



---

# **ESTIMADOR DE INGRESOS**

## **MANUAL DEL USUARIO**

## Tabla de contenido

<b>HISTÓRICO DE CAMBIOS.....</b>	<b>3</b>
<b>1. INTRODUCCIÓN.....</b>	<b>4</b>
<b>SECCIÓN 1: ESTIMACIÓN .....</b>	<b>5</b>
1.1. INTRODUCCIÓN .....	5
1.2. TIPOS DE ESTIMACIÓN .....	5
1.3. PROCESO DE DESARROLLO DEL ESTIMADOR DE INGRESOS.....	6
1.3.1. ETAPA 1 – OBJETIVOS DE LA ESTIMACIÓN.....	6
1.3.2. ETAPA 2 – RECOPIACIÓN DE DATOS.....	7
1.3.3. ETAPA 3 – ANÁLISIS INICIAL.....	8
1.3.4. ETAPA 4 – DESARROLLO DEL MODELO .....	8
1.3.5. ETAPA 5 – MODELACIÓN .....	9
1.3.6. ETAPA 6 – VALIDACIÓN DEL MODELO .....	10
2. RESUMEN DEL CAPÍTULO .....	10
<b>SECCIÓN 2: ESTIMADOR DE INGRESOS .....</b>	<b>10</b>
2.1. MODELO PARA CLIENTES QUE TIENEN HISTORIAL CREDITICIO.....	11
2.1.1. MUESTRAS.....	11
2.1.2. INGRESO POR PERIODOS .....	11
2.1.3. FRACCIONAMIENTO DE LA POBLACIÓN.....	14
2.1.4. ESCALAMIENTO DE LOS RANGOS DE ANÁLISIS.....	14
2.2. MODELO DE ESTIMACIÓN POR RANGOS PUNTUAL .....	15
2.2.1. INTERVALOS DE PREDICCIÓN Y CÓDIGOS DE RAZÓN .....	15
2.2.2. VALIDACIÓN DEL MODELO .....	16
2.3. MODELO PARA CLIENTES QUE NO TIENEN HISTORIAL CREDITICIO (NOREF) .....	16
2.3.1. ESTIMADOR DEL INGRESO MENSUAL PROMEDIO POR CÓDIGO POSTAL.....	17
2.4. MODELO DE ENSAMBLE.....	20
2.4.1. FICHA TÉCNICA .....	21
<b>SECCIÓN 3. EJEMPLOS DE USO .....</b>	<b>22</b>
1.1. UTILIZACIÓN DEL ESTIMADOR DE INGRESOS.....	22
1.2. COLOCACIÓN DE PRODUCTOS CREDITICIOS PARA OTORGANTES .....	22
1.3. ASIGNACIÓN DE LÍMITES DE CRÉDITO .....	23
1.4. COBRANZA .....	24
1.5. RESUMEN .....	24
<b>SECCIÓN 4: APÉNDICES .....</b>	<b>26</b>
APÉNDICE 1: PREGUNTAS FRECUENTES.....	26
APÉNDICE 2: SALIDA DEL REPORTE DEL ESTIMADOR DE INGRESOS .....	29
APÉNDICE 3: EXCLUSIONES DEL SCORE.....	29
APÉNDICE 4: CÓDIGOS DE RAZÓN .....	30
APÉNDICE 5: VALIDACIÓN DEL ESTIMADOR DE INGRESOS .....	30
APÉNDICE 6: GLOSARIO .....	30
<b>CONTACTOS DE BURÓ DE CRÉDITO.....</b>	<b>35</b>

## Histórico de Cambios

Historial de Cambio			
Fecha Liberación	Tipo de Cambio	Resumen del Cambio con respecto a la Información Anterior	Página
Mayo 2018		Creación de Documento	Todas

## **1. Introducción**

El ingreso en la población mexicana ha sido de interés para las instituciones de crédito en la planificación y diversificación de la economía nacional. Para el sector financiero se ha vuelto de gran interés tener un ingreso estimado para cada uno de sus prospectos y clientes existentes, siendo esta estimación una referencia para el análisis de sus portafolios y la posibilidad de nuevas estrategias del manejo de carteras.

