



**FACTORES QUE INCIDEN EN LA PERCEPCIÓN DE LOS ALUMNOS DE  
LICENCIATURA HACIA LA CALIDAD DE LOS SERVICIOS DEL SUBA**

**POR:**

**I. I. ALMA ANGELINA HOLGUÍN AGUIRRE**

**Tesina presentada como requisito parcial para obtener el grado de  
Maestría Profesional en Estadística Aplicada**

**Universidad Autónoma de Chihuahua  
Facultad de Zootecnia y Ecología  
Secretaría de Investigación y Posgrado**

Factores que inciden en la percepción de los alumnos de licenciatura hacia la calidad de los servicios del SUBA. Tesina presentada por Alma Angelina Holguín Aguirre como requisito parcial para obtener el grado de Maestría Profesional en Estadística Aplicada, ha sido aprobada y aceptada por:

---

M. A. Luis Raúl Escárcega Preciado  
Director de la Facultad de Zootecnia y Ecología

---

M. C. Antonio Humberto Chávez Silva  
Secretario de Investigación y Posgrado

---

D. Ph. Pablo Fidel Mancillas Flores  
Coordinador Académico

---

Dr. Juan Ángel Ortega Gutiérrez  
Presidente

SEPTIEMBRE 02 - 2016

Fecha

Comité:

Dr. Juan Ángel Ortega Gutiérrez  
Dr. Nicolás Callejas Juárez  
D. Ph. Joel Domínguez Viveros

© Derechos Reservados

**Alma Angelina Holguín Aguirre**  
PERIFÉRICO FRANCISCO R.  
ALMADA KM. 1, CHIHUAHUA,  
CHIH., MÉXICO C.P. 31453

SEPTIEMBRE 2016

## **AGRADECIMIENTOS**

A Dios, pues todas las cosas por Él fueron hechas. A mis maestros, D. Ph. Víctor Manuel Santana Rodríguez, Dr. Juan Ángel Ortega Gutiérrez. A mis compañeros y a esta Institución por haberme otorgado beca para mis estudios.

## **DEDICATORIA**

A Dios, Isaí y Anaí, mis hijos, que siempre me alentaron a alcanzar mi meta. A mi padre José Inés, que con su vida me inspiró a ser siempre una mejor persona, a mi madre Jesusita, por su invariable apoyo, hermanos, familia y amigos.

## CURRICULUM VITAE

La autora nació el 19 de abril de 1973 en la Ciudad de Hidalgo del Parral,  
Chihuahua, México.

1992 - 2009	Técnico Programador en el Departamento de Control Escolar de la Facultad de Ciencias Químicas de la Universidad Autónoma de Chihuahua.
2000 - 2004	Estudios de Licenciatura en el Instituto Tecnológico de Chihuahua, Chihuahua. Obteniendo el Título de Ingeniera Industrial.
2006 - 2010	Docente por horas en el Curso Propedéutico que ofrece la Facultad de Ciencias Químicas de la Universidad Autónoma de Chihuahua.
2009 a la fecha	Jefe de Unidad de Servicios Escolares de la Facultad de Ciencias Químicas de la Universidad Autónoma de Chihuahua.

## RESUMEN

FACTORES QUE INCIDEN EN LA PERCEPCIÓN DE LOS ALUMNOS DE  
LICENCIATURA HACIA LA CALIDAD DE LOS SERVICIOS DEL SUBA

POR:

I. I. ALMA ANGELINA HOLGUÍN AGUIRRE

Maestría Profesional en Estadística Aplicada

Secretaría de Investigación y Posgrado

Facultad de Zootecnia y Ecología

Universidad Autónoma de Chihuahua

Presidente: Dr. Juan Ángel Ortega Gutiérrez

Con el objetivo de identificar y evaluar aquellos factores que influyen en la satisfacción de los estudiantes de Licenciatura hacia la calidad de los servicios que presta el Sistema Universitario de Bibliotecas Académicas (SUBA) en la Universidad Autónoma de Chihuahua (UACH), se realizó el presente estudio que comprende: diseño de la encuesta, diseño del muestreo y análisis de los datos. Se eligieron 14 variables independientes y una variable respuesta dicotómica. Mediante el software SPSS se aplicaron pruebas de chi-cuadrada y se calcularon odds ratio para probar la independencia y grado de asociación entre variables. Se probó asociación significativa ( $P < 0.05$ ) entre la variable respuesta con la dependencia de educación superior (DES), género (gen), visita (vis), frecuencia (fre), otras bibliotecas (ob), curso (cur), servicios (svs) y horario (hro). Seleccionándose estas variables para realizar el análisis de regresión logística binaria, del cual la DES, gen, fre, ob, cur y svs mostraron una asociación ( $P < 0.05$ ) con la variable respuesta, se obtuvieron los coeficientes de

regresión para la ecuación del modelo. La precisión del modelo para predecir probabilidades fue de 93.2 %. Finalmente la regresión logística permitió identificar seis principales factores que generan diferencias en la satisfacción y evaluar su importancia. Las probabilidades más altas de obtener una buena calificación de los servicios del SUBA la otorgaron aquellos estudiantes de la DES de agropecuarias, que son mujeres, que asistieron al curso de inducción, que acuden con menor frecuencia, que visitan también bibliotecas externas, e identifican los servicios que ofrece el SUBA.

## **ABSTRACT**

### **FACTORS AFFECTING THE PERCEPTION OF UNDERGRADUATE STUDENTS TO THE QUALITY OF SERVICES SUBA**

**BY:**

**ALMA ANGELINA HOLGUIN AGUIRRE**

With the object of identify and evaluate those influent factors in the satisfaction of undergraduate students about the service that “Sistema Universitario de Bibliotecas Academicas” (SUBA) offers in “Universidad Autonoma de Chihuahua” (UACH), it has been made a study that contains: survey design, sampling design and data analysis. Firstly 14 independent variables and dichotomous response variable were defined. After that, chi-square tests were applied and Odds ratios were calculated to test the independence and degree of association between variables. Significant association ( $P < 0.05$ ) was tested between the response to dependence on higher education (DES), gender (gen), visits (vis), frequency (fre) , other libraries (ob) , course (cur), services (svs) and service hours (hro). These variables were selected for analysis of binary logistic regression, in which: DES, gen, fre, ob, cur and svs showed an association ( $P < 0.05$ ) with the variable response. Finally the regression coefficients were obtained for the model equation. The accuracy of the model to predict probability was 93.2 %. Logistic Regression allowed identifying six principal factors that generate differences in the satisfaction of users of SUBA. The highest probabilities of getting a good grade of the SUBA services is given by those students who belong to the “DES” of agricultural, who are women who attended the induction course, those who use to visit with less

frequently SUBA, and those who use to attend external libraries and identify the services offered by the SUBA.



## CONTENIDO

	<b>Página</b>
RESUMEN.....	vi
ABSTRACT.....	viii
LISTA DE CUADROS.....	xiii
LISTA DE GRÁFICAS.....	xiv
LISTA DE FIGURAS.....	xv
INTRODUCCIÓN.....	1
REVISIÓN DE LITERATURA.....	3
Certificación Bajo la Norma ISO 9001:2008.....	3
Diseño de Muestreo.....	4
Muestreo Estratificado Aleatorio.....	4
Selección del Tamaño de Muestra y Asignación de la Muestra para Estimar Proporciones.....	5
Análisis de Datos Categóricos.....	7
Medidas de asociación entre variables.....	7
Tablas de contingencia.....	7
Prueba chi-cuadrada de Pearson ( ).....	8
Razón de verosimilitud de chi-cuadrado.....	10
Odds, riesgo relativo y odds ratio.....	10
Control de confusión o influencia de otros factores en la variable respuesta.....	12
Regresión logística.....	18
Comprobación de la ecuación.....	20



MATERIALES Y MÉTODOS.....	33
Diseño de la encuesta.....	33
Diseño de Muestreo.....	33
Elemento de muestreo.....	33
Tipo de diseño.....	34
Tamaño de Muestra.....	34
Selección de la muestra aleatoria por estrato .....	35
Aplicación de la encuesta.....	35
Instrucciones para encuestadores.....	35
Análisis de los Datos.....	37
Depuración de los datos.....	37
Prueba y odds ratio.....	39
Regresión logística binaria.....	42
RESULTADOS Y DISCUSIÓN.....	43
Datos Descriptivos de la Población que Visita el SUBA.....	43
Pruebas de y Odds Ratio.....	45
Regresión Logística Binaria.....	47
CONCLUSIONES Y RECOMENDACIONES.....	53
LITERATURA CITADA.....	54
ANEXOS.....	56



## LISTA DE CUADROS

<b>Cuadro</b>		<b>Página</b>
1	Ejemplo de una tabla de contingencia sobre distribución de las personas según la respuesta respecto a suscribirse a una revista y a su género.....	9
2a	Distribución de las personas según género, decisión respecto a suscribirse y grupos de edad.....	13
2b	La variable respuesta (suscripción) en función del género y edad (categorizada) de las personas.....	15
3	La variable respuesta (suscripción) en función del género y edad en años de las personas.....	23
4	Probabilidades de suscripción.....	31
5	Alumnos inscritos, muestra requerida y encuestas recabadas por Facultad.....	36
6	Lista de variables.....	38
7	Distribución de la muestra de estudio (n = 818 estudiantes).	40
8	Proporciones del uso del SUBA con las variables explicatorias significativas ( $\alpha=0.05$ ).....	44
9	Dependencias encontradas mediante las pruebas Chi-cuadrada y Odds Ratio a la calificación de los servicios al SUBA.....	46
10	Variables para predecir la probabilidad de una buena/mala calificación a los servicios del SUBA.....	48
11a	Ejemplos de los valores más altos de probabilidad predicha de acuerdo a la ecuación logística binaria de obtener una buena calificación de los servicios del SUBA.....	51
11b	Ejemplos de los valores más bajos de probabilidad predicha de acuerdo a la ecuación logística binaria de obtener una buena calificación de los servicios del SUBA.....	52



## LISTA DE GRÁFICAS

<b>Gráfica</b>		<b>Página</b>
1	Problema del modelo lineal para predecir probabilidades de un evento.....	21
2	Probabilidades estimadas de suscripción por género y edad con el modelo de regresión logística binaria.....	32



## LISTA DE FIGURAS

<b>Figura</b>		<b>Página</b>
1	Salida del programa SPSS. Control de confusión o influencia de otros factores en la variable respuesta.....	17
2	Salida del programa SPSS. Regresión logística.....	25
3	Salida del programa SPSS. Continuación de regresión logística.....	27
4	Salida del programa SPSS. Variables en la ecuación de regresión logística.....	29



## INTRODUCCIÓN

Nadie duda de la importancia de la biblioteca universitaria en la formación académica, concebida como un espacio de aprendizaje donde el estudiante encuentra las herramientas que utilizará durante su formación universitaria. El alumno accede, desde la biblioteca, no sólo a bibliografía básica recomendada por los profesores, sino también a revistas impresas y electrónicas, materiales audiovisuales y multimedia, diccionarios, enciclopedias y otras obras de consulta en distintos soportes y en línea (Alfaro y Sánchez, 2005).

El SUBA apoyándose en un Sistema de Gestión de Calidad establece: proveer material documental, capacitar a los usuarios en el uso de los recursos de información y satisfacer sus necesidades para lograr su óptimo desempeño académico y profesional. El SUBA consta de una coordinación general, la Biblioteca Central, las bibliotecas de las DES de Ingeniería y de Salud, además de 12 bibliotecas de las unidades académicas que conforman la UACH. En agosto del 2011, el SUBA obtuvo la certificación de sus procesos de acuerdo a la norma ISO 9001:2008. Señalándose por parte del organismo acreditador, algunas deficiencias encontradas al aplicar sus encuestas en el año 2010 y 2011, ya que el instrumento era demasiado extenso y se dirigieron únicamente a los alumnos que asistieron a alguna de sus bibliotecas sin realizar algún tipo de muestreo. Para el refrendo de la certificación obtenida, el SUBA requirió demostrar que lograba estimar y alcanzar periódicamente el grado de satisfacción de los usuarios requerido. Se dio la necesidad de diseñar un instrumento de medición con fundamentos estadísticos adecuados para



estudiar la opinión de sus usuarios en sus modos de uso presencial y virtual.

Por lo cual el objetivo de esta tesis además del análisis de datos para identificar y evaluar aquellos factores que influyen en la satisfacción de los estudiantes de Licenciatura hacia la calidad de los servicios del SUBA; fue también el diseño de la encuesta y diseño del muestreo. El diseño de la encuesta se terminó en junio del 2013, y para su aplicación se realizó un muestreo aleatorio estratificado a través de las Facultades, donde el elemento del muestreo fueron los alumnos inscritos en el ciclo escolar agosto-diciembre 2013 en modalidad presencial de nivel Licenciatura; el levantamiento de campo se terminó en noviembre del 2013.

La encuesta generó principalmente datos de tipo categóricos por lo que se utilizaron en el análisis herramientas definidas por Agresti (2002) como tablas de contingencia, Prueba Chi-Cuadrada de Pearson, Odds Ratio. Las relaciones encontradas con estas primeras pruebas fueron aprovechadas para aplicar regresión logística, la cual permitió identificar y evaluar los factores que influyen en la satisfacción de los estudiantes hacia la calidad de los servicios que presta el SUBA.

Este estudio facilita a los administrativos del SUBA la oportunidad de conservar su certificación, provee información para conocer áreas de oportunidades y generar estrategias de mejora continua para atraer más usuarios y mantener la satisfacción de los mismos. Proporciona también información en el proceso de: elaboración de la encuesta, selección de la muestra y herramientas para el análisis de datos categóricos.



