



UNIVERSIDAD  
DE PIURA

**FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES**

**Evaluación crediticia aplicando un modelo de credit  
scoring para el otorgamiento de tarjetas de crédito**

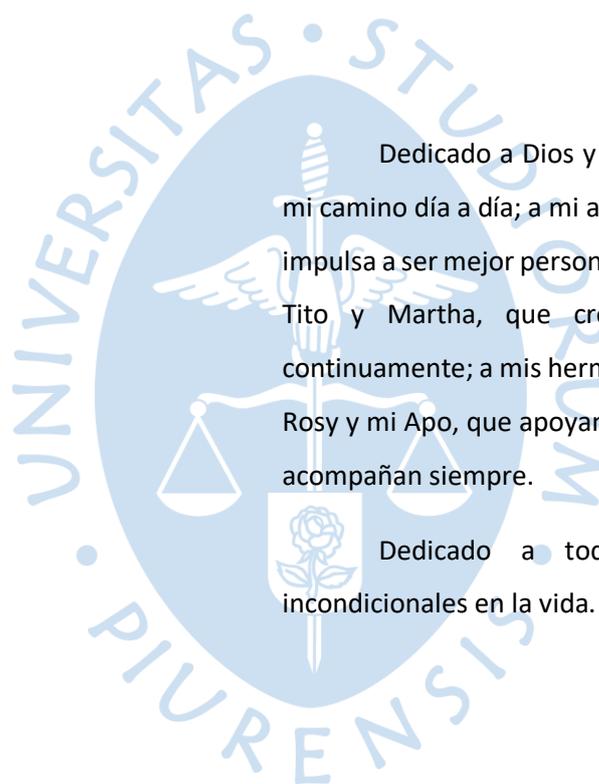
Trabajo de Suficiencia Profesional para optar el Título  
de Economista

**Victor Armando Ruiz Ayón**

**Revisor(es):**  
Mgtr. Harry Omar Patrón Torres

**Piura, febrero de 2021**





Dedicado a Dios y la Virgen María, que guían mi camino día a día; a mi amada esposa Lucía, que me impulsa a ser mejor persona y profesional; a mis papás Tito y Martha, que creen y apuestan por mí continuamente; a mis hermanos Wen y Jorge, mi abue Rosy y mi Apo, que apoyan todas mis decisiones y me acompañan siempre.

Dedicado a todos ellos por ser mis incondicionales en la vida.



## Resumen

El incremento en la utilización de las tarjetas de crédito como medio de pago, trae consigo que los bancos y entidades financieras busquen ofrecer con mayor intensidad este producto; sin embargo, es necesario que se realice una evaluación adecuada de las solicitudes recibidas, para que el riesgo de incumplimiento por parte del tarjetahabiente sea menor.

La metodología aplicada en este trabajo de investigación es un modelo paramétrico logit. Se estableció una variable dependiente dicotómica y, se evaluaron y seleccionaron las variables independientes según varios criterios, como el nivel de significancia; para obtener un modelo final de regresión.

Como resultado, se logró estimar el modelo con variables exógenas significativas al 99%. Además, para una mejor interpretación de los parámetros, se hallaron los efectos marginales de cada variable independiente sobre la variable dependiente. Y finalmente, se evaluó el desempeño del modelo a través de la matriz de confusión.





## Tabla de contenido

<b>Introducción</b> .....	13
<b>Capítulo 1. Informe de experiencia profesional</b> .....	15
1.1 Resumen de la experiencia profesional.....	15
1.2 Aspectos generales.....	16
1.2.1 Descripción de la empresa .....	16
1.2.2 Descripción general de la experiencia.....	18
<b>Capítulo 2. Evaluación crediticia aplicando un modelo de credit scoring para el otorgamiento de tarjetas de crédito</b> .....	21
2.1 Marco teórico.....	21
2.1.1 Aspectos generales de las tarjetas de crédito .....	21
2.2 Credit scoring .....	23
2.2.1 Definición de credit scoring .....	23
2.2.2 Antecedentes del credit scoring .....	24
2.2.3 Beneficios y críticas al credit scoring.....	24
2.2.4 Modelos más comunes del credit scoring .....	25
2.3 Aplicación del credit scoring .....	30
2.3.1 Descripción de la data .....	30
2.3.2 Variables seleccionadas.....	30
2.3.3 Modelo de estimación.....	32
2.3.4 Análisis de resultados.....	35
<b>Conclusiones</b> .....	39
<b>Lista de referencias</b> .....	41
<b>Apéndices</b> .....	43
Apéndice 1. Organigrama IEP Virgen del Carmen .....	45
<b>Anexos</b> .....	47
Anexo 1. Certificado de trabajo emitido por la IEP Virgen del Carmen. ....	49
Anexo 2. Descripción de las variables y signos esperados. ....	50
Anexo 3. Estimación de regresiones logísticas con cada variable independiente.....	51
Anexo 4. Resultado de estimación de modelo logit incluyendo 21 variables. ....	63
Anexo 5. Resultado de correlación entre variables. ....	64
Anexo 6. Resultado de estimación de modelo logit incluyendo las variables definidas.....	65



## Lista de tablas

Tabla 1. Comparación de modelos paramétricos de credit scoring .....	26
Tabla 2. Tabulación de variable dependiente .....	30
Tabla 3. Resumen de estimación de regresiones con una variable.....	32
Tabla 4. Relación de variables correlacionadas.....	33
Tabla 5. Estimación de cambios marginales de las variables seleccionadas.....	34
Tabla 6. Reporte de clasificación .....	35





## Lista de figuras

Figura 1. Gráfica del total de tarjetas de crédito de bancos y financieras en Perú .....	22
Figura 2. Gráfica de comparación entre la función de distribución acumulada normal y logística .....	29
Figura 3. Matriz de confusión.....	34





## Introducción

Es un hecho que el efectivo sigue siendo el medio de pago favorito para gran parte de los consumidores de bienes y servicios en el país; sin embargo, también es necesario dejar en claro que, el uso de otros medios de pago, como las tarjetas de crédito, han ido adquiriendo gran notoriedad durante los últimos años.

El aumento en la cantidad de consumos con tarjetas de crédito, se puede deber a muchos motivos diferentes, uno de los principales es el incremento en el nivel de otorgamiento de tarjetas al público peruano. Esto quiere decir que hay más tarjetas circulando en el mercado y, de la misma manera, los niveles de crédito que obtiene cada cliente, son mayores.

Sin embargo, existe una cuestión subyacente, muy importante, que es necesario tener en cuenta: las tarjetas de crédito, como cualquier otro instrumento de crédito, traen consigo una responsabilidad económica que, el banco o el organismo financiero otorgador, debe asegurar. Es por ello que todas las solicitudes presentadas deben pasar por un proceso de evaluación; para intentar reducir la probabilidad de que un cliente no cumpla con el pago.

Uno de los procedimientos de evaluación, que mayor relevancia ha tomado durante los últimos años, y que ha sido objeto de estudio de este trabajo, es la metodología bajo los modelos de credit scoring.

El presente trabajo tiene una estructura dividida en dos capítulos: El primer capítulo es una revisión de la experiencia profesional del autor, desde el término de sus estudios universitarios. Se presenta una breve descripción de todas sus posiciones laborales, centrándose en la que, el autor considera fue la más importante y retadora de su vida profesional; se indican los aspectos generales de la empresa, las funciones determinadas que tenía a cargo, y, por último, los aportes brindados a la empresa que, demuestran el alto nivel de análisis, organización y compromiso del autor. En el segundo capítulo se realiza una revisión bibliográfica centrada en el entendimiento de los criterios de evaluación para el otorgamiento de tarjetas de crédito, además de elaborar un modelo de evaluación bajo la metodología logarítmica o también llamada logit, y por último se describen los resultados y las conclusiones finales del trabajo.



## Capítulo 1

### Informe de experiencia profesional

En el presente capítulo, el autor, realiza una descripción de su experiencia profesional a partir de la culminación de sus estudios superiores, hasta la actualidad; haciendo una breve descripción de las funciones desempeñadas en el puesto de trabajo que desarrollaba en cada empresa.

#### 1.1 Resumen de la experiencia profesional.

La experiencia profesional del autor empieza en el año 2016:

- Empresa: Distribuidora Comercial Álvarez Bohl.
  - Cargo: Auxiliar administrativo de almacén.
  - Período: junio 2016 – octubre 2016.
  - Principales funciones:
    - Cuadre diario y recepción de facturas a transportistas y distribuidores.
    - Elaboración de notas de crédito sobre productos no recibidos por los clientes.
    - Elaboración de reportes y semaforizados para verificar el nivel de venta y rotación de cada producto.
- Empresa: Institución Educativa Particular Virgen del Carmen.
  - Cargo: Asistente administrativo.
  - Período: noviembre 2016 – enero 2018.
  - Principales funciones:
    - Elaboración de planillas mensuales del personal administrativo, docente y de servicio.
    - Elaboración del presupuesto para el año 2017.
    - Evaluación de solicitudes de becas por parte de los padres de familia.
    - Organización y realización de talleres relacionados al curso de Economía.
- Empresa: Enseña Perú.
  - Cargo: Profesional de Enseña Perú.
  - Período: marzo 2018 – setiembre 2018.
  - Principales funciones:
    - Docente del área de matemáticas de estudiantes de 2do de secundaria.
    - Tutor de una sección de 2do de secundaria de una comunidad educativa en Ancash.

- Empresa: Chifa San Zi.
  - Cargo: Administrador general.
  - Período: enero 2019 – noviembre 2019.
  - Principales funciones:
    - Elaboración y análisis de reportes semanales y mensuales de ventas y gastos.
    - Elaboración de informes presentados a la gerencia de la empresa.
    - Realización de requerimientos semanales a empresas proveedoras.
  
- Empresa: Centro Internacional de la Papa.
  - Cargo: Coordinador de estaciones experimentales de Huancayo y San Ramón.
  - Período: diciembre 2019 – actualidad.
  - Principales funciones:
    - Encargado del área de finanzas de las estaciones: realización de registros contables, manejo de la cuenta corriente, manejo y cuadre de caja chica, gestión de pago a proveedores.
    - Encargado del área de administración y recursos humanos: elaboración de requisiciones para el área correspondiente, coordinación de personal eventual para estaciones y para proyectos *off-station*.
    - Encargado del área de almacén y logística: proceso de requisiciones, gestión de proveedores, realización de órdenes de compra, entrega de artículos a usuarios.

A continuación, el autor, hará una descripción detallada sobre su experiencia laboral en la Institución Educativa Particular Virgen del Carmen, ya que después de realizar una evaluación concienzuda, considera que es la experiencia laboral en la que mayores retos afrontó y en donde pudo aplicar útilmente los conocimientos académicos adquiridos en los años de universidad.

## **1.2 Aspectos generales**

### **1.2.1 Descripción de la empresa**

La Institución Educativa Particular Virgen del Carmen, en adelante IEP Virgen del Carmen, es una empresa dedicada al sector educación. Surgió como iniciativa de una pareja de esposos en el año 1991, en respuesta a la necesidad de muchos padres de familia que buscaban un colegio en donde sus hijos empiecen sus estudios del nivel inicial. Con el paso de los años, la IEP Virgen del Carmen, amplió sus servicios incluyendo los niveles de primaria y secundaria.

