

MODELOS ESTADÍSTICOS PARA EL ESTUDIO DE VARIABLES QUE PUEDEN INCIDIR EN LA PROBABILIDAD DE OBTENER EMPLEO EN ECUADOR

García-Bustos Sandra¹, Plaza Andrés², Mite Mónica³.

Resumen: El propósito de este trabajo es conocer que tan influyentes son ciertos factores demográficos en la obtención de empleo para el ciudadano ecuatoriano cuya edad está comprendida entre los 18 y 70 años. Para tal efecto se consideran ciertas variables y características de la persona, que influyen de manera positiva o negativa en su inserción en el mercado laboral (tener o no empleo). Con la ayuda del software denominado SPSS (Statistical Package for the Social Sciences) y a través de técnicas estadísticas, tales como, Regresión Logística y Árboles de decisión, se determinó que el género, la edad y nivel de instrucción de la persona son características influyentes al momento de participar en el mercado laboral ecuatoriano.

Palabras Claves: Participación laboral, Regresión Logística, Árboles de Decisión.

Abstract: The purpose of this work is to know how influential are certain demographic factors in obtaining employment for Ecuadorians whose age is between 18 and 70 years old. For this purpose, we considered certain variables and characteristics from the person that can affect positively or negatively their insertion on the laboral force. With the help of SPSS (software) and statistical techniques such as Logistic Regression and Decision Trees was determined that gender, age and education of the person can affect at the moment of entering the Ecuadorian market

Keywords: Market, Logistic Regression, Decision Trees.

Recibido: Junio 2016

Aceptado: Septiembre 2016

1. INTRODUCCIÓN

El motor de un país es su fuerza laboral, por ello es importante estudiar las características que influyen en la probabilidad de pertenecer a ella. Tanto las mujeres como los hombres han participado en el desarrollo socioeconómico del Ecuador por distintas vías, y debido a que las condiciones económicas, sociales y culturales han cambiado, el mercado laboral se está volviendo cada vez más competitivo, por lo que los aspirantes a un empleo han tenido que adaptarse a esas nuevas condiciones. Por lo antes expuesto es necesario analizar cuáles son los factores que inciden para tener éxito en la búsqueda de empleo en el Ecuador. Algunas investigaciones relacionadas se mencionan a continuación:

García y Cortez¹ han realizado un estudio acerca del análisis de la participación laboral de la mujer en el mercado ecuatoriano, utilizando como metodologías el modelo ocio-consumo y regresión logística. En esta investigación se analiza el comportamiento de la probabilidad de participación laboral femenina en las diferentes provincias del territorio continental ecuatoriano durante el periodo comprendido entre los años 1990 y 2011 y se llega a la conclusión que el progreso alcanzado por las mujeres no es

uniforme y que depende de la edad, la educación y las características familiares.

La probabilidad de encontrar empleo se reduce cuando en su núcleo familiar se cuenta con ingresos económicos distintos de los que ella aporta.

Otro estudio realizado por Laura de Escobar² acerca de la participación laboral femenina y el análisis de algunos factores que condicionan la entrada de las mujeres en el mercado laboral, analiza si las medidas públicas tomadas en España ayudaban a conseguir una mayor incorporación de la mujer a la vida laboral, y la conclusión obtenida fue que al sexo femenino se le aduce la necesidad de realizar tareas domésticas y cuidados de personas dependientes, abandonando sus puestos de trabajo por el hecho de contraer matrimonio, incluso si tienen un nivel alto de educación dedican un número elevado de horas a la realización de trabajos no remunerados.