## REVISIÓN DE LITERATURA

### **Certificación Bajo la Norma ISO 9001:2008**

La norma ISO 9001:2008 contiene los requisitos para que los sistemas de gestión de calidad de las organizaciones sean capaces de proporcionar productos y/o servicios que cumplan las necesidades del cliente y los requisitos legales aplicables y puedan ser demostrados. Según esta norma, es fundamental entender las necesidades actuales y futuras del cliente, para cumplir sus requisitos y exceder sus expectativas; como lo explica en el primero de sus principios: La organización enfocada al cliente (URL1).

De acuerdo a este principio, la política de calidad del SUBA establece “Satisfacer los requisitos de información de los usuarios para su desempeño académico, ofreciéndoles material documental pertinente, suficiente y actualizado, sustentado en un sistema de gestión de calidad”. Un sistema de gestión de calidad es el que funciona basado en una cultura e ideología que busca la mejora continua del sistema completo de la organización que produce bienes y/o servicios para tratar de satisfacer al cliente.

Sapién-Aguilar *et al.* (2011) evaluaron el grado de satisfacción del sistema del SUBA a través de 340 alumnos de la Facultad de Contaduría y Administración. En una escala del uno al 10, donde 10 indica totalmente satisfecho; reportaron como notorias dos calificaciones, donde el 22 % calificaron con cinco y el 26 % calificó con ocho, concluyeron que esto manifiesta deficiencias en el sistema y recomendaron tomar en cuenta las opiniones del alumnado para mejorar la estructura, el idioma, la rapidez en las búsquedas así como cursos de inducción sobre el uso de dicho sistema.



## **Diseño de Muestreo**

El muestreo consiste en la selección de alguna parte de la población que refleje una o más características de la población completa. Un aspecto primordial en la selección de la muestra consiste en que todos los elementos de la población deben tener igual probabilidad de ser seleccionados. Las muestras aleatorias son las únicas que pueden estar sujetas al análisis estadístico, de tal manera que puede evaluarse la incertidumbre en las inferencias. De esta manera refleja más fielmente las características de la población al evitar el sesgo.

El interés del muestreo es estimar los parámetros de la población con las muestras más pequeñas posibles, pero que los estimadores estén lo más cerca posible al valor del parámetro. Esto se logra diseñando el muestreo lo más apropiado posible, de acuerdo a la heterogeneidad entre los elementos de la población será el diseño del muestreo (aleatorio simple, aleatorio estratificado o por conglomerados).

### **Muestreo Estratificado Aleatorio**

El muestreo debe aplicarse donde todos los elementos sean homogéneos para la variable de interés, por lo cual es necesario observar si existen características en las unidades de muestreo que tengan una fuerte relación con la variable de interés; estas se conocen como variables auxiliares. Cuando una variable auxiliar es conocida de antemano para cada elemento de la población, esta variable puede ser usada para clasificar los elementos de la población en subgrupos o estratos. Ya que la variable auxiliar está relacionada con la variable de interés, los subgrupos con poca variabilidad en la variable



auxiliar, serán homogéneos en la variable de interés, en tanto que la mayor parte de la variabilidad en la población se dará entre estratos.

En el muestreo estratificado (Cochran, 1986), la población de  $N$  unidades es primero dividida en subpoblaciones o estratos donde la sub división atiende un criterio de clasificación y/o diferenciación, cada unidad de muestreo pertenece a uno y solo un estrato, los estratos no se sobreponen y en su conjunto forman la población total. Una vez que han sido determinados los estratos, se toma una muestra simple aleatoria de cada estrato, asegurándose que estas muestras sean independientes. Los tamaños de muestra dentro de los estratos son representados por  $n_h$  respectivamente.

Una de las principales razones para utilizar el muestreo estratificado es incrementar la precisión en los estimadores de las características de toda la población. Si cada uno de los estratos es homogéneo, en el sentido de que las medidas varían muy poco de una unidad a otra, se puede obtener un estimador muy preciso de cualquiera de las medidas de los estratos derivados de una muestra pequeña en ese estrato. Debido a la independencia entre las muestras de cada uno de los estratos la varianza de los estimadores para estratos individuales se suman para obtener la varianza de los estimadores de la población total.

### **Selección del Tamaño de Muestra y Asignación de la Muestra para Estimar Proporciones**

La muestra de tamaño  $n$  puede ser distribuida de manera diferente a los  $L$  estratos, dependiendo de la información que se tenga acerca de ellos. Si solo se tiene información acerca del tamaño de los estratos, la asignación



proporcional es la única que puede aplicar. Como su nombre lo indica, la muestra se distribuye de acuerdo a los tamaños de estratos.

Si en adición al tamaño de cada estrato, se conoce la varianza de cada uno de ellos, y si existen notables diferencias, entre ellas, entonces la distribución se debe hacer de tal forma que se asigne mayor tamaño de muestra a estratos con mayor varianza, en comparación con los de menor varianza. Esto se logra mediante la asignación óptima.

En algunas situaciones el costo de muestreo, medido en términos de tiempo o dinero, difiere de estrato a estrato; entonces es razonable asignar menor tamaño de muestra a aquellos estratos donde el costo sea mayor por unidad de observación, en comparación con los estratos donde el costo sea menor. Lo que se conoce como asignación de costo mínimo.

Scheaffer (1987) mencionó que para estimar una proporción poblacional, primero hay que indicar qué tanta información se desea mediante la especificación del tamaño de límite del error de estimación (B); de acuerdo con esto es elegido el tamaño de la muestra. Si se reduce el tamaño de muestra se tendrá un incremento en el límite del error de estimación. La fórmula para el tamaño de muestra n (para un límite dado B del error de estimación) está dada por:

$$\frac{\sum \frac{N_i^2 S_i^2}{n_i}}{\sum S_i^2}$$

donde: (L) número de estratos definidos en la población, ( ) número de elementos en el estrato i, ( ) número de elementos en la población, ( ) tamaño



total de muestra. La ( ) ponderación para el estrato  $i$  bajo una asignación proporcional se define por  $h_i = \frac{N_i}{N}$  y el tamaño de muestra para el correspondiente estrato es  $n_i = n h_i$  donde  $\sum h_i = 1$ . Para un nivel de confianza del 95 % ( ), es la probabilidad de éxito fijada en 0.5 para maximizar el tamaño de muestra, ya que está dada por  $h_i = \frac{1}{2}$  y donde  $n = \frac{Z^2 \sum h_i p_i q_i}{e^2}$ . En una situación práctica se desconoce  $p_i$ . Un tamaño de muestra aproximado puede determinarse al reemplazar por un valor estimado. Frecuentemente, tal estimación puede ser obtenida de encuestas anteriores similares. Sin embargo, si no se cuenta con información anterior, se puede substituir  $p_i = 0.5$  para obtener un tamaño de muestra conservador (uno que será probablemente mayor que el requerido).

### **Análisis de Datos Categóricos**

**Medidas de asociación entre variables.** A través de un ejemplo práctico, se ilustra el uso de las medidas de asociación: tablas de contingencia, de Pearson, odds ratio, control de confusión, entre otras hasta llegar a la regresión lineal. A una serie de personas se les cuestionó (sí o no) se suscribirían a una nueva revista. Para cada persona las variables género (hombre=1, mujer=0), edad en años categorizada como menor a 44 y mayor o igual a 44 y codificada edad=0 y edad=1 respectivamente y suscripción (1=sí, 0=no) se registraron.

**Tablas de contingencia.** De acuerdo a Vicéns y Medina (2005) las tablas de contingencia (TC) tienen dos objetivos fundamentales: 1) Organizar la información contenida en un experimento cuando ésta es de carácter



bidimensional, es decir, cuando está referida a dos factores (variables cualitativas); 2) A partir de la TC se puede además analizar si hay dependencia entre los niveles de las variables cualitativas objeto de estudio. El hecho de que dos variables sean independientes significa que la frecuencia en los niveles de una de ellas no depende de la modalidad o nivel que adopte la otra.

En el Cuadro 1 se muestra una tabla de contingencia, donde 14 son varones y 26 son mujeres de los cuales 20 se suscribirían a la revista y 20 no. La TC permite tener información cruzada sobre ambas variables: de los hombres, 11 se suscribirían y 3 no, mientras que en el caso de las mujeres, 9 se suscribirían y 17 no.

**Prueba chi-cuadrada de Pearson ( ).** La prueba estadística Chi-cuadrado (Sánchez, 1989) se aplica para contrastar la Hipótesis nula : las variables “x” e “y” son independientes. Si la significación asociada a este estadístico es menor o igual a 0.05 se rechaza la hipótesis de independencia.

: P 0.05 variables independientes.

: P 0.05 variables dependientes.

La base de cálculo de la son las diferencias entre las frecuencias observadas y = esperadas:

$$\sum \frac{(\quad)}{\quad}$$

Cálculo de las : , , ,

. Sustituyendo valores del Cuadro 1: ,

,

Y luego en la fórmula:



Cuadro 1. Ejemplo de una tabla de contingencia sobre distribución de las personas según la respuesta respecto a suscribirse a una revista y a su género

Suscripción	Género		Total
	Hombre	Mujer	
Sí	11	9	20 ( )
No	3	17	20 ( )
Total	14 ( )	26 ( )	40 ( )

SAS (2009)



( ) ( ) ( ) ( )

El valor obtenido es ( $P < 0.008$ ) por lo que se rechaza , significa que existe alguna relación entre el hecho de suscribirse y el género de la persona. Uno de los inconvenientes de la prueba es que está influenciada por el tamaño de muestra, es decir, que a mayor número de casos analizados (a mayor N) el valor de tiende a aumentar, por lo que cuanto mayor sea la muestra más fácil será que rechacemos la hipótesis nula de independencia, cuando a lo mejor podrían ser independientes, lo que se conoce como una relación espuria, es decir, que no es genuina (Sánchez, 1989).

**Razón de verosimilitud de chi-cuadrado.** La razón de verosimilitud es una alternativa al estadístico Chi-cuadrado cuando el objetivo es contrastar la hipótesis de independencia entre las variables. La diferencia estriba en que se obtiene como el cociente entre las frecuencias observadas y esperadas. También se rechaza la hipótesis de independencia entre las variables cuando la significación de este estadístico sea menor o igual a 0.05 (Sánchez, 1989).

**Odds, riesgo relativo y odds ratio.** De acuerdo a Silva y Barroso (2004), el odds asociado a cierto suceso ( $O(E)$ ) se define como la razón entre la probabilidad de que dicho suceso ocurra ( $P(E)$ ) y la probabilidad de que no ocurra ( $1-P(E)$ ); es decir, expresa la probabilidad de ocurrencia de un evento y que no ocurra:

$$() \frac{()}{()}$$

Se suele hablar del odds que tiene de ganar un equipo de baloncesto



antes de un juego; con ello se alude a cuántas veces más probable es que gane frente a que pierda. El riesgo relativo (RR) es especialmente relevante en la investigación epidemiológica actual. Si ( ) denota el riesgo de que se produzca una enfermedad E cuando está presente la condición A y B es otra condición, de manera que ( ) denota el riesgo que se corre cuando es ella la que rige. Entonces el RR es el cociente que resulta entre estas probabilidades:

$$\frac{( )}{( )}$$

La ecuación anterior expresa el riesgo relativo de padecer la enfermedad cuando se está en la condición A respecto de cuando se está en la condición B. Dicho de otro modo, sintetiza cuánto más probable es desarrollar la enfermedad en el primer caso que en el segundo. Del mismo modo que el RR expresa la razón entre dos probabilidades, tiene sentido considerar la razón de dos odds. Finalmente, el odds ratio se obtiene como la razón del odds de un evento bajo cierta condición entre el odds que le corresponde al mismo suceso bajo otra condición. Retomando el ejemplo de suscripción de personas según su género (Cuadro 1), el odds de suscribirse para los hombres ( ) se estima mediante el cociente determinado por la proporción de hombres que se suscriben entre la proporción de los que no lo hacen y el odds para las mujeres

( ) se estima similarmente. La asociación encontrada con la prueba chi-cuadrada, entre género y suscripción puede ahora ser estimada mediante el cálculo de la razón de odds ratio.

$$( ) \frac{( )}{( )} \frac{/}{/} \text{-----}$$



A esta estimación se denomina razón de productos cruzados ya que involucra la multiplicación de las celdas opuestas en las diagonales de la TC de . Si el género no guardara relación con el hecho de suscribirse, el odds de suscribirse sería el mismo para ambos géneros, y el OR sería igual a 1. Un OR mayor que uno, como ocurre en el ejemplo, indica que los hombres tienden aproximadamente siete veces más a suscribirse que las mujeres.

**Control de confusión o influencia de otros factores en la variable respuesta.** Si además se considera que se quiere evaluar si la edad del suscriptor constituye una variable de confusión entre la relación que pudiera existir entre el género y el hecho de suscribirse. Para valorarlo, los datos tienen que ser categorizados, por lo que se dividieron en dos categorías de edades ( 44 y 44 años) codificándose como edad=0 y edad=1 (Cuadro 2a).