Actualmente, el plantel educativo y las oficinas principales se encuentran ubicadas en la Urbanización Jardín Manzana B3 Lote 1 - III Etapa, Sullana – Piura.

La misión de la IEP Virgen del Carmen es:

Ser una Institución Educativa que cuenta con personal idóneo e identificado con el óptimo desarrollo docente, que brinda a la niñez y adolescencia de Sullana un servicio educativo de calidad de acuerdo al avance tecnológico y a la modernidad, cumpliendo sus roles con responsabilidad social en el marco de una convivencia pacífica y solidaria.

Y su visión es:

Ser líderes en la promoción de una educación de calidad y trascendencia basada en la difusión, defensa y práctica de valores democráticos y respeto de derechos humanos a finde optimizar la formación biopsicosocial y espiritual de la persona como producto del arduo trabajo en equipo de los agentes educativos que contribuya al desarrollo del país.

La IEP Virgen del Carmen, se organiza en cuatro departamentos: departamento educativo, departamento administrativo y contable, departamento legal, y departamento de mantenimiento. Para mayor detalle, se muestra el organigrama de la IEP Virgen del Carmen en el apéndice (Ver apéndice 1).

El departamento educativo, está conformado por la dirección general, la subdirección y las coordinaciones de inicial, primaria y secundaria. Tiene a cargo todo lo relacionado con brindar el servicio académico a los estudiantes: la elaboración de sesiones de aprendizaje, revisión y supervisión de las clases, organización de escuelas para padres, labores relacionadas a pastoral educativa, entre otras funciones.

El departamento administrativo y contable, el área en la que laboró el autor de este trabajo, está conformado por el administrador, el contador, dos asistentes, y el auditor externo. Este departamento tiene a cargo el registro de ingresos y egresos mensuales, la elaboración de boletas de venta por la realización del servicio, la realización y pagos de planillas, elaboración de PDT mensuales, entre otras cosas; además, junto con la dirección general, realiza la evaluación de las solicitudes de becas y pensiones escalonadas para la institución.

El departamento de legal, está conformado por el abogado y el asesor legal. Tiene a cargo la emisión y recepción de documentos legales, la elaboración de contratos de servicio con los padres de familia, la elaboración de contratos del personal.

El departamento de mantenimiento, está conformado por el encargado de mantenimiento, los porteros, el vigilante, y el equipo de limpieza. Tiene a cargo la limpieza, mantenimiento y cuidado del plantel educativo y los bienes muebles.

## 1.2.2 Descripción general de la experiencia

El autor de este trabajo se desempeñó en el cargo de asistente administrativo durante el período de noviembre del 2016 hasta enero del 2018. En el anexo 1, se presenta el certificado de trabajo emitido por la dirección de la IEP Virgen del Carmen.

**1.2.2.1 Actividad profesional desempeñada.** Las principales funciones y actividades desarrolladas, fueron las siguientes:

- Elaboración de planillas mensuales del personal administrativo, docente y de servicio: Se verificaba el monto a pagar a cada trabajador, se determinaban las retenciones que se debían realizar al trabajador y los aportes del empleador, según sea el caso.
- Elaboración de informes de ingresos y egresos mensuales: Se registraban los ingresos por grado y los comprobantes de compra pagados durante el mes. En este informe, se incluía una evaluación del nivel de morosidad y según los resultados, se enviaba el reporte al asistente de contabilidad para el registro en el sistema de Equifax.
- Elaboración del presupuesto para los años 2017 y 2018: Anualmente, durante el mes de diciembre, la gerencia solicitaba la elaboración de un informe detallado y un presupuesto proyectado al siguiente año con todos los gastos en los que se esperaba incurrir en cada departamento, para hacer la división de los montos según los centros de costos disponibles y si era necesario, generar nuevos centros de costos para cargar dichos gastos.
- Gestión y liquidación de caja chica: En el departamento administrativo y contable, se manejaba la caja chica con un monto suficiente para cubrir gastos pequeños y emergencias.
- Evaluación de solicitudes de pensiones escalonadas y becas: Se contaba con un sistema de pensiones escalonadas que, empezaba con el 10% de descuento cuando había hermanos matriculados, hasta la beca completa que se otorgaba en caso de orfandad. Se verificaba que la documentación necesaria esté completa y que la familia solicitante de la pensión y/o beca cumpla con los requisitos que justifiquen la asignación del porcentaje solicitado. Una vez que pasaba la primera revisión, se realizaba una junta con la dirección general, y en consenso, se evaluaban los casos y se determinaba la aprobación o desaprobación de la solicitud. Para finalizar el proceso, se enviaba una carta con la respuesta correspondiente a la familia solicitante.
- Apoyo en el área de Historia, Geografía y Economía: Debido a los estudios en economía del autor, se le asignó el apoyo en la organización y realización de talleres con temas básicos de economía dependiendo del grado que cursaban los estudiantes del nivel secundario.

**1.2.2.2 Propósito del puesto.** El propósito principal de este cargo era brindarle apoyo y soporte al administrador general de la institución, y realizar una correcta recolección de datos que permitan generar informes y reportes adecuados para que la gerencia tome mejores decisiones.

Además, debido a que la IEP Virgen del Carmen es una empresa en crecimiento, se creó este puesto con el fin de apoyar en aspectos importantes tales como: organización, procesos logísticos, y otros temas relacionados.

**1.2.2.3 Aporte académico al puesto laboral.** Las funciones y actividades realizadas por el autor en este puesto de trabajo, requerían conocimientos y habilidades adquiridos durante los años de la carrera de economía, sobre todo en las siguientes áreas: contabilidad, finanzas e informática.

El problema identificado que es objeto del presente informe es que no había una correcta asignación del presupuesto para los cuatro departamentos que forman parte de la organización de la institución. Esto quiere decir que, había montos asignados a un departamento que no lo necesitaba, y de la misma manera, existía un déficit en el presupuesto asignado a otro departamento que si necesitaba el dinero.

Una vez identificado el problema descrito, el autor, presentó ante la gerencia un proyecto que proponía la evaluación, revisión y nueva distribución del presupuesto para el año 2017, y un seguimiento luego de la aplicación del proyecto. El proceso fue el siguiente:

En primer lugar, se realizó una evaluación previa. Con el apoyo de informes de años anteriores, la evaluación de los gastos determinó que, el departamento que mayor presupuesto requería era el departamento educativo, ya que los gastos de planilla de toda la plana docente y directiva eran cargados a los centros de costos de este departamento. Y el segundo departamento con mayor necesidad de presupuesto, era el departamento de mantenimiento, ya que los gastos de renovación y mantenimiento del plantel educativo, de los bienes muebles y de los equipos de cómputo, robótica y del laboratorio, eran cargados a los centros de costos correspondientes a este departamento.

Cuando se terminó la fase de evaluación, se revisaron los montos de presupuesto asignados inicialmente, y junto con el equipo de finanzas, se realizó una nueva distribución según las necesidades de cada departamento. Este informe detallado, se presentó a gerencia y fue aprobado para su aplicación.

Para conocer la efectividad del proyecto puesto en marcha, el autor, solicitó llevar a cabo un seguimiento durante el año que consistía en tres pasos: planificación, revisión y registro. La planificación consistía en solicitar a cada departamento, trimestralmente, la proyección y planificación de gastos; la revisión consistía en verificar que, los requerimientos elaborados por

cada departamento habían sido planificados, y se corroboraba que los centros de costos consignados eran los correctos; y el registro consistía en subir toda la información de los requerimientos en un *template* o plantilla de Excel detallada que se manejaba en la oficina de administración.

Lo que se logró con la aplicación de este proyecto fue lo siguiente:

- Los departamentos pudieron priorizar y organizar sus requerimientos para la adquisición de productos o servicios, según las necesidades que debían cubrir. Esta planificación, permitió que disminuyan los gastos por departamento.
- Hubo una mejor distribución de funciones para cada miembro del personal de la Institución, y redujeron las contrataciones de personal por temporadas.
- Se gestionó de mejor manera los centros de costos de cada departamento; es decir, permanecieron vigentes los que eran necesarios y se eliminaron aquellos que no.
- Se redujeron los gastos de caja chica, ya que las compras por emergencias disminuyeron.
- Con la redistribución de los montos asignados, se logró que el departamento de mantenimiento realice la compra de computadoras para mejorar los centros de cómputo de la institución; y se hizo la proyección para la compra e instalación de televisores para las aulas durante el siguiente año.
- El presupuesto se cumplió en todos los departamentos.

## Capítulo 2

### Evaluación crediticia aplicando un modelo de *credit scoring* para el otorgamiento de tarjetas de crédito

#### 2.1 Marco teórico

##### 2.1.1 Aspectos generales de las tarjetas de crédito

Las tarjetas de crédito son muy importantes en la evolución y desarrollo de la economía, ya que éstas representan un mecanismo de crédito y una posibilidad de incrementar el nivel de ventas; esto quiere decir que es un beneficio tanto para los consumidores, como para las empresas. Es un beneficio para los consumidores, porque les permite aumentar sus posibilidades de adquisición de bienes y servicios; y es un beneficio para las empresas, debido a que la mayor oportunidad de compra, aumenta la demanda de los bienes y servicios ofertados (Yopo, 2012).

**2.1.1.1 Beneficios y críticas al uso de tarjetas de crédito.** La utilización de las tarjetas de crédito, según Yopo (2012), tiene muchos beneficios; sin embargo, también existen muchas críticas con respecto a ello.

Algunos de los beneficios mencionados por Yopo (2012), son los siguientes:

- No es necesario manejar efectivo para realizar transacciones. Esto permite anular los intercambios de dinero y que se tenga que dar un cambio o vuelto.
- El proceso de pago es mucho más sencillo.
- Las tarjetas de crédito son un medio de pago universal, ya que no se necesita hacer cambio de monedas.
- Existe la oportunidad de consumir o realizar la compra en el momento en que se necesita y realizar el pago después.
- Dependiendo del monto consumido, es posible solicitar el pago en cuotas; lo que hace posible la adquisición de mejores productos o servicios, y que el pago sea más llevadero.