Por otro lado, una de las técnicas de regresión más utilizada actualmente para predecir una probabilidad de obtener éxito es la Regresión Logística. Bartlett³, en 1937 usó la transformación $\log[y/(1-y)]$ para estudiar proporciones. Al año siguiente, Fisher y Yates⁴ también consideraron el uso de esa transformación para analizar datos binarios. Sin embargo, fue en 1944 cuando recién se empezó a usar el término *logit* por Berkson⁵ para referirse a esta transformación, y fue uno de los autores que más utilizó la función logística, usando por primera vez la expresión *logit model* en uno de sus artículos. No obstante, en Estadística el principal difusor de la regresión

Centro de Estudios e Investigaciones Estadísticas, Facultad de Ciencias Naturales y Matemáticas, Escuela Superior Politécnica del Litoral, Campus Gustavo Galindo Velasco Km. 30.5 Vía Perimetral, Guayaquil, Ecuador email(s): ¹slgarcia@espol.edu.ec ²andplaz@espol.edu.ec;

Universidad de Guayaquil, e-mail: monica.mite@ug.edu.ec ³mitemonica@gmail.com

logística fue David R. Cox⁶ en 1970 con su libro *The Analysis of Binary Data*.

La introducción de los modelos lineales generalizados GLM, por los estadísticos británicos John Nelder y R.W.M. Wedderburn⁷, en 1972, teoría en la cual la media de toda variable que pertenezca a la familia exponencial puede ser explicada por variables independientes mediante un modelo de regresión. Entre las variables que pertenecen a esta familia se encuentran las variables Binomial y Bernoulli con lo cual, la regresión logística forma parte de estos modelos lineales generalizados. Además, el algoritmo que se suele usar para ajustar estos modelos lineales generalizados se denomina "Fisher scoring", el cual fue introducido por Fisher⁸ en 1935 para ajustar modelos probit de máxima verosimilitud.

Otra de las técnicas estadísticas de interés en éste estudio es aquella denominada árboles de decisión con el algoritmo de segmentación CHAID (Chi-Squared Automatic Interaction Detector) desarrollada por Cellard, Labbe y Cox⁹ (1967); Bouroche y Tennenhaus¹⁰ (1972); Kass¹¹ (1980) y Magdison¹² (1989) quien finalmente lo adaptó para el programa computacional SPSS, tiene como principal característica que se pueden formar segmentos con más de dos categorías al mismo tiempo. Aunque el algoritmo CHAID es el más conocido y usado, existen otros tipos de segmentaciones como el C&RT (Classification and Regression Trees) de Breiman, Friedman, Olshen y Stone¹³, (1984), el CHAID Exhaustivo de Biggs, De Ville y Suen¹⁴ (1991) y el QUEST (Quick, Unbiased Efficient Statistical Tree) de Loh y Shih¹⁵ (1997). Todas ellas se encuentran programadas en sistemas informáticos como SPSS, SAS (Statistical Analysis System), entre otros.

En este trabajo, se utilizan las dos técnicas antes expuestas para entender qué factores determinan o influyen en la probabilidad de obtener empleo en el Ecuador, para cumplir con este objetivo, este artículo está estructurado de la siguiente manera: La sección 1, presenta una breve introducción al tema. La sección 2, trata sobre el marco teórico de estas dos metodologías. La sección 3, describe los datos usados en el análisis. La sección 4, presenta los resultados relevantes y por último la sección 5, muestra las conclusiones.

2. MARCO TEÓRICO

Debido que el objetivo central de este trabajo está destinado a mostrar un conjunto de posibles causas por las que el ecuatoriano tiene o no empleo, es preciso estudiar las metodologías que utilizaremos en el análisis:

2.1. *Regresión logística*: El análisis de regresión logística es una técnica multivariada de uso explicativo y predictivo que permite modelar la presencia o ausencia de un determinado evento, el cual será explicado a través de un conjunto de variables independientes o explicativas X_i de tipo categórico o continuo, mientras que la variable de respuesta Y_i será una variable dicotómica o binaria, es decir con dos posibles opciones de respuesta opuestas entre sí (Sí-No, 0-1, Verdadero-Falso, etc.).

2.1.1 *Construcción del modelo*: El modelo de regresión clásica se expresa de la siguiente manera:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

Donde β_i , $i = 0, 1, \dots, n$, son los parámetros asociados al modelo; miden la influencia que tienen las variables explicativas sobre el problema planteado; β_0 es el término constante, adicionalmente, e es el error aleatorio con varianza constante, sin embargo como la variable dependiente es binaria, esto representa varios inconvenientes, por ejemplo:

- Es posible obtener predicciones fuera del intervalo (0,1)
- El término de perturbación e no cumple la hipótesis de homocedasticidad
- Cuando la variable de respuesta es binaria todos los valores de Y , se encontrarán en las horizontales 0 y 1, por lo que no se puede esperar un buen ajuste.