En México, el ingreso ha sido estimado a partir de encuestas muestrales de ingresos y egresos de hogares (ENIGH) con resultados a nivel estatal y hasta el momento ha sido un reto el poder estimar el ingreso a nivel individuo con información pública, e incluso utilizando información de crédito y de transacciones. En este sentido, la estimación del ingreso de un individuo deberá realizarse utilizando técnicas y herramientas robustas de modelación que se acerquen lo más posible a la realidad.

El ingreso estimado de un individuo debe considerar todas las fuentes disponibles ya sea del sector privado o del público para maximizar el desempeño de una estimación. Para ello Buró de Crédito cuenta con los resultados e información obtenida de encuestas nacionales, condiciones sociodemográficas del domicilio y características crediticias del individuo. Utilizando esta información con las técnicas y herramientas desarrolladas en Buró de Crédito se construye el Estimador de Ingresos.

Utilizando el Estimador de Ingresos los otorgantes podrán tener una referencia del ingreso mensual de cada uno de los clientes para poder ordenarlos y tener una referencia estadística del ingreso real. El uso del Estimador de Ingresos aumenta el número de atributos disponibles en la evaluación del otorgamiento de un producto crediticio.

El estimador se desarrolló considerando la base de datos de Buró de Crédito e información socio demográfica de la población mexicana. Utilizando modelos estadísticos y computacionales se conjuntaron para generar una estimación puntual del ingreso en miles de pesos; y un intervalo de predicción con un nivel de confianza que oscila entre el 80% y 90%, al ingreso real.

Para exponer la información, métodos y validación del modelo, los objetivos de este documento se dividen de la siguiente manera:

1. Proporcionar la metodología para el desarrollo del Estimador de Ingresos incluyendo los detalles técnicos de las herramientas estadísticas utilizadas.
2. Mostrar los resultados obtenidos con el Estimador de Ingresos.
3. Ejemplo de perfilamiento a través de la segmentación proporcionada por el Estimador de Ingresos.

En cumplimiento de lo anterior el documento posee tres secciones en su estructura:

- **Sección 1:** Estimación de indicadores

Se provee de una breve y concisa explicación sobre los modelos de estimación. El Estimador de Ingresos no es un score de crédito, por lo tanto, es importante familiarizarse con el concepto de “Estimación de un indicador” en general.

- **Sección 2:** Estimador de Ingresos

En esta sección se describen detalles sobre el desarrollo del Estimador de Ingresos y sus características. Se explican los detalles que se utilizaron en el desarrollo, la validación del modelo; además de los atributos utilizados en el modelo y finalmente se presentan las medidas de desempeño y validación.

- **Sección 3:** Estimador de Ingresos – Uso e Integración en los procesos de otorgamiento, administración y cobranza de créditos.

Se presentan ejemplos de su integración y uso del Estimador de Ingresos para la gestión de cartera a lo largo del ciclo de vida del crédito.

## **Sección 1: Estimación**

### **1.1. Introducción**

Estimación es un término estadístico usualmente utilizado cuando, mediante el estudio de una muestra de la población, se desea generalizar los resultados a través de conclusiones aplicables y consistentes sobre toda la población. En general, el trabajo de estimación es la aplicación de una serie de metodologías estadísticas y/o matemáticas utilizadas para la aproximación de una variable objetivo, el cual puede ser una variable discreta o continua.

La variable objetivo para este modelo es el ingreso, el cual se desea describir, explicar y estimar mediante características propias de los individuos, su situación sociodemográfica y sus hábitos crediticios. El análisis de estas características se realiza con herramientas estadísticas para encontrar las relaciones causales sobre el ingreso individual que es conocido como minería de datos.

