





LA DISTRIBUCIÓN GEOMÉTRICA CENSURADA APLICADA A DATOS DE POSTULANTES A LA UNIVERSIDAD NACIONAL AGRARIA LA MOLINA.

Geometric censored distribution applied to data of applicants to the Universidad Nacional Agraria La Molina.

 Clodomiro Fernando Miranda Villagómez

 Jesús Eduardo Gamboa Unsihuay*

Universidad Nacional Agraria la Molina, Facultad de Economía y Planificación, Departamento de Estadística e Informática, Lima, Perú.

*jgamboa@lamolina.edu.pe

RESUMEN

Esta investigación aborda el uso de la distribución geométrica con datos censurados aleatoriamente. El conjunto de datos, proporcionado por la Universidad Nacional Agraria La Molina (UNALM), corresponde a postulantes durante los años 2018 al 2021, de quienes se registró si lograron el ingreso o no. Los datos se censuraron aleatoriamente y se consideró que una censura se presenta cuando el postulante no ingresó y no postuló más. Los parámetros que se estimaron fueron la probabilidad de que ingrese el postulante y la probabilidad de que desista en sus intentos, la cual vendría a ser la probabilidad de censura. Ambas cantidades fueron estimadas con el método de máxima verosimilitud y también los intervalos de confianza fueron obtenidos a partir la matriz de información de Fisher. Así, se encontró que la probabilidad de ingresar a la universidad es 0,117, y esta es mayor en hombres. Además, se estimó una probabilidad de censura igual a 0,296.

Palabras Clave: Admisión, censura aleatoria, probabilidad de ingreso, probabilidad de abandono, distribución discreta.

ABSTRACT

This research deals with the geometric distribution with randomly censored data. The data set, provided by the Universidad Nacional Agraria La Molina (UNALM), corresponds to applicants during the years 2018 to 2021, for whom it was recorded whether they achieved admission or not. Data was randomly censored, and it was considered that a censorship was presented when the applicant did not enter and did not apply again. The estimated parameters were the probability that the applicant would enter and the probability that he would give up his attempts, which would be the probability of censorship. Both quantities were estimated with the maximum likelihood method and the confidence intervals were obtained from the Fisher information matrix. Thus, it was found that the probability of entering the university is 0,117, and this is higher in men. In addition, a probability of censorship was estimated as 0,296.

Keywords: Admission, random censorship, probability of admission, probability of abandonment, discrete distribution.

Fecha de recepción: 20-12-2021

Fecha de aceptación: 10-02-2022

Fecha de publicación: 01-08-2022

I. INTRODUCCIÓN

Existen estudios de tiempo de vida que pueden tener un alto costo de tiempo y dinero, de modo que para no incurrir en ese alto costo se suele considerar una censura intencional o aleatoria. Hay varios tipos de censura: por la izquierda, por la derecha, tipo I, tipo II, y aleatoria. Se dice que se da por la izquierda cuando el suceso ya ocurrió antes ingresar al estudio, pero no se sabe hace cuánto tiempo, siendo poco frecuente este tipo de censura. Cuando se da por la derecha lo que se sabe es que, de haber ocurrido el evento, este se presentaría después del tiempo observado, siendo esta la situación de censura más común.

La censura tipo I se presenta cuando se ha fijado un tiempo máximo de observación de un evento, entonces los individuos que al término de este tiempo máximo de observación no hayan presentado el evento son considerados como observaciones censuradas. Por otro lado, la censura tipo II ocurre cuando el investigador decide aumentar la observación de los individuos hasta que ocurran k eventos de n posibles ($k \leq n$); una razón para determinar el número de eventos que deben observarse es la potencia que se requiere para el estudio, así los individuos que no presentan el evento cuando se han completado las k primeras observaciones, son considerados como censurados.

A diferencia de las censuras tipo I y tipo II, que son controladas por el investigador, la censura tipo III o aleatoria, ocurre sin su control (1-3), pudiendo presentarse por abandono del individuo, y en consecuencia se le pierde el seguimiento, o por muerte dada alguna causa no relacionada al evento de interés; esto último aplica para estudios de supervivencia. "La censura aleatoria es una situación en la que un elemento en estudio se pierde o se elimina aleatoriamente del experimento antes de que falle. En otras palabras, algunos sujetos del estudio no han experimentado el evento de interés al final del estudio. En un estudio social, algunos sujetos se pierden durante el seguimiento en la mitad de la encuesta" (4).