Para evitarlas, hay que transformar la variable respuesta a partir de una función f estrictamente creciente, para lo cual se utilizará el *modelo logit* de la forma:

$$Y = f(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n) + e$$

Donde f es la función logística, i.e.:

$$f(z) = \exp(z)/(1 + \exp(z))$$

Por tanto, tendremos que:

$$E[Y] = P(Y = 1) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)}$$

El método de máxima verosimilitud es usado para la estimación de los parámetros del modelo *logit*. En cuanto a la interpretación de los coeficientes beta, es necesario fijarse en los signos de los estimadores.

Si el signo del estimador es positivo, indicará que cuando la variable explicativa asociada a este coeficiente crece entonces la variable respuesta $P(Y = 1)$ también crece. Por otro lado, si el estimador del coeficiente presenta un signo negativo, cuando la variable explicativa asociada a este coeficiente aumenta entonces la variable respuesta disminuye.

En el modelo *logit* se suelen usar otros dos conceptos para la interpretación de los

estimadores: El primero se denomina odds, el cual es un cociente de probabilidades definido de la siguiente manera:

$$Odds = \frac{P(Y = 1)}{1 - P(Y = 1)}$$

$$Odds = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n)$$

Tomando logaritmos en la expresión anterior, obtenemos una expresión lineal para el modelo:

$$Logit[P(Y = 1)] = \ln\left(\frac{P(Y = 1)}{1 - P(Y = 1)}\right)$$

$$Logit[P(Y = 1)] = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n$$

Cuando nos referimos al incremento unitario en una de las variables explicativas, manteniendo las demás constantes, tenemos el segundo concepto de odds-ratio, que es el cociente entre los dos odds asociados (el obtenido tras realizar el incremento y el anterior al mismo). Así, la variación que sufre el odds-ratio cuando hay un incremento unitario en la variable X_i , está dada por esta expresión:

$$Odds\ ratio = \frac{odds2}{odds1} = \exp(\beta_i)$$

2.2 Test de Hosmer y Lemeshow: Esta prueba es muy usada en Regresión Logística para determinar la bondad de ajuste de un modelo, mediante la diferencia entre los valores observados y los valores esperados.

El Test es una constante de distribución, y consiste en dividir la distancia de los valores de la variable dependiente Y_i en una serie de intervalos que contengan un número de observaciones suficientemente grande (5 o más). La prueba evalúa si la tasa de los valores observados es igual o coinciden con la tasa de los valores esperados en los diferentes subgrupos formados del modelo propuesto. La prueba sigue asintóticamente una distribución ji-cuadrado.

2.3 Algoritmo de segmentación CHAID: Consiste en un algoritmo de árbol estadístico y multidireccional que estudia los datos para crear segmentos y perfiles con respecto a una variable respuesta que se desea estudiar. En cada paso, se elige la variable independiente (predictora) que presenta la relación más fuerte con la variable dependiente o respuesta. Las categorías de cada predictor se funden si no produce resultados estadísticamente distintos respecto a la variable dependiente.

Los árboles de decisión son utilizados para la estratificación, segmentación, la predicción, la reducción de datos y filtrado de variables, entre otras. La metodología CHAID realiza la segmentación en base al estadístico χ^2 , desarrollándose en su mayoría con el enfoque de Pearson:

$$\chi^2 = \sum \sum \frac{(f_{ij} - f_{ij}^2)^2}{(f_{ij}^2)}$$

Donde f_{ij} es la frecuencia observada y f_{ij}^2 es la frecuencia esperada si hay independencia.

3. DATOS

Los datos con la que se trabajó fueron tomados de la Encuesta de Empleo, Desempleo y Subempleo (ENEMDU) del año 2014, la cual es elaborada a nivel nacional por el Instituto Nacional de Estadística y Censos (INEC), de la cual sólo se consideró a la Población en edad de trabajar (PET), que comprende a todas las personas entre 18 y 70 años, estableciendo un tamaño de 34.558 registros. La variable empleo la cual registra si una persona trabaja o no, fue modificada a una variable binaria para fines prácticos de este trabajo.