Teniendo este marco de modelación propia de la minería de datos, se aplican modelos estadísticos supervisados paramétricos para encontrar las relaciones más estables y correlacionadas al ingreso individual. Además, se aplican técnicas avanzadas de modelación como son los ensambles de modelos y optimizaciones algorítmicas para mejorar estos patrones encontrados, y así, minimizar el error de estimación. En particular para este modelo, se creó una nueva metodología de ensamble de modelos para mejorar aun más la estimación de este indicador para el ingreso.

Las técnicas estadísticas usadas durante todo el proceso de modelación corresponde a diferentes regresiones multivariadas, las cuales han demostrado un poder de estimación y validación de modelos bastante alto en problemas con dispersiones altas en la variable objetivo, como en el caso de los ingresos individuales.

### **1.2. Tipos de Estimación**

El trabajo de estimación está dirigido según el tipo de variable objetivo de interés. Esta variable se puede transformar para aumentar el número de técnicas a aplicar para su validación y estimación; y así, tener varias perspectivas de la posible solución. Con estas perspectivas se crea un solo modelo ensamblando los mejores aportes de cada una de estas perspectivas de solución.

Para el ingreso se utilizaron dos perspectivas; la puntual, que se centró en la estimación de una función lineal para el ingreso individual, y la perspectiva por intervalos, que transformó al ingreso individual en grupos de más estables y con particularidades poblacionales bien separadas que ayudaron a la estimación del ingreso puntual.

Cada una de estas etapas de estimación se describe a continuación:

- **Puntual:** Corresponde a la estimación de una ecuación que depende de los valores obtenidos de una muestra, para realizar cálculos de valores exactos. Lo que se pretende obtener es el valor de un elemento u objeto de interés.
- **Por intervalos:** El intervalo de confianza es un factor que nos permite afirmar con cierta probabilidad que el verdadero valor de un elemento está dentro de un rango de posibles valores. Este intervalo contiene al elemento estimado con alguna certeza o nivel de confianza.

Para el caso particular de la estimación por intervalos, si admitimos un error mayor, esto es, aumentar el tamaño del intervalo de confianza, tenemos también una mayor probabilidad de éxito en nuestra estimación. Este efecto en la estimación por intervalos fue intensivamente trabajado para minimizarlo y crear grupos de ingresos más estables.

### 1.3. Proceso de desarrollo del Estimador de Ingresos

El proceso de desarrollo de un estimador implica un conjunto de etapas que combinan los tipos de estimación de valores descritos anteriormente, que derivan en un algoritmo robusto y con gran poder predictivo.

#### 1.3.1. Etapa 1 – Objetivos de la estimación

Según los datos de INEGI en su último estudio de ingresos y gastos del 2016 (ENIG 2016) la distribución trimestral de ingresos por hogares se presenta en la Tabla 2.1.

Tabla 2.1. Ingreso corriente promedio trimestral por hogar según deciles. INEGI, ENIGH 2016

Déciles de hogares	ENIGH 2016	Límite Inferior/Límite superior
<b>Nacional</b>	46,521	44,353
		48,688
<b>1</b>	8,166	8,034
		8,299
<b>2</b>	14,206	14,043
		14,369
<b>3</b>	18,918	18,736
		19,100
<b>4</b>	23,556	23,343
		23,768
<b>5</b>	28,812	28,556
		29,069
<b>6</b>	34,837	34,525
		35,149
<b>7</b>	42,431	42,025
		42,837
<b>8</b>	53,383	52,788
		53,979
<b>9</b>	72,041	71,070
		73,011

Déciles de hogares	ENIGH 2016	Límite Inferior/Límite superior
10	168,855	149,046 188,665

Utilizando estimaciones sobre las personas económicamente activas en hogares se infirió el ingreso promedio de las personas según el INEGI, Esta información se contrastó con lo observado en la información de ingresos de nómina de algunos usuarios y se observó lo siguiente:

Tabla 2.2. Distribución por deciles del ingreso reportado en Buró de Crédito y la ENIGH