Además, "en los experimentos de la vida real, nos encontramos con situaciones en las que los datos del tiempo de falla son discretos, ya sea a través de la agrupación de datos continuos debido a una medición imprecisa o porque el tiempo en sí es discreto, por ejemplo, días, semanas o meses. En tales circunstancias, se mide la vida útil de un dispositivo en una escala discreta. Un modelo de vida útil discreto también puede considerar el

número de ciclos, pruebas u operaciones exitosas antes de la falla de un dispositivo. En los modelos de vida útil discretos, la distribución geométrica de un parámetro tiene una posición importante" (5).

La variable aleatoria discreta (v.a.d.) Geométrica X se define como el número de fracasos previos al primer éxito en una secuencia de pruebas de Bernoulli donde $\pi = P(\text{"Éxito"})$ y $1 - \pi = P(\text{"Fracaso"})$ (6-7), además la función de probabilidad de la v.a.d. es:

$$f(x) = \pi(1 - \pi)^x I_{\{0,1,2,\dots\}}(x) \quad (1)$$

Así, puesto que $X \sim \text{Geométrica}(\pi)$ entonces:

$$E(X) = \frac{1 - \pi}{\pi}, \text{Var}(X) = \frac{1 - \pi}{\pi^2} \text{ y } \Psi_x(t) = \frac{\pi}{1 - (1 - \pi)e^t} \quad (2)$$

Además, la distribución acumulada de la variable geométrica viene dada por:

$$P(X \leq x) = \sum_{i=0}^x (1 - \pi)^i \pi = \pi \sum_{i=0}^x (1 - \pi)^i = \pi \left[\frac{1 - (1 - \pi)^{x+1}}{1 - (1 - \pi)} \right] = 1 - (1 - \pi)^{x+1}, x = 0, 1, 2, 3, \dots \quad (3)$$

Se pueden citar investigaciones que han hecho uso de la distribución geométrica censurada, por ejemplo abordando la estimación de los parámetros de la distribución geométrica con censura tipo I, mediante inferencia clásica (8), o a través del desarrollo de modelos de regresión para datos de conteo, incluyendo censuras en las distribuciones Poisson y Binomial Negativa, siendo la geométrica un caso particular de esta última, a fin de estimar el número de años que transcurren desde que se lanza una tecnología y esta es adoptada por las personas (9).

Desde otro ángulo, se utiliza una mixtura de distribuciones geométricas censuradas a la derecha para modelar distancias en datos de genomas (10). Una aplicación distinta de la distribución geométrica censurada consiste en emplearla para estimaciones en modelos de captura – recaptura (11). También existe el modelo de regresión binomial negativa (que es el caso general de la distribución geométrica) Hurdle con datos de conteo censurados a la derecha (12).

Por su parte, la aplicación de la distribución geométrica censurada en registros históricos de exámenes de ingreso a la universidad no ha sido abordada. Se ha reportado que la participación continua en exámenes de ingreso ejerce una presión considerable sobre los postulantes de bajo rendimiento, lo que influye en sus estilos de

aprendizaje y futuro perfil profesional (13). Desde un punto de vista psicológico, ha sido señalado que, entre quienes reprobaron el examen de ingreso, presentaron mayor satisfacción en sus vidas aquellos que desarrollaron un control de compensación secundaria, dicho de otro modo, desistir un objetivo puede ser más adaptativo que esforzarse de forma persistente (14). Por otro lado, se ha encontrado que a medida que se rinden más exámenes de admisión, el uso de estrategias de evitación es menos frecuente (15).

La cantidad de postulaciones está asociada directamente con la resiliencia e inteligencia emocional del postulante (16), por lo que conocer la dinámica del número de postulaciones es de utilidad para ofrecer un servicio de orientación psicopedagógica en las charlas vocacionales de carreras universitarias, a fin de no generar desmotivación, confusión, baja autoestima y, en consecuencia, un exiguo rendimiento académico. Además, la preparación para el ingreso a la universidad conlleva un costo en tiempo y dinero: el 37% de los postulantes a una universidad pública se preparó más de un año para postular a la universidad, siendo este porcentaje superior en varones, y de manera particular, el tiempo de preparación para la UNALM es uno de los más altos, pues el 57% lo realiza durante más de un año, con un gasto mensual promedio que fluctúa entre 1000 y 2000 soles (17).