A continuación se presentan las variables que serán parte de los distintos modelos a considerar:

3.1 Variable dependiente binaria (Y_i): La cual identifica la pertenencia de un individuo a una de las dos posibles categorías, identificando con el número 1 si el individuo pertenece a la característica de interés y 0 al elemento que no posee la característica de interés. De acuerdo al trabajo planteado, el caso 1 implica el tener empleo y 0 el caso contrario.

3.2 Variables independientes (X_i): son aquellas variables que permiten discriminar entre los grupos y que determinan la pertenencia de un elemento a un grupo u otro. Las variables independientes especificadas en este trabajo son:

3.2.1 Edad: Aquella que especifica la edad de la persona, con una edad mínima y máxima de 18 y 70 años respectivamente.

3.2.2 Sexo: Indica el sexo del individuo, esta variable fue codificada de la siguiente manera:

- 1 = Hombre
- 2 = Mujer

3.2.3 Estado civil: Se refiere a la Condición de una persona según el registro civil estas pueden ser:

- 1 = Casado
- 2 = Separado
- 3 = Soltero

3.2.4 Nivel de instrucción: Es el grado más elevado de estudios realizados o en curso, esta variable esta codificada como:

- 1 = Alfabetización
- 2 = Primaria
- 3 = Secundaria
- 4 = Superior
- 5 = Postgrado

4. RESULTADOS

Como método inicial se utilizó regresión logística para conocer la influencia que tiene la edad y el sexo de la persona en su inserción en

el mercado laboral, para ello se elaboró un primer modelo que contenga éstas variables como predictoras obteniendo el siguiente resultado:

$$\frac{\widehat{\text{Empleo}}}{1 - \widehat{\text{Empleo}}} = \exp(-0.464 + 0.017 * \text{Edad} + 1.630 * \text{Hombre})$$

TABLA 4.1

Coefficientes en el primer modelo con su respectivo valor p

	Coef.	Exp(Coef)	Valor p
Intercepto	-0.464	0.629	0.00
Edad	0.017	1.017	0.00
Hombre	1.630	5.102	0.00

En el cual puede interpretarse (Véase Tabla 4.1) que la edad y el sexo de la persona son significativas para tener empleo, en donde, por cada año de edad que se tenga, las posibilidades de participar en el mercado laboral aumentan muy poco, mientras que, si una persona pertenece al género masculino, la oportunidad de obtener empleo es mayor en 4.102 veces en relación a una mujer.

$$\frac{\widehat{\text{Empleo}}}{1 - \widehat{\text{Empleo}}} = \exp(0.561 - 0.007 * \text{Edad} + 1.386 * \text{Hombre} - 0.030 * \text{Separado} - 1.191 * \text{Soltero})$$

TABLA 4.2

Coefficientes en el segundo modelo con su respectivo valor p

	Coef.	Exp(Coef)	Valor p
Intercepto	0.561	1.752	0.00
Edad	-0.007	0.993	0.00
Hombre	1.386	3.998	0.00
Separado	-0.030	0.971	0.39
Soltero	-1.191	0.304	0.00

El estado civil en la que se encuentra el individuo, puede ser un factor que influya para estar o no participando en el mercado laboral, en donde, para la prueba de ello, se estableció un segundo modelo incluyendo dicha variable. Podemos notar (Véase Tabla 4.2) que una persona separada tiene las mismas oportunidades de tener empleo que una casada. Mientras que una persona soltera reduce sus posibilidades de tener empleo un 69.6%, en relación a una casada. Además es importante destacar que a un aumento de la edad de la persona disminuye la oportunidad de participar en el mercado laboral.