Sin embargo, como se mencionó anteriormente, de acuerdo a Yopo (2012), también existen críticas al uso de las tarjetas de crédito:

- Existen un alto riesgo de fraude que puede perjudicar al propietario de la tarjeta. Dentro de estos riesgos están: la clonación de tarjetas, el cobro indebido de consumos no realizados, entre otros.
- Si el uso de tarjetas de crédito, genera la posibilidad de un aumento de la demanda, esto podría traer consigo un aumento en los niveles de precios fijados en el mercado por parte de los ofertantes de un determinado bien o servicio. Sin embargo, hay que tener muy en

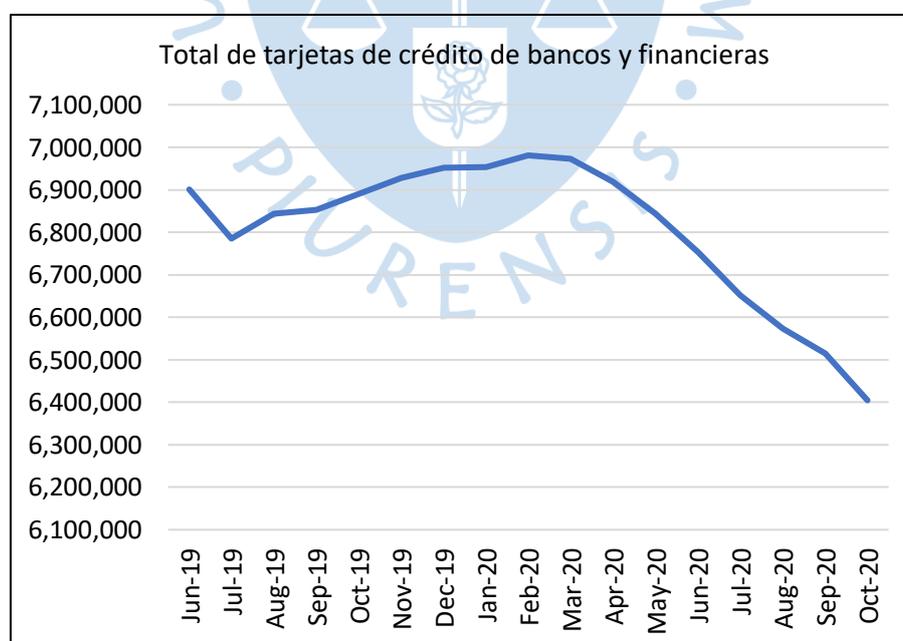
cuenta que, con el uso de las tarjetas de crédito, no se produce un exceso de demanda, por lo que no se puede fijar un precio inadecuado; sino que, es la demanda la que se ajusta a la oferta que ya existe en el mercado.

**2.1.1.2 Evolución del uso de tarjetas de crédito.** La emisión y colocación de tarjetas de crédito en el Perú, durante los últimos años, mostró un proceso de aumento y crecimiento constante hasta el primer cuarto del año 2019, en donde este aumento se revertió y, a partir del mes de abril aproximadamente, se empezó a registrar un declive (Banco Central de Reserva del Perú [BCRP], 2019).

Del 2019, Julio, fue el mes con el total de tarjetas de crédito activas más bajo, lo que significa que no solo disminuyó el nivel de emisión de tarjetas, sino que además hubo desafiliaciones de clientes. A partir de agosto, nuevamente, las colocaciones muestran un aumento constante. En el año 2020, el comportamiento de la emisión de tarjetas de crédito en el Perú, ha sido acorde a la coyuntura mundial que se vive a causa de la pandemia por el coronavirus; esto quiere decir que, a partir del mes de marzo de 2020, se empezó nuevamente, una caída en el proceso de colocación de tarjetas de crédito y hubo un aumento en las desafiliaciones de clientes (Asociación de Bancos del Perú [ASBANC], 2020). Ver la siguiente figura:

**Figura 1**

*Gráfica del total de tarjetas de crédito de bancos y financieras en Perú*



*Nota.* Estadísticas del sistema financiero (ASBANC, 2020).

**2.1.1.3 Preferencias de pago.** Conforme a Arango et al. (2011), las preferencias de los consumidores con respecto al pago de la adquisición de bienes o servicios, se relacionan con el valor del mismo. Esto quiere decir que mientras mayor sea el valor de la transacción, los consumidores, preferirán realizar el pago con una tarjeta de crédito o débito, o con una transferencia bancaria; y, por el contrario, si la transacción no tiene mucho valor, les será más sencillo desembolsar el monto en efectivo.

## **2.2 Credit scoring**

### **2.2.1 Definición de credit scoring**

Según Gutiérrez (2007), al credit scoring, se le conoce también como score-card o classifier, y es un procedimiento que busca encontrar y evaluar el nivel de riesgo de un solicitante de crédito o de un individuo que ya es cliente de una institución financiera o no financiera.

Puertas y Martí (2013) dicen que es un método de valoración crediticia, que hace posible la medición y evaluación automática del riesgo que trae consigo una petición de crédito por parte de un solicitante.

Para Mester (1997), es una forma de evaluar el riesgo crediticio de una solicitud de préstamo o crédito. Para ello, con el apoyo de datos históricos, se trata de encontrar las relaciones y efectos de las características de los solicitantes sobre la posibilidad de que el crédito fracase.

El Credit Scoring tiene dos perspectivas: una perspectiva personal porque se centra en las particularidades propias de cada solicitante, es decir, en las cualidades personales del individuo; y tiene una perspectiva desde el ámbito financiero, ya que, se busca obtener información sobre préstamos u otros instrumentos de crédito pasados, en donde se pueda verificar su nivel de compromiso y responsabilidad en el cumplimiento de las condiciones del crédito (Gutierrez, 2007).

La información obtenida, junto con los datos adicionales brindados en la petición del crédito, permiten tomar una decisión positiva o negativa con respecto a la solicitud. Por lo tanto, se puede concluir, que el objetivo principal de la aplicación de esta herramienta de evaluación es lograr establecer una distinción entre quien es un buen solicitante y quien es un mal solicitante de crédito (Millán y Caicedo, 2018).

### **2.2.2 Antecedentes del credit scoring**

Según Mester (1997), este método de evaluación fue introducido durante los años cincuenta. Sin embargo, varios años antes durante los años cuarenta, Durand (1941), fue una de las primeras investigaciones en la que se logró aplicar un modelo basado en resultados estadísticos.

Lo que se buscaba en Durand (1941) era generar una diferenciación entre lo que podría resultar en un buen o mal crédito; partiendo de la teoría que, un buen crédito es aquel que puede generar un beneficio bruto capaz de cubrir todos los gastos en los que se incurre, inclusive una posible pérdida.

Desde entonces, muchos investigadores han centrado todos sus estudios en la búsqueda de un modelo que permita obtener mejores resultados. Para esto, han aplicado nuevas metodologías y métodos, e inclusive se han desarrollado nuevos programas de computación que sirven para facilitar la aplicación de estos modelos (Kim, 2005).

Durante los años noventa, la aplicación del credit scoring, estaba centrada sobre todo en la evaluación de créditos hipotecarios y tarjetas de crédito (Mester, 1997). Sin embargo, la utilidad y aplicación de este método ha ido evolucionando en el tiempo, y en la actualidad, muchas de las instituciones financieras y no financieras, utilizan el credit scoring para evaluar a sus posibles clientes. Pueden evaluar las solicitudes de los distintos tipos de créditos: personales, para negocios, hipotecarios; e inclusive, el otorgamiento de una tarjeta de crédito (Cruz y Villalta, 2017).

El credit scoring presenta una alta capacidad de predicción, lo que puede permitir realizar un perfeccionamiento en el desarrollo de la evaluación crediticia (Schreiner, 2002).

### **2.2.3 Beneficios y críticas al credit scoring**

Según Schreiner (2002), algunos de los beneficios que tiene este proceso son los siguientes:

- Es objetivo: Si se hace una evaluación de dos solicitudes de personas distintas pero que tienen características parecidas, el pronóstico va a ser el mismo para ambas.
- Es claro y preciso: El proceso de pronóstico del riesgo, se sabe y puede manifestarse, ya que es exacto.
- Considera varios factores o variables: Se puede realizar una mejor evaluación, ya que se puede incluir varias características del solicitante al mismo tiempo.
- Puede testarse: Existe la posibilidad de realizarse ensayos y pruebas con préstamos actuales, y haciéndose uso únicamente de las características conocidas durante el desembolso.

- Genera una reducción en la cantidad de atrasos en los pagos: El resultado de la evaluación va a disminuir la cantidad de préstamos aprobados a solicitantes riesgosos.
- Puede estimarse el efecto del uso del credit scoring sobre la rentabilidad: Se conocería el costo del desembolso de un préstamo que corresponde a una mala solicitud, pero también se sabría la utilidad del desembolso de un préstamo que corresponde a una buena solicitud.

Sin embargo, Schreiner (2002), también indica que existen varias críticas que se pueden hacer contra el credit scoring, y es necesario conocerlas porque de lo contrario puede generarse un fallo en el proceso. Dentro de las críticas, están las siguientes:

- Se necesita información de gran cantidad de préstamos y mucha información acerca de cada préstamo en específico: Es necesario contar con una amplia base de datos.
- Requiere un profesional a cargo, ya que el proceso de evaluación debe ser verificado y ajustado según las necesidades de la entidad que lo utilice.
- Existe información importante que no está incluida dentro de los elementos de riesgo o registrada en la base de datos. Por este motivo, muchas veces es también necesaria una visión subjetiva del proceso de evaluación.
- No aporta en la reducción del tiempo usado en la recolección y evaluación de información brindada por un solicitante: El proceso de evaluación del credit scoring puede descartar una solicitud, sin embargo, no puede aprobarlas.
- Se trabaja bajo el supuesto de que el riesgo depende de ciertas características que son cuantificables: Incluye características como el género, edad, lugar de residencia.
- Se trabaja bajo el supuesto de que el comportamiento del futuro será el mismo del pasado: Los retrasos en pagos de préstamos antiguos son incluidos dentro de las características que se evalúan en el proceso y pueden generar fallos en los pronósticos.

#### **2.2.4 Modelos más comunes del credit scoring**

Existen varias técnicas o metodologías para el desarrollo de modelos credit scoring. Estas técnicas se pueden diferenciar en dos grupos: técnicas paramétricas y técnicas no paramétricas (Cruz y Villalta, 2017).

Las técnicas paramétricas, son aquellas que muestran una probabilidad para el análisis de la relación de los datos, y tienen ciertos parámetros para evaluar y calcular la probabilidad de que un cliente pague o no el crédito. Sin embargo, estas metodologías tienen limitaciones cuando se utilizan muestras pequeñas (Cruz y Villalta, 2017).

Las técnicas no paramétricas, o también llamadas de distribución libre, no tienen limitaciones con respecto al tamaño de la muestra, y por lo tanto se supone que resultan mucho más sencillas de aplicar. Sin embargo, si las variables no son cualitativas, estos métodos resultan poco eficientes (Cruz y Villalta, 2017).

A pesar que no existe un modelo que sea mejor que otro, la forma más avanzada de utilizar los modelos credit scoring, es a través de metodologías no paramétricas o también llamadas de respuesta cualitativa (Herrán, 2009). En el presenta trabajo de investigación, se profundizará en estos modelos.

En la Tabla 1 se muestran los beneficios y críticas de utilizar las técnicas paramétricas de credit scoring.

**Tabla 1**

*Comparación de modelos paramétricos de credit scoring*

Metodología	Beneficios	Críticas
Técnicas Paramétricas	Modelo lineal de probabilidad	<ul style="list-style-type: none"> <li>• Estimadores ineficientes.</li> <li>• La probabilidad estimada puede quedar fuera del rango [0,1].</li> </ul>
	Modelo logit	<ul style="list-style-type: none"> <li>• Dificultad de interpretación de los parámetros.</li> </ul>
	Modelo probit	<ul style="list-style-type: none"> <li>• Dificultad de interpretación de los parámetros.</li> <li>• Proceso de estimación relativamente complicado.</li> </ul>

*Nota.* Adaptado de A credit risk model for agricultural loan portfolios under the new Basel capital accord (Kim, 2005).