Déciles de hogares	Buró de Crédito 2017 (Miles de pesos)			ENIGH 2016 (Miles de pesos)		
	Mínimo	Media	Máximo	Límite Inferior *	Media	Límite Superior*
1	3	5	6	3	3	3
2	7	8	8	5	5	5
3	9	10	10	6	6	6
4	11	12	12	8	8	8
5	13	14	15	10	10	10
6	16	17	18	12	12	12
7	19	20	22	14	14	14
8	23	25	29	18	18	18
9	30	36	48	24	24	24
10	49	94	500	50	56	63

\* Límite inferior y Límite superior con un 90% de confianza

Se observó que las muestras de los usuarios tienen rangos superiores a los que reporta el INEGI, esto se debe a que el ingreso utilizado para este modelo proviene de personas que perciben ingresos vía nómina sin incluir los ingresos por actividades informales como ventas por catalogo, comercios particulares, etc. Entonces, para poder crear un modelo aplicable a toda la población mexicana se alinea la base de ingresos a partir de una transformación, para tener un rango de ingresos aplicable a la mayoría poblacional.

De la alineación anterior, se definió y limitó el monto de ingreso a estimar para abarcar al mayor rango de ingresos posibles en la población mexicana. El establecimiento de este límite sobre el ingreso a estimar, se da ya que el ingreso es un objetivo con mucha variabilidad por su alta dispersión puntual que es cambiante en el tiempo. Por todo lo anterior, para este modelo se trabajó en la estimación puntual del ingreso (sin impuestos aplicados) mensual de hasta \$60,000 MXN durante un periodo de 12 meses. Para completar la estimación del ingreso, se crea un índice para especificar e identificar ingresos superiores a este umbral.

Para ambos casos de estimación, puntual y mediante el índice, se crean intervalos de predicción en los cuales el ingreso real está considerado para montos de hasta \$300,000 MXN mensuales.

### 1.3.2. Etapa 2 – Recopilación de datos

Ya definido el objeto de estudio, es necesario explorar las fuentes de información que pueden aportar detalles en la explicación del ingreso, buscando incluir el mayor número de características de la población estudiada. La búsqueda de información no sólo se refiere a la disponible en las bases de Buro de Crédito, sino incluyó explorar varias fuentes de información relacionada y disponible en fuentes externas.

La información de la población analizada por grupo de características se tiene de la siguiente manera:

- Detalles socio-demográficos: Información del reporte de crédito y datos del INEGI
- Detalles de condiciones demográficas: información de encuestas, ENIGH, ENSANUT, CONAPO, PROSPERA, SEDESOL, entre otros.
- Detalles de estructura de crédito: Información del reportes de crédito y en específico de las cuentas de crédito
- Hábitos crediticios: Información del reporte de crédito con respecto a utilidades, compras, pagos y consultas.

Para toda la información anterior se validó su consistencia y se limpió para poder utilizarse en el proceso de modelación.

### 1.3.3. Etapa 3 – Análisis inicial

Se realizaron diferentes análisis estadísticos con el objetivo de desarrollar el modelo preliminar:

#### ■ Análisis de clasificación o segmentación

La investigación ha demostrado que segmentar una población por ciertas características y desarrollar clasificaciones adaptadas a estas subpoblaciones, mejora significativamente la capacidad de predicción del modelo final. La técnica de usar subpoblaciones tiene como objetivo general optimizar y mejorar la búsqueda de características que determinen de mejor manera el ingreso puntual para estos grupos.

El crear segmentación para la estimación aumenta el grado de complejidad del modelo ya que involucra un proceso de alienación y ensamble para poder obtener la estimación final.

#### ■ Análisis descriptivo

El análisis descriptivo favorece la toma de decisiones ya que ayuda a interpretar de mejor manera las variables/características de la población que define el ingreso. En este análisis se comienzan a encontrar patrones sobre el ingreso que nos van indicando cuál(es) variable(s) utilizar para su estimación. Además, durante esta etapa se asegura la correcta utilización de la información para evitar utilizar información irrelevante o que no tenga sentido para el objetivo del modelo.