Esta situación debe motivar a las autoridades universitarias, de modo que aprovechen la potencial fortaleza de resiliencia e inteligencia emocional en los nuevos estudiantes y las refuercen en aquellos nuevos estudiantes que lo requieran, siendo esta una manera de retribuir los esfuerzos por el ingreso a la universidad. Esta situación cobra mayor relevancia en el escenario de la pandemia por COVID-19, pues el aislamiento social y la soledad han causado un deterioro en la salud mental de niños y adolescentes (18), siendo relevante este último grupo en los exámenes de admisión de la UNALM pues representa al menos el 30% del total de postulantes.

En este contexto, se presenta una aplicación de la distribución geométrica con datos censurados aleatoriamente al histórico de postulación e ingreso de los aspirantes a la Universidad Nacional Agraria La Molina (UNALM) en los ciclos 2018-2 al 2021-1. Cabe resaltar que un postulante en los ciclos mencionados pudo haber intentado el ingreso en períodos previos como el 2018-1 o durante el 2017. Se considera que un dato está censurado aleatoriamente cuando la persona

desistió de postular y por ello no logró el ingreso a la universidad. El ingreso a la universidad se da entre quienes aprueban un examen de admisión y alcanzan una vacante por orden de mérito.

El objetivo de esta investigación consiste en estimar de forma puntual y por intervalos la probabilidad de ingreso a la universidad, así como la probabilidad de censura, utilizando la distribución geométrica, en postulantes a la UNALM en los ciclos 2018-2 al 2021-1. El método de estimación utilizado es el de máxima verosimilitud (19-21).

II. MATERIALES Y MÉTODOS

La población de postulantes a la UNALM durante los ciclos 2018-2 al 2021-1 ascendió a $N=9675$. Fueron excluidos los registros de ingreso directo por el Centro Pre Universitario. A fin de realizar la estimación de parámetros de las distribuciones se extrajo una muestra de tamaño $n=4820$, compuesta por 2505 mujeres y 2315 hombres, de modo que la proporción por sexo se mantiene respecto a la población. Este tamaño resultó de considerar una probabilidad de ingreso conservadora ($p=0.5$), margen de error del 1% y nivel de confianza del 95%. Asimismo, puesto que se conoce el tamaño de la población, se aplica el ajuste para poblaciones finitas (22).

Para cada uno de los registros en la muestra, se examinó en retrospectiva si había postulado e ingresado en alguno de los ciclos considerados (2017-1 al 2021-1, excepto 2020-2 pues no se llevó a cabo examen de admisión debido a la pandemia por COVID-19). Una muestra de cinco datos se presenta en la tabla 1, a partir de la cual se puede apreciar que los registros 1 y 4 presentan censura aleatoria, ya que los postulantes no ingresaron y desistieron luego del primer y tercer intento, respectivamente. Sucede lo contrario con los registros 2, 3 y 5, puesto que han postulado hasta lograr el ingreso a la universidad. Finalmente, las celdas en blanco implican que el postulante no se inscribió en el examen de admisión para ese ciclo.

Se considera que las variables aleatorias X_1, \dots, X_n son los tiempos discretos de falla o de supervivencia, que en este caso viene a ser el número de veces que la persona postula antes de ingresar. Estas variables aleatorias son independientes e idénticamente distribuidas (i.i.d.) según una Geométrica (θ), siendo θ la probabilidad de que suceda el evento de interés (que el postulante ingrese a la UNALM). Además, los postulantes pueden ser censurados en los momentos T_1, \dots, T_n , es decir se registra el número de veces que la persona postula sin lograr

el ingreso. Así, la variable está censurada pues no se obtiene una realización del evento de interés (el ingreso), el cual podría ocurrir en un tiempo futuro, mayor a T_i , denotado como T_i^+ . También se asume que las T_i son variables aleatorias independientes e igualmente distribuidas i.i.d. según una Geométrica (λ), siendo λ la probabilidad de que suceda el evento de censura de un postulante. Otro supuesto es que las X_i y las T_i son independientes. En la muestra de ejemplo de la tabla 1, los datos observados se representan de la siguiente manera: 1+, 4, 2, 3+, 0.