Un método usual para valorar la confusión, mencionan Silva y Barroso (2004) consiste en comparar de forma directa el estimado bruto del efecto con su estimación una vez controlado el presunto factor de confusión. Para ello, se debe obtener una estimación del efecto global a partir de los datos estratificados, mediante una media ponderada de las estimaciones de los efectos de los estratos. Esto es lo que provee el odds ratio de Mantel Haenszel ( ), que se obtiene a través del análisis de tablas de estratificadas supóngase que hay  $r$  estratos y , se obtiene de la siguiente manera:



Cuadro 2a. Distribución de las personas según género, decisión respecto a suscribirse y grupos de edad

Edad	Género	Suscripción	
		sí	no
<44	Hombre	3 ( )	2 ( )
	Mujer	2 ( )	8 ( )
≥44	Hombre	8 ( )	1 ( )
	Mujer	7 ( )	9 ( )



$$\frac{\sum ( \quad / \quad )}{\sum ( \quad / \quad )}$$

Sustituyendo los valores del Cuadro 2a:

$$\frac{\quad \quad}{\quad \quad}$$

La correspondiente prueba Chi-cuadrada de Mantel y Haenszel se realiza usando el estadístico de la  $\chi^2$  con un grado de libertad.

$$\chi^2 = \frac{(\sum \quad | \quad )}{\sum \frac{(\quad)(\quad)(\quad)(\quad)}{(\quad)}} \frac{(\quad)(\quad)}{(\quad)}$$

Los límites de confianza inferior y superior al 95 % para el  $\chi^2$  se obtienen respectivamente a través de:

$$\frac{\quad}{\sqrt{\quad}} \quad \frac{\quad}{\sqrt{\quad}}$$

El intervalo de confianza para el  $\chi^2$  es [1.58, 42.28]. Para corroborar los cálculos anteriores, se ingresaron los datos del Cuadro 2b al programa estadístico SPSS (Anexo 2), en las opciones: ANALIZAR / ESTADÍSTICOS DESCRIPTIVOS/TABLAS DE CONTINGENCIA. En fila(s) ingresar el primer factor (género) y en Columna(s) ingresar la variable respuesta (suscribe), Capa 1 de 1 ingresar el factor confusor (edad categorizada), en [Estadísticos] se



Cuadro 2b. La variable respuesta (suscripción) en función del género y edad (categorizada) de las persona

<b>ed</b>	<b>sc</b>	<b>gen</b>	<b>ed</b>	<b>sc</b>	<b>gen</b>	<b>ed</b>	<b>sc</b>	<b>gen</b>
<44	si	mujer	≥44	no	mujer	≥44	si	mujer
<44	no	mujer	≥44	no	mujer	<44	si	hombre
<44	no	mujer	≥44	si	mujer	≥44	no	hombre
<44	no	mujer	≥44	no	mujer	≥44	si	mujer
<44	no	mujer	≥44	no	mujer	≥44	si	hombre
<44	no	mujer	<44	no	hombre	≥44	si	Hombre
<44	no	mujer	≥44	si	mujer	≥44	si	Hombre
<44	no	mujer	≥44	no	mujer	≥44	si	Hombre
<44	no	mujer	≥44	no	mujer	≥44	si	Hombre
<44	si	mujer	≥44	no	mujer	≥44	si	Hombre
≥44	no	mujer	≥44	si	mujer	≥44	si	Hombre
≥44	no	mujer	≥44	si	mujer	≥44	si	Hombre
<44	no	hombre	<44	si	hombre			
≥44	si	mujer	<44	si	hombre			

sc=suscripción, gen=género, ed=edad en años



selecciona Estadísticos de Cochran y Mantel-Haenszel y Riesgo, CONTINUAR y ACEPTAR. De la Fuente-Fernández (2011) reportó que primero se verifican los OR de las categorías en el cuadro de salida “Estimación de Riesgo” (Figura 1). La categoría de edad menor a 44 años (edad=0) genera un OR = 6.0 con límites de confianza [0.563, 63.984] que por contener a la unidad no es significativo y la categoría de mayor o igual a 44 años (edad=1) con OR=10.28 y límites de confianza [1.030, 102.753] el cual resulta significativo.

Se observa en “pruebas de independencia condicional” el valor de  $\chi^2$  que coincide con el cálculo previo. Luego se debe verificar si estos OR's se distribuyen de manera homogénea. En el cuadro de salida “Pruebas de Homogeneidad de la razón de las ventajas”. Significancia=0.747 indica homogeneidad de acuerdo a:

- = Los OR's se distribuyen de manera homogénea ( $P > 0.05$ ).
- = Los OR's no se distribuyen de manera homogénea.

Como en este caso existe homogeneidad, entonces se puede interpretar el OR global, es decir el  $\chi^2$  que aparece en la salida “Estimación de la razón de las ventajas común de Mantel-Haenszel” Estimación=8.195 (valor que coincide con el cálculo anterior) y cuyos límites de confianza [1.58, 42.28] no contiene a la unidad por lo que es significativo. Nótese que el OR obtenido al considerar sólo el género, indicó que los hombres eran casi siete veces más propensos a suscribirse que las mujeres. Con este nuevo valor de OR se puede afirmar que los hombres con edad igual o mayor a 44 años serán los más propensos a suscribirse y se concluye que tanto la edad como el género están



**Pruebas de independencia condicional**

	Chi-cuadrado	gl	Sig. asintótica (bilateral)
Cochran's	7.287	1	.007
Mantel-Haenszel	5.277	1	.022

**Estimación de Riesgo**

Edadcat	Valor	Intervalo de Confianza al 95 %		
		Inferior	Superior	
0	Razón de las ventajas para género (Hombre / Mujer)	6	0.563	63.984
	Para la cohorte susc = si	3	0.717	12.553
	Para la cohorte susc = no	0.5	0.164	1.528
	N de casos válidos	15		
1	Razón de las ventajas para género (Hombre / Mujer)	10.286	1.03	102.753
	Para la cohorte susc = si	2.032	1.113	3.708
	Para la cohorte susc = no	0.198	0.03	1.318
	N de casos válidos	25		
Total	Razón de las ventajas para género (Hombre / Mujer)	6.926	1.529	31.377
	Para la cohorte susc = si	2.27	1.252	4.115
	Para la cohorte susc = no	0.328	0.116	0.928
	N de casos válidos	40		

**Pruebas de Homogeneidad de la razón de las ventajas**

	Chi-cuadrado	gl	Sig. asintótica (bilateral)
Breslow-Day	0.104	1	0.747
De Tarone	0.104	1	0.0748

**Estimación de la razón de las ventajas común de Mantel-Haenszel**

Estimación		8.195	
ln(Estimación)		2.104	
Error típ. de ln(Estimación)		.837	
Sig. asintótica (bilateral)		.012	
Intervalo de confianza asintótico al 95 %	Razón de ventajas común	Límite inferior	1.588
		Límite superior	42.287
	ln(Razón de ventajas común)	Límite inferior	.463
		Límite inferior	3.744

La estimación de la razón de las ventajas común de Mantel-Haenszel se distribuye de manera asintóticamente normal bajo el supuesto de razón de las ventajas común igual a 1,000. Lo mismo ocurre con el log natural de la estimación.

Figura 1. Salida del programa SPSS. Control de confusión o influencia de otros factores en la variable respuesta.



asociados al hecho de suscribirse o no hacerlo. Cuando los OR's no se distribuyan de manera homogénea de acuerdo a la edad, se puede utilizar, por ejemplo, una regresión logística.

Silva y Barroso (2004) mencionaron que el análisis estratificado y el enfoque de Mantel y Haenszel ofrecen una vía simple para la eliminación de la influencia que pueda tener una variable de confusión sobre una posible asociación. Sin embargo, este no se puede llevar a cabo con un número elevado de categorías, salvo que los tamaños de las muestras sean grandes. Además, exige categorizar aquellas variables de confusión que sean continuas. Para el control de múltiples variables de confusión (categóricas y continuas) los modelos de regresión, en especial la regresión logística, aportan una solución óptima.

**Regresión logística.** Silva y Barroso (2004) reportaron que existen situaciones en las que resulta impropio usar la regresión simple o múltiple, lineal o no, para caracterizar la relación funcional entre cierta variable de respuesta y un conjunto de variables independientes. La más común de las circunstancias invalidantes es la que se da cuando la variable de respuesta expresa la ocurrencia o no de cierto suceso, tal como que un estudiante de ingeniería desaprobe o apruebe un examen de álgebra, que el producto de un embarazo sea o no un bebé con determinada malformación congénita, o que un trasplante renal sea aceptado o rechazado por el organismo. En situaciones como estas suele interesar la evaluación del efecto de uno o más antecedentes sobre el hecho de que el acontecimiento se produzca. Si  $Y$  es la variable dependiente que refleja la ocurrencia o no del suceso, puesto que  $Y$  es



dicotómica, se admite que puede asumir los dos valores siguientes:

{

La situación más simple es aquella en que se trata de evaluar el efecto de un solo factor, al que se representa mediante la variable sobre el desenlace  $Y$ . Retomando el ejemplo de suscripción a la revista en donde se registró suscripción (1=sí, 0=no) y la edad en años, para la situación en que se tiene una sola variable explicativa, podría considerarse el siguiente modelo lineal, con valores para  $\beta_0$  y edad continua:

Y su ecuación:

$\hat{Y}$

La ecuación puede obtenerse con la ayuda de Excel (Anexo 3), realizando lo siguiente: seleccionar los 40 datos de suscripción y edad, insertar gráfico de dispersión, sobre los puntos con el botón derecho del mouse, seleccionar agregar línea de tendencia, seleccionar lineal, presentar ecuación en el gráfico y cerrar.

Si se pudiera interpretar  $\hat{Y}$  como la probabilidad de suscripción o la probabilidad de éxito en general. La interpretación de la pendiente sería entonces obvia, por cada año adicional de edad, la probabilidad de suscribirse sube 2.8 %. Los problemas de esta ecuación surgen cuando se intenta hacer pronósticos con ella (Pérez *et al.*, 2004). La probabilidad de que una persona de 35 años se suscriba a esta revista,  $\hat{Y}$  es de 23 %. La probabilidad para una persona de 26 o de 63 años, es de



respectivamente, tal como se muestra en la Gráfica 1, los cuales son valores incoherentes, hablando de probabilidades donde los valores deben ser entre 0 y 1. Para forzar que los pronósticos estén siempre en el intervalo {0,1} se tienen dos condiciones que se deben cumplir: valores siempre positivos y que no sean mayores a uno. Claramente esta función no puede ser lineal. Se obtienen valores siempre positivos usando una función exponencial:

$$( \quad )$$

Así también se puede lograr que no sea mayor de uno, dividiendo entre algo un poco más grande:

$$\frac{( \quad )}{( \quad )}$$

Siendo esta la ecuación que se estima con la regresión logística expresada en términos de probabilidad.

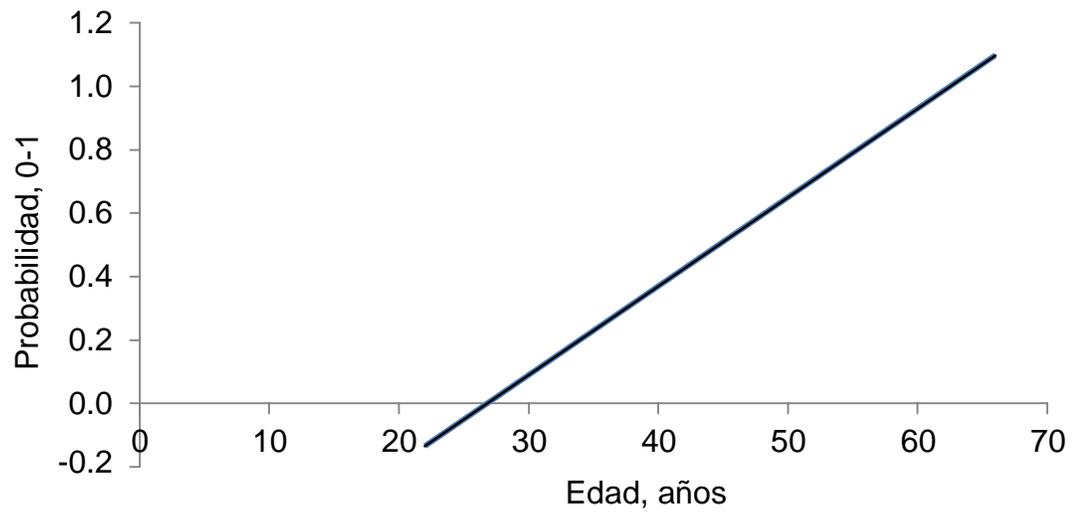
**Comprobación de la ecuación.** Lo anterior se puede explicar algebraicamente de la siguiente manera, partiendo de un modelo lineal:

Dadas las condiciones para forzar las predicciones de estar en el intervalo {0,1} que se logran mediante la transformación logística de la probabilidad P de que cierto suceso ocurra o no y aplicando el logaritmo al odds asociado a esta probabilidad:

$$( \text{---} )$$

Igualando ambas

$$\text{---}( \quad )$$



Gráfica 1. Problema del modelo lineal para predecir probabilidades de un evento.



Eliminando ln aplicando exp del otro lado de la ecuación

$$\frac{(\quad)}{(\quad)}$$

Despejando P

$$\frac{(\quad)}{(\quad)}$$

$$\frac{(\quad)}{(\quad)}$$

$$\frac{(\quad)}{(\quad)}$$

$$\frac{(\quad)}{(\quad)}$$

$$\frac{(\quad)}{(\quad)}$$

No es posible interpretar directamente las estimaciones de los parámetros , ya que no es un modelo lineal. Lo que se hace en la práctica es fijarse en el signo de los estimadores. Si el estimador es positivo, significará que incrementos en la variable asociada causan incrementos en la probabilidad de suscripción. Por el contrario, si el estimador muestra un signo negativo, supondrá disminuciones en dicha probabilidad (Pérez *et al.*, 2004).

