

UNIVERSIDAD NACIONAL DEL ALTIPLANO - PUNO
FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA



**Factores que determinan el otorgamiento de crédito de la
financiera Credinka en la ciudad de Ayaviri, 2015**

TESIS

PRESENTADA POR:

Bach. MARIBEL GIOVANA SARCO YAMPASI

**PARA OPTAR EL TÍTULO PROFESIONAL DE:
INGENIERO ESTADÍSTICO E INFORMÁTICO**

PUNO – PERÚ

2017

UNIVERSIDAD NACIONAL DEL ALTIPLANO - PUNO

FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA

ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA

**Factores que determinan el otorgamiento de crédito de la
financiera Credinka en la ciudad de Ayaviri, 2015****TESIS**

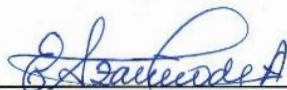

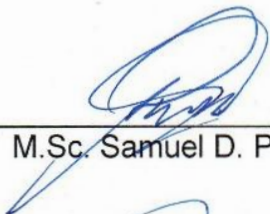

PRESENTADA POR:

Bach. MARIBEL GIOVANA SARCO YAMPASI

Para optar el Título Profesional de:

INGENIERO ESTADÍSTICO E INFORMÁTICO

APROBADA POR:

PRESIDENTE DE JURADO : 
Mg. Emma O. Azañero de AguirrePRIMER MIEMBRO : 
M.Sc. Alejandro Apaza TarquiSEGUNDO MIEMBRO : 
M.Sc. Remo Choquejahu AceroDIRECTOR DE TESIS : 
M.Sc. Samuel D. Perez QuispeASESOR DE TESIS : 
Dr. Edgar E. Carpio VargasÁREA : Estadística
TEMA : Análisis multivariado

DEDICATORIA

Esta tesis la dedico con todo mi amor a:

En primer lugar a Dios por darme salud, vida, fortaleza para salir adelante día a día; y una familia muy maravillosa que me apoya y me brinda mucho cariño.

A mis padres Manuel y Gumerinda que me dieron la vida y han estado conmigo en todo momento. Ustedes han sido y serán la base de toda mi vida. Gracias por enseñarme a ser quien soy, por todo el amor y apoyo que me han brindado y por darme una carrera profesional. Toda la vida les estaré muy agradecida por el buen trabajo de padres que han hecho conmigo y con mis hermanos. Espero haber aprendido muy bien de ustedes los buenos valores.

A mis hermanos, gracias por estar conmigo, por su apoyo, por las bromas, gracias por todo lo que son y por lo que me dan incondicionalmente. Los quiero mucho.

A los docentes, que me enseñaron a que todo se logra con esfuerzo; muchas gracias.

Los quiero a todos y este trabajo es para ustedes.

AGRADECIMIENTO

A las instituciones financieras de la ciudad de Ayaviri que me brindaron sus instalaciones para poder recabar información, facilitándome la mayor libertad posible.

A mis profesores de la carrera que me enseñaron no sólo letras y números; me enseñaron a ser más valiente para buscar algo mejor en la vida a nivel profesional. Gracias por ayudarme a darme cuenta que tengo el conocimiento y el valor para emprender mi carrera sin obstáculo.

A todos aquellos que me dieron el apoyo y me abrieron las puertas para poder recabar la información, para efectuar la presente tesis.

ÍNDICE

RESUMEN	10
ABSTRACT	11
INTRODUCCIÓN	12
CAPÍTULO I	14
PLAN DE INVESTIGACIÓN	14
1.1. El Problema	14
1.2. Descripción del Problema	14
1.3. Formulación del Problema	15
1.4. Objetivos	15
1.4.1. Objetivo General	15
1.4.2. Objetivos Específicos	16
1.5. Hipótesis	16
1.5.1. Hipótesis general	16
1.5.2. Hipótesis específicas	16
1.6. Justificación de la Investigación	16
1.7. Limitaciones de la Investigación	17
CAPÍTULO II	18
MARCO TEÓRICO	18
2.1. Antecedentes de la Investigación	18
2.1.1. En el aspecto mundial	18
2.1.2. En el aspecto nacional	22

2.1.3. En el aspecto local	24
2.2. Base teórica	25
2.2.1. Financiera	25
2.2.2. Crédito financiero	25
2.2.3. Requisitos para un préstamo	26
2.2.4. Procedimiento para otorgar créditos	26
2.2.5. El Modelo de Regresión Logística.....	30
2.2.6. Modelo de Regresión Logística múltiple	37
2.2.7. Estimación de los Parámetros.....	39
2.2.8. Pruebas de Significación	40
2.2.9. Interpretación de los Resultados	43
2.2.10. Selección de Variables.....	44
CAPÍTULO III.....	48
MATERIALES Y MÉTODOS	48
3.1. Población	48
3.2. Diseño de la Muestra	48
3.3. Recolección de la Información	49
3.3.1. Cuestionario	49
3.3.2. Operacionalización de variables	51
3.4. Metodología	54
3.5. Tratamiento de datos	56
3.5.1. El programa SPSS	56

CAPÍTULO IV	58
RESULTADOS Y DISCUSIÓN	58
4.1. Determinantes del Otorgamiento de Crédito.	58
4.2. Modelo para Otorgamiento de Crédito	58
4.2.1. Proceso de Selección de variables	59
4.2.2. Parámetros Estimados del Modelo Completo	59
4.2.3. Bondad de Ajuste del modelo con 7 variables	60
4.2.4. Interpretación de los parámetros estimados del Modelo – ODDS Ratio 62	
4.2.5. Validación del Modelo Logístico	63
4.3. Modelo para la Hipótesis	64
4.3.1. Selección de las Variables	64
4.3.2. Parámetros	64
4.3.3. Bondad de Ajuste	64
4.3.4. Interpretación de los parámetros Estimados – ODDS Ratio	65
4.4. Con respecto al ingreso mínimo y gasto máximo	66
4.4.1. Ingreso mínimo curva ROC	66
4.4.2. Gasto máximo curva ROC	70
CONCLUSIONES	72
RECOMENDACIONES Y SUGERENCIAS	73
BIBLIOGRAFÍA	74

ÍNDICE DE TABLAS

Tabla 1: Operacionalización de Variables.....	51
Tabla 2: Variables y coeficientes del modelo completo.....	60
Tabla 3: Resumen del modelo	61
Tabla 4: Test de Hosmer y Lemeshow	61
Tabla 5: Variables en el Modelo	62
Tabla 6: Cruce de otorgamiento de crédito estimados y originales	63
Tabla 7: Variables en la ecuación.....	64
Tabla 8: Resumen del modelo	65
Tabla 9: Prueba de Hosmer y Lemeshow	65
Tabla 10: Parámetros en la variable	66
Tabla 11: Área bajo la curva.....	68
Tabla 12: Coordenadas de la Curva ROC	69
Tabla 13: Área bajo la curva.....	70
Tabla 14: Coordenadas de la curva.....	71

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Información para la entidad financiera.....	29
Ilustración 2: Clase de garantías	30
Ilustración 3: Newton-Raphson	36
Ilustración 4: Función de verosimilitud.....	37
Ilustración 5: Modelo de Regresión Logística	39

ÍNDICE DE GRÁFICOS

Gráfico 1: Curva ROC de Ingreso	67
Gráfico 2: Curva ROC de Gasto.....	70

RESUMEN

Las empresas de microfinanza al otorgar créditos financieros asumen riesgos de no devolución del préstamo, existen metodologías estadísticas que identifican factores que deben ser evaluados para no correr riesgos, una de ellas es la regresión logística binaria. El objetivo del presente trabajo fue identificar los factores determinantes del otorgamiento de crédito de la Financiera Credinka en la ciudad de Ayaviri, así mismo determinar el ingreso mínimo y el gasto máximo requeridos para acceder al crédito financiero. Para obtener respuesta a nuestros objetivos se usó muestreo no aleatorio tomando en cuenta 61 clientes como muestra de un total de 110 (población de un solo mes), como metodología de investigación se usó la regresión logística binaria y las Curvas ROC. Los resultados fueron: los factores influyentes para el otorgamiento de crédito son: el ingreso, gasto, número de hijos, edad, seguro, vivienda y material de la vivienda. El modelo estadístico es el siguiente:

$$P(\text{crédito}) = \frac{1}{1 + e^{-(-5.960 + 0.004\text{Ingreso} - 0.003\text{Gasto} - 0.205\text{Hijos} + 0.124\text{Edad} - 0.299\text{Seguro} + 0.566\text{Vivienda} + 0.640\text{Material})}}$$

El cual explica en un 78.7%. su fiabilidad.

PALABRAS CLAVES: Financiera Credinka Ayaviri, Crédito, Regresión Logística.

ABSTRACT

Microfinance companies in providing financial loans assume risks of non-repayment of the loan, there are statistical methodologies that identify factors that must be evaluated in order to avoid risks, one of them being binary logistic regression. The objective of the present work was to identify the determinants of the granting of credit of the Credinka Financier in the city of Ayaviri, as well as to determine the minimum income and the maximum expenditure required to access the financial credit. To obtain a response to our objectives, non-random sampling was used, taking into account 61 clients as a sample of a total of 110 (single-month population). As a research methodology, binary logistic regression and ROC curves were used. The results were: the factors influencing the granting of credit are: income, expenditure, number of children, age, insurance, housing and housing material. The statistical model is as follows:

$P(\text{crédito})$

$$= \frac{1}{1 + e^{-(5.960 + 0.004\text{Income} - 0.003\text{expenditure} - 0.205N.\text{children} + 0.124\text{Age} - 0.299\text{Insurance} + 0.566\text{housing} + 0.640\text{H.Material})}}$$

Which explains in 78.7%. Its reliability.

KEYWORDS: Financiadora Credinka Ayaviri, Credit, Logistic Regression.

INTRODUCCIÓN

Con la presente investigación de tesis se hizo con el propósito de Identificar los factores que determinan el otorgamiento de crédito de la Financiera Credinka en la ciudad de Ayaviri,2015.

En lo que respecta a los factores a considerar para un crédito, la mayoría de las financieras considera la capacidad de pago que tengan los clientes como principal factor de crédito. Para nuestra investigación consideramos la experiencia de los clientes con otras financieras crediticias para el otorgamiento de crédito, en vista de que ya entienden el modo de trabajo.

La estructura del presente trabajo en su primer capítulo define el problema de investigación; los objetivos de la investigación a alcanzar, justificación del estudio y limitaciones del problema.

El segundo capítulo corresponde al marco de referencia, que define los conceptos y referencias teóricas pertinentes al estudio. Además, se hace mención de algunos antecedentes del problema, la hipótesis que se plantea probar, identificación de variables y su forma de tratamiento.

El tercer capítulo hace referencia a la metodología que, del estudio, definiendo así, el tipo de estudio, la población y muestra, el tipo de diseño, las técnicas e instrumentos de recolección de información.

El cuarto capítulo corresponde a los resultados obtenidos y discusiones que se ha llegado a tener después del método de tratamiento de datos.

En la parte final tenemos las conclusiones de la investigación, seguido de las respectivas recomendaciones y para concluir se hace mención de la bibliografía utilizada para realizar el estudio; finalmente los anexos que permitirán un aporte al trabajo.

CAPÍTULO I

PLAN DE INVESTIGACIÓN

1.1. El Problema

En la entidad financiera Credinka existe la necesidad de identificar los posibles factores que determinan el otorgamiento de crédito de la financiera Credinka en el distrito de Ayaviri.

1.2. Descripción del Problema

La entidad financiera Credinka agencia Ayaviri representa el 4.0% de riesgo en crédito según la evaluación del SBS 2016, por lo cual la institución lo considera de mayor riesgo de mora.

Por lo cual es importante realizar una mejor evaluación de los clientes solicitantes de crédito para reducir el riesgo de endeudamiento y mora.

El presente trabajo pretende aportar en este sentido; identificar los posibles factores que determinan el otorgamiento de crédito de la financiera Credinka en el distrito de Ayaviri.

Teneindo en cuenta que la cantidad de clientes con experiencia crediticia que solicitan un crédito en la entidad son evaluadas con menos rigurocidad que un clente nuevo, lo que lleva al otorgamiento automatico del crédito financiero por lo que se comete el error ya que se representaria en el incumplimiento de pagos si se tiene muchos creditos y bajo capacidad de pago.

Asi que el presente trabajo trata de proporcionar una opción para poder evaluar a los clientes de una forma en que se identifiquen los factores que se tienen que tomar en mayor concideracion para poder tomar la aprobación del crédito, que beneficiará a la financiera Credinka.

Con el método de la regresión logística se permite determinar que factores estan relacionados con el otorgamiento de crédito.

Por todo lo expuesto anteriormente nos planteamos el problema a ineestigiar, haciéndonos la siguiente presgunta.

1.3. Formulación del Problema

¿Cuáles son los factores que determinan en el otorgamiento de crédito de la financiera Credinka de la ciudad de Ayaviri?

1.4. Objetivos

1.4.1. Objetivo General

“Identificar los factores que determinan significativamente en el momento de aprobar el otorgamiento de crédito de la Financiera Credinka en la ciudad de Ayaviri”

1.4.2. Objetivos Específicos

- Analizar los factores con mayor influencia que intervienen en el otorgamiento de crédito de la financiera Credinka en la ciudad de Ayaviri.
- Determinar ingreso mínimo y gasto máximo requeridos para acceder al crédito financiero.

1.5. Hipótesis

1.5.1. Hipótesis general

Los factores que determinan significativamente en la aprobación de otorgamiento de crédito de la financiera Credinka en la ciudad de Ayaviri son: ingreso, gastos, tipo de trabajo y educación.