#### ■ Análisis de regresión para grupos de datos

Durante el análisis de las variables para la predicción se aplicaron enfoques de análisis univariado y multivariado simultáneamente, esto debido al trabajo en paralelo para los subgrupos poblacionales creados para la estimación del ingreso. En cada uno de los grupos se realizaron análisis univariado y multivariados para seleccionar/crear las principales características útiles en la estimación del ingreso; pero además, se tuvo que realizar los análisis considerando los efectos encontrados en otros grupos poblacionales para determinar cuáles tenían un efecto real y diferenciado que mejoren la estimación puntual del ingreso.

El proceso anterior dio como resultado una serie de características que son propias de los subgrupos poblacionales de ingreso que mejora la estimación dentro del subgrupo, pero además ayuda a la estimación dentro de otros subgrupos poblacionales.

### 1.3.4. Etapa 4 – Desarrollo del modelo



En este proceso de desarrollo del modelo se consideran los subgrupos poblacionales y características creadas durante la etapa anterior. Los modelos corresponden a diferentes regresiones multivariadas de las cuales se hacen selecciones para obtener la mejor función que estime el ingreso puntual y por rangos. Este trabajo se realizó en tres fases descritas a continuación:

- **Estimación**

Para cada subgrupo poblacional se tomaron las variables y características optimizadas durante el proceso anterior y se utilizaron para ajustar varios modelos de regresión. La modelación se diseñó considerando el monto de los ingresos subpoblacionales y las probabilidades construidas para la determinación de dichos subgrupos. Con estas consideraciones se realizó un proceso iterativo de ajuste de parámetros y selección de variables para minimizar el error de la estimación

- **Ponderación**

Una vez que se han determinado las características y los atributos, es necesario ponderarlos sobre todos los subgrupos. La ponderación es el proceso de selección de las combinaciones más predictivas de las características dentro de alguna subpoblación, de tal manera que se maximiza la capacidad de hallar un buen estimador. Como ejemplo, una característica podría ser el número de obligaciones activas, con valores de atributo potenciales de cero, uno, dos a cuatro y más de cuatro. A cada uno de estos atributos se le asignó un cierto número de score para maximizar su capacidad de predicción. Generalmente, los ponderadores se originan a través de una serie de etapas iterativas para mejorar la capacidad de predicción buscando un equilibrio entre el error y los coeficientes.

- **Alineación**

Teniendo los modelos por subpoblaciones se trabajó en la alineación de las predicciones, esto debido a que la estimación final no debe de depender de los subgrupos creados por rangos de ingresos. Entonces, se crearon reglas de selección de rangos con las características de todos los subgrupos a la vez para tener un sistema de predicción solamente utilizando las características de la población analizada.

Además durante este proceso se crearon los intervalos de predicción para cada estimación puntual para los ingresos de hasta \$60,000 MXN mensuales.

### **1.3.5. Etapa 5 – Modelación**

Los resultados del modelo se revisan mostrando el cumplimiento y supuestos matemáticos necesarios. Los algoritmos del modelo y demás elementos de evaluación también son revisados para asegurar la aplicabilidad correcta del modelo en la toma de decisiones.

- **Modelo Final**

Después de concluir un modelo, se genera el sustento metodológico con las especificaciones técnicas del modelo final y se completa la documentación del proceso de desarrollo. Se validan los algoritmos de las predicciones y se comprueba con información adicional a la utilizada para el desarrollo. En esta validación se realiza un proceso de reiterativo para obtener un modelo aplicable y consistente en el tiempo.

Finalmente se crea una forma de reporte del resultado que sea adecuada para que los otorgantes reciban, procesen y utilicen la estimación directa en sus sistemas internos. Esto ya que se reporta el ingreso en

formato de Score donde los valores del estimador van desde 1 hasta 61; indicando el ingreso en miles de pesos mensuales hasta el valor 60 y utilizando el valor 61 como el indicador para ingresos superiores.

### **1.3.6. Etapa 6 – Validación del Modelo**

Después de desarrollar el Estimador de Ingresos, se deben llevar a cabo varias actividades para monitorear y calibrar el estimador. Las actividades de monitoreo son importantes para garantizar la estabilidad y eficacia del estimador, ya que las características poblacionales pueden cambiar.