A través de la experiencia se ha comprobado que el nivel de instrucción influye en la persona para su inserción en el mercado laboral, donde se elaboró un tercer y último modelo que contenga éstas variables como predictoras, obteniendo lo siguiente:

$$\frac{\widehat{\text{Empleo}}}{1 - \widehat{\text{Empleo}}} = \exp(-0.924 + 0.017 * \text{Edad} + 1.631 * \text{Hombre} + 0.465 * \text{Primaria} + 0.372 * \text{Secundaria} + 0.546 * \text{Superior} + 2.118 * \text{Postgrado})$$

TABLA 4.3

Coefficientes en el tercer modelo con su respectivo valor p

	Coef.	Exp(Coef)	Valor p
Intercepto	-0.924	0.397	0.00
Edad	0.017	1.018	0.00
Hombre	1.631	5.110	0.00
Primaria	0.465	1.593	0.00
Secundaria	0.372	1.450	0.00
Superior	0.546	1.726	0.00
Postgrado	2.118	8.316	0.00

Es evidente que el nivel de instrucción ha sido significativo para la obtención de empleo (Véase Tabla 4.3), en donde por ejemplo, un hombre de 30 años de edad, con nivel de instrucción Superior tiene una probabilidad del 85.35% de participar en el mercado laboral, esto indica, que para aquellas personas con mayor edad y un alto nivel de instrucción, las posibilidades de obtener empleo son muy altas, en el cuál es evidente que para una persona con un nivel de instrucción de postgrado, la oportunidad de participar en el mercado laboral es mayor en 7.684 veces a diferencia de una persona que solo tenga conocimientos de lectura y escritura.

Aplicando el test de Hosmer y Lemeshow para éste último modelo, se obtuvo un valor p de 0.00, donde no existe evidencia estadística para indicar que el modelo es adecuado. Para establecer con más detalle cuales fueron en porcentaje los errores de estimación con los valores observados, se procedió a realizar una matriz de confusión (Véase Tabla 4.4).

TABLA 4.4

Matriz de confusión para el tercer modelo

	Sin empleo	Con empleo
Sin empleo	2813	7651
Con empleo	1964	22130

Podemos notar entonces que para el total de personas que cuentan con un empleo el modelo estimó muy bien esta condición un 91.20% y un

26.90% para aquellos que no lo tienen, existiendo un error de estimación del 8.80% y 73.10% para aquellas personas que trabajan y aquellas que no lo hacen, respectivamente.

Utilizando árboles de decisión con el algoritmo de segmentación CHAID, para un primer modelo donde sólo se consideren como variables predictoras la edad y el género, ésta última es el mejor predictor para participar o no en el mercado laboral, donde la probabilidad más alta de obtener empleo (96.7%) se da para aquellas personas de género masculino con edades que comprende entre 41 y 46 años. La probabilidad más baja de obtener empleo (26.3%) se establece para personas del género femenino con edades menores o iguales a 20 años.

Al igual que en regresión logística, se estableció un segundo modelo utilizando como variables predictoras la edad, el género y el estado civil, donde se concluye que la probabilidad más alta de participar en el mercado laboral (97.7%) se da para las personas de género masculino con edad comprendida entre 41 y 46 años y con estado civil Casado o Separado. Es interesante mencionar que las personas de género femenino, que comprende edades entre 20 y 24 años con estado civil Casada tienen una probabilidad muy baja de obtener empleo, con apenas un 37.3%, aquella que aumenta en un 44.6% si mantiene un estado civil soltero o separado.

Como último modelo a estudiar se analizó, como varias predictoras, a la edad, el género y el nivel de instrucción, para lo cual mediante el árbol de decisión propuesto (Véase en Anexo) se obtuvo que la mayor probabilidad de tener empleo (96.7%) es aquella persona de sexo masculino y con edades comprendidas entre 41 y 46 años, donde el nivel de instrucción no ha sido significativo para ésta segmentación. Es importante indicar que la variable género sigue siendo el mejor predictor para obtener o no empleo, donde al existir un mayor porcentaje de desempleados por parte de las mujeres, según este modelo aumenta las probabilidades de no participar en el mercado laboral en un 83.6%, si además ésta persona tiene una edad menor o

igual a 20 años con nivel de instrucción de alfabetización o superior.

De los distintos modelos propuestos, tanto regresión logística y árboles de decisión, todos clasifican de manera correcta un porcentaje aceptable a los individuos en general (Véase Tabla 4.5), donde observamos que no existe una considerable diferenciación entre cada técnica estadística, en el cuál los árboles de decisión son aquellos que obtuvieron mejores porcentajes correctos de clasificación.

