacional Enrique Guzmán y Valle.
- Popper, K. (2007). *Conocimiento objetivo* (5a. ed. ed.). Madrid: Tecnos.
- Rodríguez Rivas, M. Á. (2007). *Proceso de la Ciencia. Operaciones básicas: problemas e hipótesis*. Lima: Universidad Inca Garcilaso de la Vega.
- Shmueli, G., & Koppius, O. (2011). Predictive Analytics in Information Systems Research. *Mis Quarterly*, 35(3), 553-572. Obtenido de <https://www.jstor.org/stable/23042796>
- Supo, J. (2015). *Cómo empezar una tesis*. Paucarpata, Arequipa, Perú: Bioestadístico EIRL.
- Tafur, R., & Izaguirre, M. (2015). *Cómo hacer un proyecto de investigación*. Bogotá, Colombia: Alfaomega Colombiana.