Con base en Cruz y Villalta (2017), y Herrán (2009), se presenta una breve descripción y evaluación de cada modelo comparado en la tabla anterior.

**2.2.4.1 Modelo lineal de probabilidad.** Para estimar este modelo, se utiliza una regresión simple por mínimos cuadrados.

La variable dependiente es una variable dicotómica, o también llamada dummy, lo que quiere decir que solo puede tomar dos valores: 0 y 1.

$$Y_i = \begin{cases} 1, & \text{si el cliente no cumple con el pago del crédito.} \\ 0, & \text{si el cliente sí cumple con el pago del crédito.} \end{cases}$$

La regresión es una función lineal de un conjunto de variables explicativas. Por temas de simplicidad, se asumirá que existe una sola variable  $X_i$ .

$$Y_i = \beta_0 + \beta_i X_i + \mu_i$$

El beneficio principal de este modelo, es que la estimación e interpretación de los resultados es muy sencilla. Sin embargo, es necesario aclarar que, la utilización de estos modelos ha ido disminuyendo debido a dos razones principales:

- La probabilidad puede tomar valores fuera de  $[0,1]$ , lo que dificulta la evaluación del solicitante, debido a que lo que se busca estimar es justo la probabilidad de que el cliente no cumpla con el pago.
- No existe homocedasticidad. Esto quiere decir que hay una dispersión en los valores de la varianza de los errores para todas las observaciones de la muestra. Además, la homocedasticidad es un requisito esencial de la estimación de los modelos lineales.

Para hacerle frente al primer problema que presenta el modelo lineal de probabilidad, descrito en el párrafo anterior, surgen los modelos logit y probit. Estos modelos tienen la cualidad de restringir la probabilidad en un rango de valores determinado entre  $[0,1]$ .

**2.2.4.2 Modelo logit.** Este modelo considera un grupo de variables explicativas o independientes  $X = \{X_1, X_2, \dots, X_i\}$ ; y una variable dependiente, que en el marco del credit

scoring puede ser definida como la predisposición del cliente a incumplir con el pago, y es una variable binaria porque puede tomar valores de 0 y 1.

El modelo inicial es una regresión como la siguiente:

$$y_i^* = \beta_0 + \sum_{j=1}^n \beta_j x_{ij} + \mu_i$$

El modelo logit se define según la siguiente función de distribución logística acumulada:

$$P_i = E\left(y_i = \frac{1}{x_i}\right) = \frac{1}{1 + e^{-(\beta_0 + \sum \beta_j x_{ij})}}$$

Si  $Z_i = \beta_0 + \sum \beta_j x_{ij}$ , entonces:

$$P_i = E\left(y_i = \frac{1}{x_i}\right) = \frac{1}{1 + e^{-Z_i}} = \frac{e^{Z_i}}{1 + e^{Z_i}}$$

En este modelo, la importancia de los coeficientes encontrados radica en el signo que tienen, ya que se busca hallar una relación de causalidad de las variables explicativas, sobre la variable dependiente; es decir, si el signo es positivo, la relación es directa.

**2.2.4.3 Modelo probit.** Al igual que el modelo logit, este modelo, permite hallar probabilidades estimadas dentro del rango de [0,1].

El modelo probit, se basa en una función de distribución normal acumulada, como la siguiente:

$$P_i = \Pr(y_i = 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_i} e^{-\frac{t^2}{2}} dt$$

Donde:

$$Z_i = \beta_0 + \sum \beta_j x_{ij}$$

$$t = \frac{Z_i - \mu_Z}{\sigma} \sim N(0,1)$$

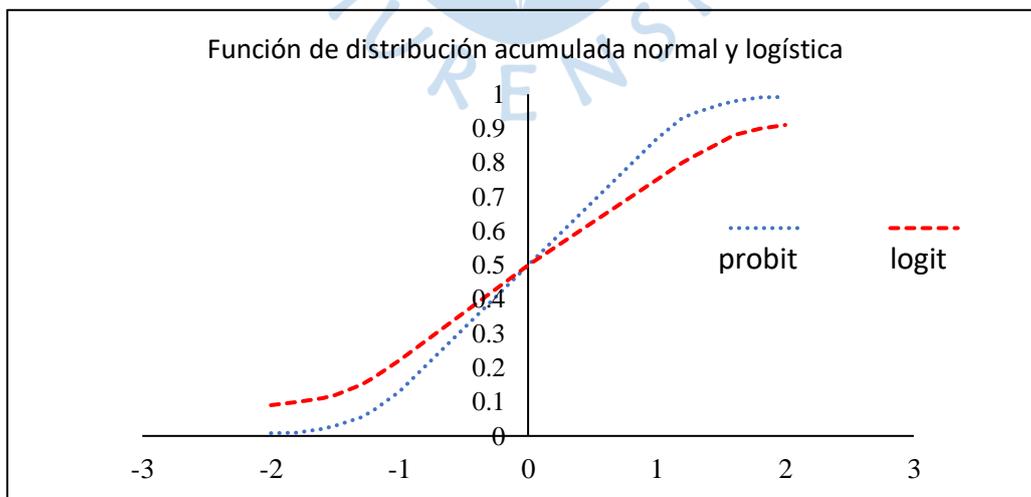
$$t = Z_i, \text{ dado que } \mu_Z = 0 \text{ y } \sigma = 1$$

La estimación de los coeficientes del modelo probit, se realiza a través del método de máxima verosimilitud.

En general, el modelo logit y el probit, son muy semejantes, y tienen funciones de probabilidad muy cercanas entre sí; sin embargo, si existen muchas observaciones en los extremos, los resultados encontrados, empiezan a separarse (Herrán, 2009). Ver la siguiente figura:

**Figura 2**

*Gráfica de comparación entre la función de distribución acumulada normal y logística*



*Nota.* Lara Rubio (2010), citado por Cruz y Villalta (2017).

Con respecto al nivel de dificultad de aplicación, evaluación e interpretación de los datos, el modelo logit es más sencillo que el modelo probit; pero, hay que tener en cuenta que, actualmente, con la tecnología, esta diferencia se hace menos notoria (Herrán, 2009).

## 2.3 Aplicación del *credit scoring*

### 2.3.1 Descripción de la data

Para este trabajo de investigación, se ha utilizado la base de datos de Yeh y Lien (2009). Esta base de datos cuenta con 30,000 observaciones; de las cuales, el 77.88% (23,364 observaciones) corresponde a tarjetahabientes que cumplen responsablemente con las condiciones de pago de sus tarjetas de crédito, y el 22.12% (6,636 observaciones) restante, corresponde a tarjetahabientes que no lo hacen. Ver la siguiente tabla:

**Tabla 2**

*Tabulación de variable dependiente*

Variable dependiente	Frec.	Porc.	Acum.
0	23,364	77.88	77.88
1	6,636	22.12	100.00
Total	30,000	100.00	

*Nota.* Elaboración propia.

### 2.3.2 Variables seleccionadas

Tal como se mencionó en la definición, al inicio de la segunda sección de este capítulo; la metodología de *credit scoring*, tiene una perspectiva personal y otra financiera (Gutiérrez, 2007). Es por ello, que es necesario incluir variables que sean cualidades propias del individuo y variables que demuestren su comportamiento en solicitudes anteriores de crédito.

Con base en Yeh y Lien (2009), la variable dependiente será una dicotómica que toma el valor de 1 cuando se incumple con el pago, y toma el valor de 0 cuando se cumple con el pago a tiempo. Con respecto a las variables independientes, se presenta una tabla en la que se hace una descripción de las 23 variables incluidas en la base de datos y los signos esperados para cada una de ellas (Ver anexo 2).

Como se puede observar en la tabla (Anexo 2), las variables que representan características personales del solicitante tales como el género o el nivel de educación, tienen en la casilla de signo

esperado un +/- . Esto se debe a que los efectos esperados de estas variables son ambiguos, y por lo tanto pueden interpretarse de formas diferentes.

A continuación, se presenta una breve explicación de los signos y efectos esperados de las variables explicativas sobre la probabilidad de que un tarjetahabiente no cumpla con el pago bajo las condiciones:

- SEX: Es una dummy que indica cual es la probabilidad de que el género del solicitante, influya sobre el incumplimiento del pago de la tarjeta de crédito.
- EDUCATION: Se esperaría que, un solicitante con un mayor nivel de preparación, en cuanto a educación, sea más responsable con sus pagos, lo que significa que tiene menor probabilidad de incumplimiento. De la misma manera, se esperaría que, un solicitante con menor nivel académico, sea menos responsable con sus pagos y tenga una mayor probabilidad de incumplimiento.
- MARRIAGE: Puede presumirse que, un solicitante comprometido o casado tiene más gastos que uno soltero y, por lo tanto, tiene una mayor probabilidad de incumplimiento del pago de la tarjeta de crédito.
- AGE: Puede interpretarse de la misma manera que la variable education, es decir que, un solicitante con más años de edad será más responsable que uno menor y, por lo tanto, tendría una menor probabilidad de incumplimiento del pago.
- LIMIT\_BAL: Mientras mayor es el monto de crédito que ha recibido el solicitante, quiere decir que, ha tenido varias evaluaciones previas, y que ha cumplido adecuadamente con los pagos; por lo tanto, existe menor probabilidad de incumplimiento.
- PAY\_MONTH: Si un solicitante ya ha mostrado, en meses anteriores, que se atrasa en los pagos; esto quiere decir que, la probabilidad de no cumplir con los pagos a futuro es mayor.
- BILL\_AMT\_MONTH: Mientras mayor es el monto en el estado de cuenta del solicitante, quiere decir que tiene la liquidez para solventar los pagos a futuro; por lo tanto, la probabilidad de no cumplir con los pagos disminuye.
- PAY\_AMT\_MONTH: Mientras mayor sea el monto amortizado o pagado en meses anteriores, quiere decir que el solicitante tiene la capacidad económica para hacerlo; por lo tanto, la probabilidad de incumplir con los pagos disminuye.

Es importante resaltar que, las variables incluidas en la base de datos son muy limitadas y únicamente reflejan el comportamiento de pago del individuo durante seis meses (desde abril hasta setiembre de 2005).

### 2.3.3 Modelo de estimación

Para realizar la estimación del modelo de credit scoring, se utilizará la metodología de regresión logística o logit ya que, como se describió anteriormente, no presenta errores en los parámetros.

Previo a la estimación del modelo final, es necesario hacer algunas pruebas adicionales que permitan escoger las variables que se van a incluir. El proceso que se realiza consta de varias etapas, que serán explicadas a continuación:

En primer lugar, se evaluará cada variable por separado. Para esto, se estima un modelo logit incluyendo únicamente una variable independiente para comparar los signos de los coeficientes y verificar el nivel de significancia (Ver anexo 3).