TABLA 4.5
Porcentajes correctos clasificación

	Regresión Logística	Árboles de Decisión
Modelo 1 (Edad y Género)	72.4%	73.4%
Modelo 2 (Edad, Género y Est. Civil)	69.0%	73.5%
Modelo 3 (Edad, Género y Niv. Inst.)	72.2%	73.9%

5. CONCLUSIONES

- En el Ecuador, es más probable que una persona del género masculino obtenga empleo a diferencia de la mujer, que en un estado civil Casado y con un nivel de instrucción bajo, reduce aún más sus posibilidades de participar en el mercado laboral.
- El nivel de instrucción es un factor importante para la empleabilidad, donde las probabilidades de obtener empleo, a mayor nivel de instrucción, aumentan considerablemente.
- Tanto el método de regresión logística, como los árboles de decisión clasifican aceptablemente a todos los individuos en general, siendo ambas metodologías útiles para otros enfoques de estudio.

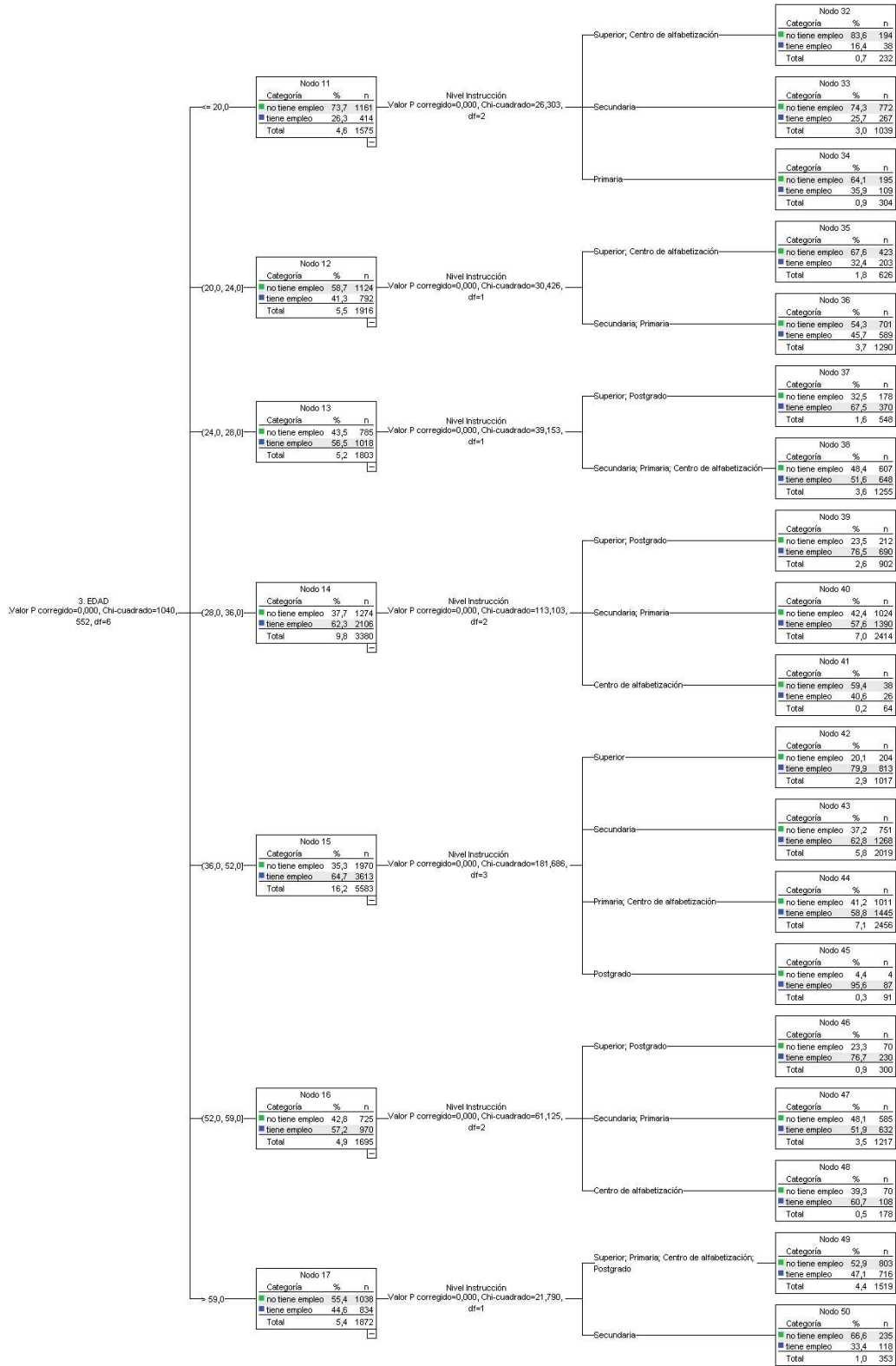
6. REFERENCIAS BIBLIOGRÁFICAS

- [1]. García, J. C., & Cortez, P. (2012). Análisis de la participación laboral de la mujer en el mercado ecuatoriano.
- [2]. Laura de Pablos Escobar. (2004). PARTICIPACIÓN LABORAL FEMENINA. ANÁLISIS DE LOS FACTORES DETERMINANTES. LA IMPORTANCIA DE LA EDUCACIÓN. In *XI Encuentro de Economía Pública: [los retos de la descentralización fiscal ante la globalización]* (p. 24).
- [3]. Bartlett, M. (1937). Properties of sufficiency and statistical tests. *Proceedings of the royal society of London series a-mathematical and physical sciences*, 160, A901, 0268-0282. ISSN 0080-4630.
- [4]. Fisher, Ronald A.; Yates, Frank (1948) [1938]. *Statistical tables for biological, agricultural and medical research (3rd edición)*. Londres: Oliver & Boyd. pp. 26-27. OCLC 14222135. Nota: la 6ª edición, ISBN 0-02-844720-4,