En el Cuadro 3 se incluye la edad en años de las personas, así como los datos considerados anteriormente. A través del programa estadístico SPSS se ajustó un modelo de regresión logística para la respuesta de suscripción en función de las variables género (hombre=1, mujer=0) y edad en años.

$$\frac{(\quad)}{(\quad)} \text{ o bien } \frac{(\quad)}{(\quad)}$$

Una vez ingresados los datos en SPSS (Anexo 4), la regresión logística se realiza en las siguientes opciones: ANALIZAR/ REGRESIÓN/ LOGÍSTICA



Cuadro 3. La variable respuesta (suscripción) en función del género y edad en años de las personas

<b>ed</b>	<b>sc</b>	<b>gen</b>	<b>ed</b>	<b>sc</b>	<b>gen</b>	<b>ed</b>	<b>sc</b>	<b>gen</b>
30	si	mujer	46	no	mujer	56	si	mujer
32	no	mujer	47	no	mujer	42	si	hombre
34	no	mujer	47	si	mujer	44	no	hombre
35	no	mujer	47	no	mujer	59	si	mujer
35	no	mujer	48	no	mujer	45	si	hombre
35	no	mujer	34	no	hombre	46	si	hombre
39	no	mujer	50	si	mujer	47	si	hombre
39	no	mujer	51	no	mujer	49	si	hombre
39	no	mujer	51	no	mujer	50	si	hombre
43	si	mujer	51	no	mujer	54	si	hombre
45	no	mujer	52	si	mujer	58	si	hombre
45	no	mujer	52	si	mujer	59	si	hombre
31	no	hombre	38	si	hombre			
46	si	mujer	39	si	hombre			

sc=suscripción, gen=género, ed=edad en años



BINARIA e ingresar en DEPENDIENTES la variable respuesta (suscribe) y en COVARIABLES (edad y género), además en CATEGÓRICA ingresar (género) ya que esta variable cuenta con dos categorías, no es el caso de edad por ser una variable continua, seleccionar CONTINUAR. En GUARDAR marcar únicamente PROBABILIDADES seleccionar CONTINUAR. En OPCIONES marcar BONDAD DE AJUSTE DE HOSMER- LEMESHOW, IC para  $\exp(B)$  e INCLUIR CONSTANTE EN MODELO, en MÉTODO seleccionar ATRÁS:WALD, CONTINUAR y ACEPTAR.

De la Fuente-Fernández (2011) indica el visor de resultados de SPSS (Figura 2) muestra un cuadro con el número de casos introducidos, para este ejemplo el número de registros 40 corresponde al 100 %, los seleccionados para el análisis y los excluidos 0 casos perdidos, por tener algún valor faltante. La codificación de la variable dependiente dicotómica, internamente el programa asigna el valor 0 al menor de los dos códigos, y el valor 1 al mayor.

La codificación de las categorías de la variable independiente (género) y la frecuencia absoluta de cada valor. Si en el cuadro de definir variables categóricas se ha seleccionado en contraste indicador y en categoría de referencia última (opciones que da el programa por defecto), la categoría codificada con el valor interno más bajo (0) será la de referencia la 'última' para el SPSS. La tabla de clasificación permite evaluar el ajuste del modelo de regresión (hasta este momento, con un solo parámetro en la ecuación), comparando los valores predichos con los valores observados. Por defecto se ha empleado un punto de corte (0.5) de la probabilidad de Y para clasificar a los individuos. Esto significa que aquellos sujetos para los que la ecuación, con



**Resumen del procedimiento de los casos**

Casos no ponderados		N	Porcentaje
	Incluidos en el análisis	40	100.0
Casos seleccionados	Casos perdidos	0	.0
	Total	40	100.0
Casos no seleccionados		0	.0
Total		40	100.0

**Codificación de la variable dependiente**

Valor Original	Valor Interno
No	0
Si	1

**Codificación de las variables categóricas**

		Frecuencia	Codificación de parámetros
			(1)
Género	Hombre	14	1.000
	Mujer	26	.000

**Tabla de clasificación<sup>a,b</sup>**

Observado			Pronosticado		Porcentaje correcto
			suscribe		
			No	si	
Paso 0	suscribe	No	0	20	.0
		Si	0	20	100.0
Porcentaje global					50.0

<sup>a</sup> La constante es incluida en el modelo, <sup>b</sup> El valor de corte es .500

**Variables en la ecuación**

		B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 0	Constante	.000	.316	.000	1	1.000	1.000

**Variables que no están en la Ecuación**

		Puntuación	gl	Sig.	
Paso 0	Variables	edad	7.483	1	.006
		género(1)	7.033	1	.008
	Estadísticos globales		13.640	2	.001

Figura 2. Salida del programa SPSS. Regresión logística.



éste único término, calcula una probabilidad  $< 0.5$  se clasifican como suscribe=0 (no se suscriben), mientras que si la probabilidad resultante es  $\geq 0.5$  se clasifican como suscribe=1 (se suscribe). En este primer paso el modelo ha clasificado correctamente a un 50 % de los casos, y ningún caso de 'no se suscribe' ha sido clasificado correctamente.

Variables en la ecuación, en este primer bloque, en la ecuación de regresión sólo aparece el parámetro estimado , el error estándar E.T.= 0.316 y la significación estadística con la prueba de Wald, que es un estadístico que sigue una distribución Chi-cuadrado con 1 grado de libertad y la estimación de la OR = = . Variables que no están asociadas en la ecuación figura la significación estadística asociada al índice de Wald.

La prueba Ómnibus (Figura 3) ofrece tres entradas: la primera fila (PASO) es la correspondiente al cambio de verosimilitud (de -2LL) entre pasos sucesivos en la construcción del modelo, contrastando la hipótesis nula de que los coeficientes de las variables añadidas en el último paso son cero. La segunda fila (BLOQUE) es el cambio en -2LL entre bloques de entrada sucesivos durante la construcción del modelo. Si como es habitual en la práctica, se introducen las variables en un solo bloque, la Chi-Cuadrado del Bloque es el mismo que la Chi-Cuadrado del modelo. Seguidamente, tres medidas resumen de los modelos se utilizan para evaluar su validez de forma global.

El logaritmo de la verosimilitud (-2LL) mide hasta qué punto el modelo se ajusta bien a los datos. El resultado de esta medición recibe también el nombre



**Pruebas ómnibus sobre los coeficientes del modelo**

		Chi-cuadrado	gl	Sig.
Paso 1	Paso	16.471	2	.000
	Bloque	16.471	2	.000
	Modelo	16.471	2	.000

**Resumen de los modelos**

Paso	-2 Log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	38.981 <sup>a</sup>	.338	.450

<sup>a</sup> La estimación ha finalizado en la iteración número 5 porque las estimaciones de los parámetros han cambiado en menos de .001.

**Prueba de Hosmer y Lemeshow Test**

Paso	Chi-cuadrado	gl	Sig.
1	11.307	8	0.185

**Tabla de contingencia para la prueba de Hosmer y Lemeshow**

	suscribe = no		suscribe = si		Total	
	Observado	Esperado	Observado	Esperado		
Paso 1	1	2	2.84	1	0.16	3
	2	3	2.753	0	0.247	3
	3	3	3.306	1	0.694	4
	4	4	3.312	1	1.688	5
	5	3	2.387	1	1.613	4
	6	4	2.355	1	2.645	5
	7	0	1.522	4	2.478	4
	8	1	0.856	3	3.144	4
	9	0	0.509	4	3.491	4
	10	0	0.159	4	3.841	4

**Tabla de clasificación**

Observado		Pronosticado		
		suscribe		Porcentaje correcto
		no	si	
Paso 1	suscribe no	16	4	80
	si	4	16	80
Porcentaje global				80

<sup>a</sup> El valor de corte es .500

Figura 3. Salida del programa SPSS. Continuación de regresión logística.



de desviación. Cuanto más pequeño sea el valor, mejor será el ajuste. La de Cox y Snell es un coeficiente de determinación generalizado que se utiliza para estimar la proporción de varianza de la variable dependiente explicada por las variables predictoras. La de Cox y Snell se basa en la comparación del logaritmo de la verosimilitud (LL) para el modelo respecto al logaritmo de la verosimilitud (LL) para un modelo de línea base. Los valores oscilan entre 0 y 1. La de Nagelkerke es una versión corregida de la de Cox y Snell. La de Cox y Snell tiene un valor máximo inferior a 1, incluso para un modelo "perfecto". La de Nagelkerke corrige la escala del estadístico para cubrir el rango completo de 0 a 1.

La prueba de Hosmer y Lemeshow con una significancia de 0.185, valor que permite no rechazar la hipótesis nula de un buen ajuste del modelo y su tabla de contingencia para contrastar los valores observados con los predichos, se observa que coinciden el 80 % de los valores.

SPSS ofrece las variables de la ecuación, los coeficientes de regresión y sus errores estándar (E.T.), el valor del estadístico de Wald para evaluar la hipótesis nula ( $\pi = 0$ ), la significación estadística asociada, y el valor de la  $OR = \exp(\beta)$  con sus intervalos de confianza (Figura 4). Los valores estimados para los coeficientes del modelo fueron:  $\beta_1$  además sus valores de p-asociados 0.008, 0.011 y 0.011 respectivamente, que fueron inferiores a 0.05. Al nivel de significación  $\alpha = 0.05$  se rechaza la hipótesis nula de que dichos coeficientes son nulos lo que significa que son válidos. Así como los  $\beta_2$  de género ( ) y edad ( ), de acuerdo a sus



Variables en la Ecuación								
	B	E. T.	Wald	gl	Sig.	Exp(B)	I.C. 95 % para EXP(B)	
							Inferior	Superior
género(1)	2.422	.956	6.422	1	.011	11.273	1.731	73.401
Paso 1 <sup>a</sup> edad	.165	.065	6.399	1	.011	1.179	1.038	1.340
Constante	-8.184	3.096	6.986	1	.008	.000		

□<sup>a</sup> Variable(s) introducida(s) en el paso 1: género, edad.

Figura 4. Salida del programa SPSS. Variables en la ecuación de regresión logística.

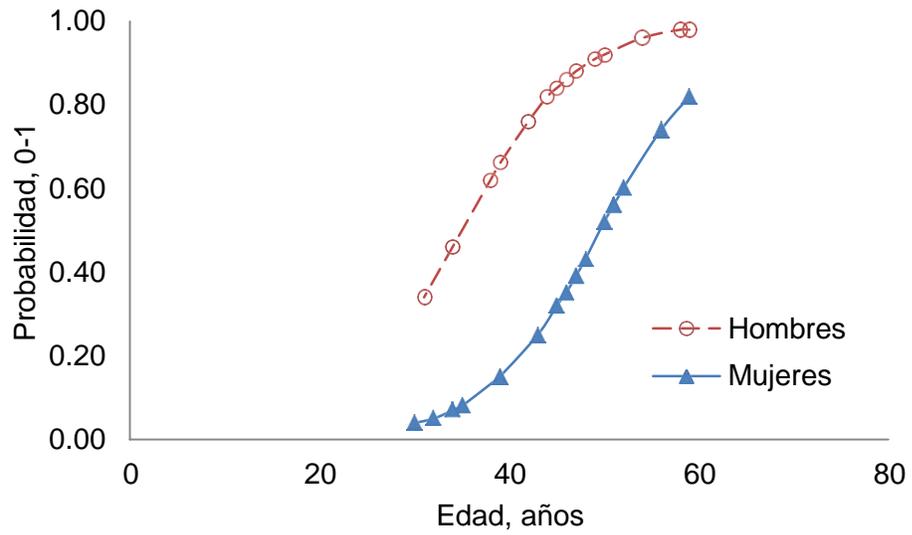




Cuadro 4. Probabilidades estimadas de suscripción por género y edad

sc	gen	ed	P	sc	gen	ed	P	sc	gen	ed	P
si	m	59	0.98	si	f	52	0.6	no	m	31	0.34
si	m	58	0.98	si	f	52	0.6	no	f	45	0.32
si	m	54	0.96	no	f	51	0.56	no	f	45	0.32
si	m	50	0.92	no	f	51	0.56	si	f	43	0.25
si	m	49	0.91	no	f	51	0.56	no	f	39	0.15
si	m	47	0.88	si	f	50	0.52	no	f	39	0.15
si	m	46	0.86	no	m	34	0.46	no	f	39	0.15
si	m	45	0.84	no	f	48	0.43	no	f	35	0.08
si	f	59	0.82	no	f	47	0.39	no	f	35	0.08
no	m	44	0.82	si	F	47	0.39	no	f	35	0.08
si	m	42	0.76	no	F	47	0.39	no	f	34	0.07
si	f	56	0.74	si	F	46	0.35	no	f	32	0.05
si	m	39	0.66	no	F	46	0.35	si	f	30	0.04
si	m	38	0.62								

sc=suscripción, gen=género (f=femenino, m=masculino), ed=edad, P=valor de probabilidad



Gráfica 2. Probabilidades estimadas de suscripción por género y edad con el modelo de regresión logística binaria.



## MATERIALES Y MÉTODOS

### Diseño de la Encuesta

El desarrollo del instrumento o encuesta, se llevó a cabo en varias reuniones entre personas representantes del SUBA y de estadística de la Facultad de Zootecnia y Ecología de la UACH, en las cuales se revisaron y ajustaron las preguntas y su estructura. Se emplearon las preguntas que responden a las necesidades del SUBA para sus procesos de mejora continua y/o sugerencias del organismo acreditador. Se limitó el uso de preguntas abiertas, empleándose preguntas cerradas dicotómicas y de respuestas en la escala conocida como Likert que permite conocer el grado de conformidad del encuestado.

La encuesta (Anexo 1) inició con una invitación al encuestado, garantizando confidencialidad de los datos. Fueron 38 reactivos que se dividieron en tres secciones: Información general. Donde se identifican algunos antecedentes y rasgos básicos (seis reactivos del 1.1 al 1.6); Servicios Bibliotecarios. Para evaluar qué proporción de la población hace uso o no de los servicios y con qué frecuencia (seis reactivos del 2.1 al 2.6) y calificación de los servicios bibliotecarios. Se califican los servicios como préstamo interno y externo, servicios electrónicos de información y servicios de apoyo (26 reactivos del 3.1 al 3.18). Al final se agradece al encuestado y se agregó un espacio para que expresaran posibles sugerencias.