Con los resultados de las estimaciones obtenidas, se presenta el resumen de la tabla 3. En el resumen podemos observar que, de las 23 variables iniciales, hay 17 variables cuyo signo del coeficiente estimado es el mismo que el esperado, y que cumplen con un nivel de significancia del 99%. Las variables a descartarse son: AGE, BILL\_AMT\_AUG, BILL\_AMT\_JUL, BILL\_AMT\_JUN, BILL\_AMT\_MAY y BILL\_AMT\_APR. Ver tabla 3:

**Tabla 3**

*Resumen de estimación de regresiones con una variable*

Variable	Beta	Prob.
SEX	-0.194931	0.0000
EDUCATION	0.083947	0.0000
MARRIAGE	-0.112376	0.0000
AGE	0.003607	0.0162
LIMIT_BAL	-3.30E-06	0.0000
PAY_SEP	0.737194	0.0000
PAY_AUG	0.518176	0.0000
PAY_JUL	0.460205	0.0000
PAY_JUN	0.431796	0.0000
PAY_MAY	0.420324	0.0000
PAY_APR	0.377344	0.0000
BILL_AMT_SEP	-6.67E-07	0.0007
BILL_AMT_AUG	-4.94E-07	0.0140
BILL_AMT_JUL	-5.04E-07	0.0148
BILL_AMT_JUN	-3.88E-07	0.0786
BILL_AMT_MAY	-2.72E-07	0.2417

Variable	Beta	Prob.
BILL_AMT_APR	-2.20E-07	0.3521
PAY_AMT_SEP	-3.64E-05	0.0000
PAY_AMT_AUG	-3.22E-05	0.0000
PAY_AMT_JUL	-2.10E-05	0.0000
PAY_AMT_JUN	-2.01E-05	0.0000
PAY_AMT_MAY	-1.82E-05	0.0000
PAY_AMT_APR	-1.40E-05	0.0000

*Nota.* Elaboración propia con los resultados de las estimaciones obtenidas en el anexo 3.

Luego, se evalúan las variables en conjunto. Para esto, se estima un modelo logit incluyendo las 17 variables independientes, para nuevamente, hacer una comparación del signo de los coeficientes estimados y verificar el nivel de significancia (Ver anexo 4).

Como resultado de la estimación conjunta, podemos descartar siete variables que no cumplen con lo explicado en el párrafo anterior: PAY\_JUN, PAY\_MAY, PAY\_APR, PAY\_AMT\_JUL, PAY\_AMT\_JUN, PAY\_AMT\_MAY y PAY\_AMT\_APR.

El siguiente paso, consiste en revisar la correlación de las variables explicativas (Ver anexo 5), a fin de eliminar las multicolinealidad. En la tabla 4, se presentan los resultados con un valor de covarianza mayor al 50%.

**Tabla 4**

*Relación de variables correlacionadas*

Variable 1	Variable 2	Correlación
PAY_AUG	PAY_JUL	0.766552
PAY_AUG	PAY_SEP	0.672164
PAY_JUL	PAY_SEP	0.574245

*Nota.* Elaboración propia con los resultados de la matriz de correlación.

Por lo tanto, también vamos a descartar las variables: PAY\_SEP, PAY\_AUG y PAY\_JUL.

Ahora, se estimará el modelo logit, con las variables finales: SEX, EDUCATION, MARRIAGE, LIMIT\_BAL, BILL\_AMT\_SEP, PAY\_AMT\_SEP y PAY\_AMT\_AUG (Ver anexo 6).

Los coeficientes estimados en la regresión del modelo final de logit, no pueden ser directamente interpretados o comparados; ya que, tal como se mencionaba en la Tabla 1, la interpretación de los parámetros del modelo logit, así como del modelo probit, es más compleja. Por eso, es necesario hallar los cambios marginales de las variables seleccionadas.

A continuación, en la tabla 5 se presentan los resultados de las estimaciones de los cambios marginales de las variables explicativas.

**Tabla 5**

*Estimación de cambios marginales de las variables seleccionadas*

Variable	Cambio marginal
SEX	-0.2211874
EDUCATION	-0.0876728
MARRIAGE	-0.2627907
LIMIT_BAL	-0.4416996
BILL_AMT_SEP	0.0810416
PAY_AMT_SEP	-0.1085136
PAY_AMT_SEP	-0.0862661

*Nota.* Elaboración propia.

Finalmente, se evaluará el desempeño del modelo elegido; para lo cual, se utilizará la matriz de confusión. En figura 3 se pueden ver los resultados obtenidos.

**Figura 3**

*Matriz de confusión*

[	[23364	0]	
	[6636	0]	]

*Nota.* Elaboración propia.

La matriz de confusión, o también llamada de clasificación, es una herramienta muy útil porque permite visualizar el nivel de desempeño de una estimación de probabilidad. Cada fila de la matriz

representa el valor real de cada clase, y en cada columna, se representa el valor estimado (Recuero, 2018). En este caso en particular, las clases son dos: el individuo cumple o incumple con el pago.

En la tabla 6 se muestra el reporte de clasificación, que nos permitirá evaluar el modelo utilizado.

**Tabla 6**

*Reporte de clasificación*

Nombre de la variable: DEFAULT_PAYMENT_NEXT_MONTH				
	precision	recall	f1-score	support
Non-Default	0.78	1	0.88	23364
Default	0	0	0	6636
Accuracy			0.78	30000
macro avg	0.39	0.5	0.44	30000
weighted avg	0.61	0.78	0.68	30000

*Nota.* Elaboración propia.

### 2.3.4 Análisis de resultados

- Por medio de la primera evaluación de variables, que consistió en regresiones de cada variable por separado, se descartaron: AGE, BILL\_AMT\_AUG, BILL\_AMT\_JUL, BILL\_AMT\_JUN, BILL\_AMT\_MAY y BILL\_AMT\_APR; por no cumplir con el nivel de significancia esperado del 99%.
- La segunda evaluación, consistió en estimar un modelo incluyendo todas las variables restantes, de las cuales fueron eliminadas: PAY\_JUN, PAY\_MAY, PAY\_APR, PAY\_AMT\_JUL, PAY\_AMT\_JUN, PAY\_AMT\_MAY y PAY\_AMT\_APR.
- Para evitar tener un nivel de multicolinealidad elevado, se decidió descartar las siguientes variables: PAY\_SEP, PAY\_AUG y PAY\_JUL.
- Para el modelo final se decidió incluir siete variables explicativas: SEX, EDUCATION, MARRIAGE, LIMIT\_BAL, BILL\_AMT\_SEP, PAY\_AMT\_SEP y PAY\_AMT\_AUG.
- Las variables SEX, EDUCATION y MARRIAGE; es importante incluirlas, porque brinda información personal del usuario de la tarjeta de crédito. Esto es muy necesario, porque tal

como se indica al inicio del capítulo, para los modelos de estimación de credit scoring, se debe incluir una perspectiva para la evaluación financiera, pero también una perspectiva personal que brinde información particular de cada individuo o solicitante (Gutiérrez, 2007).

- La variable LIMIT\_BAL, debe incluirse en el modelo porque es la muestra del historial crediticio del solicitante; ya que, incorpora el monto de los créditos obtenidos por el usuario tanto de manera personal, como familiar. Además, se puede inferir que mientras mayor es el monto de los créditos anteriores, el usuario ha pasado por buenos filtros y evaluaciones previas.
- Las variables BILL\_AMT\_SEP, PAY\_AMT\_SEP y PAY\_AMT\_AUG, muestran el comportamiento del usuario durante los meses previos, lo cual es muy importante porque con la experiencia del pasado, se puede anticipar el futuro comportamiento del usuario en los meses próximos.
- Los cambios marginales estimados indican que, la única variable independiente que tiene una relación positiva con la variable dependiente es BILL\_AMT\_SEP; lo que significa que mientras mayor es la cantidad del monto del estado de cuenta en el mes de setiembre, la probabilidad de incumplir con el pago de la tarjeta de crédito, aumenta. Además, ésta variable presenta el menor cambio marginal, lo que significa que tiene menor efecto sobre la probabilidad de incumplimiento.
- Las seis variables restantes: SEX, EDUCATION, MARRIAGE, LIMIT\_BAL, PAY\_AMT\_SEP y PAY\_AMT\_AUG, tienen una relación negativa con la variable dependiente. Estas variables pueden ser interpretadas de dos formas, según sean discretas o continuas.

En el caso de EDUCATION, que es una variable discreta, se interpreta de la siguiente manera: si una persona tiene un mayor grado de educación, la probabilidad de incumplimiento disminuye. De la misma manera se puede interpretar las variables SEX y MARRIAGE.

En el caso de LIMIT\_BAL, que es una variable continua, se interpreta de la siguiente manera: mientras mayor es la cantidad de crédito total obtenido por el solicitante (personal y familiar), la probabilidad de incumplir con el pago de la tarjeta de crédito, disminuye. Además, ésta variable presenta el mayor cambio marginal, lo que significa que tiene un efecto mayor sobre dicha probabilidad.

- La matriz de confusión (Figura 3) muestra que, 23364 predicciones han sido negativos reales o verdaderos; y que solo 6636 predicciones han sido falsos negativos. Además, muestra que el modelo no ha podido predecir ningún positivo verdadero, ni ningún falso positivo.
- En el reporte de clasificación (Tabla 6), se puede observar el valor de una medida llamada accuracy o exactitud, que se puede definir como el número total de predicciones correctas, con respecto del total de predicciones; lo que quiere decir que usualmente buscamos que el valor de esta medida sea lo mayor posible. En este caso, la exactitud es de 0.78; sin embargo,

a pesar de ser un buen valor, no es representativo, ya que no ha podido clasificar ninguno de los valores en la columna de positivos verdaderos o falsos positivos.

- La precisión es la relación entre el total de verdaderos positivos sobre el total de positivos predichos; y la sensibilidad o también llamada recall, es la relación entre el total de positivos verdaderos sobre el total de positivos reales.





## Conclusiones

El uso de tarjetas de crédito, en nuestro país, sigue en aumento, por lo tanto, es necesario mejorar y, de la misma manera, facilitar los procesos de evaluación para el otorgamiento de líneas de crédito adecuadas a las posibilidades de cada solicitante o usuario.

Además de las desventajas de la utilización de las tarjetas de crédito mencionadas por Yopo (2012), es importante incluir la mala educación financiera que existe en el país, ya que, esto trae consigo que los usuarios o tarjetahabientes hagan mal uso de las tarjetas de crédito y muchas veces, generan un nivel de endeudamiento que finalmente no pueden asumir.

Con respecto al modelo, basado en la estimación logit final, se pudo obtener resultados que no indican un buen ajuste de este, ya que el  $R^2$  de Mc Fadden es de 0.036, que es un valor muy pequeño, lo que significa que el nivel de predicción es muy bajo.

Por medio de la matriz de confusión y del reporte de clasificación, se hallaron métricas con valores deseables, sin embargo, no representativos, ya que el modelo únicamente lograba clasificar los valores en una de las clases.

Los resultados pueden ser mejorados; sin embargo, ya que en este caso en especial, hubo una falta importante de información, es necesario obtener una base de datos adecuada que incluya variables relevantes para la evaluación de las solicitudes, además de características personales del individuo, que permitan mejorar las predicciones. Las posibles variables que pueden incluirse son: la tasa de interés de la tarjeta de crédito, el nivel de ingreso mensual del usuario, el nivel de otros ingresos en el mes, la experiencia o productividad del usuario, si es dueño de su propia empresa o es empleado, si es un empleado fijo o no, el nivel de endeudamiento con otras instituciones financieras, si tiene otras tarjetas de crédito con la misma entidad u otras instituciones financieras, la línea de crédito con la que cuenta en otras entidades, entre otras.