1.5.2. Hipótesis específicas

- Los factores con mayor influencia determinan en el otorgamiento de crédito de la financiera Credinka en la ciudad de Ayaviri.
- El ingreso mínimo y gasto máximo requeridos determinan el acceso al crédito financiero.

1.6. Justificación de la Investigación

El presente trabajo es importante y relevante porque contribuirá a que la microfinanciera alcance niveles de evaluación de la clientela mejor o igual en comparación de las empresas líderes del sector. Para que mediante la regresión logística tengan una herramienta en la consideración de los factores de evaluación crediticia con respecto a los deferentes aspectos socio-

económicos de la clientela y de esta manera mejorar reducir riesgos crediticios.

1.7. Limitaciones de la Investigación

Las limitaciones que se ha encontrado son las siguientes:

- Desconfianza por parte de las clientes que pretenden un crédito a la hora de realizar la entrevista.
- El tiempo reducido por parte de los entrevistados.

CAPÍTULO II

MARCO TEÓRICO

2.1. Antecedentes de la Investigación

Tenemos a continuación resultados de trabajos para el análisis y comparación, en diferentes lugares.

2.1.1. En el aspecto mundial

Forero, L. (2015), *“Propuesta de modelo para la evaluación y predicción de riesgo del insolvente de pequeñas y medianas empresas manufactureras en Colombia”*. Escuela de Estudios Industriales y Empresariales, Universidad Nacional de Santander, Bucaramanga Colombia.

El objetivo principal del trabajo es diseñar una propuesta de modelo basado en el análisis de los ratios financieras, para la evaluación y predicción del riesgo de insolvencia financiera de pequeñas y medianas empresas manufactureras en Colombia.

Concluyendo que los indicadores más representativos a la hora de definir el estado posterior de las empresas son la rentabilidad del activo, rotación del activo y nivel de endeudamiento. Las dos primeras variables son inversas al resultado de probabilidad de cancelación y la tercera directamente al mismo. La variable que mejor predice el nivel de endeudamiento por la menor significancia en el estadístico de wald.

El modelo elaborado por medio de la regresión logística binaria predice correctamente un porcentaje sin zona de incertidumbre de 92.22% para las 90 empresa iniciales, 90% para 10 empresas de validación y 92% para las 100 empresas con un corte de probabilidad de cancelación de 0.5; mientras que el modelo elaborado por medio de análisis discriminantes predice correctamente un porcentaje de 87% para las 60 empresas iniciales con un 26% de zona de incertidumbre, 80% para las 10 empresas de validación con un 40% de zona de incertidumbre y 86% para las 100 empresas con zona de incertidumbre del 28%; con un corte discriminante de 0.

Zapata, D. (2009), *“Caracterización de las Variables Determinantes del Riesgo en el Microcrédito Rural”*. Escuela de Ingeniería de la Organización, Facultad de Minas, Universidad Nacional de Colombia Medellín.

El objetivo del presente trabajo pretende aportar en ese sentido; investigando si existen variables determinantes del riesgo crediticio en el microcrédito rural colombiano, partiendo del análisis de las bases de datos de información crediticia de un banco, que recogen las características de los préstamos que adjudica y de su comportamiento.

Concluyendo que el Análisis discriminante permite contrastar las variables en su conjunto y al observar sus interacciones reconocer algunas de sus dinámicas, que no son de fácil deducción. La regresión logística es una herramienta que puede ayudar en la toma de decisiones; reflejar las dinámicas de los componentes de las variables y realizar lecturas y aproximaciones fundamentadas a casos específicos, su valor en la calificación de clientes y evaluación de crédito puede ser confirmado en este trabajo.

Nieto, S. (2010); *“Crédito al Consumo: La Estadística Aplicada a un problema de Riesgo Crediticio”, Universidad Autónoma Metropolitana, Mexico.*

El objetivo de este trabajo es hacer uso de una colección de técnicas que se utilizan en el Credit Scoring y que usan las empresas consultoras para generar una scorecard; esto es, el puntaje (score) que se les asigna a las características de los clientes en base a sus datos demográficos y de buró de crédito. Haremos una comparación de la scorecard que obtengamos con una generada por una empresa consultora sobre un segmento de población. Ya que las empresas consultoras sólo entregan resultados como cajas negras, sin indicar que desarrollos llevan a cabo, y además tienen costos elevados para quienes consultan estos servicios.

Concluyendo que las variables directamente observables elegidas miden correctamente las variables latentes subyacentes de ellas; Dado que el p-valor para el modelo en la tabla del Análisis de la Varianza es inferior a 0.01, hay una relación estadísticamente significativa entre las variables al 99 % de nivel de confianza. Además, el p-valor para los residuos es mayor o igual a 0.10, indicando que el modelo no es significativamente peor que el mejor modelo

posible para estos datos al 90 % de nivel de confianza o superior. La ventana también muestra que el porcentaje e desviación en Clase explicado por el modelo es igual a 3.78213. El porcentaje ajustado más adecuado para comparar modelos con diferentes números de variables independientes, es 3.7074 %. Determinando si el modelo puede simplificarse, observe que el p-valor más alto para los test de proporción de probabilidad es 0.0167, perteneciendo a Seguro. Dado que el p-valor es inferior a 0.05, ese término es estadísticamente significativo al 95 % de nivel de confianza. Por consiguiente, probablemente no querrá eliminar ninguna variable del modelo.

Andréu Abela,Jaime. (2011):” El análisis multinivel: una revisión actualizada en el ámbito sociológico”,Departamento de Sociología. Universidad de Granada.

El objetivo de estudio es, modelizar estadísticamente la influencia de variables contextuales sobre las actitudes o los comportamientos medidos a nivel individual. En este sentido, nos permite, en ciencias sociales, tener en cuenta el efecto de las variables de la estructura social y económica sobre el individuo.

Concluyendo al final del proyecto que: La aplicación del análisis multinivel abre un abanico de posibilidades en el ámbito sociológico, puesto que posibilita la utilización en un mismo modelo analítico de datos de encuestas (individuales) y datos agregados (contextuales). Esto en el ámbito del análisis de los indicadores sociales permite interrelacionar la realidad subjetiva de los individuos con la realidad objetiva, lo cual admite la posibilidad de modelos analíticos integrales en el análisis de la realidad social.

2.1.2. En el aspecto nacional

Bullon, C. y Vergara, A. (2009);” *Análisis multinivel: Regresión logística binaria con dos niveles*”, Universidad Nacional Agraria La Molina, Perú.

El objetivo del presente trabajo es proporcionar, a nivel introductorio, la metodología de la regresión logística multinivel o jerárquica con dos niveles, la misma que es aplicada a los datos de una encuesta sobre consumo de drogas legales en la región Lima, donde los alumnos son tratados como micro observaciones y los colegios como macro observaciones.

Concluye que la salida para el modelamiento multinivel cuando la variable respuesta es binaria arroja dos modelos: modelo de unidad específica y modelo población promedio. El uso de uno de estos modelos depende sobre todo de lo que se tiene como interés. La elección del modelo de unidad específica depende de las preguntas siguientes: ¿Cuál es el efecto en un grupo típico?, ¿Cómo este efecto varía a través de los grupos?

Vela L., Uriol J., Paima M., Palacios F. y Velasco P., Leider E. (2012); “*Los factores que determinan la calidad de la cartera crediticia de las entidades micro financieras de la Amazonía peruana en el periodo 2008-2011*”, Universidad Nacional Pedro Ruiz Gallo Lambayeque Perú.

El objetivo principal es: Determinar cuáles son los factores que determinan la calidad de la cartera crediticia en las entidades micro financieras en la Amazonía peruana.

La calidad de la cartera está explicada por un conjunto de variables de carácter macroeconómico relacionados con el ciclo del producto y

microeconómico tales como el nivel de solvencia de las entidades, la eficiencia y gestión de sus costos operativos, la tasa de crecimiento de los activos rentables. La relevancia global está corroborada con el coeficiente de determinación significativo encontrado en el modelo, sin embargo, la prueba de hipótesis “f”, demuestra que se deben incluir más variables en la especificación del modelo.

Quispe, S. (2014); “Rotación de personal: Predicción con modelo de regresión logística multinivel”, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, E.A.P. de Estadística. Perú.

El objetivo de esta tesis es identificar las características de los trabajadores que se retiran de la empresa antes del término del periodo de prueba, de modo que se pueda mejorar el proceso de selección, implementar programas de retención y comprobar la diferencia de perfiles de trabajadores que cesan durante el periodo de prueba, de acuerdo al área donde trabajaron.

Concluye que debido a que la finalidad era identificar un modelo que permita predecir la decisión de retirarse de la empresa durante el periodo de prueba considerando variables del trabajador, en función de postulante, y del área a la cual pertenece, se puede sugerir incorporar variables como tiempo transcurrido desde el último trabajo, formación, estudios actuales, actividades adicionales (hobbies). Las variables que actualmente se tienen de un ingresante son básicas. Respecto a las áreas, se puede considerar el tipo de liderazgo del jefe, las características y funciones del puesto. Las variables señaladas no han sido incluidas en el modelo debido a que actualmente no se cuenta con los datos al no ser solicitados al trabajador al ingresar a laborar o

no se encuentran registradas en los sistemas de información que se utilizan. La variabilidad encontrada entre las áreas nos indicó que es correcto considerar esta agrupación para analizar la decisión de retiro o permanencia.

2.1.3. En el aspecto local

Caxi, F. (2016); *“Factores Individuales y Escolares que Influyen en el Ingreso a la Universidad Nacional del Altiplano Puno en los Estudiantes del Cepreuna 2016”*, Perú.

El objetivo principal del estudio es Determinar los factores individuales y escolares que inciden significativamente en el ingreso a la Universidad Nacional del Altiplano en los estudiantes del cepreuna 2016.

Concluyendo que se ha identificamos 6 factores relevantes en el ingreso a la universidad, estos son: el género, el gasto mensual, la procedencia, el área donde se ubica la IES, el nivel educativo del padre (superior), y las horas que destina el/la postulante a sus estudios, todos ellos con un nivel menor al 5% de significancia. Se destaca la importancia del factor gasto mensual para el modelo a un nivel del 5% de significancia, eso quiere decir que si el/la postulante destina un sol más en sus estudios el logaritmo de la odds de la variable dependiente ingresa incrementa en 0.144, permaneciendo el resto de los factores constantes. Cabe mencionar que el gasto promedio de los ingresantes fue de S/. 396.83 y de los que no ingresaron fue de S/. 378.15, notamos una diferencia de S/. 18.68 en ambos, y por ello deducimos que los que ingresan están destinando mayores recursos económicos con respecto a sus pares que no ingresaron.

2.2. Base teórica

2.2.1. Financiera

Claramente se inscriben dentro del nuevo paradigma en la provisión de servicios a las microempresas y a los sectores más necesitados de la población. En términos del principio del crédito, por ejemplo, las cajas municipales al estar ubicadas en localidades específicas están cerca de su mercado objetivo. Más aún, al cargar tasas de interés comerciales que son más elevadas que la tasa de interés bancarias, pero que son más bajas que las tasas informales (el costo de oportunidad de los pobladores de estas zonas), están tomando en cuenta que el poblador de la zona busca primordialmente la conveniencia del préstamo. Asimismo, los préstamos de estas instituciones son usualmente de corta duración, no mayores a un año, siendo su objetivo apoyar al capital de trabajo de los microempresarios de la zona de su influencia.

2.2.2. Crédito financiero

(ESAN) El crédito es una operación financiera en la que se pone a nuestra disposición una cantidad de dinero hasta un límite especificado y durante un período de tiempo determinado.

En un crédito nosotros mismos administramos ese dinero mediante la disposición o retirada del dinero y el ingreso o devolución del mismo, atendiendo a nuestras necesidades en cada momento. De esta manera podemos cancelar una parte o la totalidad de la deuda cuando creamos conveniente, con la consiguiente deducción en el pago de intereses.

Además, por permitirnos disponer de ese dinero debemos pagar a la entidad financiera unas comisiones, así como unos intereses de acuerdo a unas condiciones pactadas. En un crédito sólo se pagan intereses sobre el capital utilizado, el resto del dinero está a nuestra disposición, pero sin que por ello tengamos que pagar intereses. Llegado el plazo del vencimiento del crédito podemos volver a negociar su renovación o ampliación.

El propósito del crédito es cubrir los gastos, corrientes o extraordinarios, en momentos puntuales de falta de liquidez. El crédito conlleva normalmente la apertura de una cuenta corriente.

2.2.3. Requisitos para un préstamo

Según la entrevista realizada a la administradora de la financiera Credinka, los requisitos para el otorgamiento de un préstamo son los siguientes:

- Documentos personales (D.N.I. o Carné de extranjería en caso de personas naturales; Certificado de Vigencia de Poder y Testimonio de Constitución en caso sea persona jurídica).
- Documentos del domicilio y/o la vivienda (recibo de algún servicio, ficha RUC).
- Documentos que acrediten los ingresos (estados financieros, boleta de pago de remuneraciones, recibo por honorarios).

2.2.4. Procedimiento para otorgar créditos

(ESAN) Comprende un conjunto de procedimientos previos (recolección y análisis de información de los potenciales prestatarios) y posteriores

(cobranza y monitoreo) al otorgamiento del crédito. Tienen por finalidad estimar la capacidad y la voluntad de pago del prestatario. Toma en cuenta los siguientes aspectos metodológicos:

a) Evaluación y otorgamiento de crédito.

Analiza lo siguiente:

- El perfil del cliente respecto de su capacidad técnica, valores, cultura y experiencia.
- El capital y el proceso de acumulación del empresario.
- Las condiciones del entorno de la actividad económica.
- La verificación in situ de la unidad económica: familia y empresa.
- La comprobación de la información dada por el prestatario.
- El riesgo de crédito, tanto cualitativo como cuantitativo (situación financiera y operación del negocio).
- La diversificación de cartera (monto, cliente, actividad económica y área geográfica).
- También se suele emplear el scoring cuando se dispone de información.
- Historia crediticia de clientes con similares características.

b) Recuperaciones y relaciones de largo plazo.