Como el experimento que se está considerando es censurado en forma aleatoria, será de interés observar el $\min(X_i, T_i)$ para $i=1,2,\dots,n$. Con la información anterior se puede definir la variable aleatoria Bernoulli $D_i=1(0)$ si $X_i < T_i$ ($X_i \geq T_i$) con probabilidad $p=P[X_i < T_i]$, es decir toma el valor de 1 si no se da la censura. Así, se puede definir la v.a.d. $D_i \sim \text{Bernoulli}[P(X_i < T_i)]$, donde el parámetro de probabilidad de éxito en esta distribución está dado por $P(X_i < T_i) = \theta(1-\lambda)/(1-(1-\theta)(1-\lambda))$, por lo tanto $D_i \sim \text{Bernoulli}(\theta(1-\lambda)/(1-(1-\theta)(1-\lambda)))$ se expresa como en (4).

$$f_D(d) = \left[\frac{\theta(1-\lambda)}{1-(1-\theta)(1-\lambda)} \right]^d \left[\frac{1-(1-\lambda)(1-\theta+\theta)}{1-(1-\theta)(1-\lambda)} \right]^{1-d} = \frac{\theta^d (1-\lambda)^d \lambda^{1-d}}{1-(1-\theta)(1-\lambda)} \quad (4)$$

Nº	SEXO	2017-1	2017-2	2018-1	2018-2	2019-1	2019-2	2020-1	2021-1
1	MASCULINO				NO				
2	FEMENINO		NO	NO	NO	NO	SÍ		
3	FEMENINO					NO		NO	SÍ
4	FEMENINO			NO	NO		NO		
5	MASCULINO				SÍ				

Tabla 1. Estructura de datos (muestra aleatoria) indicando si el postulante logró o no el ingreso.

Por otro lado, se obtiene la distribución del valor observado $Y = \min(X_i, T_i)$, de modo que $Y_i \sim \text{Geométrica}[1-(1-\theta)(1-\lambda)]$, ya que:

$$\begin{aligned} P(Y_i \leq y_i) &= 1 - P(Y_i > y_i) = 1 - P(X_i > y_i)P(T_i > y_i) = \\ &= 1 - [1 - P(X_i \leq y_i)][1 - P(T_i \leq y_i)] \\ &= 1 - (1-\theta)^{y_i+1} (1-\lambda)^{y_i+1} = \\ &= 1 - [(1-\theta)(1-\lambda)]^{y_i+1}, y_i = 0, 1, 2, 3, \dots \end{aligned} \quad (5)$$

Luego, las variables aleatorias Y y D son independientes (dado que X y T lo son), entonces:

$$\begin{aligned} f_{Y,D}(y_i, d_i / \theta, \lambda) &= f_Y(y_i / \theta, \lambda) f_D(d_i / \theta, \lambda) \\ &= [1-(1-\theta)(1-\lambda)] [(1-\theta)(1-\lambda)]^{y_i} \frac{\theta^{d_i} (1-\lambda)^{d_i} \lambda^{1-d_i}}{1-(1-\theta)(1-\lambda)} \\ &= \theta^{d_i} (1-\theta)^{y_i} \lambda^{1-d_i} (1-\lambda)^{y_i+d_i}, \\ &= y_i = 0, 1, 2, 3, \dots; d_i = 0, 1; 0 < \theta < 1; 0 < \lambda < 1 \end{aligned} \quad (6)$$