### Diseño de Muestreo

**Elemento de muestreo.** Se definió como elemento de muestreo al estudiante de nivel licenciatura en modalidad presencial inscrito en alguna de



las 16 unidades académicas en el ciclo escolar enero-junio 2013.

**Tipo de diseño.** Se determinó que la variable auxiliar más fuertemente relacionada a la variable de interés analizada en el presente estudio fue Facultad. Los elementos de la población se observaron heterogéneos debido a Facultad, por lo que fue necesario estratificar. Se utilizó el muestreo aleatorio estratificado en donde Facultad constituyó un estrato. Luego, dentro de cada estrato los elementos quedaron homogéneos para la variable de interés, se redujo la variabilidad y se obtuvo una mayor precisión de los estimadores. Todos los elementos de la población tuvieron igual probabilidad de ser seleccionados La Unidad de Servicios Escolares de la Dirección Académica proporcionó 23,541 registros con: matrícula, nombre y Facultad correspondientes al total de alumnos inscritos en nivel licenciatura de modalidad presencial en el ciclo escolar enero-junio 2013. Con los que se procedió a calcular el tamaño de muestra total y por estrato.

### Tamaño de Muestra

Entre las tres técnicas definidas para distribuir la muestra de tamaño n a los L estratos, se eligió la asignación proporcional ya que se solo se conocía el tamaño de los estratos, es decir el número de estudiantes de cada Facultad. De acuerdo a Scheaffer (1987) se detallaron las formulas:

$$\frac{n}{L} = \frac{\sum N_h}{\sum 1}$$

De acuerdo a esta asignación se calcularon primero ( ), conocidos ( ) número de elementos en cada estrato y , luego el tamaño de muestra



total ( ) con una confiabilidad del 95 %, ( ) y un margen de error deseado no mayor al 3 % ( ), valores para y (se fijó la probabilidad de éxito en 0.5 para maximizar el tamaño de muestra, ya que no se contó con un valor anterior),  $L=16$  estratos, y finalmente se calcularon ( ). En el Cuadro 5 se presenta el tamaño de la muestra total ( ) estratificada y el tamaño de muestra por Facultad ( ) o estrato.

**Selección de la muestra aleatoria por estrato ( )** Se ingresaron los registros por Facultad (matrícula, nombre) al paquete de software Minitab-16, en el cual se seleccionó de manera aleatoria (Scheaffer, 1987) la lista de nombres de los estudiantes inscritos seleccionados por Facultad. El proceso se realizó dos veces para generar una lista de reemplazos a usarse en caso de no encontrar alguno. Estas listas generadas se entregaron a cada uno de los jefes de biblioteca en las Facultades.

**Aplicación de la encuesta.** Se solicitó el apoyo al(a) Secretario(a) Académico(a) en cada Facultad para localizar a los estudiantes seleccionados en la muestra y acordar con ellos quiénes aplicarían la encuesta. Se acordó que los jefes de biblioteca y bibliotecarios aplicaran la encuesta a los estudiantes seleccionados en un lugar distinto a la biblioteca.

**Instrucciones para encuestadores.** Se les pidió presentarse con el alumno e indicarle que fue seleccionado para participar de manera voluntaria en el estudio, pedirle leer detenidamente la encuesta, contestar con letra clara y legible y entregarla a los encuestadores. Señalar que la información registrada se considerará en forma anónima y solo para los objetivos del curso. Se aplicó



Cuadro 5. Tamaños poblacionales de los estratos, tamaños de muestra estratificada y muestra aleatoria por estrato, y el número de encuestas recuperadas

DES		Facultad (estrato)	Alumnos Inscritos	Muestra Calculada	Encuestas recabadas
			( )	( )	
EC	1	Artes	704	31	13
AG	2	Ciencias Agrícolas	510	22	20
AG	3	Ciencias Agrotecnológicas	725	31	29
*	4	Ciencias Políticas y Sociales	2152	93	44
ING	5	Ciencias Químicas	1387	60	55
CAE	6	Contaduría y Administración	3856	167	151
CAE	7	Contaduría y Administración ext. Delicias	591	26	24
DE	8	Derecho	1976	86	78
CAE	9	Economía Internacional	117	5	5
SA	10	Ciencias de la Cultura Física	2482	108	96
SA	11	Enfermería y Nutriología	1396	61	55
EC	12	Filosofía y letras	760	33	30
ING	13	Ingeniería	3497	152	137
SA	14	Medicina	1155	50	46
SA	15	Odontología	1047	45	41
AG	16	Zootecnia y Ecología	1186	51	38
Total			23541	1021	862

DES: Dependencias de Educación Superior, AG: Agropecuarias, CAE: Contaduría y Administración, Economía Internacional, D:Derecho, EC: Educación y Cultura, ING: Ingeniería, SA: Salud, \*sin clasificar, : número de elementos en cada estrato, : tamaño de muestra por estrato.



la encuesta a cada estudiante por separado para evitar ser influenciados unos a otros. El período de aplicación de la encuesta fue del 28 de mayo al 30 de agosto del 2013. Las encuestas recabadas fueron entregadas en la Coordinación General del SUBA, en donde se realizó la captura de datos.

### **Análisis de los Datos**

**Depuración de los datos.** Se recabaron y capturaron 862 encuestas (Cuadro 5). Se procedió a verificar la uniformidad y consistencia de los datos capturados encontrándose que debido a un error de impresión de la encuesta en la Facultad de Ciencias Políticas las respuestas no se completaron de manera correcta, por lo que fueron invalidadas 44. En la primera sección, se registró el nombre de la Facultad y se registró en la captura. Para los objetivos de esta tesis se decidió analizar 17 variables sobre diferentes aspectos.

La selección de variables (Cuadro 6) se realizó de acuerdo a Hair *et al.* (2001) quienes sugirieron que si el 30 % de las encuestas no fueron contestadas en un reactivo, éste sea eliminado del análisis. Por ejemplo en la pregunta 2.5 el 65 % de los encuestados no respondieron (por error de secuencia en la redacción, en el reactivo 2.3 menciona "Si contestaste No, pasa al 2.6 y debió haber enviado a la 2.5), y la 3.15 ya que el 34 % no la respondieron.

Con fines del análisis se recategorizaron seis variables (Cuadro 6) para formar tablas de contingencia sin celdas vacías y valores menores a cinco (Agresti, 2002). Las 16 facultades de la UACH se clasificaron en seis DES (Cuadro 5 y 6). Edad de los estudiantes en dos categorías, de 17-24 (1) y de 25 años en adelante (2). De una lista de 86 lugares de origen se categorizaron los



Cuadro 6. Variables seleccionadas para el análisis donde algunas se recategorizaron y la considerada variable respuesta

Variable	No. de Reactivo	Descripción	Categorías	
			Antes	Después
gen	Sin	Facultad a DES	16	6
	1.1	Género	2	2
	1.2	Edad en años	Continua	2
sem	1.3	Semestre	3	3
	1.4	Lugar de origen	86	4
	1.5	Sistema de bachillerato	9	3
emp	1.6	Ser empleado	2	2
vis	2.1	Haber visitado SUBA	2	2
	2.2	Frecuencia de visita al SUBA	5	2
	2.3	Acudir a otras bibliotecas	2	2
ob	2.5	Frecuencia de actividades realizadas en SUBA	5	5
	3.2	Asistencia al curso de inducción	2	2
	3.6	Identificar servicios del SUBA	2	2
svs	3.7	Calificación a los servicios del SUBA	5	2
	3.8	La biblioteca cumple el horario	2	2
	web	3.11	Conocer y haber usado la página del SUBA	2
3.15		Acceder a las bases de datos	2	2

a=Datos ausentes en más del 30 % del total de registros, b= Variables recategorizadas, c=Variable respuesta



diferentes sistemas de educación media superior se agruparon cuatro categorías: municipio de Chihuahua (1), otro municipio (2), otro Estado (3) y otro país (4). Los diferentes sistemas de educación media superior se agruparon en tres categorías: COBACH (1); CBTIS, CECYT, CONALEP, CBTA y CETIS (2); preparatoria particular y otros (3). La frecuencia con que los estudiantes acuden al SUBA se registró en cinco categorías y se reagrupó en dos: una vez al día, varias veces a la semana y varias veces al mes (1); pocas veces y sólo en periodo de exámenes (2).

Debido a que el objetivo de este estudio fue evaluar la satisfacción de los usuarios del SUBA se definió la calificación a los servicios del SUBA (califbin) como la variable respuesta para el análisis de regresión logística binaria. Se recategorizó de cinco categorías a dos: buena (buena y muy buena) y mala (muy mala, mala y regular).

Después de la depuración quedaron 818 registros válidos, de los 862 capturados. Al reducirse el tamaño de muestra calculada de 1021 a 818 (Cuadro 5) se incrementó a un 3.4 % el margen de error deseado inicialmente no mayor al 3 %. Con la depuración de variables se redujeron de 17 a 15. En el Cuadro 7 se muestra una descripción general de la integración de la muestra de estudio (n=818).

**Prueba y odds ratio.** La información fue tabulada y analizada en el paquete SPSS Vs20. Se crearon tablas de contingencia, se realizaron las pruebas estadísticas ( $p \leq 0.05$ ) y Odds Ratio para analizar la asociación entre dichas variables y la variable respuesta.



Cuadro 7. Distribución de la muestra estratificada de estudio (n = 818 estudiantes) en las variables seleccionadas y categorías definidas

	<b>Variable</b>	<b>Categorías</b>	<b>n</b>	<b>%</b>
1	DES	Educación y cultura	43	5.3
		Agropecuarias	87	10.6
		Ingeniería	192	23.5
		Derecho	78	9.5
		CAE	180	22
		Salud	238	29.1
2	Género	Hombre	371	45.4
		Mujer	445	54.4
		Sin respuesta	2	0.2
3	Edad (años)	17-24	755	92.3
		25- en adelante	57	7
		Sin respuesta	6	0.7
4	Semestre	1° a 3°	330	40.3
		4° a 6°	238	29.1
		7° a 10°	246	30.1
		Sin respuesta	4	0.5
5	Lugar de origen	Mpo. de Chih.	509	62.2
		Otro mpo. del Edo. de Chih.	250	30.6
		Otro Edo.	38	4.6
		Otro país	4	0.5
		Sin respuesta	17	2.1
6	Sistema de bachillerato	COBACH	387	47.3
		CBTIS, CECYT, CONALEP, CBTA, CETIS	212	25.9
		particular u otros	212	25.9
		Sin respuesta	7	0.9
7	Trabaja	si	204	24.9
		no	583	71.3
		Sin respuesta	31	3.8



Cuadro 7 Continuación. Distribución de la muestra estratificada de estudio (n = 818 estudiantes) en las variables seleccionadas y categorías definidas

	<b>Variable</b>	<b>Categorías</b>	<b>n</b>	<b>%</b>
8	Haber visitado el SUBA	si	760	92.9
		no	57	7
		Sin respuesta	1	0.1
9	Frecuencia de visita	mayor	340	41.6
		menor	416	50.9
		Sin respuesta	62	7.6
10	Acudir a otras bibliotecas	si	201	24.6
		no	570	69.7
		Sin respuesta	47	5.7
11	Curso de inducción	si	456	55.7
		no	323	39.5
		Sin respuesta	39	4.8
12	Identificar los servicios SUBA	si	658	80.4
		no	153	18.7
		Sin respuesta	7	0.9
13	Calificación a los servicios	Muy Buena	387	47.3
		Buena	354	43.3
		Regular	53	6.4
		Mala	8	1
		Muy Mala	8	1
		Sin respuesta	8	1
14	La biblioteca cumple el horario	si	771	94.3
		no	34	4.2
		Sin respuesta	13	1.6
15	Conocer y usar la web SUBA	si	580	70.9
		no	233	28.5
		Sin respuesta	5	0.6

CAE: Contaduría y Administración, Economía Internacional, COBACH: Colegio de Bachilleres del Estado de Chihuahua, CBTIS: Centro de Bachillerato Tecnológico Industrial y de Servicios, CECYT: Centro de Estudios Científicos Y tecnológicos, CONALEP: Colegio Nacional de Educación Profesional Técnica, CBTA: Centro de Bachillerato Tecnológico Agropecuario, CETIS: Centro de Estudios Tecnológicos Industrial y de Servicios.



**Regresión logística binaria.** Luego se procedió a identificar la existencia de una dependencia multivariada. Se utilizó el paquete SPSS Vs20 ajustándose el modelo de regresión logística (Anexo 4), para (califbin) en función de las variables que resultaron dependientes en las pruebas anteriores. De acuerdo a Silva y Barroso (2004) la ecuación del modelo es la siguiente:

$$P(Y=1|X) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}$$

donde:  $P(Y=1|X)$  es la probabilidad de que tome el valor de 1 (presencia de una “buena calificación” dado el valor de las covariables  $X$ ).

$X$  = es un conjunto de covariables que forman parte del modelo

$\beta_0$  = es la constante del modelo o término independiente

$\beta_1, \dots, \beta_k$  = los coeficientes de las covariables

De esta manera se obtuvo la predicción de la variable respuesta al estimar la probabilidad de tener una buena calificación. Los valores de probabilidad estimados se ordenaron para describir los factores que influyeron significativamente la calificación de los servicios del SUBA.