Comprende:

- El seguimiento preventivo y oportuno que identifique los potenciales riesgos.
- Monitoreo y negociación.
- El mantenimiento de una relación de largo plazo con el sujeto de crédito.

Según lo declarado por la mencionada administradora en la entrevista, el flujo del proceso de evaluación del crédito se resume en: • Realizar una exposición ante un comité de crédito integrado por un grupo de analistas comerciales. Este procedimiento tarda aproximadamente tres días.

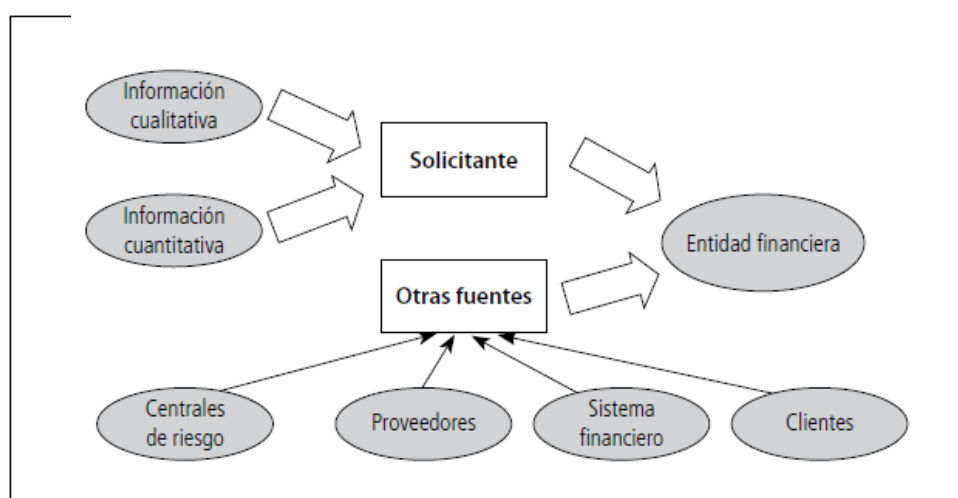
- El analista evaluará cuáles son las garantías que respaldan al solicitante, cuál es su capacidad de endeudamiento, cuál es su liquidez. Esto puede llevar tres días.
- Análisis por el área de riesgos, para aprobación. Tarda dos días.
- Firma de declaración de salud mediante un examen médico.
- Se eleva el contrato y el testimonio a registros públicos.
- Todo el proceso suma alrededor de siete días hábiles.

Actualmente, las fuentes de financiamiento de las pymes se basan en fondos propios, créditos de proveedores y de bancos.

(Bloch y Granato) que la falta de acceso a fuentes institucionales como el crédito bancario o el capital de riesgo se traduce en una dependencia excesiva sobre el crédito comercial de corto plazo y un exceso de capitalización que limita las posibilidades de expansión de las Pymes y les resta competitividad

(2007: párr. 15). Las entidades del sistema financiero reúnen información relevante para la evaluación de riesgo crediticio del solicitante; lo hacen a través de diferentes fuentes de información cualitativa y cuantitativa, como centrales de riesgo, proveedores, el sistema financiero, entre otras (ver figura 1).

Ilustración 1: Información para la entidad financiera



Fuente:

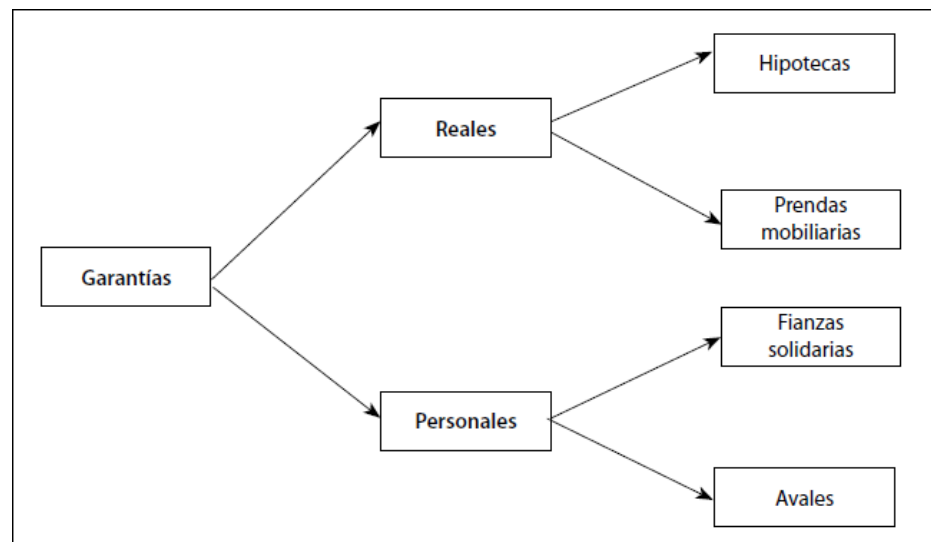
Weinberger, 2009.

Según la presentación de Paul Lira, de USAID, en finanzas y financiamiento existen dos fuentes para conseguir información. La más importante es aquella que proporciona el propio cliente, y que brinda información cualitativa y cuantitativa, como la que corresponde a los aspectos general y legal, operacional, patrimonial, crediticio, a los estados financieros y comprobantes de pagos de IGV e impuesto a la renta, lo cual está normado por la Superintendencia de Banca, Seguros y AFP (SBS). (Lira Briceño, 2009: 35).

En cuanto a las otras fuentes de información, dependerá del monto del crédito, pero siempre se recurrirá a las centrales de riesgo para recabar datos acerca del cumplimiento de aquellas obligaciones contraídas anteriormente.

Las empresas financieras solicitan garantías, las cuales se hacen efectivas al préstamo impago producto de la venta o adjudicación hasta donde este alcance (ver ilustración 2).

Ilustración 2: Clase de garantías



2.2.5. El Modelo de Regresión Logística

Es un modelo no lineal, los datos no se ajustan a una línea recta, a las variables explicativas no se les exige una distribución determinada. Permite construir modelos donde las variables dependientes pueden ser cuantitativas o cualitativas, además éstas últimas pueden ser dicotómicas o politómicas, y dentro de éstas las variables pueden ser ordinales o nominales.

Las variables explicativas pueden ser cualitativas y/o cuantitativas.

La función logística es una curva sigmoidea en forma de letra S, utilizando los logaritmos es posible linealizar el modelo.

Si la variable dependiente es dicotómica, se utiliza para predecir la probabilidad estimada de que la variable dependiente “Y” presente uno de los valores posibles (1=sí o 0=no) en función de los diferentes valores que adoptan el conjunto de variables independientes “X”.

Las variables predictoras “X” pueden ser categóricas, medidas a nivel nominal u ordinal, y de intervalo o razón.

Jovell plantea que los objetivos del modelo logístico son “determinar la existencia o ausencia de relación entre una o más variables independientes y la variable dependiente; medir la magnitud de dicha relación y estimar o predecir la probabilidad de que se produzca un suceso $Y=1$ en función de los valores que adopten las variables independientes “ X_i ”.

En la mayoría de los casos las regresiones logísticas además pueden utilizarse para medir el rendimiento de los parámetros alternativos de determinación de beneficiarios de programas y transferencias sociales. Se trata de un tema importante, ya que las autoridades suelen utilizar variables representativas para evaluar el nivel de renta o de consumo de un hogar o individuo con el objeto de determinar si les corresponde o no ser incluidos en dichos programas. Las regresiones logísticas pueden ayudar a seleccionar las mejores variables representativas para identificar a pobres y no pobres, o bien más ampliamente para seleccionar beneficiarios

Definición:

Sea Y una variable dependiente binaria que toma dos valores posibles (0 y 1).

Sean X_1, \dots, X_k un conjunto de variables independientes observadas con el fin de explicar y/o predecir el valor de Y .

El objetivo es determinar:

$$P[Y=1/ X_1, \dots, X_k], \text{ donde } P \text{ indica probabilidad}$$

$$\text{Así } P[Y=0/ X_1, \dots, X_k], = 1 - P[Y=1/ X_1, \dots, X_k].$$

Se construye un modelo:

$$P[Y=1/ X_1, \dots, X_k] = p(X_1, \dots, X_k; \beta)$$

Donde $p(X_1, \dots, X_k; \beta)$ es una función que recibe el nombre de función de enlace (función de probabilidad) cuyo valor depende de un vector de parámetros.

$$\beta = (\beta_1, \dots, \beta_k)'$$

El modelo de regresión logística parte de la hipótesis de que los datos siguen el siguiente modelo:

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k + u = xb + u$$

Con el fin de simplificar la notación, definimos Z :

$$z = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k$$

Por lo tanto, el modelo se puede representar como:

$$\ln\left(\frac{p}{1-p}\right) = z + u$$

Donde p es la probabilidad de que ocurra el suceso de estudio.

Operando algebraicamente sobre el modelo:

$$\ln\left(\frac{p}{1-p}\right) = z$$

$$\frac{p}{1-p} = e^z$$

$$p = (1-p) * e^z$$

$$p = e^z - p * e^z$$

$$p + p * e^z = e^z$$

$$p(1 + e^z) = e^z$$

$$p = \frac{e^z}{1 + e^z}$$

Como la función de distribución logística es:

$$F(x) = \frac{e^x}{1 + e^x}$$

Por tanto, podemos reescribir el modelo de forma mucho más compacta:

$$p = \frac{e^x}{1 + e^x} = F(z) = F(x * b)$$

De donde se deduce que el modelo de regresión logística es, en principio, un modelo de **regresión no lineal**, pero es lineal en escala logarítmica atendiendo a su definición original:

$$\ln\left(\frac{p}{1-p}\right) = z$$

$$\ln(p) - \ln(1-p) = z$$

$$\ln(p) - \ln(1-p) = b_0 + b_1x_1 + b_1x_2 + \dots + b_kx_k$$

Es decir, la diferencia de la probabilidad de que ocurra un suceso respecto de que no ocurra es lineal, pero en escala logarítmica. Por tanto, el significado de los coeficientes, aunque guardando una cierta relación con el modelo de regresión lineal, va a ser algo más complejo de interpretar.

Recordemos las dos formas más importantes de expresar el modelo de regresión logística:

$$\ln(p) - \ln(1-p) = b_0 + b_1x_1 + b_1x_2 + \dots + b_kx_k$$

$$\frac{p}{1-p} = e^{b_0} * e^{b_1x_1} * e^{b_2x_2} * \dots * e^{b_kx_k}$$

La primera expresión se llama **logit** y a la segunda **Odds ratio** o cociente de probabilidades.

Estimación de los parámetros.

Brevemente, vamos a ver en esquema el problema que ofrece, en el caso de regresión logística, la estimación de los parámetros.

Sea una muestra de **n** elementos, donde se ha observado la variable respuesta Y (que sólo puede tomar dos valores: cero y uno) y la variable X.

La función de probabilidad de una observación cualquiera es:

$$P(Y = 1/x) = p$$

$$P(Y = 0/x) = 1 - p$$

Por tanto:

$$P(Y/x) = p^y * (1 - p)^{1-y}$$

Por tanto, la función de probabilidades de la muestra es:

$$P(y_1, y_2, \dots, y_n) = \prod_i p_i^{y_i} * (1 - p_i)^{1-y_i}$$

Esta expresión recibe el nombre de verosimilitud de la muestra (likelihood).

Tomando logaritmos:

$$\log P(Y) = \sum_i^n y_i * \text{Log} \left(\frac{p_i}{1 - p_i} \right) + \sum_i^n \log(1 - p_i)$$

Expresando P_i en función de los parámetros que deseamos estimar:

$$L(B) = \sum_i^n y_i * x_i * b - \sum_i^n \text{Log}(1 + e^{x_i * b})$$

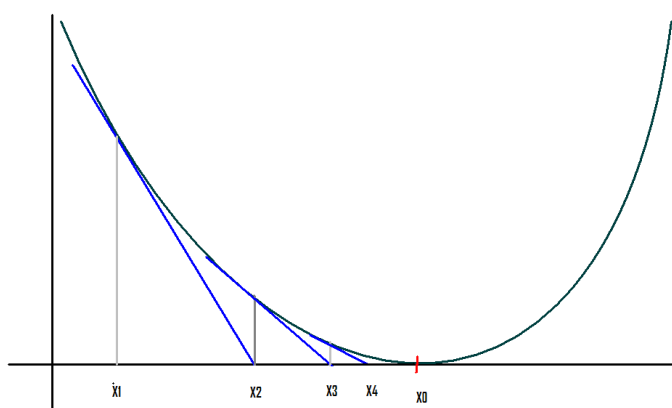
Resulta obvio que, aunque derivemos y establezcamos la condición de máximo, no vamos a poder despejar los coeficientes B.

La solución que vamos a obtener es:

$$B_a = B_0 + \left(-\frac{\partial^2 L(B)}{\partial B * \partial B'} \right)^{-1} * \left(\frac{\partial L(B)}{\partial B} \right)$$

Esta solución establece cómo encontrar una solución a partir de un punto próximo cualquiera, denominado. Por lo tanto, deberemos de hacer una estimación inicial del valor de los verdaderos parámetros y mediante un procedimiento recursivo encontrar el verdadero valor de los mismos. Para encontrar los verdaderos valores se suele utilizar el algoritmo de Newton-Raphson.