- [5]. Berkson, J. (1944). Application to the Logistic Function to Bio-assay. *Journal of the American Statistical Association*, 39 (1944) 357-365.
- [6]. Cox, D. (1970). *The Analysis of Binaria Data*. Methuen, London.
- [7]. Nelder, J. A. and R.W.M. Wedderburn (1972). Generalized Linear Models. *Journal of the Royal Statistical Society, Series A* 135: 370-384
- [8]. Monahan, J. (2011). *Numerical Methods of Statistics*, Second Edition. Cambridge, New York.
- [9]. Cellard, J. C., Labbe, B. & Cox, G. (1967). Le programme Elisée. Présentation et application. *Metra*, 3, 511-519.
- [10]. Bourouche, J. M. & Tennenhaus, M. (1972). Some segmentation methods. *Metra*, 7, 407-418.
- [11]. Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29, 119-127.
- [12]. Madgison, J. (1989). SPSS/PC+ CHAID. Chicago: SPSS Inc
- [13]. Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.I. (1984). Classification and regression trees. *Belmont, Calif.: Wadsworth*.
- [14]. Biggs, D. B., B. de Ville and E. Suen (1991). "A method of choosing multiway partitions for classification and decision trees,." *Journal of Applied Statistics* 18: 49.
- [15]. Loh WY, Shih YS (1997). "Split Selection Methods for Classification Trees." *Statistica Sinica*, 7, 815-840.
- [16]. Espin-García, O., & Rodríguez-Caballero, C. V. (2013). Metodología para un scoring de clientes sin referencias crediticias. *Cuadernos de Economía*, 32(59), 137-162.
- [17]. INEC. (2014). la Encuesta de Empleo, Desempleo y Subempleo (ENEMDU). Febrero 5 del 2016, de INEC Sitio web: www.ecuadorencifras.gov.ec