## RESULTADOS Y DISCUSIÓN

### Datos Descriptivos de la Población que visita el SUBA

Los resultados obtenidos indican que el 93 % de la población estudiantil han visitado las bibliotecas del SUBA, valor que revela que la biblioteca es un servicio conocido y utilizado por los estudiantes. Resultado que concuerda con los datos aportados por otros estudios como el de Bancroft (1998) para la Universidad Estatal de Washington, donde se señala que 90.8 % visitó alguna vez la biblioteca y más recientemente, Hayden *et al.* (2005), señalan que 98 % de los estudiantes del Instituto Tecnológico de Waterford visitó alguna vez la biblioteca durante el curso 2002-2003. Esto muestra que la proporción de los estudiantes que nunca han visitado la biblioteca tiende a ser baja.

Se encontraron asociaciones significativas ( $P < 0.05$ ) de acuerdo con la prueba chi-cuadrada, con el hecho de “haber visitado al SUBA” (sí/no) y las variables: DES, género, semestre, sistema de bachillerato, curso de inducción, identificar los servicios y conocer y/o usar la web (Cuadro 8).

En las Facultades de Artes y Filosofía y Letras el 100 % de los estudiantes visitaron las bibliotecas del SUBA, por el contrario la DES de salud (Ciencias de la Cultura Física, Enfermería y Nutriología, Medicina y Odontología) 86.6 % las visitaron. Considerando el género, 95.5 % de las mujeres visitaron las bibliotecas del SUBA mientras que los hombres 90.3 %. Estos resultados coinciden con Grimes y Charters (2000), quienes señalan que los hombres tienden a estar menos horas en la biblioteca que las mujeres. Los alumnos de primero a tercer semestre son los que menos han visitado (89.1 %), incrementándose a 96 % los de cuarto en adelante. Del sistema de bachillerato



Cuadro 8. Proporciones del uso del SUBA con las variables explicatorias significativas ( )

<b>Variable</b>	<b>Categorías</b>	<b>%</b>
DES: Facultades	EC: Artes, Filosofía y Letras	100.0
	AG: Ciencias Agrícolas, Ciencias Agrotecnológicas, Zootecnia y Ecología	97.7
	ING: Ciencias Químicas, Ingeniería	97.4
	CAE: Contaduría y Administración, Economía Internacional	91.1
	SA: Ciencias de la Cultura Física, Enfermería y Nutriología, Medicina y Odontología	86.6
	Género	Mujer
Semestre	Hombre	90.3
	1° a 3°	89.1
	4° a 6°	96.2
Sistema de bachillerato	7° a 10°	95.9
	COBACH	96.6
	CBTIS, CECYT, CONALEP, CBTA, CETIS particular u otros	91.5
Curso de inducción	si	88.7
	no	96.9
Identificar los servicios SUBA	si	89.2
	no	94.7
Conocer y usar la web SUBA	si	85.5
	no	95.0
		87.9

COBACH: Colegio de Bachilleres del Estado de Chihuahua, CBTIS: Centro de Bachillerato Tecnológico Industrial y de Servicios, CECYT: Centro de Estudios Científicos Y tecnológicos, CONALEP: Colegio Nacional de Educación Profesional Técnica, CBTA: Centro de Bachillerato Tecnológico Agropecuario, CETIS: Centro de Estudios Tecnológicos Industrial y de Servicios.



particular u otros son los que reflejan menor porcentaje de visita (88.7 %). Es notoria la importancia el hecho de haber tomado el curso de inducción, identificar los servicios y el de conocer y/o usar la web SUBA, pues más del 94 % de ellos han visitado el SUBA.

### **Prueba de y Odds Ratio**

Con respecto a la variable respuesta (califbin) se observó que el 90.6 % de la muestra obtuvo una calificación satisfactoria (“buena” y “muy buena”, en la escala de Likert utilizada en la encuesta) a los servicios que ofrece el SUBA, 8.4

% “mala” y 1 % restante no contestó. La asociación dada a la calificación (buena/mala) con ocho de las variables explicatorias se muestra en el Cuadro 9.

Al menos una de las seis DES ( $P < 0.009$ ) tiene relación con califbin. De acuerdo a los valores de OR; hro, resultó ser el factor más importante para obtener una calificación satisfactoria a los servicios del SUBA, seguida de sv, cur, vis, fre, ob y gen. Indica que las personas que consideran que se cumple el horario son cinco veces más afectas a los servicios del SUBA, que las que opinan que existe incumplimiento.

Los que identifican los servicios están casi cinco veces más a favor de los servicios del SUBA, que los que no lo hacen. Las personas que asistieron al curso de inducción, son dos veces más tendientes a dar una buena calificación que aquellas que no acudieron. Las personas que han visitado, tienen dos veces más interés por estos servicios que las que no han visitado. Se encontró que los que acuden de manera poco frecuente (pocas veces y/o solo en periodo de exámenes), son casi dos veces más propensos a otorgar una buena calificación a los servicios, que los que acuden más frecuentemente (diario,



Cuadro 9. Dependencias probadas mediante la prueba Chi-cuadrada y Odds Ratio a la calificación de los servicios del SUBA

<b>Variable</b>	<b>Categorías</b>		<b>P-valor</b>	<b>OR [ IC 95 % ]</b>
DES	Seis	15.269	0.009	*
Gen	fem/mas	4.602	0.032	1.719 [1.043, 2.833]
Vis	sí/no	4.387	0.036	2.120 [1.034, 4.731]
Fre	menor/mayor	5.603	0.018	1.925 [1.111, 3.333]
Ob	sí/no	5.103	0.024	1.865 [1.078, 3.225]
Cur	sí/no	9.510	0.002	2.260 [1.331, 3.836]
Svs	sí/no	38.760	0.000	4.568 [2.736, 7.628]
Hro	sí/no	18.786	0.000	5.142 [2.273, 11.63]

P-valor: significancia estadística, OR: valor de odds ratio, IC 95 %: Límites de confianza para los valores de OR. \* Solo se determinó dependencia por el valor de chi-cuadrada ya que el OR se calculó para tablas de contingencia de 2X2. Se calculó este OR en el análisis de regresión logística.



varias veces a la semana y/o varias veces al mes) y que las personas que acuden a bibliotecas externas, tienden a dar casi dos veces más una buena calificación al SUBA que las que no lo hacen. Por último indica que las mujeres son casi dos veces más afectas a los servicios de la biblioteca que los hombres.

### Regresión Logística Binaria

De acuerdo a los resultados de la regresión logística se probó una asociación estadística con seis de las ocho variables estudiadas. Se ajustó el siguiente modelo de regresión logística para predecir la probabilidad de una calificación satisfactoria a los servicios del SUBA:

$$\frac{\left( \right)}{\left( \right)}$$

Donde:

Dependencias de Educación Superior

Género

Frecuencia de visita al SUBA

Visitar otras bibliotecas

Haber asistido al curso de inducción

Identificar los servicios que ofrece el SUBA

En el Cuadro 10 se muestran los valores estimados para los coeficientes de la ecuación del modelo, además sus valores P-asociados ( $P < 0.05$ ) al nivel de significación, se rechaza la hipótesis nula de que dichos coeficientes son nulos. Así como los OR y que de acuerdo a sus intervalos de confianza, resultaron ser significativos al no contener a la unidad (Barón y Téllez, 2004). Se probó que hay un efecto debido a seis factores: DES, gen, fre,



Cuadro 10. Estimación de los coeficientes del modelo y de los OR en las variables importantes

Variable	Categoría		P	OR	IC95 %
DES	AG	1.865	0.016	6.458	1.415 - 29.487
	SA	Categoría de referencia			
gen	Fem	1.083	0.002	2.954	1.488 - 5.866
	Mas	Categoría de referencia			
fre	Men	0.689	0.041	1.991	1.029 - 3.852
	May	Categoría de referencia			
ob	si	0.876	0.013	2.401	1.202 - 4.796
	no	Categoría de referencia			
cur	si	0.794	0.021	2.212	1.126 - 4.346
	no	Categoría de referencia			
svs	si	1.532	<0.001	4.626	2.386 - 8.968
	no	Categoría de referencia			
C		-0.83	0.073	0.435	

: Coeficientes de la ecuación de regresión logística, P-valor: significancia estadística, OR: valor de odds ratio, IC 95 %: Límites de confianza para los valores de OR.



ob, cur, sv. Adicional al hecho de saber cuáles son los factores relacionados a la calificación; los coeficientes estimados de cada uno de ellos indican el orden de importancia de los mismos. En orden descendente: DES (1.865) se encontró como el factor más importante para obtener una buena calificación. En términos prácticos indica que se requiere principal atención a los estudiantes DES de salud, por lo que se sugiere revisar más de cerca sus sugerencias, necesidades o inconformidades ya que este estudio representa a la población estudiantil y todas las DES en conjunto conforman la calificación que se obtenga en posteriores evaluaciones.

Identificar los servicios (1.532). Los estudiantes que identifican los servicios, opinan mejor que los que no los conocen por lo que se sugiere realizar una mayor difusión (véase Cuadro 8 los que menos han visitado) principalmente en los estudiantes de género (1.083) masculino para lograr que la población estudiantil total conozcan todos los servicios que ofrece el SUBA. Se encontró que al visitar otras bibliotecas (0.876) los estudiantes están comparando la calidad de los servicios por lo que se concluye que encontraron mejor al SUBA La importancia de que los estudiantes tomen el curso de inducción (0.794) resalta nuevamente ya que los que asistieron califican mejor que los que no asistieron. Se reitera positivo el impartirlo y se sugiere buscar estrategias para incrementar el número de estudiantes que asistan.

Respecto a la frecuencia (0.689) de visita al SUBA. Los estudiantes que asisten con menor frecuencia (“pocas veces” y/o “solo en periodos de exámenes”) califican mejor que los que asisten con mayor frecuencia (“una vez al día”, “varias veces a la semana”, “varias veces al mes”). La explicación es



que por ir menos veces es más probable que todo lo haya encontrado satisfactorio a comparación del que acude muchas veces, puede ser más probable encontrar alguno de los servicios como insatisfactorio.

Se hace referencia a los valores más altos (Cuadro 11a) de estas probabilidades predichas obtenidos a través de la ecuación de regresión logística; en donde la probabilidad más alta para obtener una calificación satisfactoria (“buena” y “muy buena” ,en la escala de Likert utilizada en la encuesta) de los servicios del SUBA está en aquellos estudiantes que pertenecen a la DES de agropecuarias, que son mujeres, que visitan con menor frecuencia las bibliotecas del SUBA, que acuden a bibliotecas externas, que asistieron al curso de inducción y que identifican los servicios que ofrece el SUBA.

La probabilidad más baja (Cuadro 11b) para obtener una calificación satisfactoria (“buena y “muy buena”), está en aquellos estudiantes que pertenecen a la DES de salud, que son hombre, que visitan con mayor frecuencia las bibliotecas del SUBA, que no acuden a bibliotecas externas, que no asistieron al curso de inducción y no identifican los servicios que ofrece el SUBA.



Cuadro 11a. Ejemplos de los valores más altos de probabilidad predicha de acuerdo a la ecuación logística binaria de obtener una buena calificación de los servicios del SUBA

DES	Variables					P > 0.9945
	gen	fre	ob	cur	svs	
DER	fem	men	SI	SI	SI	0.9976
DER	fem	men	SI	SI	SI	0.9976
EC	fem	men	SI	SI	SI	0.9976
EC	fem	men	SI	SI	SI	0.9976
EC	fem	men	SI	SI	SI	0.9976
AG	fem	men	SI	SI	SI	0.9975
AG	fem	men	SI	SI	SI	0.9975
AG	fem	men	SI	SI	SI	0.9975
DER	fem	may	SI	SI	SI	0.9953
DER	fem	may	SI	SI	SI	0.9953
DER	fem	may	SI	SI	SI	0.9953
EC	fem	may	SI	SI	SI	0.9952
EC	fem	may	SI	SI	SI	0.9952
AG	fem	may	SI	SI	SI	0.9951
AG	fem	may	SI	SI	SI	0.9951
AG	fem	may	SI	SI	SI	0.9951
DER	fem	men	SI	NO	SI	0.9947
DER	fem	men	SI	NO	SI	0.9947
DER	fem	men	SI	NO	SI	0.9947
DER	fem	men	SI	NO	SI	0.9947
DER	fem	men	SI	NO	SI	0.9947
EC	fem	men	SI	NO	SI	0.9947
EC	fem	men	SI	NO	SI	0.9947
AG	fem	men	SI	NO	SI	0.9946
AG	fem	men	SI	NO	SI	0.9946
AG	fem	men	SI	NO	SI	0.9946
AG	fem	men	SI	NO	SI	0.9946
AG	fem	men	SI	NO	SI	0.9946
AG	fem	men	SI	NO	SI	0.9946



Cuadro 11 b. Ejemplos de los valores más bajos de probabilidad predicha de acuerdo a la ecuación logística binaria de obtener una buena calificación de los servicios del SUBA

DES	Variables					P < 0.6900
	gen	fre	ob	cur	svs	
SA	mas	men	SI	NO	NO	0.6752
SA	mas	men	SI	NO	NO	0.6752
SA	mas	men	SI	NO	NO	0.6752
SA	mas	may	NO	NO	SI	0.6680
SA	mas	may	NO	NO	SI	0.6680
SA	mas	may	NO	NO	SI	0.6680
SA	mas	may	NO	NO	SI	0.6680
ING	fem	may	NO	NO	NO	0.6625
CAE	mas	may	SI	NO	NO	0.6360
CAE	mas	may	NO	SI	NO	0.6168
CAE	mas	may	NO	SI	NO	0.6168
ING	mas	may	SI	NO	NO	0.6147
ING	mas	may	SI	NO	NO	0.6147
ING	mas	may	SI	NO	NO	0.6147
ING	mas	may	SI	NO	NO	0.6147
ING	mas	men	NO	NO	NO	0.5695
ING	mas	men	NO	NO	NO	0.5695
SA	fem	may	NO	NO	NO	0.5623
SA	fem	may	NO	NO	NO	0.5623
SA	fem	may	NO	NO	NO	0.5623
SA	fem	may	NO	NO	NO	0.5623
SA	fem	may	NO	NO	NO	0.5623
SA	mas	may	SI	NO	NO	0.5109
SA	mas	may	SI	NO	NO	0.5109
SA	mas	may	SI	NO	NO	0.5109
SA	mas	may	NO	SI	NO	0.4903
SA	mas	may	NO	SI	NO	0.4903
SA	mas	men	NO	NO	NO	0.4640
SA	mas	may	NO	NO	NO	0.3031



## **CONCLUSIONES Y RECOMENDACIONES**

Se logró cuantificar la satisfacción del usuario de los servicios del SUBA a través de la metodología estadística propuesta para ello.