Para encontrar los estimadores máximo-verosímiles (EMV) $\hat{\theta}_{MV}$ y $\hat{\lambda}_{MV}$, se define en (7) la función de verosimilitud y en (8) la log-verosimilitud, a partir de la cual se obtienen los estimadores máximo-verosímiles, así como sus varianzas estimadas en (9) y (10):

$$\begin{aligned} L(\theta, \lambda; \underline{Y}, \underline{D}) &= \prod_{i=1}^n \theta^{d_i} (1-\theta)^{y_i} \lambda^{1-d_i} (1-\lambda)^{y_i+d_i} = \\ &= \theta^{\sum_{i=1}^n d_i} (1-\theta)^{\sum_{i=1}^n y_i} \lambda^{n - \sum_{i=1}^n d_i} (1-\lambda)^{\sum_{i=1}^n y_i + \sum_{i=1}^n d_i} \end{aligned} \quad (7)$$

$$\begin{aligned} \log[L(\theta, \lambda; \underline{Y}, \underline{D})] &= \sum_{i=1}^n d_i \times \log(\theta) + \sum_{i=1}^n y_i \times \log(1-\theta) + \\ &+ \left[n - \sum_{i=1}^n d_i \right] \times \log(\lambda) + \left[\sum_{i=1}^n y_i + \sum_{i=1}^n d_i \right] \times \log(1-\lambda) \end{aligned} \quad (8)$$

$$\hat{\theta}_{MV} = \frac{\sum_{i=1}^n d_i}{\sum_{i=1}^n d_i + \sum_{i=1}^n y_i} \rightarrow V(\hat{\theta}_{MV}) = \frac{\hat{\theta}_{MV} (1-\hat{\theta}_{MV}) [1-(1-\hat{\theta}_{MV})(1-\hat{\lambda}_{MV})]}{n(1-\hat{\lambda}_{MV})} \quad (9)$$

$$\hat{\lambda}_{MV} = \frac{n - \sum_{i=1}^n d_i}{n + \sum_{i=1}^n y_i} \rightarrow V(\hat{\lambda}_{MV}) = \frac{\hat{\lambda}_{MV} (1-\hat{\lambda}_{MV}) [1-(1-\hat{\theta}_{MV})(1-\hat{\lambda}_{MV})]}{n} \quad (10)$$

El hecho de que la distribución geométrica pertenezca a la familia exponencial permite que la distribución muestral de los estimadores θ y λ sea asintóticamente normal (4). Finalmente, los estimadores puntuales y sus respectivos errores estándar fueron obtenidos a partir de los registros reportados según la estructura de la tabla 1. Para el preprocesamiento y análisis de datos, se utilizaron los programas R 4.1.1 y RStudio 1.4.1717 y los paquetes readxl, dplyr, tidyr y purrr.

III. RESULTADOS

Al realizar un análisis exploratorio, la tabla 2 muestra que solo entre el 12% y 19% de los postulantes alcanza una vacante por examen de admisión. Por otro lado, el 48% de los postulantes fue de sexo masculino y el 52% de sexo femenino, no obstante, estos porcentajes se invierten y se acercan al 50% en el grupo de ingresantes, ya que

el 48,9% fueron mujeres y 51,1% fueron hombres.

La figura 1 muestra la distribución de postulantes considerando si lograron o no el ingreso, y la cantidad de intentos. Entre aquellos que han logrado ingresar a la universidad, el 20% lo hizo postulando una vez, un 28.3% postuló dos veces, un 27,7% lo hizo tres veces, y el 24% restante, realizó entre cuatro y siete postulaciones. Por otro lado, entre quienes no han ingresado a la universidad, la mayoría (63,2%) sólo realizó un intento, luego del cual desistió. Al agrupar por cantidad de postulaciones, la proporción de ingresantes en la primera postulación es la más baja (8.1%), y va ascendiendo a medida que se incrementan los intentos, de modo que se alcanza una tasa de ingreso cercana al 50% para quienes postularon al menos 3 veces.

Ciclo	2018-2	2019-1	2019-2	2020-1	2021-1
Postulantes	1449	1551	1562	1630	1238
Número de Ingresantes	199	190	233	199	231
% de ingresantes	13,7%	12,3%	14,9%	12,2%	18,7%

Tabla 2. Evolución de los ingresantes por ciclo.

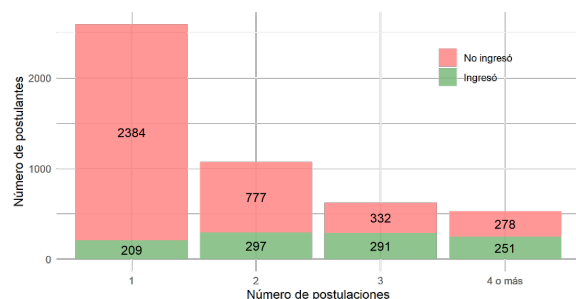


Figura 1. Distribución observada de postulantes.