ANEXO. ÁRBOL DE DECISIÓN

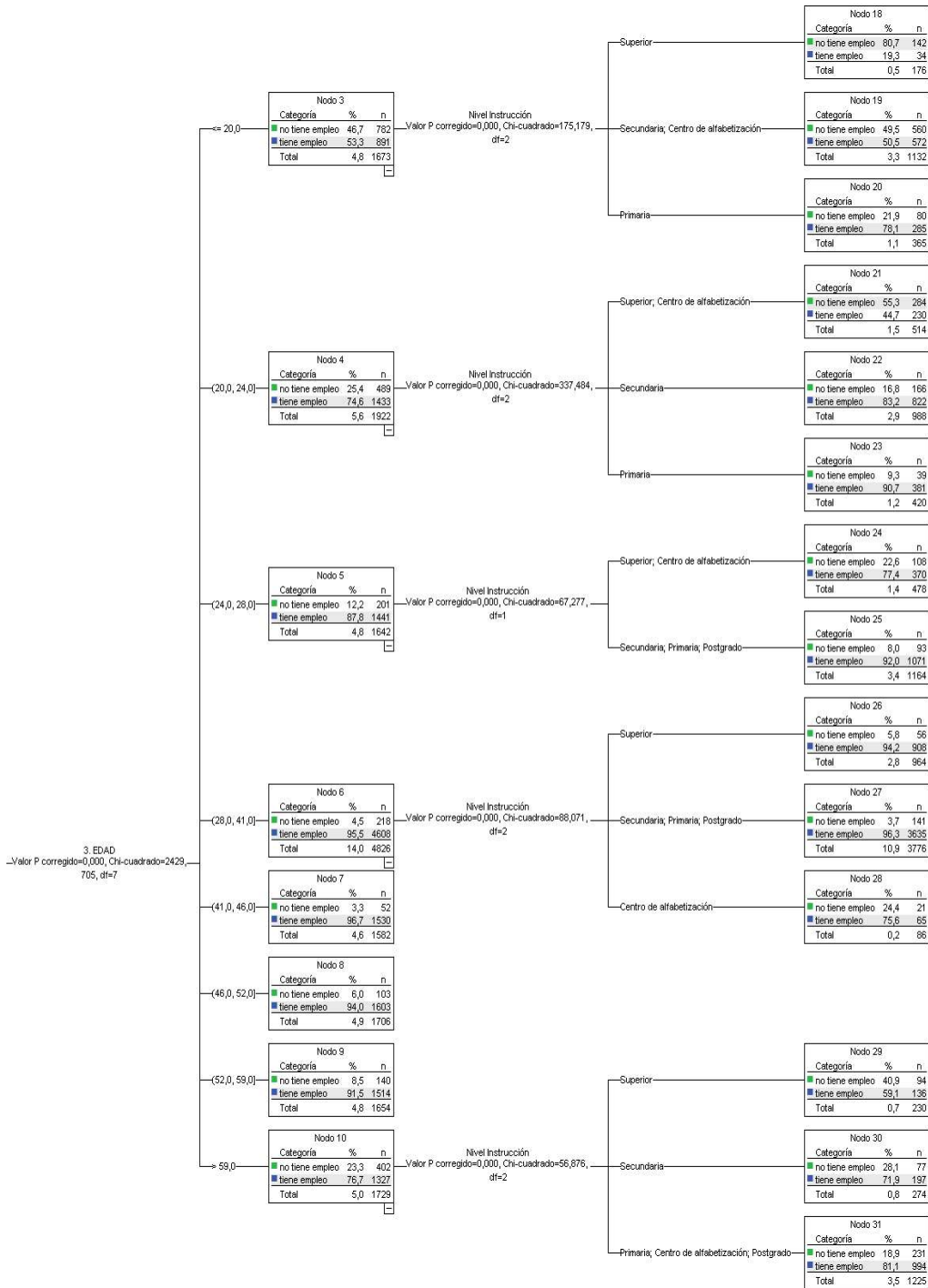


GRÁFICA A 1. Primera segmentación, árbol de decisión para el tercer modelo.



MODELOS ESTADÍSTICOS PARA EL ESTUDIO DE VARIABLES QUE PUEDEN INCIDIR EN LA PROBABILIDAD DE OBTENER EMPLEO EN ECUADOR

GRÁFICA A 2. Segunda y tercera segmentación del género femenino, árbol de decisión para el tercer modelo.



GRÁFICA A 3. Segunda y tercera segmentación del género masculino, árbol de decisión para el tercer modelo.