Ilustración 3: Newton-Raphson



Función de Verosimilitud

Con el fin de estimar β y analizar el comportamiento modelo, observamos una muestra aleatoria simple de tamaño n dada por

$\{(X_i^j), Y_i ; i=1, \dots, n\}$ donde $X_i = (X_{i1}, \dots, X_{ik})$, es el valor de las variables independientes e $Y_i = \{0, 1\}$ es el valor observado de Y en el i -ésimo elemento de la muestra.

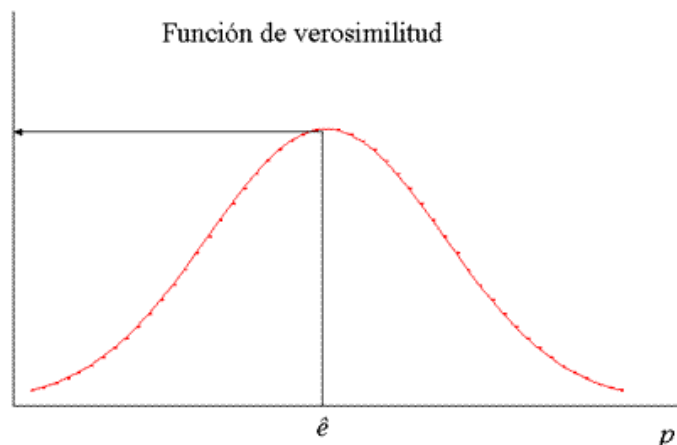
$Y/(X_1, \dots, X_k)$ se distribuye Binomial $(1, p(Y=1/ X_1, \dots, X_k; \beta))$.

La función de verosimilitud es:

$$L(\beta/(X_1, Y_1) \dots \dots \dots, (X_n, Y_n)) = \prod_1^n P_i^{y_1} (1 - P_i)^{1-y_1} \quad (1)$$

Donde $p_i = p(x_1', \beta) = p(x_{i1}, \dots, x_{ik}; \beta)$ con $i = 1, \dots, n$

Ilustración 4: Función de verosimilitud



Se demuestra que la mejor estimación del parámetro \hat{e} es aquel valor que maximiza esta función de verosimilitud, ya que son estimadores consistentes (conforme crece el tamaño muestral, la estimación se aproxima al parámetro desconocido), suficientes (aprovechan la información de toda la muestra), asintóticamente normales y asintóticamente eficientes (con mínima varianza), si bien no siempre son insesgados (no siempre la media de las estimaciones para diferentes muestras tenderá hacia el parámetro desconocido).

2.2.6. Modelo de Regresión Logística múltiple

El modelo matemático se construye en base a probabilidades, las cuales se obtienen considerando la probabilidad de que ocurra un suceso determinado $P(Y)$ en relación con la dependencia de que dicha probabilidad no ocurra $1 - P(Y)$.

La probabilidad proporciona predicciones consistentes y fáciles de los resultados en términos del "Odds" del evento $Y=1$.

$$Odds(Y = 1) = \frac{P(Y)}{1 - P(Y)} \quad (2)$$

El modelo de regresión logística múltiple, relaciona la probabilidad de que ocurra un determinado suceso denotado por el vector

$X' = (X_1, \dots, X_k)$ con probabilidad condicional $P(Y=1| X)$ en función de k variables independientes que pueden ser cuantitativas, cualitativas o ambas según sea el tipo de diseño de estudio.

El modelo logístico múltiple es:

$$L = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i \quad (3)$$

O también:

$$P_i = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_i X_i}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_i X_i}} = \frac{1}{1 + e^{-z}} \quad (4)$$

Dónde:

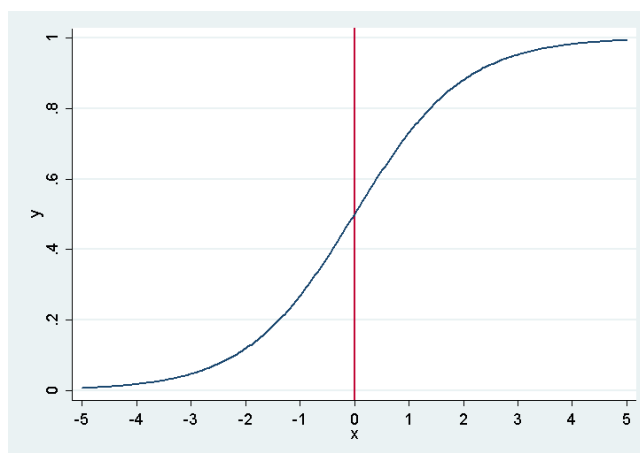
- \ln es el logaritmo natural de "Odds" también denominado "Logit" o "L"
- $\beta_0, \beta_1, \dots, \beta_k$, son constantes
- X una variable explicativa que puede ser continua o discreta.

Como los coeficientes del modelo logístico no tienen restricciones éstos son fácilmente interpretables en términos de independencia o asociación entre las variables.

Gráficamente la función es simétrica

- Continua y creciente sobre el intervalo 0 y 1.
- Sigue una curva sigmoidea.

Ilustración 5: Modelo de Regresión Logística



2.2.7. Estimación de los Parámetros

Sea una muestra de n observaciones independientes definido por $(X_{i1}, X_{i2}, \dots, X_{ik}; Y_i)$, $i=1, \dots, n$; se elige el vector $\beta' = (\beta_0, \beta_1, \dots, \beta_k)$ el método más usado es el de verosimilitud, con $k+1$ ecuaciones de verosimilitud que se obtienen derivando el \ln de la función de verosimilitud respecto a $k+1$ coeficientes.

Las ecuaciones de verosimilitud son:

$$\sum_{i=1}^n (y_i - P_i) = 0$$

$$\sum_{i=1}^n X_{ij} (y_i - P_i) = 0, j = 1, \dots, k; i = 1, \dots, n$$

Mediante cálculo diferencial se encuentran las soluciones a este conjunto de ecuaciones, actualmente existen software estadístico para estimar los parámetros.

2.2.8. Pruebas de Significación

Una vez estimado los coeficientes del modelo, se tiene que verificar si el modelo predice de manera adecuada a la variable dependiente. Para evaluar la bondad del modelo se utiliza el logaritmo del cociente de verosimilitud y la prueba de Hosmer-Lemeshow.

Modelo

Se procede de la siguiente manera:

1. Se prueba la hipótesis de que el modelo encontrado es el que mejor se ajusta a través de la razón de verosimilitud.

$$D = -2 \ln \frac{\text{verosimilitud del modelo sin la variable (modelo analizado)}}{\text{verosimilitud del modelo con la variable (modelo saturado)}}$$

La diferencia entre estos dos valores de $-2 \ln$ se llama **Devianza** y prueba si la o las variables X_i son significativas.

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_k = 0$$

$$H_1 : \text{Algún } \beta_i \text{ distinto de } 0.$$

Estadístico de prueba $D \sim \chi^2$ con $n-k-1$ grados de libertad.

Región de Rechazo: si $D > \chi^2_{\alpha, (n-k-1)}$ si se rechaza H_0 , significa que al menos uno de los coeficientes es diferente de cero y la variable correspondiente

es significativa con respecto a la probabilidad de que ocurra o no el suceso en estudio.

2. Cuando el número de variables del modelo es grande y/o tiene variables independientes cuantitativas, se debe utilizar la prueba de Hosmer-Lemeshow, como consecuencia de la gran cantidad de variables, el número de pautas de variación existente entre ellas es tan elevado que puede invalidar la utilidad de los estadísticos de bondad de ajuste clásicos.

Evalúa la bondad del modelo construyendo una tabla de contingencia, divide la muestra en 10 grupos a partir de los deciles de las probabilidades estimadas (a los 10 grupos resultantes se les denomina deciles de riesgo). En cada decil de riesgo se calcula el número de casos que pertenecen a cada categoría de la variable dependiente (número observado) y el número de casos que el modelo pronostica que pertenecen a cada categoría de la variable dependiente (número esperado). A continuación, se compara el número de casos observado con el número de casos esperado (esta comparación se realiza en cada una de las 20 casillas definidas por la combinación de las 2 categorías de la variable dependiente con los 10 deciles de riesgo).

Este estadístico permite contrastar la hipótesis nula de igualdad de distribuciones, es decir, la hipótesis de que la variable dependiente se distribuye de la misma manera en los 10 deciles de riesgo o, que no existen diferencias entre las frecuencias observadas y las esperadas. El estadístico de contraste tiene siempre k grados de libertad.

Las hipótesis que se contrastan son:

H_0 : El modelo es adecuado

H_1 : El modelo ajustado no es adecuado

Decisión si el estadístico de prueba es mayor o igual que

$\chi^2_{\alpha, (n^{\circ} \text{ de grupos} - 2)}$, se rechaza H_0 y se concluye que el modelo no es adecuado con un nivel de significación α .

Pruebas de coeficientes

Una vez encontrado el mejor conjunto de variables explicativas que predicen la variable dependiente Y , se debe evaluar cada coeficiente para determinar cuál o cuáles ingresan al modelo, este proceso se realiza mediante el estadístico Wald.

El estadístico de Wald

Contrasta la hipótesis de que un coeficiente aislado es distinto de 0, y sigue una distribución normal de media 0 y varianza 1. Su valor para un coeficiente concreto viene dado por el cociente entre el valor del coeficiente y su correspondiente error estándar. La obtención de significación indica que dicho coeficiente es diferente de 0 y merece la pena su conservación en el modelo. En modelos con errores estándar grandes, el estadístico de Wald puede proporcionar falsas ausencias de significación (es decir, se incrementa el error tipo II). Tampoco es recomendable su uso si se están empleando variables de diseño.

La hipótesis que se plantea es la siguiente:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \text{ no es igual a } 0$$

Para contrastar la hipótesis señalada se usa el estadístico de Wald:

$$WALD = \frac{\beta_i^2}{S_{\beta_i}^2} \sim X^2_{\alpha,1} \quad (5)$$

En el caso de la regresión logística multivariada corresponde a un vector donde cada celda es la división entre el coeficiente β_i y el error estándar de éste. Donde S_{β_i} , es el error estándar del coeficiente de regresión logística muestral y k es el número de variables independientes.

Se formula la hipótesis

$$H_0 : \text{La variable independiente no influye sobre } p_i.$$

$$H_1 : \text{La variable independiente influye sobre } p_i.$$

2.2.9. Interpretación de los Resultados

La interpretación de los resultados obtenidos se realiza a partir de los coeficientes del modelo. Para ello basta tener en cuenta que, si el modelo ajustado es adecuado, entonces se dice que el modelo es significativo. Además, se debe analizar el grado de asociación estadística que existe en sus parámetros.

Si:

$\beta_1 > 0$ el factor de riesgo será mayor que 1 y $p(X_1, X_2, \dots, X_k; \beta)$ aumentará

$\beta_1 < 0$ el factor de riesgo será menor que 1 y $p(X_1, X_2, \dots, X_k; \beta)$ disminuirá.

$\beta_1 = 0$ la variable X_1 no ejerce ningún efecto sobre la probabilidad de riesgo.

Si el modelo de Regresión Logística es significativo y una de variables independientes es dicotómica con valores de 0 y 1, el número e^{β_i} es el OR, denominado factor de riesgo o protección que implica un aumento unitario de la variable independiente.

En el caso de una variable cuantitativa, e^{β_i} es el número de veces que aumenta la "chance" de que ocurra el suceso, en este caso de ser pobre, por cada unidad de aumento de la variable independiente.

2.2.10. Selección de Variables

Pero, del conjunto de variables que pueda tener un estudio, ¿qué variables deben introducirse en el modelo? El modelo debe ser aquél más reducido que explique los datos (principio de parsimonia), y que además sea congruente e interpretable. Hay que tener en cuenta que un mayor número de variables en el modelo implicará mayores errores estándar.

Deben incluirse todas aquellas variables que se consideren importantes para el modelo, con independencia. Por otro lado, no debería dejarse de incluir toda variable que en un análisis univariado previo demostrara una relación "suficiente" con la variable dependiente.

Una cuestión importante a tener en cuenta es el correcto manejo de las variables cualitativas transformadas en varias variables ficticias.

Siempre que se decida incluir (o excluir) una de estas variables, todas sus correspondientes variables ficticias deben ser incluidas (o excluidas) en bloque. No hacerlo así implicaría que se habría recodificado la variable, y por tanto la interpretación de la misma no sería igual.

Otro aspecto de interés es la significación que pudiera tener cada variable ficticia. No siempre todas las variables ficticias de una covariable son significativas, o todas no significativas. En estos casos es recomendable contrastar el modelo completo frente al modelo sin la covariable mediante la prueba de razón de verosimilitud (es decir, se sacarían del modelo en bloque todas las variables ficticias de la covariable de interés).

Una vez que se dispone de un modelo inicial debe procederse a su reducción hasta obtener el modelo más reducido que siga explicando los datos. Para ello se puede recurrir a métodos de selección paso a paso, bien mediante inclusión "hacia adelante" o por eliminación "hacia atrás", o a la selección de variables por mejores subconjuntos de covariables. Estos métodos se encuentran implementados en numerosos paquetes estadísticos, por lo que son muy populares. Dado que para la comprensión de los métodos de selección paso a paso se requiere un conocimiento previo acerca del ajuste del modelo, éste es un aspecto que debe ser tratado en otro momento; se sugiere al lector que se introduzca en este aspecto una vez tenga conocimientos sobre el análisis del ajuste del modelo. No obstante, hay que advertir que su uso nunca puede sustituir a la valoración juiciosa de los

modelos que van surgiendo de forma seriada en cada paso y del modelo final. No hacerlo así puede llevar a dar por bueno un modelo surgido de forma automática (por criterios preestablecidos por el paquete estadístico muchas veces mal conocidos por el usuario del software), con escaso valor clínico.