## Lista de referencias

- Aparicio, C., Gutiérrez, J., Jaramillo, M., y Moreno, H. (2013). Indicadores alternativos de riesgo de crédito en el Perú: matrices de transición crediticia condicionadas al ciclo económico. Superintendencia de Banca, Seguros y Administradoras Privadas de fondos de Pensiones.
- Arango, C., Huynh, K., y Sabetti, L. (2011) How do you pay? the role of incentives at the point of sale. *Working Paper Series*. Recuperado de: <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1386.pdf>
- Azteca Perú. (s. f.). Estadísticas: ASBANC. Recuperado 4 de enero de 2021, de <https://www.asbanc.com.pe/Paginas/Estadistica/Estadisticas.aspx?posTabActivo=2>
- Caicedo, E. y Millán, J. (2018). Modelos para otorgamiento y seguimiento en la gestión de riesgo de crédito. *Revista de métodos cuantitativos para la economía y la empresa*, (25). pp. 23-41. Recuperado de: <https://www.upo.es/revistas/index.php/RevMetCuant/article/view/2370/2709>
- Cheng, I., y Che, L. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients, *Expert Systems with Applications*. *Science Direct*, 36(1), 2473-2480. <https://doi.org/10.1016/j.eswa.2007.12.020>
- Cruz, D. y Villalta, A. (2017). *A credit scoring como componente de los sistemas de evaluación de riesgo en microcrédito dentro de la banca cooperativa salvadoreña*. (Tesis de maestría). Universidad de El Salvador. Ciudad Universitaria, El Salvador.
- Durand, D. (1941). *Risk elements in consumer instalment financing*, technical edition. Chicago, Estados Unidos: NBER.
- Gutiérrez, M. (2007). Modelos de Credit Scoring: Qué, Cómo, Cuándo y Para Qué. Recuperado de: <https://mpa.ub.uni-muenchen.de/16377/>
- Guzmán, M. y Ruiz, J. (2010). *Análisis de riesgo de crédito y evidencia empírica en Chile*. (Tesis de licenciatura). Universidad de Chile. Santiago de Chile, Chile.
- Hernández, L. Meneses, L. y Benavides, J. (2005). Desarrollo de una metodología propia de análisis de crédito empresarial en una entidad financiera. *Estudios Gerenciales*, 21(97), pp.129-165. Recuperado

de:[http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S012359232005000400007&lng=en&tlng=es](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S012359232005000400007&lng=en&tlng=es).

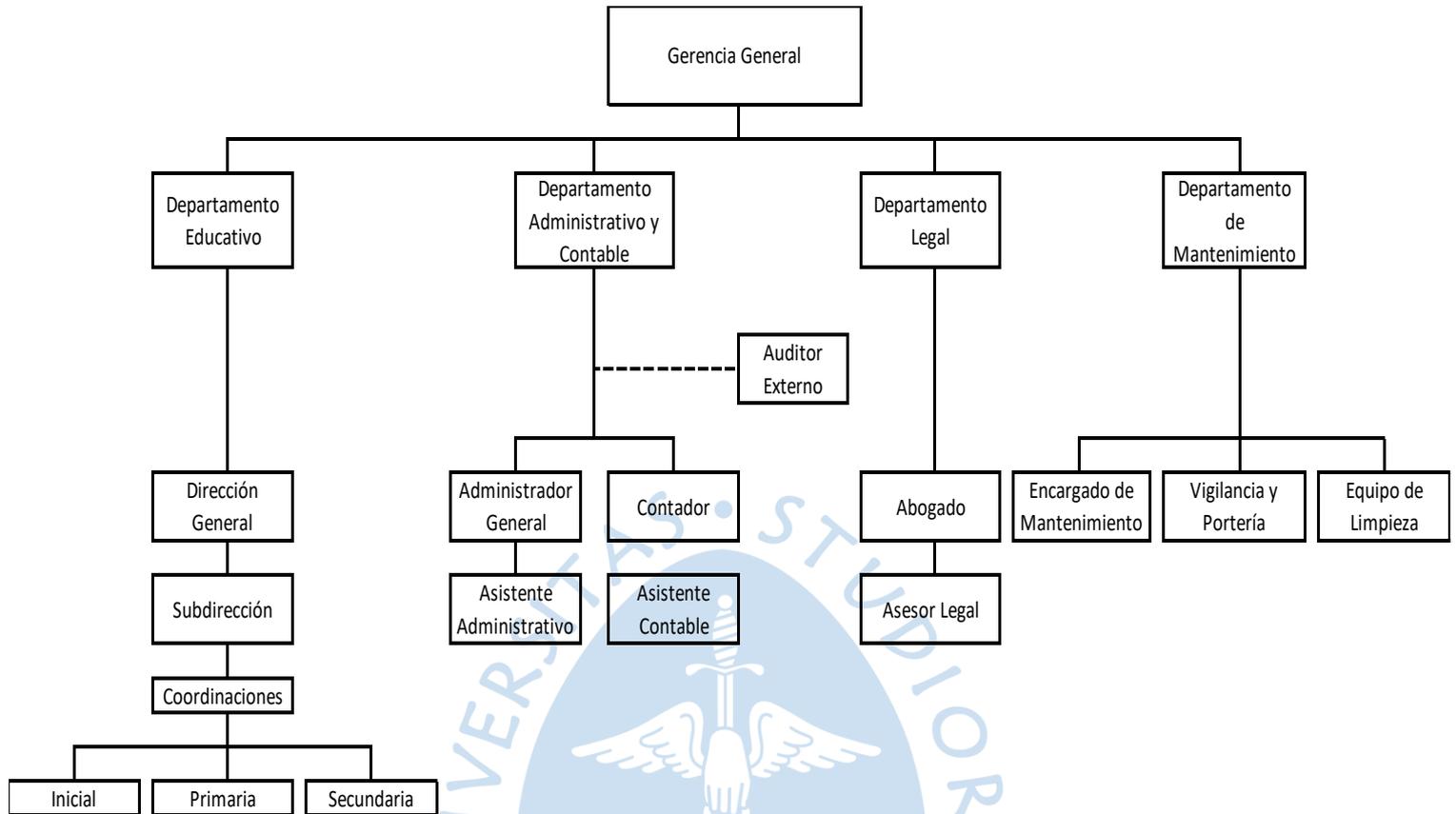
- Kim, J. (2005). *A credit risk model for agricultural loan portfolios under the new basel capital Accord*. (Tesis doctoral) Universidad de Texas A y M. Texas, Estados Unidos.
- Lara, A. (2008). *Medición y control de riesgos financieros* (3.ª ed.). Ciudad de México, México: Limusa.
- Mester, L. (1997). What's the point of credit scoring? *Business Review*. pp.3–16. Recuperado de: [https://www.researchgate.net/publication/5051659\\_What\\_Is\\_the\\_Point\\_of\\_Credit\\_Scoring](https://www.researchgate.net/publication/5051659_What_Is_the_Point_of_Credit_Scoring)
- Núñez, A. (2012). *Diseño de un manual de políticas y procedimientos de crédito y cobranza para disminuir la cartera vencida de la empresa autopolo de la ciudad de Ambato* (Tesis de licenciatura). Pontificia universidad católica del ecuador, Ambato, Ecuador
- Pedrosa, J. (2020, 27 julio). Heterocedasticidad. Recuperado 4 de enero de 2021, de [https://economipedia.com/definiciones/heterocedasticidad.html#:~:text=La%20heterocedasticidad%20es%2C%20en%20estad%C3%ADstica,t%C3%A9rmino%20es%20contrario%20a%20homocedasticidad.&text=La%20palabra%20heterocedasticidad%20se%20puede,%20y%20cedasticidad%20\(dispersi%C3%B3n](https://economipedia.com/definiciones/heterocedasticidad.html#:~:text=La%20heterocedasticidad%20es%2C%20en%20estad%C3%ADstica,t%C3%A9rmino%20es%20contrario%20a%20homocedasticidad.&text=La%20palabra%20heterocedasticidad%20se%20puede,%20y%20cedasticidad%20(dispersi%C3%B3n)
- Puertas, R. y Martí, M. (2013). Análisis del credit scoring. *Revista de Administración de Empresas*, 53(3), pp.303-315. <https://doi.org/10.1590/s0034-75902013000300007>
- Recuero, P. (2018, 23 enero). Machine Learning a tu alcance: La matriz de confusión. Recuperado 30 de febrero de 2021, de <https://empresas.blogthinkbig.com/ml-a-tu-alcance-matriz-confusion/>
- Schreiner, M. (2002). Ventajas y desventajas del scoring estadístico para las microfinanzas. Recuperado de: [http://www.microfinance.com/Castellano/Documentos/Scoring\\_Ventajas\\_Desventajas.pdf](http://www.microfinance.com/Castellano/Documentos/Scoring_Ventajas_Desventajas.pdf)
- Yeh, I. y Lien, C. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications*, 36(2), 2473-2480.
- Yopo, N. (2012). *Responsabilidad en los casos de fraude por extravío, hurto o robo de la tarjeta de crédito*. (Tesis de licenciatura). Universidad de Chile. Santiago de Chile, Chile.

Apéndices

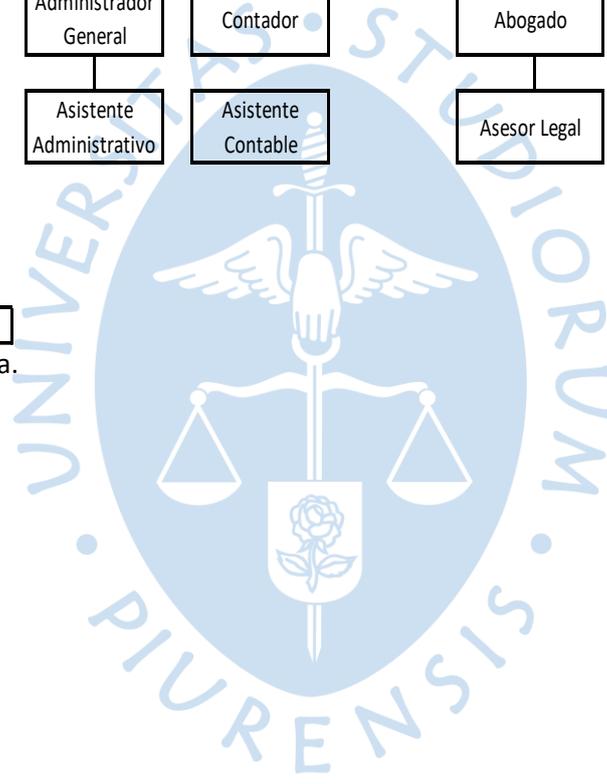




### Apéndice 1. Organigrama IEP Virgen del Carmen



Nota. Elaboración propia.