El análisis de los datos categóricos aplicado a la información generada con el instrumento diseñado para el muestreo permitió identificar y cuantificar los factores principales que inciden en la satisfacción del usuario del SUBA.

Es importante que en todo proceso de mejora continua se diseñe la metodología adecuada para mostrar la satisfacción de los usuarios sobre los servicios que ofrece una institución.



## LITERATURA CITADA

- Agresti, A. 2002. *Categorical Data Analysis*. John Wiley & sons. 2ª ed. University of Florida, Gainesville, Florida.
- Alfaro-Torres P. y S. Sánchez García. 2005. La importancia de la biblioteca universitaria en la formación académica. Universidad de Castilla La Mancha. Real, España.
- Bancroft, A. F., V. F. Croft, R. Speth y D. M. Phillips. 1998. A forward-looking library user survey. *WSU libraries in the 21st Century. J. of Academic Librarianship*. 24: 216-224.
- Barón-López F. y F. Téllez Montiel. 2004. Apuntes de Bioestadística. Universidad de Málaga. 52-57.
- Cochran, W. 1986. *Técnicas de muestreo*. 6ª Edición. Editorial Cecsca. México.
- De la Fuente-Fernández, S. 2011. *Regresión Logística*. Facultad de Ciencias Económicas y Empresariales. Universidad Autónoma de Madrid, España.
- Grimes, P. W. y M. F. Charters. 2000. Library use and the undergraduate economics student. *College Student Journal*. 24, 4, 557-570.
- Hair Jr. F. Joseph, E. Rolph Anderson, L. Ronald, T. y W. Black 2001. *Análisis Multivariante*. 5a. Edición. Editado por Prentice Hall.
- Hayden, H., T. O'brien y M. O. Rathaille. 2005. User survey at Waterford Institute of Technology Libraries. *New Library World*. 106, 1208/1209, 43-57.
- Pérez, A. A. J., R. Kizys y L. M. Manzanedo. 2004. *Regresión logística binaria*.
- Sánchez-Carrión, J. 1989. *Análisis de tablas de contingencia*. Centro de Investigaciones Sociológicas. Madrid, España.
- Sapién-Aguilar, A.L., L.C. Piñón-Howlet y M.C. Gutiérrez-Diez. 2011. *Análisis de la satisfacción del Sistema Universitario de Bibliotecas Académicas*. Caso de estudio: Facultad de Contaduría y Administración de la Universidad Autónoma de Chihuahua, México.
- SAS. 2009. *SAS/STAT User's Guide*. SAS Inst. Inc., Cary NC, USA.
- Scheaffer, R. L., W. Mendenhall y L. Ott. 1987. *Elementos de muestreo*. Grupo editorial Iberoamérica. México.
- Silva, L. C. e I. M. Barroso Utra. 2004. *Regresión Logística*. Editorial La Muralla



Hespérides. España.

URL1: <http://es.slideshare.net/garciamarcelli/requisitos-iso-90012008?related=1>.  
Consultado el 2 de julio de 2015.

Vicéns-Otero, J. y E. Medina-Moral, 2005. Análisis de datos cualitativos.



## **ANEXOS**



## Anexo 1. La encuesta, pagina 1 de 7



### Evaluación de la calidad del Sistema Universitario Bibliotecas Académicas (SUBA)

Estimado estudiante, te invitamos a responder esta encuesta con el objetivo de evaluar los servicios que presta el Sistema Universitario de Bibliotecas Académicas (SUBA). La información que nos proporciones será utilizada exclusivamente para el objetivo expuesto, y manejada en forma confidencial. Te pedimos seas lo mas objetivo posible en tus respuestas, de ello dependerá la confiabilidad de los indicadores que pretendemos obtener.

Instrucciones: Contesta las siguientes preguntas, marcando con una "X" la opción para la respuesta que consideres adecuada.

#### I. Información General

- 1.1. Género  Hombre  Mujer
- 1.2. Edad en años: \_\_\_\_\_
- 1.3. Indica el semestre que actualmente cursas:  
 Primero a tercero  Cuarto a sexto  Séptimo en adelante.
- 1.4. Escribe el nombre de tu lugar de origen: \_\_\_\_\_
- 1.5. Señala el sistema de educación media superior del cual egresaste:  
 Colegio de Bachilleres  CONALEP  
 CBTIS  CbTa  
 Preparatoria particular  Telebachillerato  
 CECYT  Otros \_\_\_\_\_
- 1.6. En este semestre ¿Tienes un trabajo con un horario formal?  SI  NO

#### II. Servicios Bibliotecarios

- 2.1. ¿Has visitado la biblioteca de la facultad donde estudias o alguna otra del SUBA?  
 SI  NO

**Nota:** Si contestaste que NO pasa a la pregunta 2.6

- 2.2. ¿Con qué frecuencia acudes a las del SUBA?  
 Una vez al día  Varias veces a la semana  Varias veces al mes  
 Pocas veces  Solo en período de exámenes

- 2.3. Además del SUBA ¿acudes a otras bibliotecas?  
 SI  NO

Si contestaste NO, pasa a la pregunta 2.6.



### Anexo 1. La encuesta, pagina 2 de 7



2.4. Señala las razones por las cuales acudes a otras bibliotecas diferentes al SUBA.

- Por cercanía a mi domicilio
- Por un horario más accesible
- Por un servicio que no encuentro en mi biblioteca
- Por la rapidez de sus servicios
- Por literatura especializada con mi carrera
- Por literatura general
- Por la atención especial de su personal
- Por el costo de sus servicios

2.5. ¿Con qué frecuencia acudes a la biblioteca a realizar las siguientes actividades?

(S=Siempre; MF= Muy Frecuente; F= Frecuentemente; E= Esporádicamente; N= Nunca)

Actividad	S	MF	F	E	N
Para hacer tareas					
Para usar el internet					
A leer libros, revistas, periódicos, etc.					
A reunirme con mis amigos.					
Para solicitar el préstamo de un libro.					
A tener un rato de esparcimiento					
A realizar búsquedas en las bases de datos del SUBA					
Para realizar una investigación con referencias bibliográficas					
Para solicitar un servicio de fotocopias, de impresión, de escaneo, entre otros.					
Para usar los sanitarios					
Para usar un cubículo					

2.6. Señala las razones por las que no has requerido los servicios de tu biblioteca

- Porque compro los libros de texto
- Horario inadecuado
- El personal nos trata mal
- No encuentro los libros que necesito
- Está muy lejos de mi casa
- Prefiero buscar en internet en mi casa

### III. Calificación de los servicios bibliotecarios

3.1. ¿Desde tu ingreso a la Universidad te informaron sobre los cursos de inducción a Biblioteca?

- SI
- NO

3.2. Después de invitado al curso ¿Asististe a él?:

- SI
- NO



### Anexo 1. La encuesta, pagina 3 de 7



Si tu respuesta fue SI pasa a la pregunta 3.3

Si tu respuesta fue NO en la pregunta anterior, escoge una o varias de las siguientes causas:

- Falta de interés de tu parte.
- El horario fue inadecuado.
- Desconocías el lugar donde lo ofrecieron.
- Te inscribiste después del curso.

3.3. ¿Cuándo te gustaría que se ofreciera el curso de inducción?

- Inicio del semestre
- Medios del semestre
- Finales del semestre

3.4. ¿Consideras que en el curso de inducción se te proporcionó la información general sobre los servicios que ofrece el SUBA?

- SI
- NO

Si contestaste NO selecciona una de las siguientes opciones:

- Poco tiempo

Faltaron temas como:

- Ubicación, horarios y servicios que ofrecen las bibliotecas
- Requisitos para darse de alta en el sistema
- Requisitos para solicitar libros en préstamo
- Períodos de préstamo externo.
- Uso del Catálogo en línea.

3.5. De las siguientes opciones ¿Cuál consideras que te ayudaría a conocer mejor el uso de los servicios que ofrece el SUBA?

- Folleto / tríptico
- Archivo electrónico para descargar (PDF)
- Cartel / poster / lona
- Página Web UACH / FACULTAD / SUBA
- Video

3.6. ¿Puedes identificar los servicios que se ofrecen en la biblioteca de la facultad en la que estudias?

- SI
- NO

3.7. De acuerdo a la siguiente escala. ¿Cómo evaluarías los servicios que se ofrecen en el SUBA?

(MB= Muy Bien; B= Bien; N= Regular; M= Mal; MM= Muy Mal.)

Actividad	MB	B	N	M	MM
Trato del personal					



### Anexo 1. La encuesta, pagina 4 de 7



Orientación o asesoría por parte del personal					
Acceso al catálogo en línea (OPAC)					
Facilidad para localizar los libros en las estanterías					
Préstamo externo					
Préstamo de cubículos					
Préstamo de computadoras					
Préstamo de salas audiovisuales					
Servicio de fotocopiado, impresiones y escáner.					
Confort que te brinda el área de estantería.					
Comodidad que te brindan las áreas de lectura, cubículos, audiovisuales					
Condiciones en las que se encuentran los sanitarios					
Valoración general de las instalaciones bibliotecarias					

3.8 ¿En la biblioteca a la que asistes se cumple el horario de servicio que tienen establecido?

SI                       NO

3.9. ¿El horario de servicio que ofrece la biblioteca a la que asistes te parece adecuado?

SI                       NO

Si no te parece adecuado, sugiere un horario y el período académico en el que se aplicaría:

Horario	Todo el semestre	Fin de semestre	Período de exámenes
_____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3.10 ¿El personal bibliotecario te orienta y/o asesora en forma adecuada para la búsqueda de información?

SI     NO

3.11.- ¿Conoces y has usado la Página del SUBA?

SI     NO

Si tu respuesta es afirmativa, evalúa (marcando con una X) el acceso a los siguientes servicios que se ofrecen a través de la página, de acuerdo a la siguiente escala:

(MB= Muy Bien; B= Bien; N= Regular; M= Mal; MM= Muy mal.)

	MB	B	N	M	MM
Califica la estructura (diseño) de la página web					
Acceso al buzón de sugerencias					
Visibilidad de los recursos que se ofrecen en la					



### Anexo 1. La encuesta, pagina 5 de 7



página.					
Facilidad de acceso al catálogo en línea					
Facilidad para realizar el apartado de un libro a través del Catálogo en línea.					
Facilidad de acceso y visibilidad de los reglamentos y políticas de uso de los servicios.					
Facilidad de acceso a sitios de interés relacionados con el servicio bibliotecario.					

3.12.- Cuando accedes a la página del SUBA ¿encuentras todo lo que necesitas?

SI                       NO

Si tu respuesta es NO, ¿qué es lo que no has encontrado? \_\_\_\_\_

3.13.- Sólo en caso de que hayas realizado alguna queja o sugerencia a través de la página del SUBA contesta lo siguiente:

¿Le han dado seguimiento?

SI             NO

¿Han resuelto tus dudas?

SI             NO

3.14. ¿Te han informado que tienes derecho a acceder a Bases Electrónicas de Datos a través de la página del SUBA?

SI             NO

3.15. ¿Alguna vez has accedido a éstas?

SI             NO

Si contestaste NO pasa a la pregunta 3.18

Si tu respuesta anterior fue afirmativa, valora los siguientes aspectos de acuerdo a la escala sugerida:

	Poco		Mucho			
¿Con qué frecuencia accedes a ellas?	①	②	③	④	⑤	⑥
Cuando accedes a ellas, ¿localizas la información que buscas?	①	②	③	④	⑤	⑥
¿Te resulta fácil acceder a los artículos en las bases de datos?	①	②	③	④	⑤	⑥



### Anexo 1. La encuesta, pagina 6 de 7



Para tus exigencias académicas consideras que la información que ofrecen es suficiente.	①	②	③	④	⑤	⑥
Consideras que la calidad de las publicaciones que contienen las Bases de Datos cubre tus necesidades.	①	②	③	④	⑤	⑥
Consideras que para utilizar las bases de datos una limitante sería tu dominio del idioma inglés	①	②	③	④	⑤	⑥

3.16.- Marca con una X, cuales de estas características tienen la información que te ofrecen las Bases Electrónicas de Datos de acuerdo a lo siguiente:

<input type="checkbox"/> Fácil de obtener el contenido a texto completo.	<input type="checkbox"/> Es posible descargar o guardar los artículos que necesito.
<input type="checkbox"/> Te es fácil comprender el contenido de los artículos en el idioma en que están escritos.	<input type="checkbox"/> El contenido de los artículos, imágenes y tablas es legible.
<input type="checkbox"/> Se puede exportar o copiar la cita bibliográfica.	<input type="checkbox"/> Puedes de compartir la información.
<input type="checkbox"/> Es fácil conseguir la información que requieres para citar correctamente los artículos que utilizaste.	<input type="checkbox"/> Es fácil ordenar el resultado de la búsqueda por relevancia o por fecha.

3.17.- Marque con una X, ¿Cuáles son las bases de datos que más utilizas de las siguientes opciones?