Cada vez que se encuentre ante un modelo de regresión logística (el inicial, cualquiera de los intermedios o el final), se tendrá que contrastar su significación global, mediante las pruebas de ajuste global del modelo.

Una vez se dispone un modelo preliminar, se podrían incluir factores de interacción, es decir, estudiar cómo la asociación de dos o más covariables puede influir en la variable dependiente. Existen estrategias de desarrollo de modelos de regresión por las que se recomienda la inclusión en el modelo inicial de todas las covariables necesarias más las interacciones de las mismas, o por lo menos, las interacciones de primer orden (tomadas las covariables dos a dos), a los que se les llama modelos saturados. Interacciones de mayor orden suelen ser de difícil interpretación. En cualquier caso, siempre hay que tener presente las limitaciones de tamaño muestral (que se verán luego), y de interpretación desde el punto de vista clínico (no se deberían incluir interacciones de significado incierto).

Otra estrategia en el desarrollo del modelo final es el diseño y ajuste de un modelo final preliminar sin interacciones, en el que luego se ensayarían la inclusión, uno por uno, de términos de interacción que pudieran tener traducción clínica (Hosmer y Lemeshow), y valorar su significación respecto del modelo previo sin interacciones.

Una vez se haya decidido la inclusión de un factor de interacción, se tendrá en cuenta que siempre deberán estar incluidas también de forma aislada en el modelo las covariables que componen la interacción (principio jerárquico).

Por otra parte, y en relación con la inclusión de interacciones, hay que tener en cuenta que la inclusión de las mismas puede generar multicolinealidad, tanto más probable cuanto mayor sea el número de interacciones. Siempre debe considerarse la suficiencia del tamaño muestral para el número de covariables que se desea incluir en el modelo: modelos excesivamente grandes para muestras con tamaños muestrales relativamente pequeños implicarán errores estándar grandes o coeficientes estimados falsamente muy elevados (sobreajuste). En general se recomienda que por cada covariable se cuente con un mínimo de 10 individuos por cada evento de la variable dependiente con menor representación (Peduzzi). Por ejemplo, si la variable dependiente Y es “muerte” y en los datos hay 120 sujetos vivos y 36 sujetos muertos, el evento de Y menos representado es “muerte”, con 36 sujetos; de esta forma el modelo no debería contener más de $36/10 \sim 3$ covariables.

Lo anterior es válido siempre que se trate de covariables cuantitativas o cualitativas con distribuciones bien equilibradas. La situación se complica si una o más de las covariables cualitativas no tiene una distribución equilibrada (uno de sus dos valores tiene una mínima representación); en ese caso se recomienda que, en su tabla de contingencia respecto a la variable dependiente, en cada celda haya un mínimo de 10 observaciones.

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. Población

Según el estudio, la población se define como el promedio de clientes solicitantes de crédito en un mes del año 2015, que es de 110 clientes.

3.2. Diseño de la Muestra

Cálculo del Tamaño de la Muestra

El tipo de muestra que se utilizó es no probabilístico para población finita, considerando un nivel de confianza de 95%.

Tamaño de la muestra para la población finita y conocida:

Donde:

$$n = \frac{Z_{\alpha}^2 \cdot N \cdot p \cdot q}{i^2(N - 1) + Z_{\alpha}^2 \cdot p \cdot q}$$

n: tamaño muestral

N: tamaño de la población

Z: valor correspondiente a la distribución de gauss, $Z\alpha = 0.05 = 1.96$ y $Z\alpha = 0.01 = 2.58$ (en este caso $0.05 = 1.96$)

p: proporción esperada (en este caso 10% según criterio financiero = 0.10).

q: $1 - p$ ($1 - 0.10 = 0.90$)

i: error que se prevé cometer si es del 5%, $i = 0.05$

Reemplazando en la fórmula anterior.

$Z\alpha = 0.05 = 1.96$

$N = 110$

$p = 0.10$

$q = 1 - p = 1 - 0.10 = 0.90$

$i = 5\% = 0.05$

Reemplazando los valores nos resulta: $n = 61$

Considerando una muestra de 61 clientes.

3.3. Recolección de la Información

Para la recolección de datos se utilizó una entrevista tipo cuestionario con preguntas cerradas y específicas. La entrevista está adecuada de manera que las respuestas sean concretas. (Anexos pag.77)

3.3.1. Cuestionario

El cuestionario cuenta con varias partes socioeconómicas.

- Parte económica: Preguntas que incluyen, número de perceptores de ingreso, ingreso mensual y gasto mensual.
- Parte del trabajo: Preguntas sobre el tipo de trabajo de los clientes.

- Parte social: Preguntas relativas al nivel educativo, número de hijos, edad, estado civil y tipo de seguro de salud.
- Parte de la condición de la vivienda: Preguntas relativas a estado de la vivienda que ocupa el cliente como, condición de la vivienda y tipo del material de construcción.
- Parte de los servicios con los que cuenta la vivienda: Preguntas relacionadas con, acceso a los servicios básicos y accesorios en la vivienda.
- Parte respecto a la experiencia crediticia: Preguntas relacionadas a si tiene experiencia en préstamo, etc.

3.3.2. Operacionalización de variables

Tabla 1: Operacionalización de Variables

VARIABLES	DIMENSIÓN	INDICADOR	ÍNDICE		
Variable X (independiente): factores socio económicos	Económico	Número de perceptores de ingresos	Cuantitativo		
		Ingreso mensual	Cuantitativo		
		Gasto mensual	Cuantitativo		
	Trabajo	Categoría de ocupación	Empleador Asalariados Trabajador independiente Trabajador familiar no remunerado		
	Sociales	Nivel de educación alcanzado	Sin estudios Primaria Secundaria Sup.No universitaria Sup.Universitaria		
			Número de hijos	Cuantitativo	
			Edad	Cuantitativo	
			Estado civil	Soltera Casada Viuda Conviviente	
		Seguro de salud		Únicamente SIS Únicamente EsSalud Con otros seguros de salud	
				Vivienda	Propia Alquilada Otra forma
		Material de construcción	Ladrillo y cemento Adobe Madera Otro material		
	Servicios		Acceso a servicios básicos	Servicio agua y desagüe Servicios de alumbrado eléctrico	
		Accesorios en la vivienda		Televisor Computadora Vehículo	
		Variable Y (dependiente): Otorgamiento de Crédito	Experiencia en crédito	Crédito	1
	No crédito			0	

Fuente:Elaboración propia

El Modelo Logístico es apropiado en aquellas situaciones donde se analiza una única variable dependiente categórica o nominal y varias variables independientes.

La variable dependiente bajo estudio, en este caso, es el otorgamiento de crédito, definido desde un punto de vista cualitativo.

Cuando la variable dependiente es dicótoma, se necesita utilizar una variable ficticia para definirla; dicha variable consiste en asignar dos valores (0 y 1) a la variable dependiente, según sea su categoría, de allí la denominación de variables categóricas. En este estudio se asigna a la variable dependiente – otorgamiento de crédito (OC) – dos categorías, según la experiencia con otras entidades financieras; tiene experiencia, le corresponde la categoría “1” y no tiene experiencia, se le asigna la categoría “0”.

Las variables independientes: (factores socioeconómicos) considerados en este estudio y cuyo grado de incidencia sobre el otorgamiento de crédito se pretenden determinar, pueden clasificarse en cuatro dimensiones:

– **Económico:** cuyos indicadores son:

- Ingreso mensual, se clasificó en: si es \geq básico S/. 750.00 es 1 y si es $<$ 0, para convertirlo en variable dicotómica.
- Gasto mensual, variable cuantitativa.
- Número de perceptores, variable cuantitativa.

– **Tipo de trabajo:** Indicador:

- Ocupación, se clasificó en: empleador (1), y trabajo independiente (0), que agrupa a los asalariados, trabajo independiente y trabajo en casa; lo que nos permite convertirlo en variable dicotómica.

- **Sociales:** Indicadores:
 - Nivel de educación, se clasificó en: educación media (si 1 y no 0), que agrupa a los sin estudios, primaria y secundaria; y educación superior (si 1 y no 0) que a su vez agrupa a sup. no universitaria y sup. universitaria.

 - Número de hijos, es una variable cuantitativa.

 - Edad, cuantitativa.

 - Estado civil, se clasifico en: soltero (1), que agrupa a los que son solteros, viudos y divorciados; y casados (0), que agrupa a los casados y convivientes.

 - Seguro de salud, se clasifico en: tienen seguro (1), que agrupa a los que tienen SIS, EsSalud, SIS/EsSalud y otros seguros de salud; no tienen seguro (0).

- **Vivienda:** indicadores.
 - Condición de la vivienda, se clasifica en: propia (1) y alquilada (0). Variable cualitativa.

 - Material de construcción, se clasifica en: ladrillo y cemento (1) y material rústico (0), que agrupa a los de adobe y madera. Variable cualitativa.

- **Servicios:** Indicadores:

- Acceso a servicios básicos; servicios de agua y desagüe (si 1 y no 0) y servicio de alumbrado eléctrico (si 1 y no 0).
- Vienes; televisor (si 1 y no 0), computadora (si 1 y no 0) y vehículo motorizado (si 1 y no 0).

3.4. Metodología

La metodología utilizada es un modelo Logístico para estimar la probabilidad de conseguir el otorgamiento de crédito en función de distintas variables. Esta metodología permite evaluar cuáles son los determinantes que influyen en el otorgamiento de crédito. Para este fin se utilizó una base de datos para conducir el análisis, la información de la encuesta. Para empezar, se busca identificar la cantidad de quienes tienen experiencia en crédito o los que no tienen experiencia en crédito, que tienen las clientas, para luego poder elegir cuales son las variables más adecuadas para poder regresionar y obtenerlos los determinantes.

Mediante la regresión logística:

$$P_i = \frac{e^{\beta_0 + \beta_1 x_i + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_i + \dots + \beta_k x_k}} = \frac{1}{1 + e^{-z}}$$

P_i : Representa el otorgamiento de crédito.

0= no experiencia en crédito

1= experiencia en crédito

B_0 : Es la media general de los datos

B_1, B_2, \dots, B_i : Coeficientes de los factores de crédito

X_i : Factores de crédito.

Para analizar la bondad de ajuste de un modelo Logístico, se deben considerar: el *Odds Ratio*, el Estadístico de Wald, el Valor de Verosimilitud, el Coeficiente de Determinación y el Test de Hosmer y Lemeshow; los dos primeros hacen referencia al nivel de significación individual de cada variable independiente, mientras que los demás dan cuenta del ajuste global del modelo. A continuación, se hace una breve referencia a cada uno de ellos siguiendo a Hair *et al.* (1999) y Johnson (2000).

- *Odds Ratio* (razón de posibilidades): siendo β_i el coeficiente correspondiente a la variable x_i , se define como *Odds Ratio* a la exponencial $\text{Exp}(\beta_i)$, un signo positivo en el exponente provoca un aumento en la probabilidad de ocurrencia del evento; un signo negativo reduce dicha probabilidad y un coeficiente cercano a cero da lugar a un valor próximo a la unidad, que prácticamente no incide en la probabilidad de ocurrencia (o no) del evento.
- Estadístico de Wald: hace referencia a la significación individual de cada variable y su interpretación es muy similar a la del Estadístico t para la Regresión Lineal Múltiple. Si el nivel de significación asociado es menor que el seleccionado (α), la variable en cuestión es relevante y debe ser tomada en cuenta en el modelo.
- Valor de Verosimilitud (-2LL): se obtiene como dos veces el logaritmo de la verosimilitud con signo contrario; un buen ajuste del modelo proporciona valores pequeños; un modelo ideal tendría una verosimilitud de uno, con lo cual su Valor de Verosimilitud sería nulo. Esta medida de ajuste global, compara las probabilidades predichas por el modelo con las observadas y

mediante una prueba Chi - Cuadrado, indica si existe ó no diferencia significativa en la reducción del (-2LL) entre el modelo inicial y el final. Este contraste es equivalente al Estadístico F en un modelo de Regresión Lineal Múltiple.

- Coeficiente R2 logístico: este coeficiente puede asumir valores entre cero y uno, cuanto más cercano a la unidad se encuentre, tanto mejor será el ajuste del modelo. Esta medida está directamente relacionada con el Valor de Verosimilitud, ya que también mide el grado de reducción del mismo entre los modelos inicial y final.
- Test de Hosmer y Lemeshow: es la medida final para valorar el ajuste del modelo; analiza mediante una prueba Chi – Cuadrado, si existe ó no, diferencia significativa entre los valores observados y los predichos por el modelo para la variable dependiente.

3.5. Tratamiento de datos

3.5.1. El programa SPSS

SPSS permite efectuar tanto análisis estadísticos básicos como avanzados. En la mayor parte de las ocasiones, las organizaciones necesitan reportes descriptivos del proyecto. Por ejemplo, una institución que recolectó información sobre sus estudiantes querrá tener un perfil del estudiantado que incluya una descripción del tipo de estudiante que asiste a la institución, sus características de edad, intereses, ingreso familiar, lugar de origen, etc. SPSS podrá apoyar en el desarrollo de este perfil a través de diversos análisis descriptivos básicos de su base de datos.

En otros casos, se comparan las características de dos o más grupos con respecto a diversas variables: por ejemplo, para saber si existe una diferencia en el desempeño de los estudiantes según su género. SPSS permite responder a esta pregunta a través de procedimientos más avanzados como la Prueba-T. Igualmente, si se quiere comparar el desempeño de estudiantes dependiendo de su nivel socioeconómico, existen otros procedimientos para análisis estadísticos como el ONE-way ANOVA con el cual se pueden comparar más de dos grupos.

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1. Determinantes del Otorgamiento de Crédito.