La tabla 3 muestra los estimadores máximo-verosímiles de θ y λ , así como sus errores estándar, de manera global y para cada sexo. Se aprecia que la probabilidad de ingresar a la universidad es 0.117, y al desagregar por sexos es mayor en hombres pues $IC_{hombre}(\theta)=(0,117;0,137)$ mientras que $IC_{mujer}(\theta)=(0,099;0,117)$, en ambos casos al 95% de confianza. Esto permite señalar que el número medio de postulaciones antes del ingreso es de 7.56 de manera general, 6.9 para varones y 8.3 para mujeres, mientras que la mediana es de 5 postulaciones previas al ingreso tanto en hombres como mujeres. La distribución de probabilidades del número de postulaciones antes de conseguir el ingreso se aprecia en la figura 2.

Por otro lado, la probabilidad de que un registro esté censurado, es decir de que el aspirante desista de seguir postulando a la UNALM, es 0,296 con un intervalo del 95% de confianza que va de 0,288 a 0,304, y una estimación similar para cada uno de los sexos. Así, el número medio de postulaciones antes de desistir es 2,38, su mediana es igual a 1, y en la figura 2 presenta la distribución de probabilidades del número de veces que un postulante intenta ingresar antes de desistir.

Grupo	EMV de θ (e.e.)	EMV de λ (e.e.)
Global	0,117 (0,0034)	0,296 (0,0040)
Mujeres	0,108 (0,0045)	0,295 (0,0056)
Hombres	0,127 (0,0051)	0,296 (0,0059)

Tabla 3. Estimación de parámetros.

EMV: estimadores máximo-verosímiles.

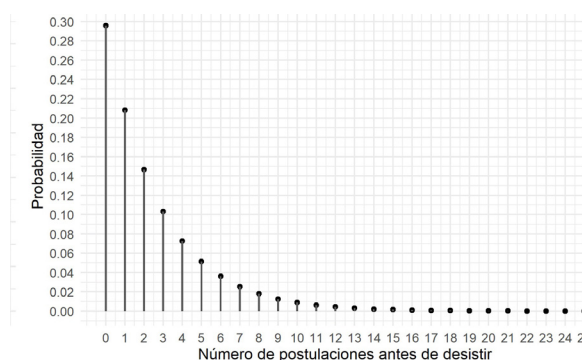
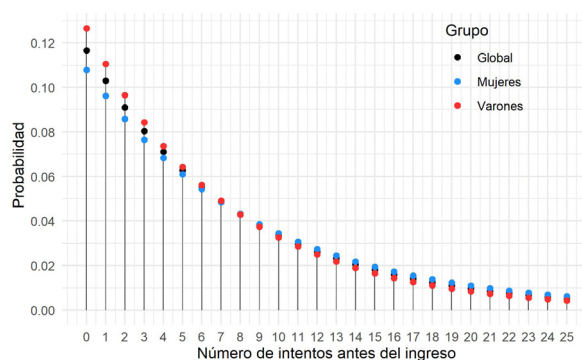


Figura 2. Distribución estimada del número de postulaciones antes del ingreso y antes de desistir

IV. DISCUSIÓN

En la presente investigación, para el 53.8% de los postulantes era la primera vez que intentaba ingresar a la UNALM, mientras que el 76.1% postuló entre una y dos veces. En la caracterización los postulantes de la Universidad Católica Santa María (Arequipa, Perú) estos porcentajes

fueron 26.25% y 49.25%, respectivamente (23), mostrando así que la tasa de abandono en las postulaciones es más alta en la UNALM, ya que solo 1 de cada 4 postuló más de 2 veces. Por otro lado, se evidenció que el 73.2% de los ingresantes a la escuela profesional de Enfermería de la Universidad Señor de Sipán lo hizo al primer intento (24). En nuestra investigación, ese porcentaje es de apenas 20%, poniendo de manifiesto que la dificultad de ingreso es mayor.