Anexos





Anexo 1. Certificado de trabajo emitido por la IEP Virgen del Carmen



I.E.P.  
"Virgen del Carmen" E.I.R.L.  
EDUCACIÓN INICIAL PRIMARIA Y SECUNDARIA

**CERTIFICADO DE TRABAJO**

La Srta. **MARÍA ANTONIETA SEMINARIO SAN MARTÍN**, identificada con DNI N° 03583699, Directora de I.E.P. VIRGEN DEL CARMEN E.I.R.L., con RUC N° 20526673302.

**CERTIFICA:**

Que el Sr. **VICTOR ARMANDO RUIZ AYÓN**, identificado con DNI N° 70579800, ha laborado en nuestra empresa como **ASISTENTE**, durante el período comprendido desde el 01 de Noviembre de 2016 hasta el 18 de Enero de 2018, demostrando durante su permanencia responsabilidad, honestidad y dedicación en las labores que le fueron encomendadas.

Se expide el presente certificado para los fines que estime conveniente.

Sullana, 19 de Enero de 2018

I.E.P. VIRGEN DEL CARMEN E.I.R.L.

*Maria Antonieta Seminario San Martín*  
Maria A. Seminario San Martín

DIRECTORA

C.M. 1003583699

*Nota.* Certificado emitido por la Dirección General de la I.E.P. Virgen del Carmen E.I.R.L.

**Anexo 2. Descripción de las variables y signos esperados**

Variable	Descripción	Signo esperado
SEX	Género: 1 = masculino. 2 = femenino.	+/-
EDUCATION	Nivel de educación o académico: 1 = posgrado. 2 = universidad. 3 = secundaria. 4 = otros.	+/-
MARRIAGE	Estado civil: 1 = casado. 2 = soltero. 3 = otro.	+/-
AGE	Edad expresada en años.	+/-
LIMIT_BAL	Cantidad total del crédito otorgado: personal y familiar.	-
PAY_SEP PAY_AUG PAY_JUL PAY_JUN PAY_MAY PAY_APR	Historial del registro de pagos en meses anteriores: -1 = pago a tiempo. 1 = pago con 1 mes de retraso. 2 = pago con 2 meses de retraso. 3 = pago con 3 meses de retraso. ... 9 = pago con 9 meses de retraso o más.	+
BILL_AMT_SEP	Monto del estado de cuenta en setiembre.	-
BILL_AMT_AUG	Monto del estado de cuenta en agosto.	-
BILL_AMT_JUL	Monto del estado de cuenta en julio.	-
BILL_AMT_JUN	Monto del estado de cuenta en junio.	-
BILL_AMT_MAY	Monto del estado de cuenta en mayo.	-
BILL_AMT_APR	Monto del estado de cuenta en abril.	-
PAY_AMT_SEP	Monto de pago en setiembre.	-
PAY_AMT_AUG	Monto de pago en agosto.	-
PAY_AMT_JUL	Monto de pago en julio	-
PAY_AMT_JUN	Monto de pago en junio.	-
PAY_AMT_MAY	Monto de pago en mayo.	-
PAY_AMT_APR	Monto de pago en abril.	-

*Nota.* Se evalúa cada variable por separado según el efecto se espera obtener.

### Anexo 3. Estimación de regresiones logísticas con cada variable independiente

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH  
 Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)  
 Sample: 1 30000  
 Included observations: 30000  
 Convergence achieved after 4 iterations  
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
SEX	-0.194931	0.028186	-6.915953	0.0000
C	-0.948602	0.046598	-20.35696	0.0000
McFadden R-squared	0.001501	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.414737	
Akaike info criterion	1.055392	Sum squared resid	5159.864	
Schwarz criterion	1.055946	Log likelihood	-15828.89	
Hannan-Quinn criter.	1.055570	Deviance	31657.77	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	47.58268	Avg. log likelihood	-0.527630	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH  
 Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)  
 Sample: 1 30000  
 Included observations: 30000  
 Convergence achieved after 4 iterations  
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
EDUCATION	0.083947	0.017317	4.847622	0.0000
C	-1.415475	0.035441	-39.93912	0.0000
McFadden R-squared	0.000734	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.414928	
Akaike info criterion	1.056203	Sum squared resid	5164.622	
Schwarz criterion	1.056757	Log likelihood	-15841.05	
Hannan-Quinn criter.	1.056381	Deviance	31682.09	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	23.26159	Avg. log likelihood	-0.528035	
Prob(LR statistic)	0.000001			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH  
 Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)  
 Sample: 1 30000  
 Included observations: 30000  
 Convergence achieved after 4 iterations  
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
MARRIAGE	-0.112376	0.026664	-4.214503	0.0000
C	-1.085253	0.043225	-25.10712	0.0000
McFadden R-squared	0.000561	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.414945	
Akaike info criterion	1.056386	Sum squared resid	5165.024	
Schwarz criterion	1.056940	Log likelihood	-15843.79	
Hannan-Quinn criter.	1.056564	Deviance	31687.58	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	17.77329	Avg. log likelihood	-0.528126	
Prob(LR statistic)	0.000025			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH  
 Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)  
 Sample: 1 30000  
 Included observations: 30000  
 Convergence achieved after 3 iterations  
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
AGE	0.003607	0.001499	2.405500	0.0162
C	-1.386982	0.055242	-25.10733	0.0000
McFadden R-squared	0.000182	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.415026	
Akaike info criterion	1.056786	Sum squared resid	5167.058	
Schwarz criterion	1.057340	Log likelihood	-15849.80	
Hannan-Quinn criter.	1.056964	Deviance	31699.59	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	5.762287	Avg. log likelihood	-0.528327	
Prob(LR statistic)	0.016374			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 4 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
LIMIT_BAL	-3.30E-06	1.26E-07	-26.14463	0.0000
C	-0.755583	0.022412	-33.71352	0.0000
McFadden R-squared	0.024414	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.409561	
Akaike info criterion	1.031176	Sum squared resid	5031.867	
Schwarz criterion	1.031730	Log likelihood	-15465.64	
Hannan-Quinn criter.	1.031354	Deviance	30931.29	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	774.0666	Avg. log likelihood	-0.515521	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 5 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_SEP	0.737194	0.014277	51.63491	0.0000
C	-1.402491	0.015596	-89.92409	0.0000
McFadden R-squared	0.100102	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.385013	
Akaike info criterion	0.951186	Sum squared resid	4446.745	
Schwarz criterion	0.951740	Log likelihood	-14265.78	
Hannan-Quinn criter.	0.951363	Deviance	28531.57	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	3173.784	Avg. log likelihood	-0.475526	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 4 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_AUG	0.518176	0.011997	43.19303	0.0000
C	-1.284615	0.014638	-87.75994	0.0000
McFadden R-squared	0.063450	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.396659	
Akaike info criterion	0.989922	Sum squared resid	4719.838	
Schwarz criterion	0.990476	Log likelihood	-14846.83	
Hannan-Quinn criter.	0.990100	Deviance	29693.66	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	2011.690	Avg. log likelihood	-0.494894	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 4 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_JUL	0.460205	0.011883	38.72933	0.0000
C	-1.257972	0.014415	-87.26737	0.0000
McFadden R-squared	0.050463	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.400836	
Akaike info criterion	1.003647	Sum squared resid	4819.775	
Schwarz criterion	1.004201	Log likelihood	-15052.71	
Hannan-Quinn criter.	1.003825	Deviance	30105.41	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	1599.942	Avg. log likelihood	-0.501757	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 3 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_JUN	0.431796	0.012175	35.46673	0.0000
C	-1.226024	0.014268	-85.92538	0.0000
McFadden R-squared	0.042571	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.402928	
Akaike info criterion	1.011987	Sum squared resid	4870.201	
Schwarz criterion	1.012541	Log likelihood	-15177.81	
Hannan-Quinn criter.	1.012165	Deviance	30355.61	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	1349.743	Avg. log likelihood	-0.505927	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 5 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_MAY	0.420324	0.012555	33.47781	0.0000
C	-1.202657	0.014223	-84.55760	0.0000
McFadden R-squared	0.037869	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.404119	
Akaike info criterion	1.016957	Sum squared resid	4899.050	
Schwarz criterion	1.017511	Log likelihood	-15252.35	
Hannan-Quinn criter.	1.017135	Deviance	30504.70	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	1200.650	Avg. log likelihood	-0.508412	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 3 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_APR	0.377344	0.012182	30.97514	0.0000
C	-1.196466	0.014184	-84.35047	0.0000
McFadden R-squared	0.031762	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.405972	
Akaike info criterion	1.023411	Sum squared resid	4944.067	
Schwarz criterion	1.023965	Log likelihood	-15349.16	
Hannan-Quinn criter.	1.023589	Deviance	30698.32	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	1007.031	Avg. log likelihood	-0.511639	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 3 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
BILL_AMT_SEP	-6.67E-07	1.96E-07	-3.399667	0.0007
C	-1.225177	0.016937	-72.33622	0.0000
McFadden R-squared	0.000374	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.414982	
Akaike info criterion	1.056583	Sum squared resid	5165.965	
Schwarz criterion	1.057137	Log likelihood	-15846.74	
Hannan-Quinn criter.	1.056761	Deviance	31693.49	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	11.86595	Avg. log likelihood	-0.528225	
Prob(LR statistic)	0.000572			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)  
 Sample: 1 30000  
 Included observations: 30000  
 Convergence achieved after 4 iterations  
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
BILL_AMT_AUG	-4.94E-07	2.01E-07	-2.457242	0.0140
C	-1.234754	0.016902	-73.05265	0.0000
McFadden R-squared	0.000194	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.415024	
Akaike info criterion	1.056773	Sum squared resid	5167.008	
Schwarz criterion	1.057327	Log likelihood	-15849.60	
Hannan-Quinn criter.	1.056951	Deviance	31699.20	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	6.152274	Avg. log likelihood	-0.528320	
Prob(LR statistic)	0.013124			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH  
 Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)  
 Sample: 1 30000  
 Included observations: 30000  
 Convergence achieved after 3 iterations  
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
BILL_AMT_JUL	-5.04E-07	2.07E-07	-2.436966	0.0148
C	-1.235322	0.016817	-73.45875	0.0000
McFadden R-squared	0.000191	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.415026	
Akaike info criterion	1.056776	Sum squared resid	5167.048	
Schwarz criterion	1.057330	Log likelihood	-15849.64	
Hannan-Quinn criter.	1.056954	Deviance	31699.29	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	6.065050	Avg. log likelihood	-0.528321	
Prob(LR statistic)	0.013788			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 3 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
BILL_AMT_JUN	-3.88E-07	2.21E-07	-1.758722	0.0786
C	-1.242073	0.016760	-74.10980	0.0000
McFadden R-squared	0.000099	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.415046	
Akaike info criterion	1.056874	Sum squared resid	5167.562	
Schwarz criterion	1.057428	Log likelihood	-15851.11	
Hannan-Quinn criter.	1.057052	Deviance	31702.22	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	3.135738	Avg. log likelihood	-0.528370	
Prob(LR statistic)	0.076594			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 3 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
BILL_AMT_MAY	-2.72E-07	2.32E-07	-1.170814	0.2417
C	-1.247818	0.016688	-74.77357	0.0000
McFadden R-squared	0.000044	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.415059	
Akaike info criterion	1.056932	Sum squared resid	5167.871	
Schwarz criterion	1.057486	Log likelihood	-15851.99	
Hannan-Quinn criter.	1.057110	Deviance	31703.97	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	1.383335	Avg. log likelihood	-0.528400	
Prob(LR statistic)	0.239534			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 4 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
BILL_AMT_APR	-2.20E-07	2.36E-07	-0.930446	0.3521
C	-1.250197	0.016608	-75.27734	0.0000
McFadden R-squared	0.000027	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.415063	
Akaike info criterion	1.056949	Sum squared resid	5167.964	
Schwarz criterion	1.057503	Log likelihood	-15852.24	
Hannan-Quinn criter.	1.057127	Deviance	31704.48	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	0.871892	Avg. log likelihood	-0.528408	
Prob(LR statistic)	0.350432			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 5 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_AMT_SEP	-3.64E-05	2.56E-06	-14.17551	0.0000
C	-1.103487	0.016695	-66.09657	0.0000
McFadden R-squared	0.011054	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.412122	
Akaike info criterion	1.045296	Sum squared resid	5095.002	
Schwarz criterion	1.045850	Log likelihood	-15677.43	
Hannan-Quinn criter.	1.045473	Deviance	31354.87	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	350.4865	Avg. log likelihood	-0.522581	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 6 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_AMT_AUG	-3.22E-05	2.43E-06	-13.27902	0.0000
C	-1.120476	0.016413	-68.26962	0.0000
McFadden R-squared	0.010132	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.412207	
Akaike info criterion	1.046271	Sum squared resid	5097.096	
Schwarz criterion	1.046825	Log likelihood	-15692.06	
Hannan-Quinn criter.	1.046449	Deviance	31384.13	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	321.2267	Avg. log likelihood	-0.523069	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 5 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_AMT_JUL	-2.10E-05	2.01E-06	-10.47186	0.0000
C	-1.171447	0.015630	-74.94709	0.0000
McFadden R-squared	0.005727	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.413356	
Akaike info criterion	1.050926	Sum squared resid	5125.565	
Schwarz criterion	1.051480	Log likelihood	-15761.88	
Hannan-Quinn criter.	1.051103	Deviance	31523.77	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	181.5851	Avg. log likelihood	-0.525396	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 5 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_AMT_JUN	-2.01E-05	1.99E-06	-10.09108	0.0000
C	-1.180374	0.015417	-76.56490	0.0000
McFadden R-squared	0.005153	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.413687	
Akaike info criterion	1.051533	Sum squared resid	5133.775	
Schwarz criterion	1.052087	Log likelihood	-15770.99	
Hannan-Quinn criter.	1.051711	Deviance	31541.99	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	163.3648	Avg. log likelihood	-0.525700	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