En relación al otorgamiento de crédito de la financiera Credinka en la ciudad de Ayaviri, es necesario mencionar que el “Otorgamiento de Crédito” se mide respecto a los que tienen experiencia en crédito su clasificación es crédito (1), a los que tienen experiencia en créditos financieros y no crédito (0), a los que nunca hicieron un crédito financiero. En efecto; la distribución del índice del Otorgamiento de Crédito no es homogéneo, existiendo disparidad en la distribución entre los clientes; los que tuvieron experiencia anteriormente cuentan con mayor posibilidad de que se les otorgue un crédito, mientras que por el otro lado, los que no tiene experiencia en crédito incurren a mayor número de veces de salir desaprobado en la solicitud de crédito.

4.2. Modelo para Otorgamiento de Crédito

$$\text{Logit } P(\text{pobreza}) = \beta' \hat{X}$$

con $X' = (1, \text{Perceptores, Ingreso, Gasto, Trabajo, Educación media, educación superior, Hijos, edad, Estado civil, Seguro, Vivienda, Material Vivienda, Servicios agua desague, Servicio Eléctrico, televisor, computadora, vehículo})$

$$\hat{\beta} = \hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_{17}$$

4.2.1. Proceso de Selección de variables

A partir del modelo completo, la selección de variables se realizó en base a la probabilidad del coeficiente de regresión estimado por cada variable con un nivel de significación $\alpha = 0.05$. Este ajuste se efectuó mediante el software estadístico SPSS versión 20.

4.2.2. Parámetros Estimados del Modelo Completo

En la Tabla 2 se presentan los parámetros del modelo completo, considerando 20 iteraciones que arrojó el software SPSS.

Se utilizó el estadístico de Wald para la prueba de hipótesis de cada uno de los coeficientes.

Las hipótesis son las siguientes:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \text{ no es igual a } 0$$

Tabla 2: Variables y coeficientes del modelo completo

	B	E.T.	Wald	Gl	Sig.	Exp(B)
Perceptores	-19,724	17889,041	,000	1	,999	,000
Gasto	-,005	,006	,801	1	,371	,995
Trabajo(1)	-2,641	2,023	1,705	1	,192	,071
Educ.Medía(1)	19,334	40192,969	,000	1	1,000	249346323,059
Educ.Superior(1)	1,485	1,199	1,534	1	,216	4,415
Hijos	-,656	,578	1,286	1	,257	,519
Edad	,163	,078	4,364	1	,037	1,177
Paso 1 ^a Estado.Civil(1)	20,349	17889,041	,000	1	,999	688044495,529
Seguro(1)	,259	1,028	,064	1	,801	1,296
Vivienda(1)	1,803	1,115	2,614	1	,106	6,069
Material(1)	,660	,766	,742	1	,389	1,934
Agua_Desague(1)	11,076	40192,970	,000	1	1,000	64591,592
Televisor(1)	-23,022	40192,970	,000	1	1,000	,000
Computadora(1)	,917	1,286	,509	1	,476	2,502
Ingreso1	,008	,005	3,479	1	,062	1,008
Constante	10,777	17889,041	,000	1	1,000	47898,103

a. Variable(s) introducida(s) en el paso 1: Perceptores, Gasto, Trabajo, Educ.Medía, Educ.Superior, Hijos, Edad, Estado.Civil, Seguro, Vivienda, Material, Agua_Desague, Televisor, Computadora, Ingreso1.

Fuente: producida por el programa SPSS.

Como se mencionó en el capítulo II, este estadístico se distribuye $X^2_{(1)}$. El valor crítico utilizado es de 5%, por lo cual se rechaza la hipótesis nula si el valor del estadístico de WALD es mayor que el valor crítico.

$$WALD = \frac{\beta^2_i}{S^2_{\beta_{1i}}} \sim X^2_{\alpha,1} = 0.004$$

$$X^2_{\alpha(n-k-1)} = X^2_{0.050(61-16-1)} = X^2_{0.050(44)} = 30.61$$

4.2.3. Bondad de Ajuste del modelo con 7 variables

Se eligió el mejor modelo logístico a través del método denominado “Adelante Condicional” de programa SPSS, el cual correspondió a un modelo con 7 variables explicativas.

Las Hipótesis son:

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_k = 0$$

H_1 : Algún β_i distinto de 0.

Este estadístico se distribuye como $X^2_{(n-k-1)}$, y se rechaza la hipótesis nula si el valor de $-2 \log - \text{Likelihood}$ es mayor que $X^2_{\alpha(n-k-1)}$, donde α es el nivel de significación, siendo en este caso de 0,05. Se rechaza H_0 , por lo tanto el modelo con 7 variables contribuye igual que el modelo saturado.

$$X^2_{\alpha(n-k-1)} = X^2_{0.05(55)} = 38.96$$

Tabla 3: Resumen del modelo

Paso	-2 log Likelihood	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	57,608 ^a	,213	,306

a. La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Fuente: producida por el programa SPSS

Si se utiliza el Test de Hosmer y Lemeshow³⁴, sugerido para un número de variables significativo, se plantea la siguiente hipótesis:

H_0 : El modelo ajustado es significativo, se ajusta correctamente

H_1 : El modelo ajustado no es significativo, no se ajusta correctamente.

Tabla 4: Test de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	4,316	8	,828

Fuente: producida por el programa SPSS.

En la tabla anterior se observa que no hay antecedentes para rechazar H_0 , por lo cual se concluye que el modelo ajustado es significativo.

El mejor modelo para el otorgamiento de crédito, se obtuvo con 7 variables, que son las siguientes:

Tabla 5: Variables en el Modelo

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)		
							Inferior	Superior	
Paso 1 ^a	Ingreso1	,004	,003	1,915	1	,166	1,004	,998	1,010
	Gasto	-,003	,005	,308	1	,579	,997	,988	1,006
	Hijos	-,205	,406	,254	1	,614	,815	,368	1,807
	Edad	,124	,062	4,030	1	,045	1,132	1,003	1,279
	Seguro(1)	-,299	,835	,128	1	,720	,741	,144	3,809
	Vivienda(1)	,566	,788	,516	1	,473	1,761	,376	8,247
	Material(1)	,640	,704	,828	1	,363	1,897	,478	7,534
	Constante	-5,960	2,691	4,905	1	,027	,003		

a. Variable(s) introducida(s) en el paso 1: Ingreso1, Gasto, Hijos, Edad, Seguro, Vivienda, Material.
Fuente: producida por el programa SPSS.

La mejor ecuación de regresión logística para predecir el otorgamiento de crédito es:

$$P_i = \frac{e^{\beta_0 + \beta_1 x_i + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_i + \dots + \beta_k x_k}} = \frac{1}{1 + e^{-z}}$$

$$P(\text{crédito}) = \frac{1}{1 + e^{-(-5.960 + 0.004 \text{Ingreso} - 0.003 \text{Gasto} - 0.205 \text{Hijos} + 0.124 \text{Edad} - 0.299 \text{Seguro} + 0.566 \text{Vivienda} + 0.640 \text{Material})}}$$

4.2.4. Interpretación de los parámetros estimados del Modelo – ODDS Ratio

El ODDS Ratio cambia cuando la i-ésima variable explicativa regresora se incrementa en una unidad, si:

$\beta_i > 0$ significa que el ODDS RATIO se incrementa.

$\beta_i < 0$ significa que el ODDS RATIO decrece.

$\beta_i = 0$ significa que el factor es igual a uno, lo cual hace que ODDS RATIO no varía.

En la columna $exp(\beta)$ de la tabla 5, se observa que:

$\beta_1, \beta_4, \beta_6$ y $\beta_7 > 0$, por lo tanto, por cada 1 unidad que aumente el ingreso, la posibilidad de otorgar un crédito aumenta en 1,004 veces si se mantiene el resto de las variables constantes. De la misma manera para edad, la probabilidad de que se otorgue crédito aumenta en 1,003 veces. Con respecto a la vivienda la probabilidad aumenta en 0,376 veces y por último el material de la vivienda se incrementa en 0,478 veces.

β_2, β_3 y $\beta_5 < 0$, por lo tanto, las variables, gasto, hijos y seguro, son factores que disminuyen la probabilidad de otorgamiento de crédito. Así, el cliente que su gasto sea alto, disminuye la probabilidad de que se le otorgue crédito en 0.997 veces si mantiene constante el resto de las variables.

4.2.5. Validación del Modelo Logístico

Para validar el modelo de regresión logística se utilizó la una encuesta, encontrándose un 78.7% de casos coincidentes entre la variable original y la estimación del modelo.

Tabla 6: Cruce de otorgamiento de crédito estimados y originales

	Observado		Pronosticado		
			Experiencia en crédito		Porcentaje correcto
			No crédito	Crédito	
Paso 1	Experiencia en crédito	No crédito	6	11	35,3
		Crédito	2	42	95,5
	Porcentaje global				78,7

a. El valor de corte es ,500
Fuente: producida por el programa SPSS.

4.3. Modelo para la Hipótesis

4.3.1. Selección de las Variables

Se corrió el programa con las variables: ingreso, gasto, tipo de trabajo y educación.

Tabla 7: Variables en la ecuación

	B	E.T.	Wald	Gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
Ingreso1	,004	,003	2,304	1	,129	1,004	,999	1,009
Gasto	-,002	,004	,312	1	,577	,998	,989	1,006
Educ.Superior(1)	1,370	,728	3,542	1	,060	3,936	,945	16,393
Constante	-2,347	1,528	2,359	1	,125	,096		

a. Variable(s) introducida(s) en el paso 1: Ingreso1, Gasto, Educ.Superior.
 Fuente: producida por el programa SPSS.

4.3.2. Parámetros

Las hipótesis son las siguientes:

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \text{ no es igual a } 0$$

El estadístico se distribuye $X^2_{(1)}$. El valor crítico utilizado es de 5%, por lo cual se rechaza la hipótesis nula si el valor del estadístico de WALD es mayor que el valor crítico.

4.3.3. Bondad de Ajuste

Las hipótesis son:

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_k = 0$$

$$H_1 : \text{Algún } \beta_i \text{ distinto de } 0.$$

Se distribuye como $X^2_{(n-k-1)}$, y se rechaza la hipótesis nula si el valor de $-2\log$ Likelihood es mayor que $X^2_{\alpha(n-k-1)}$, donde α es el nivel de significación, siendo en este caso de 0,05. Se rechaza H_0 , por lo tanto el modelo con 3 variables contribuye igual que el modelo saturado.

Tabla 8: Resumen del modelo

Paso	-2 log de la verosimilitud
1	63,228 ^a

Fuente: producida por el programa SPSS.

$$X^2_{\alpha(n-k-1)}=38.96$$

El Test de Hosmer y Lemeshow, plantea la siguiente hipótesis:

H_0 : El modelo ajustado es significativo , se ajusta correctamente

H_1 : El modelo ajustado no es significativo, no se ajusta correctamente.

Tabla 9: Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	Gl	Sig.
1	8,065	8	,427

Fuente: producida por el programa SPSS.

En la tabla anterior se observa que no hay antecedentes para rechazar H_0 , por lo cual se concluye que el modelo ajustado es significativo.

El mejor modelo es :

$$P_i = \frac{e^{\beta_0 + \beta_1 x_i + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_i + \dots + \beta_k x_k}} = \frac{1}{1 + e^{-z}}$$

$$P(\text{crédito}) = \frac{1}{e^{-(-2.347 + 0.004 \text{Ingreso} - 0.002 \text{Gasto} + 1.370 \text{Educación superior})}}$$

4.3.4. Interpretación de los parámetros Estimados – ODDS Ratio

$B_i > 0$ significa que el ODDS RATIO se incrementa.

$B_i < 0$ significa que el ODDS RATIO decrece.

$B_i = 0$ significa que el factor es igual a uno, lo cual hace que ODDS RATIO no varía.

Tabla 10: Parámetros en la variable

	B	E.T.	Wald	gl	Sig.	Exp(B)
Ingreso1	,004	,003	2,304	1	,129	1,004
Gasto	-,002	,004	,312	1	,577	,998
Educ.Superior(1)	1,370	,728	3,542	1	,060	3,936
Constante	-2,347	1,528	2,359	1	,125	,096

Fuente: producida por el programa SPSS.

En la columna $\exp(\beta)$ de la tabla, se observa que:

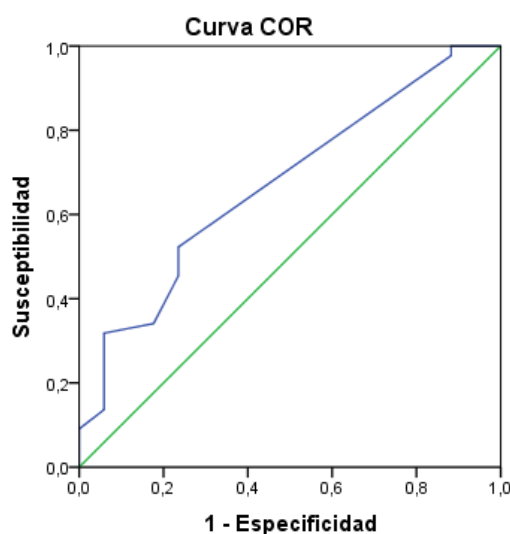
β_1 y $\beta_3 > 0$, por lo tanto, por cada 1 unidad que aumente el Ingreso, la posibilidad de un otorgamiento de crédito se incrementa en 1,004; del mismo modo ocurre con el nivel de educación, considerando educación superior con 3,936.

El gasto decrece en 0.577 la posibilidad de crédito.

4.4. Con respecto al ingreso mínimo y gasto máximo

4.4.1. Ingreso mínimo curva ROC

Dado a que lo que interesa es clasificar adecuadamente a los que se otorga crédito, con referencia cual será el ingreso mínimo, por lo cual para determinar el punto de corte se utilizó la curva de ROC basada en los cálculos de la especificidad y sensibilidad del modelo, Para identificar el mejor punto de clasificaciones utilizó el SPSS graficando la "Curva de ROC" utilizando como variable la probabilidad predicha y la observada.