Por otro lado, si bien es cierto que los datos utilizados en esta investigación pertenecen a un periodo determinado en la UNALM (2018 al 2021), es factible generalizar a más periodos, ya que los resultados obtenidos permiten a las autoridades universitarias conocer más sobre la realidad de los postulantes. Además, es posible considerar un trabajo futuro desagregando el análisis por carreras, a fin de tomar una decisión basada en datos respecto a la asignación de vacantes, o en general, establecer un modelo de regresión para datos de conteo con censura, donde la cantidad de postulaciones antes del ingreso sería la variable respuesta, y no una atributo predictivo como en (25), mientras que distintos atributos demográficos, socioeconómicos y/o educativos cumplirían el papel de variables predictoras. Asimismo, es factible que a futuro se considere agregar un truncamiento en cero intentos antes de desistir, así como un truncamiento superior

para una cantidad máxima de postulaciones, por definirse.

V. CONCLUSIONES

1. Entre los ingresantes a la universidad, el 20% solo postuló una vez, un 28,3% postuló dos veces, un 27,7% lo hizo tres veces, y el 24% restante, realizó al menos cuatro postulaciones. Por otro lado, entre quienes no han ingresado a la universidad, la mayoría sólo realizó un intento, luego del cual desistió.
2. La proporción de ingresantes en el primer intento es la más baja (8,1%), y va ascendiendo a medida que se incrementan los intentos, de modo que este porcentaje es cercano al 50% para quienes postularon al menos tres veces.
3. En cada examen de admisión, al menos 4 de cada 5 postulantes no alcanza una vacante. Así, se estimó que la probabilidad global de ingresar a la universidad es 0.117, y al desagregar por sexos es mayor en hombres.
4. La probabilidad de que un postulante desista de seguir postulando (censura) es 0,296, con una estimación similar para cada uno de los sexos. Además, el número medio de postulaciones antes de desistir es 2,38 y su mediana es 1.

VI. REFERENCIAS

1. Caizana Martínez I. Modelo de riesgos proporcionales log-lineal [Tesis para optar por el Título de Licenciatura en Estadística]. La Paz: Universidad Mayor de San Andrés; 2012. Disponible en: <https://repositorio.umsa.bo/bitstream/handle/123456789/5331/T-069.pdf>
2. Flores Flores C. Modelo de regresión de Cox usando splines [Tesis para optar por el Título Profesional de Licenciado en Estadística]. Lima: Universidad Nacional Mayor de San Marcos; 2011. Disponible en: https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/211/Flores_fc.pdf
3. Uriarte F. Determinantes del riesgo de incumplimiento: una aplicación al mercado de préstamos de consumo no revolventes [Internet]; 2016. Disponible en: https://www.sbs.gob.pe/Portals/0/jer/ddt_ano2016/20160916_SBS-DT-001-SBS_FUriarte.pdf
4. Krishna H, Goel N. Maximum Likelihood and Bayes Estimation in Randomly Censored Geometric Distribution. *Journal of Probability and Statistics*. 2017. doi: <https://doi.org/10.1155/2017/4860167>
5. Krishna H, Goel N. Randomly censored geometric distribution under koziol-green model. *International Journal of Agricultural and Statistical Sciences*. 2017; 85-95. <https://www.researchgate.net/publication/318926215>
6. James B. Probabilidad: Un curso de nivel intermedio. Lima: Instituto de Matemática y Ciencias Afines; 2004.
7. Miranda Villagómez CF. Cálculo de probabilidades con gráficos en R. 1a ed. Lima: Universidad Nacional