Included observations: 30000

Convergence achieved after 4 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_AMT_MAY	-1.82E-05	1.92E-06	-9.496884	0.0000
C	-1.187046	0.015357	-77.29897	0.0000
McFadden R-squared	0.004517	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.413789	
Akaike info criterion	1.052205	Sum squared resid	5136.304	
Schwarz criterion	1.052759	Log likelihood	-15781.08	
Hannan-Quinn criter.	1.052383	Deviance	31562.15	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	143.2016	Avg. log likelihood	-0.526036	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH

Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)

Sample: 1 30000

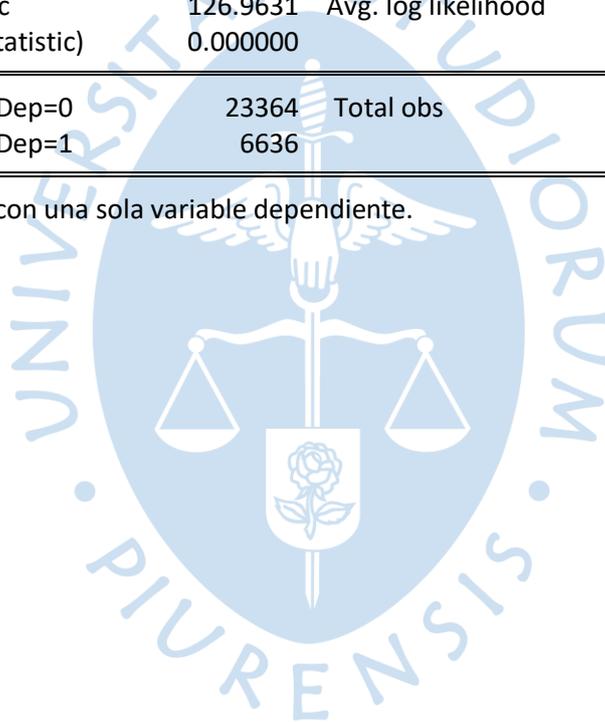
Included observations: 30000

Convergence achieved after 4 iterations

Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
PAY_AMT_APR	-1.40E-05	1.57E-06	-8.965945	0.0000
C	-1.198880	0.015031	-79.76167	0.0000
McFadden R-squared	0.004004	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.413976	
Akaike info criterion	1.052746	Sum squared resid	5140.930	
Schwarz criterion	1.053300	Log likelihood	-15789.20	
Hannan-Quinn criter.	1.052924	Deviance	31578.39	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	126.9631	Avg. log likelihood	-0.526307	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

*Nota.* Regresiones logit con una sola variable dependiente.



**Anexo 4. Resultado de estimación de modelo logit incluyendo 17 variables**

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH  
 Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)  
 Sample: 1 30000  
 Included observations: 30000  
 Convergence achieved after 7 iterations  
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
SEX	-0.122449	0.030380	-4.030647	0.0001
EDUCATION	-0.088735	0.020597	-4.308133	0.0000
MARRIAGE	-0.206379	0.029334	-7.035415	0.0000
LIMIT_BAL	-6.61E-07	1.55E-07	-4.275738	0.0000
PAY_SEP	0.581912	0.017691	32.89302	0.0000
PAY_AUG	0.077496	0.020161	3.843760	0.0001
PAY_JUL	0.076688	0.022529	3.403942	0.0007
PAY_JUN	0.027235	0.024925	1.092689	0.2745
PAY_MAY	0.036052	0.026758	1.347345	0.1779
PAY_APR	0.016095	0.021842	0.736875	0.4612
BILL_AMT_SEP	-1.66E-06	2.69E-07	-6.150991	0.0000
PAY_AMT_SEP	-1.08E-05	2.06E-06	-5.254640	0.0000
PAY_AMT_AUG	-8.45E-06	1.85E-06	-4.561485	0.0000
PAY_AMT_JUL	-2.85E-06	1.49E-06	-1.921760	0.0546
PAY_AMT_JUN	-3.71E-06	1.55E-06	-2.395512	0.0166
PAY_AMT_MAY	-3.24E-06	1.48E-06	-2.186376	0.0288
PAY_AMT_APR	-2.42E-06	1.27E-06	-1.906160	0.0566
C	-0.359454	0.088466	-4.063184	0.0000
McFadden R-squared	0.119579	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.380764	
Akaike info criterion	0.931669	Sum squared resid	4346.825	
Schwarz criterion	0.936654	Log likelihood	-13957.03	
Hannan-Quinn criter.	0.933268	Deviance	27914.06	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	3791.295	Avg. log likelihood	-0.465234	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

*Nota.* Regresión logit con 17 variables cuyo signo esperado es el mismo que el signo estimado en la regresión de una sola variable.

**Anexo 5. Resultado de correlación entre variables**

	SEX	EDUCATION	MARRIAGE	LIMIT_BAL	PAY_SEP
SEX	1.00000	0.01423	-0.03139	0.02476	-0.05764
EDUCATION	0.01423	1.00000	-0.14346	-0.21916	0.10536
MARRIAGE	-0.03139	-0.14346	1.00000	-0.10814	0.01992
LIMIT_BAL	0.02476	-0.21916	-0.10814	1.00000	-0.27121
PAY_SEP	-0.05764	0.10536	0.01992	-0.27121	1.00000
PAY_AUG	-0.07077	0.12157	0.02420	-0.29638	0.67216
PAY_JUL	-0.06610	0.11402	0.03269	-0.28612	0.57425
BILL_AMT_SEP	-0.03364	0.02358	-0.02347	0.28543	0.18707
PAY_AMT_SEP	-0.00024	-0.03746	-0.00598	0.19524	-0.07927
PAY_AMT_AUG	-0.00139	-0.03004	-0.00809	0.17841	-0.07010

	PAY_AUG	PAY_JUL	BILL_AMT_SEP	PAY_AMT_SEP	PAY_AMT_AUG
SEX	-0.07077	-0.06610	-0.03364	-0.00024	-0.00139
EDUCATION	0.12157	0.11402	0.02358	-0.03746	-0.03004
MARRIAGE	0.02420	0.03269	-0.02347	-0.00598	-0.00809
LIMIT_BAL	-0.29638	-0.28612	0.28543	0.19524	0.17841
PAY_SEP	0.67216	0.57425	0.18707	-0.07927	-0.07010
PAY_AUG	1.00000	0.76655	0.23489	-0.08070	-0.05899
PAY_JUL	0.76655	1.00000	0.20847	0.00129	-0.06679
BILL_AMT_SEP	0.23489	0.20847	1.00000	0.14028	0.09936
PAY_AMT_SEP	-0.08070	0.00129	0.14028	1.00000	0.28558
PAY_AMT_AUG	-0.05899	-0.06679	0.09936	0.28558	1.00000

*Nota.* Se evalúa el nivel de correlación entre las variables, para ello se obtiene la matriz de correlación.

**Anexo 6. Resultado de estimación de modelo logit incluyendo las variables definidas**

Dependent Variable: DEFAULT\_PAYMENT\_NEXT\_MONTH  
 Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)  
 Sample: 1 30000  
 Included observations: 30000  
 Convergence achieved after 6 iterations  
 Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
SEX	-0.175950	0.028743	-6.121579	0.0000
EDUCATION	-0.061137	0.019121	-3.197433	0.0014
MARRIAGE	-0.216491	0.027788	-7.790721	0.0000
LIMIT_BAL	-3.18E-06	1.43E-07	-22.26631	0.0000
BILL_AMT_SEP	2.00E-06	2.34E-07	8.555788	0.0000
PAY_AMT_SEP	-2.21E-05	2.57E-06	-8.608968	0.0000
PAY_AMT_AUG	-1.67E-05	2.23E-06	-7.477879	0.0000
C	0.021186	0.083173	0.254726	0.7989
McFadden R-squared	0.036530	Mean dependent var	0.221200	
S.D. dependent var	0.415062	S.E. of regression	0.406685	
Akaike info criterion	1.018772	Sum squared resid	4960.451	
Schwarz criterion	1.020987	Log likelihood	-15273.57	
Hannan-Quinn criter.	1.019483	Deviance	30547.15	
Restr. deviance	31705.35	Restr. log likelihood	-15852.68	
LR statistic	1158.207	Avg. log likelihood	-0.509119	
Prob(LR statistic)	0.000000			
Obs with Dep=0	23364	Total obs	30000	
Obs with Dep=1	6636			

*Nota.* Regresión logit con variables significativas que se incluyen en el modelo final.