Gráfico 1: Curva ROC de Ingreso

Los segmentos diagonales son producidos por los empates.

El eje de las abscisas, que corresponde a la proporción de falsos positivos, aparece etiquetado como 1-Especificidad y el eje de ordenadas, que corresponde a la proporción de aciertos positivos, aparece etiquetado como Sensibilidad. La línea curva escalonada representa los valores de 1-Especificidad y Sensibilidad para cada punto de corte ensayado (es decir, para cada una de las puntuaciones discriminantes incluidas en el análisis). Cada punto de esta curva se corresponde con un valor observado de la función.

La Tabla siguiente entrega una estimación del tamaño del área existente bajo la curva ROC. También ofrece el error típico de esa estimación (error estándar) y el nivel crítico (Sig. asintótica) resultante de contrastar la hipótesis nula de que el área existente bajo la curva vale 0,5; incluye los límites del intervalo de confianza (calculados al 95 %) correspondientes a la estimación efectuada. Si el intervalo de confianza incluye el valor 0,5 o lo que es lo mismo, si el nivel crítico es mayor que 0,05, no se podrá rechazar la hipótesis nula y,

por tanto, no se podrá afirmar que los puntos de corte muestreados generen una curva ROC bajo la cual exista un área significativamente mayor que 0,5. Por el contrario, si el intervalo de confianza no incluye el valor 0,5, o lo que es lo mismo, si el nivel crítico es menor que 0,05, se podrá rechazar la hipótesis nula y concluir que el área existente bajo la curva ROC es significativamente mayor que 0,5.

Tabla 11: Área bajo la curva

Variables resultado de contraste: ingreso

Área	Error típ. ^a	Sig. asintótica ^b	Intervalo de confianza asintótico al 95%	
			Límite inferior	Límite superior
,672	,075	,039	,524	,819

La variable (o variables) de resultado de contraste: ingreso tiene al menos un empate entre el grupo de estado real positivo y el grupo de estado real negativo.

Los estadísticos pueden estar sesgados.

a. Bajo el supuesto no paramétrico

b. Hipótesis nula: área verdadera = 0,5

Fuente: producida por el programa.

El gráfico de la curva ROC permite apreciar una curvatura muy pronunciada; y el área estimada (0,672) toma un valor próximo a 1. Además, dados los valores del nivel crítico y del intervalo de confianza, se puede concluir que esta área es significativamente mayor que 0.5, por lo cual indica una capacidad o eficacia predictiva del modelo alta.

La Tabla muestra una selección de los puntos representados (máximos y mínimos y las que se encuentran en torno al valor que corresponde al mejor punto de corte) en el gráfico, corresponde a los distintos valores que toma la función ordenados de menor a mayor, y las coordenadas correspondientes a esos valores en cada uno de los ejes del gráfico. El punto de corte va a estar dado por el cruce en el gráfico entre la especificidad y la sensibilidad.

**Tabla 12: Coordenadas de la Curva
ROC**

VARIABLES resultado de contraste: ingreso

Positivo si es mayor o igual que ^a	Sensibilidad	1 – Especificidad
499,0000	1,000	1,000
550,0000	1,000	,941
650,0000	1,000	,882
725,0000	,977	,882
825,0000	,523	,235
950,0000	,455	,235
1050,0000	,341	,176
1125,0000	,318	,059
1175,0000	,295	,059
1250,0000	,227	,059
1350,0000	,182	,059
1450,0000	,136	,059
1550,0000	,091	,000
1700,0000	,068	,000
1900,0000	,023	,000
2001,0000	,000	,000

La variable (o variables) de resultado de contraste: ingreso tiene al menos un empate entre el grupo de estado real positivo y el grupo de estado real negativo.

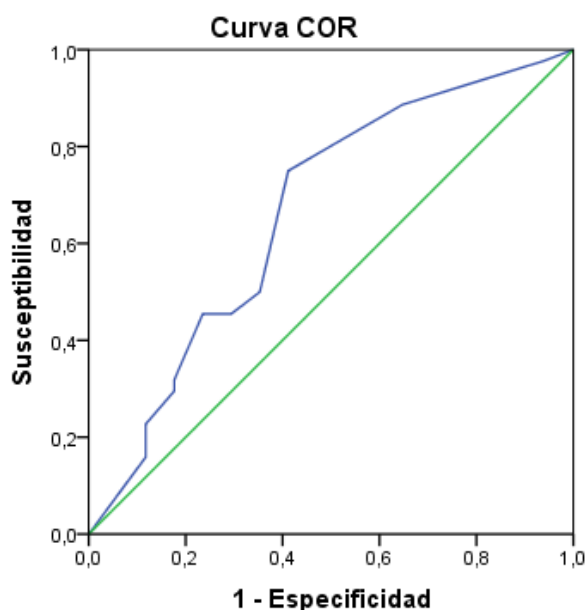
a. El menor valor de corte es el valor de contraste observado mínimo menos 1, mientras que el mayor valor de corte es el valor de contraste observado máximo más 1. Todos los demás valores de corte son la media de dos valores de contraste observados ordenados y consecutivos.

Fuente: producida por el programa

El mejor punto de corte se encuentra e 825.00.

4.4.2. Gasto máximo curva ROC

Gráfico 2: Curva ROC de Gasto



Los segmentos diagonales son producidos por los empates.

El gráfico de la curva ROC permite apreciar una curvatura muy pronunciada; y el área estimada 0,664 toma un valor próximo a 1. Además, dados los valores del nivel crítico y del intervalo de confianza, se puede concluir que esta área es significativamente mayor que 0,5.

Tabla 13: Área bajo la curva

VARIABLES resultado de contraste: Gasto Mensual

Área	Error típ. ^a	Sig. asintótica ^b	Intervalo de confianza asintótico al 95%	
			Límite inferior	Límite superior
,664	,083	,049	,501	,826

La variable (o variables) de resultado de contraste: Gasto Mensual tiene al menos un empate entre el grupo de estado real positivo y el grupo de estado real negativo. Los estadísticos pueden estar sesgados.

a. Bajo el supuesto no paramétrico

b. Hipótesis nula: área verdadera = 0,5

Fuente: Producida por el programa

La tabla muestra los puntos de corte que se hizo y que se encuentra en el gráfico. El punto de corte va a estar dado por el cruce en el gráfico entre la especificidad y la sensibilidad.

Tabla 14: Coordenadas de la curva

Positivo si es mayor o igual que ^a	Sensibilidad	1 – Especificidad
299,00	1,000	1,000
350,00	,977	,941
425,00	,886	,647
475,00	,750	,412
525,00	,500	,353
555,00	,455	,294
580,00	,455	,235
625,00	,318	,176
675,00	,295	,176
725,00	,227	,118
775,00	,159	,118
801,00	,000	,000

La variable (o variables) de resultado de contraste: Gasto Mensual tiene al menos un empate entre el grupo de estado real positivo y el grupo de estado real negativo.

a. El menor valor de corte es el valor de contraste observado mínimo menos 1, mientras que el mayor valor de corte es el valor de contraste observado máximo más 1. Todos los demás valores de corte son la media de dos valores de contraste observados ordenados y consecutivos.

Fuente: Proporcionada por el programa.

El mejor punto de corte es el de 580.00.

CONCLUSIONES

1. Los factores influyentes para la determinación de préstamos financieros en la financiera Credinka son: el Ingreso económico, Gasto, número de Hijos, Edad del cliente, Seguro de salud, tipo de Vivienda y Material de construcción de la vivienda. Los factores con mayor influencia son la edad con 0.037 y el ingreso con 0.062 de significancia en vista de que ambos se encuentran en el rango entre 0 y 1 más cercanos al cero.

2. El modelo final es:

$$P(\text{crédito}) = \frac{1}{1 + e^{-(-5.960 + 0.004\text{Ingreso} - 0.003\text{Gasto} - 0.205\text{Hijos} + 0.124\text{Edad} - 0.299\text{Seguro} + 0.566\text{Vivienda} + 0.640\text{Material})}}$$

Con una fiabilidad del 78.7%.

3. Mediante la curva ROC se obtuvo que el ingreso mínimo recomendado es de S/.825.00 y el gasto máximo recomendado es de S/.580.00 nuevos soles mensuales. El ahorro para cubrir prestamos asciende a S/. 245.00 nuevos soles mensuales.

RECOMENDACIONES Y SUGERENCIAS

De acuerdo a los resultados sugerimos incluir más índices de clasificación al indicador Accesorios en la vivienda, para conocer mejor sus posesiones y de esta manera medir mejor su posibilidad de crédito.

Se sugiere realizar un análisis sectorial en vista de que no todos viven en una misma are, separándolos por los que viven en zona urbana y rural, de esta manera se haría un mejor estudio de los factores y así poder hacer una comparación.

Se sugiere a la financiera Credinka mostrar a los clientes más oportunidades, tomando en cuenta otros aspectos para la evaluación del otorgamiento de crédito, creando una buena imagen a partir de un trabajo responsable. Asimismo, se debe incentivar a la población con limitaciones económicas, se hace necesario diseñar programas de capacitación de negocios dirigido a aquellos que no cumplen los requisitos.

BIBLIOGRAFÍA**TEXTOS**

AGAPITO, G. (2010). *Determinantes del Crédito de consumo por tipo de institución financiera en el Perú.*

DALLAS, J. (2000). *Métodos Multivariados Aplicados al Análisis de Datos* (Vol. 1a). México: Internacional Thomson Editores.

ESAM, (2014). *Factores críticos de éxito para la creación de un mercado alternativo de emisión de valores para las pymes en el Perú.* Editorial, Cecosami Prerensa e Impresión Digital S. A.

KADMIR, L. (1991). *Estadística Aplicada a Administración y Economía.* México Edit. Mc Graw Hill.

MAHIA, R. (2013). *Técnicas de modelización de variables cualitativas: regresión Logística.* ramon.mahia@uam.es

MURRAY, R (1997):” *Estadística*”. Segunda edición
Editorial, Industrias Graficas S.A. Chile: pag.186

MURRAY R. y LARRY J. (2009). “*Estadística*”. 4ta edición.
Editorial, Mc Graw-Hill. México, D.F: pag. 268

WEBGRAFÍA

<http://repositorio.ug.edu.ec/bitstream/redug/1667/1/ESTUDIO%20Y%20AN%C3%81LISIS%20DE%20LA%20POBREZA%20DE%20>

ANDRÉU ,J. (2011):”El análisis multinivel: una revisión actualizada en el ámbito sociológico”.Accesado.07/09/14

<http://universidadlamolina.edu.org>.

BULLON C. Y VERGARA, A. (2009):” Análisis multinivel: Regresión logística binaria con dos niveles”. Accesado: 07/09/14

http://www.sbs.gob.ec/medios/PORTALDOCS/downloads/articulos_financieros/Estudios%20Tecnico/Notas%20en%20Finanzas%20No.%203.pdf

CARVAJAL, A, (2005). Los determinantes del crédito: el caso ecuatoriano. Accesado.12/09/2016

<http://tangara.uis.edu.co/biblioweb/tesis/2015/159163.pdf>

FORERO, L. (2015), "Propuesta de modelo para la evaluación y predicción de riesgo del insolvente de pequeñas y medianas empresas manufactureras en Colombia". Accesado.13/02/2017

<http://mat.izt.uam.mx/mcmai/documentos/tesis/Gen.07-O/Nieto-S-Tesis.pdf>

NIETO, S. (2010); "Crédito al Consumo: La Estadística Aplicada a un problema de Riesgo Crediticio". Accesado.13/02/2017

http://cybertesis.unmsm.edu.pe/xmlui/bitstream/handle/cybertesis/4233/Quispe_ms.pdf?sequence=1

QUISPE, S. (2014); "Rotación de personal: Predicción con modelo de regresión logística multinivel". Accesado.13/02/2017

<http://www.bdigital.unal.edu.co/2005/1/70192935.2010.pdf>

ZAPATA, D. (2009), "Caracterización de las Variables Determinantes del Riesgo en el Microcredito Rural". Accesado.13/02/16

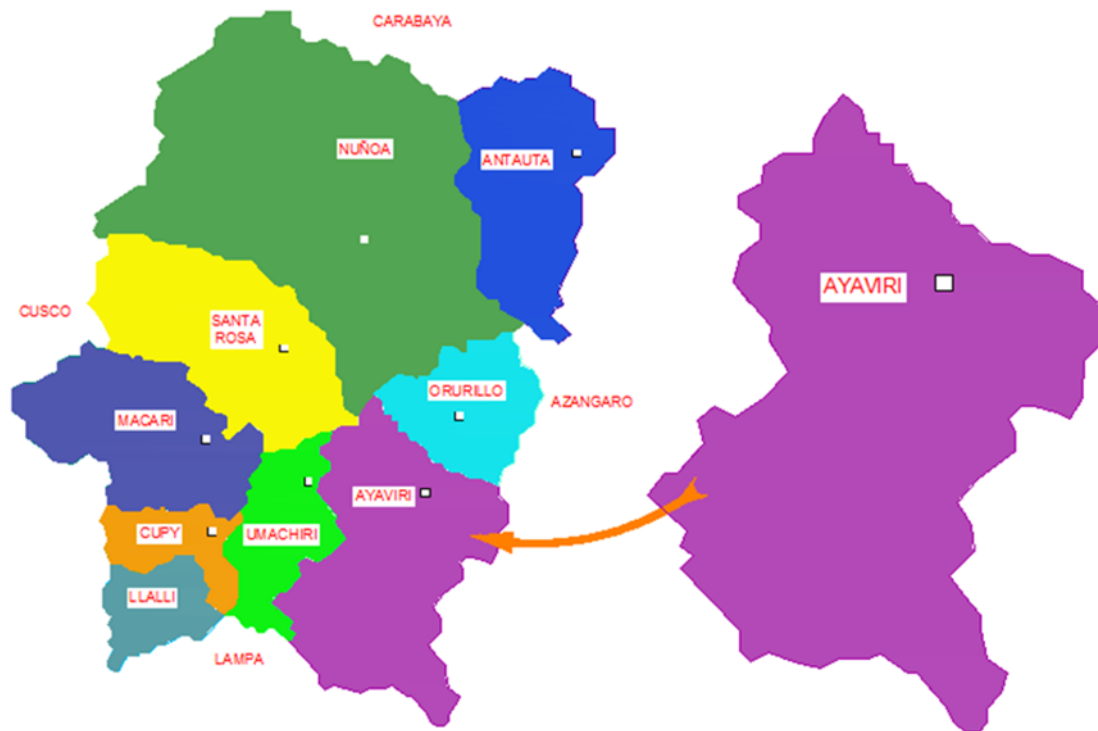
TESIS.