- |  |   |  |
|--|---|--|
| <b>Integradoras</b>                            |   | <b>Editoriales</b>                                     |
| <input type="checkbox"/> EBSCO                 | <input type="checkbox"/> Elsevier   | <input type="checkbox"/> American Institute of Physics |
| <input type="checkbox"/> Gale Cengage Learning | <input type="checkbox"/> Springer   | <input type="checkbox"/> American Physical Society     |
|  | <input type="checkbox"/> Thompson-Reuters                                 | <input type="checkbox"/> Institute of Physics          |
|  | <input type="checkbox"/> Science AAA's                                    | <input type="checkbox"/> Bio One                       |
|  | <input type="checkbox"/> American Mathematical Society                    | <input type="checkbox"/> Nature                        |
|  | <input type="checkbox"/> Institute of Electrical and Electronic Engineers | <input type="checkbox"/> Annual Review                 |
|  | <input type="checkbox"/> American Chemical Society                        |  |

3.18 Indica uno o varios de los siguientes motivos por los cuales no accedes a las Bases de Datos que ofrece el SUBA:



## Anexo 1. La encuesta, pagina 7 de 7



- Desconozco la existencia de las bases de datos
- No se cómo obtener el acceso
- Tengo acceso a otras bases de datos
- Prefiero buscar en internet
- No las requiero para mis actividades académicas

Con el fin de mejorar la calidad de los Servicios Bibliotecarios que te ofrece el SUBA ¿Qué sugieres para poderte dar un mejor servicio?

---

---

---

---

*Te agradecemos tu participación en este estudio, la información nos servirá para mejorar nuestros Servicios Bibliotecarios.*

*Gracias por tu apoyo, tu opinión nos será de mucha utilidad.*

*Orgullo de ser UACH!*



## Anexo 2. Procedimiento en el programa SPSS para el control de confusión en la variable respuesta.

Ingresar los datos (Cuadro 2b) al programa, y utilizar las opciones: ANALIZAR / ESTADÍSTICOS DESCRIPTIVOS / TABLAS DE CONTINGENCIA

The screenshot shows the SPSS interface with the 'Analyze' menu open, highlighting 'Contingency Tables'. The 'Contingency Tables' dialog box is open, showing 'género' in the 'Rows' field and 'susc' in the 'Columns' field. The 'Capa 1 de 1' section has 'edadcat' in the 'Layer' field. The 'Estadísticos' sub-dialog box is also open, showing various statistical tests. The 'Nominal' section has 'Coficiente de contingencia', 'Phi y V de Cramer', 'Lambda', and 'Coficiente de incertidumbre' checked. The 'Ordinal' section has 'Gamma', 'd de Somers', 'Tau-b de Kendall', and 'Tau-c de Kendall' checked. The 'Nominal por intervalo' section has 'Eta' checked. The 'Kappa' section has 'Riesgo' checked. The 'Estadísticos de Cochran y Mantel-Haenszel' section has 'Estadísticos de Cochran y Mantel-Haenszel' checked, and the 'Probar que la razón de ventajas común equivale a:' field is set to 1. The 'Aceptar' button is highlighted in the 'Tablas de contingencia' dialog box.

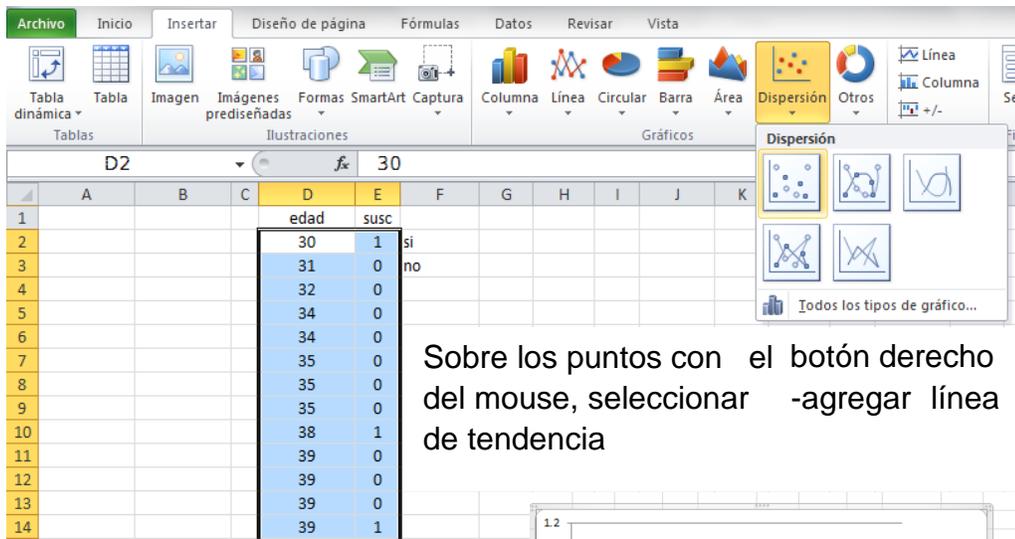
En fila(s) ingresar el primer factor (género) y en Columna(s) ingresar la variable respuesta (suscribe), Capa 1 de 1 ingresar el factor confusor (edad categorizada),

En [Estadísticos] se selecciona - Estadísticos de Cochran y Mantel-Haenszel y -Riesgo, CONTINUAR y ACEPTAR.



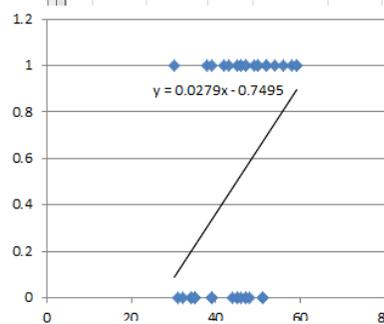
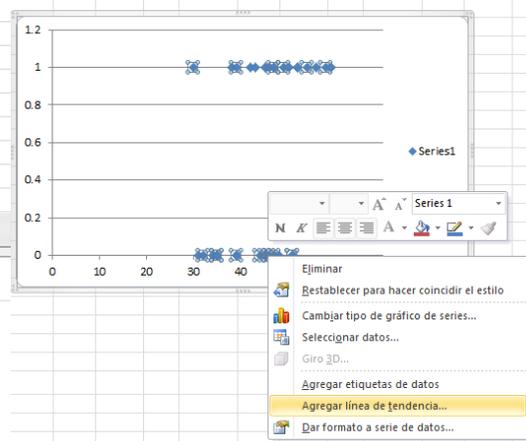
### Anexo 3. Procedimiento en el programa Excel para obtener una ecuación del modelo lineal.

Seleccionar los datos de suscripción y edad, -insertar gráfico de dispersión



Sobre los puntos con el botón derecho del mouse, seleccionar -agregar línea de tendencia

Seleccionar -lineal, - presentar ecuación en el gráfico y cerrar



Se presenta la ecuación del modelo dentro del gráfico.



### Anexo 4. Procedimiento en el programa SPSS para la regresión logística binaria.

Ingresar los datos (Cuadro 3) al programa, seleccionar las opciones: ANALIZAR/ REGRESIÓN/ LOGÍSTICA BINARIA

The screenshot shows the SPSS Statistics Editor interface. The main window displays a data table with the following data:

1. género	edad	susc	género
1	30	1	0
2	32	0	0
3	34	0	0
4	35	0	0
5	35	0	0
6	35	0	0
7	39	0	0
8	39	0	0
9	39	0	0
10	43	1	0
11	45	0	0
12	45	0	0
13	31	0	1
14	46	1	0
15	46	0	0
16	47	0	0

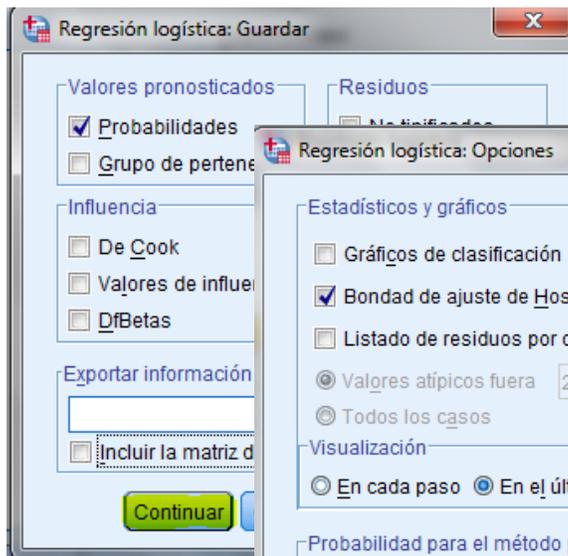
The 'Análizar' menu is open, and 'Regresión' > 'Logística binaria...' is selected. The 'Regresión logística' dialog box is open, showing 'susc' as the dependent variable and 'edad' and 'género' as covariables. The method is set to 'Atrás: Wald'. The 'Definir variables categóricas' dialog box is also open, showing 'edad' as a covariable and 'género(Indicador)' as a categorical variable. The contrast is set to 'Indicador' and the reference category is 'Último'.

Ingresar en DEPENDIENTES la variable respuesta (suscribe) y en COVARIABLES (edad y género) en MÉTODO seleccionar ATRÁS: WALD

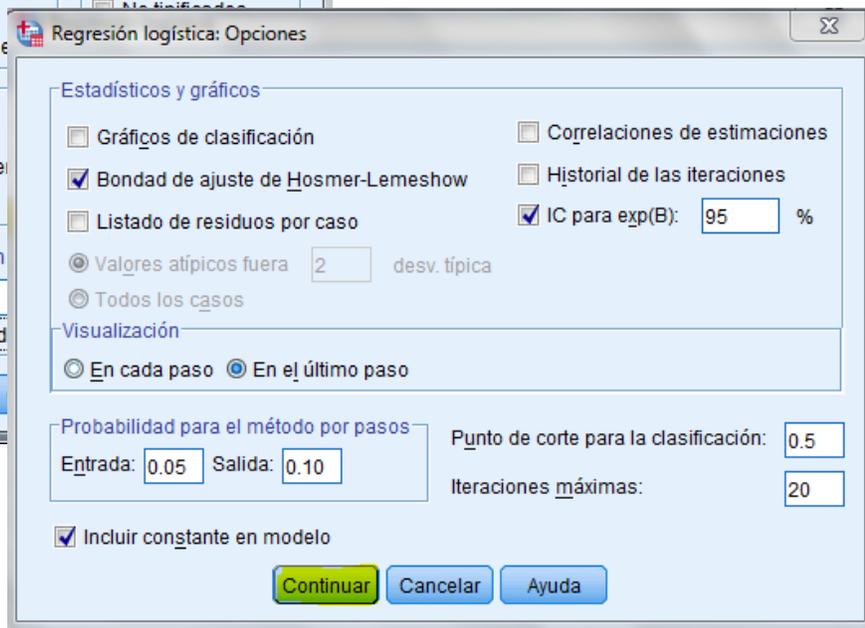
En CATEGORICA ingresar (género) ya que esta variable cuenta con dos categorías, no es el caso de edad por ser una variable continua, seleccionar CONTINUAR.



### Anexo 4. Continuación. Procedimiento en el programa SPSS para la regresión logística binaria

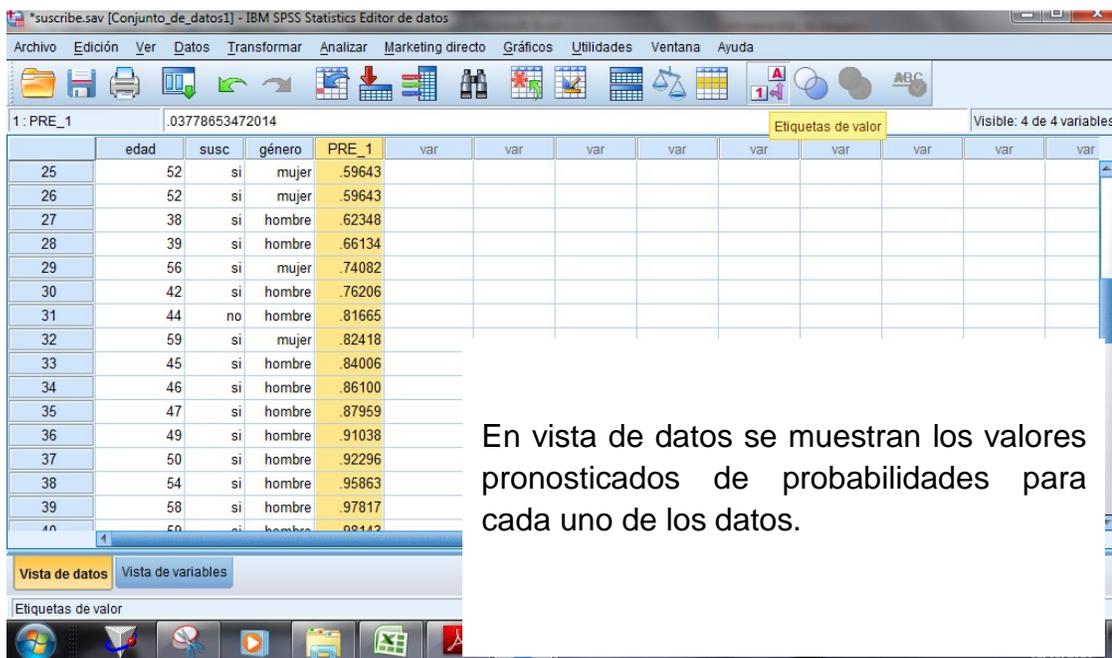


En GUARDAR marcar únicamente PROBABILIDADES seleccionar CONTINUAR



En OPCIONES marcar:

BONDAD DE AJUSTE DE HOSMER- LEMESHOW, IC para exp(B) e INCLUIR CONSTANTE EN MODELO, CONTINUAR y ACEPTAR



En vista de datos se muestran los valores pronosticados de probabilidades para cada uno de los datos.