- Agraria La Molina; 2017.
8. Gan G, Bain, LJ. Some results for type I censored sampling from geometric distributions. *Journal of Statistical Planning and Inference*. 1998. doi: [https://doi.org/10.1016/S0378-3758\(97\)00096-7](https://doi.org/10.1016/S0378-3758(97)00096-7)
 9. McWilliams B, Tsur Y, Hochman E, Zilberman D. Count-data regression models of the time to adopt new technologies. *Applied Economics Letters*. 1998; 5(6): 369-373. doi: <https://doi.org/10.1080/135048598354744>
 10. Biswas B, Lai Y. A distance-type measure approach to the analysis of copy number variation in DNA sequencing data. *BMC Genomics*. 2019; 20(2): 195. doi: <https://doi.org/10.1186/s12864-019-5491-x>
 11. Niwitpong S, Böhning D, van der Heijden PGM, Holling, H. Capture–recapture estimation based upon the geometric distribution allowing for heterogeneity. *Metrika*. 2013; 76(4): 495-519. doi: <https://doi.org/10.1007/s00184-012-0401-0>
 12. Saffari S, Adnan R, Green W. Hurdle negative binomial regression model with right censored count data. *SORT*. 2012; 36 (2): 181-194.
 13. Seddigh R, Abdollahpour E, Azarnik S, Shariati B, Keshavarz-Akhlaghi, AA. Participating in university entrance exam despite repeated failure: A qualitative study of participants' experiences. *International Journal of Medical Education*. 2016; 7:345-353. doi: <https://doi.org/10.5116/ijme.57eb.cc09>
 14. Tomasik MJ, Salmela-Aro K. Knowing when to let go at the entrance to university: Beneficial effects of compensatory secondary control after failure. *Motivation and Emotion*. 2012; 36(2):170-179. doi: <https://doi.org/10.1007/s11031-011-9246-5>
 15. Vásquez Wilson JA. Regulación emocional y estrategias de afrontamiento en alumnos preuniversitarios de una academia de Lima Metropolitana [Tesis para optar por el Título de Licenciado en Psicología]. Lima: Pontificia Universidad Católica del Perú; 2020. Disponible en: <https://tesis.pucp.edu.pe/repositorio/handle/20.500.12404/17743>
 16. Roque Caqui MK. Relación entre resiliencia e inteligencia emocional en estudiantes de un centro pre-universitario privado de Lima Metropolitana [Tesis para optar por el título de licenciado en psicología]. Lima: Universidad Peruana Cayetano Heredia. 2020. Disponible en: https://repositorio.upch.edu.pe/bitstream/handle/20.500.12866/8553/Relacion_RoqueCaqui_Maribel.pdf
 17. Ministerio de Educación del Perú. Encuesta Nacional de Estudiantes de Educación Superior Universitaria 2019 [Internet]. Repositorio MINEDU; 2021. Disponible en: <https://repositorio.minedu.gob.pe/handle/20.500.12799/7745>
 18. Loades ME, Chatburn E, Higson-Sweeney N, Reynolds S, Shafran R, Brigden A, et al. Rapid Systematic Review: The Impact of Social Isolation and Loneliness on the Mental Health of Children and Adolescents in the Context of COVID-19. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2020 noviembre 1; 59(11):1218-1239. doi: <https://doi.org/10.1016/j.jaac.2020.05.009>
 19. Casella G, Berger R. *Statistical Inference*. 2a ed. Estados Unidos: Thompson Learning; 2002.
 20. Miranda Villagómez CF. Un curso de inferencia estadística. 1a ed. Lima: Universidad Nacional Agraria La Molina; 2020.
 21. Mood F, Graybill F, Boes D. *Introduction to the theory of statistics*. 3a ed. Estados Unidos: Mc Graw Hill; 1974.
 22. Doane D, Seward L. *Applied Statistics in Business & Economy*. 3a ed. New York: McGraw-Hill; 2011.
 23. Salazar Caracela IK. Características Generales y Niveles de Ansiedad en Postulantes a la Universidad Católica de Santa María, Precatólica Periodo 2017 [Tesis para obtener el Título Profesional de Médico Cirujano]. Arequipa: Universidad Católica de Santa María; 2017. Disponible en: <http://tesis.ucsm.edu.pe/repositorio/handle/UCSM/6542>
 24. Tucto Coronado JA. Características del perfil de ingreso del estudiante de la profesión de enfermería en la Universidad Señor de Sipán, 2019 [Tesis para optar por el Título Profesional de Licenciada en Enfermería]. Pimentel: Universidad Señor de Sipán; 2019. Disponible en: <http://repositorio.uss.edu.pe/handle/20.500.12802/6506>
 25. Moína Fuentes JR. Factores importantes que determinan el ingreso a la Universidad Nacional Mayor de San Marcos, 2003. [Monografía para optar por el Título de Licenciado en Estadística]. Lima: Universidad Nacional Mayor de San Marcos; 2003. Disponible en: <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/12779>