CAXI, F. (2016); "*Factores Individuales y Escolares que Influyen en el Ingreso a la Universidad Nacional del Altiplano Puno en los Estudiantes del Cepreuna 2016*". Tesis para optar el título profesional de Licenciado en Educación, con mención en la especialidad de matemática e informática. Pag. (38).

ACURIO, L.(2013):"*Modelo de regresión logística en datos censurados y no censurados para el síndrome de Burnout de los estudiantes de la Universidad Nacional del Altiplano Puno,2013,II*".Tesis para optar el título profesional de Ingeniero Estadístico e Informático. Puno. Pag. (42).

ANEXOS

Mapa político de Ayaviri



CUESTIONARIO

Con el siguiente cuestionario recolectaremos algunos datos por motivos de investigación profesional.

Rogamos a Ud. Contestar las interrogantes. Le tomara unos cuantos minutos.
Gracias

Señale marcando con una X la alternativa que más se ajuste a Ud.

1. ¿Cuál es su edad?

2. ¿Cuál es su estado Civil?

- Soltera
- Casada
- Viuda
- Conviviente

3. ¿Cuál es su Nivel de Escolaridad?

- Primaria
- Secundaria
- Sup.No universitaria
- Sup.Universitaria
- Sin estudios

4. ¿Quién realiza la labor de jefatura de Hogar?

- Padre
- Madre
- Otro _____

5. ¿Cuánto es el número de Hijos que tiene?

6. ¿Cuál es la situación de la vivienda que ocupa?

- Propia
- Alquilada
- Otros

7. ¿Cuál es el material de la vivienda que ocupa?

- Ladrillo/cemento
- Adobe
- Madera
- Otros

8. ¿Su vivienda cuenta con servicios básicos?

- Agua
- Luz
- ninguno

9. ¿Qué ocupación laboral tiene actualmente?

- Activo
- Cesante
- Pasivo

10. ¿Cuál es el ingreso familiar que recibe?

- Más que el Ingreso Mínimo
- Menos que el Ingreso Mínimo
- Ingreso Mínimo

11. ¿Qué tipo de trabajo desempeña actualmente?

- Empleador
- Asalariados
- Trabajador independiente
- Trabajo familiar no remunerado

12. ¿Cuántos aportan en su hogar?

- De 1 a 2
- De 3 a más.

13. ¿Cuánto gasta al mes?**14. ¿Cuánto es su ingreso?**

15. ¿Con que artefactos cuenta su vivienda?

- Televisor
- Computadora
- Vehículo

16. ¿Con que tipo se seguro de salud cuenta?

- Únicamente SIS.
- Únicamente EsSalud
- Con SIS y EsSalud
- Con otro seguro de salud

Cuadro de codificación de las variables para el paquete

Credito.Experiencia	Perceptores	Ingreso1	Gasto	Trabajo	Educ.Mediana	Educ.Superior	Hijos	Edad	Estado Civil	Seguro	Vivienda	Material	Agua desague	Alumbrado	Televisor	Computadora	Vehículo
No crédito	2	1500,00	800	Trabajador independiente	Si	No	4	40	Casada	Con seguro	Propia	Ladrillo/ce mento	Si	Si	Si	Si	No
No crédito	2	750,00	450	Trabajador independiente	Si	No	3	42	Casada	Sin seguro	Propia	Rústico	Si	Si	Si	No	No
No crédito	2	1100,00	800	Trabajador independiente	Si	No	3	29	Casada	Con seguro	Alquilado	Ladrillo/ce mento	Si	Si	Si	Si	No
No crédito	1	750,00	500	Trabajador independiente	Si	Si		30	Soltera	Sin seguro	Alquilado	Rústico	Si	Si	Si	Si	No
No crédito	1	600,00	300	Trabajador independiente	Si	No	1	28	Soltera	Con seguro	Alquilado	Ladrillo/ce mento	Si	Si	Si	No	No
No crédito	1	750,00	400	Trabajador independiente	Si	No	1	31	Soltera	Con seguro	Alquilado	Rústico	Si	Si	No	No	No
No crédito	2	500,00	450	Trabajador independiente	Si	No	3	52	Casada	Con seguro	Propia	Rústico	Si	Si	Si	No	No
No crédito	2	750,00	400	Trabajador independiente	Si	No	3	35	Casada	Con seguro	Propia	Rústico	Si	Si	Si	No	No
No crédito	1	750,00	450	Trabajador independiente	Si	Si		25	Soltera	Con seguro	Alquilado	Rústico	Si	Si	Si	No	No
No crédito	2	1100,00	700	Trabajador independiente	Si	No	3	38	Casada	Con seguro	Propia	Rústico	Si	Si	Si	Si	Si
No crédito	2	750,00	400	Trabajador independiente	Si	No	1	19	Casada	Con seguro	Alquilado	Ladrillo/ce mento	Si	Si	Si	No	No

No crédito	1	750,00	400	Trabajador independiente	Si	1	21	Soltera	Con seguro	Alquilado	Rústico	Si	Si	No	No
No crédito	1	750,00	400	Trabajador independiente	Si		25	Soltera	Sin seguro	Propia	Ladrillo/ce	Si	Si	No	No
No crédito	2	750,00	450	Trabajador independiente	Si	2	40	Casada	Con seguro	Propia	Ladrillo/ce	Si	Si	No	No
No crédito	1	750,00	550	Trabajador independiente	Si	2	23	Soltera	Sin seguro	Propia	Ladrillo/ce	Si	Si	Si	No
No crédito	1	1000,00	600	Asalariado	Si		19	Soltera	Con seguro	Propia	Rústico	Si	Si	No	No
No crédito	1	750,00	560	Trabajador independiente	Si	1	43	Soltera	Con seguro	Propia	Ladrillo/ce	Si	Si	No	No
Crédito	2	1200,00	700	Trabajador independiente	Si	3	36	Casada	Con seguro	Propia	Rústico	Si	Si	Si	No
Crédito	2	750,00	400	Trabajador independiente	Si	3	45	Casada	Sin seguro	Alquilado	Rústico	Si	Si	Si	No
Crédito	2	750,00	500	Trabajador independiente	Si	3	34	Casada	Con seguro	Propia	Rústico	Si	Si	No	No
Crédito	2	1000,00	600	Trabajador independiente	Si	3	38	Casada	Con seguro	Propia	Rústico	Si	Si	No	No
Crédito	2	1800,00	600	Asalariado	Si	2	35	Casada	Con seguro	Propia	Rústico	No	Si	Si	Si
Crédito	1	750,00	400	Asalariado	Si		20	Soltera	Con seguro	Propia	Ladrillo/ce	Si	Si	No	No
Crédito	2	750,00	500	Trabajador independiente	Si	3	50	Casada	Con seguro	Propia	Rústico	Si	Si	No	No
Crédito	1	1300,00	750	Trabajador independiente	Si	8	65	Soltera	Sin seguro	Propia	Rústico	Si	Si	No	No
Crédito	1	700,00	400	Trabajador independiente	Si	3	48	Soltera	Sin seguro	Propia	Rústico	Si	Si	No	No

Crédito	2	750,00	400	Trabajador independiente	Si	No	3	34	Casada	Con seguro	Alquilado	Ladrillo/ce mento	Si	Si	No	No
Crédito	2	2000,00	800	Trabajador independiente	Si	No	3	45	Casada	Con seguro	Propia	Rústico	Si	Si	Si	No
Crédito	2	1500,00	750	Trabajador independiente	Si	No	2	40	Casada	Con seguro	Propia	Rústico	Si	Si	Si	Si
Crédito	2	1300,00	800	Trabajador independiente	Si	Si	3	50	Casada	Con seguro	Propia	Ladrillo/ce mento	Si	Si	Si	No
Crédito	2	1400,00	700	Trabajador independiente	Si	Si	3	33	Casada	Con seguro	Alquilado	Rústico	Si	Si	Si	No
Crédito	2	900,00	500	Trabajador independiente	Si	No	2	30	Casada	Con seguro	Alquilado	Rústico	Si	Si	No	No
Crédito	2	750,00	500	Trabajador independiente	Si	No	1	38	Casada	Con seguro	Alquilado	Ladrillo/ce mento	Si	Si	No	No
Crédito	1	750,00	550	Trabajador independiente	Si	No	1	27	Soltera	Con seguro	Alquilado	Ladrillo/ce mento	Si	Si	No	No
Crédito	1	750,00	500	Trabajador independiente	Si	No	1	35	Soltera	Con seguro	Alquilado	Rústico	Si	Si	Si	No
Crédito	2	1400,00	800	Trabajador independiente	Si	No	3	42	Casada	Con seguro	Propia	Ladrillo/ce mento	Si	Si	Si	Si
Crédito	2	1000,00	450	Trabajador independiente	Si	No	7	58	Casada	Sin seguro	Propia	Ladrillo/ce mento	Si	Si	No	No
Crédito	2	1200,00	800	Trabajador independiente	Si	No	1	41	Casada	Con seguro	Propia	Rústico	Si	Si	Si	Si
Crédito	2	900,00	500	Asalariado	Si	No	4	47	Casada	Con seguro	Propia	Rústico	Si	Si	No	No
Crédito	1	750,00	450	Trabajador independiente	Si	No	2	33	Soltera	Con seguro	Alquilado	Rústico	Si	Si	No	No
Crédito	2	1600,00	800	Asalariado	Si	No	2	40	Casada	Con seguro	Propia	Ladrillo/ce mento	Si	Si	Si	Si

Crédito	2	1000,00	600	Trabajador independiente	No	No	3	50	Casada	Con seguro	Propia	Ladrillo/ce miento	Si	Si	Si	Si
Crédito	1	750,00	500	Trabajador independiente	Si	No	2	40	Casada	Con seguro	Propia	Rústico	Si	Si	No	No
Crédito	2	1150,00	800	Trabajador independiente	Si	No	2	30	Casada	Con seguro	Propia	Ladrillo/ce miento	Si	Si	Si	Si
Crédito	1	1000,00	500	Trabajador independiente	Si	Si	2	33	Soltera	Con seguro	Alquilado	Rústico	Si	Si	No	No
Crédito	1	750,00	300	Trabajador independiente	Si	No	1	26	Soltera	Con seguro	Alquilado	Rústico	Si	Si	No	No
Crédito	2	1200,00	700	Trabajador independiente	Si	No	4	41	Casada	Con seguro	Propia	Ladrillo/ce miento	Si	Si	Si	No
Crédito	2	900,00	500	Trabajador independiente	Si	No	3	46	Casada	Con seguro	Propia	Rústico	Si	Si	No	No
Crédito	2	1800,00	800	Trabajador independiente	Si	No	5	60	Casada	Con seguro	Propia	Ladrillo/ce miento	Si	Si	No	Si
Crédito	2	750,00	500	Trabajador independiente	Si	No	2	34	Casada	Con seguro	Propia	Rústico	Si	Si	No	No
Crédito	2	750,00	450	Trabajador independiente	Si	No	1	33	Casada	Sin seguro	Propia	Rústico	Si	Si	No	No
Crédito	1	1500,00	750	Asariado	Si	No	3	39	Casada	Sin seguro	Propia	Rústico	Si	Si	Si	Si
Crédito	2	750,00	450	Trabajador independiente	Si	No	3	46	Casada	Con seguro	Propia	Rústico	Si	Si	No	Si
Crédito	2	1100,00	600	Trabajador independiente	Si	Si	1	31	Casada	Con seguro	Propia	Ladrillo/ce miento	Si	Si	Si	No
Crédito	1	1000,00	650	Trabajador independiente	Si	No	3	42	Soltera	Sin seguro	Propia	Ladrillo/ce miento	Si	Si	No	No
Crédito	1	750,00	550	Trabajador independiente	Si	No	4	47	Casada	Con seguro	Propia	Rústico	Si	Si	No	No

Crédito	2	750,00	600	Asalariado	Si	No	2	40	Casada	Con seguro	Propia	Rústico	Si	Si	Si	No	No
Crédito	1	750,00	500	Trabajador independiente	Si	No	3	50	Casada	Sin seguro	Propia	Rústico	Si	Si	Si	No	No
Crédito	1	750,00	450	Trabajador independiente	Si	No	1	30	Soltera	Sin seguro	Alquilado	Rústico	Si	Si	Si	No	No
Crédito	2	750,00	450	Trabajador independiente	Si	No	2	43	Casada	Con seguro	Propia	Rústico	Si	Si	Si	No	No
Crédito	1	750,00	600	Trabajador independiente	Si	No		31	Soltera	Con seguro	Alquilado	Rústico	Si	Si	Si	No	No

Fuente:Elaboración propia