El monitoreo para el Estimador de Ingresos se realiza extrayendo una muestra aleatoria de expedientes de la base de datos de Buró de Crédito y comparándola con la muestra utilizada para construir el modelo. Si la distribución del ingreso mostrará cambios significativos, se estudiará la distribución de variables incluidas en el modelo del Estimador de Ingresos con la finalidad de entender y ajustar cualquier cambio en su incremento o decremento a nivel poblacional.

Con respecto a la información externa utilizada en este modelo, para su validación se planea de forma periódica la evaluación y desempeño de los parámetros.

## **2. Resumen del capítulo**

Un estimador funciona como una referencia del ingreso real, se construyó utilizando diferentes fuentes de información y técnicas avanzadas de modelación para obtener una estimación puntual del ingreso y un intervalo de predicción. Durante el proceso de modelación se consideraron subpoblaciones que tenían características en común pero que además se podían diferenciar de otros subgrupos, haciendo así segmentos poblacionales estables para su modelación. Para tener un modelo aplicable a toda la población utilizando solamente las características mismas de la población, se trabajó en la alineación de todos los modelos y un proceso de ensamblado que resulta en la estimación puntual de ingresos que van de \$1,000 a \$60,000 MXN mensuales y un indicador para identificar ingresos mayores a este umbral.

Finalmente se diseñó el resultado de esta estimación para adecuarlo a los reportes de Buró de Crédito de modelos de valoraciones numéricas para la utilización de los usuarios en los sistemas diseñados para Score. Así, se reporta la estimación como un número con valores de 1 a 60 para los ingresos en Miles de pesos estimados puntualmente y un valor 61 para indicar los ingresos mayores. Estas estimaciones están acompañadas de un intervalo de predicción donde el ingreso real estaría en el 80% - 90% de los casos.

## **Sección 2: Estimador de Ingresos**

Buró de Crédito con el Estimador de Ingresos ofrece una herramienta de apoyo para el manejo de la cartera en todo el ciclo de crédito, cuya principal virtud será tener un clasificador de ingreso mensual para todos los clientes de cada institución.

La Estimación del Ingreso mensual tiene dos enfoques que dependen de la información crediticia disponible del prospecto. El primer enfoque corresponde **clientes que tienen historial de crédito**, donde se consideran la información crediticia de las bases de Buró de Crédito con los siguientes resultados:

- Una estimación del ingreso mensual y
- Un intervalo de predicción para el mismo.

El segundo enfoque es para aquellos **clientes que no tienen historial crediticio (NOREF)** que corresponde a una estimación con información de fuentes públicas externas a Buró de Crédito; como encuestas nacionales y censos. Los resultados de este modelo son estimaciones del ingreso promedio mensual de un individuo cuyo domicilio se encuentra en cierto código postal y considera las características particulares de los prospectos para tener una estimación puntual. El Estimador de Ingresos ofrecerá también un intervalo de predicción para estos casos.

## **2.1. Modelo para clientes que tienen historial crediticio**

Utilizando las variables del historial crediticio de un individuo se construyó un conjunto de variables explicativas del comportamiento crediticio, se analizaron aquellas variables que describieran mejor la distribución del ingreso a través de un modelo estadístico. El ingreso mensual de un individuo fue estimado con un valor en miles de pesos utilizando información de ingreso comprobado, el cual se alinee a una distribución poblacional más estable y aplicable a la población mexicana.

### **2.1.1. Muestras**

Para el desarrollo del modelo se utilizó una muestra aleatoria simple de tamaño representativo de los consumidores que se encuentran la base de Buró de Crédito, representativa sobre las 32 entidades federativas del país y con ingreso bruto comprobado.

La muestra de desarrollo incluye información general del individuo como sexo, edad, domicilio, información del ingreso comprobado para mexicanos no difuntos con historial crediticio reciente, medidos durante el 2016 y 2017 como sigue:

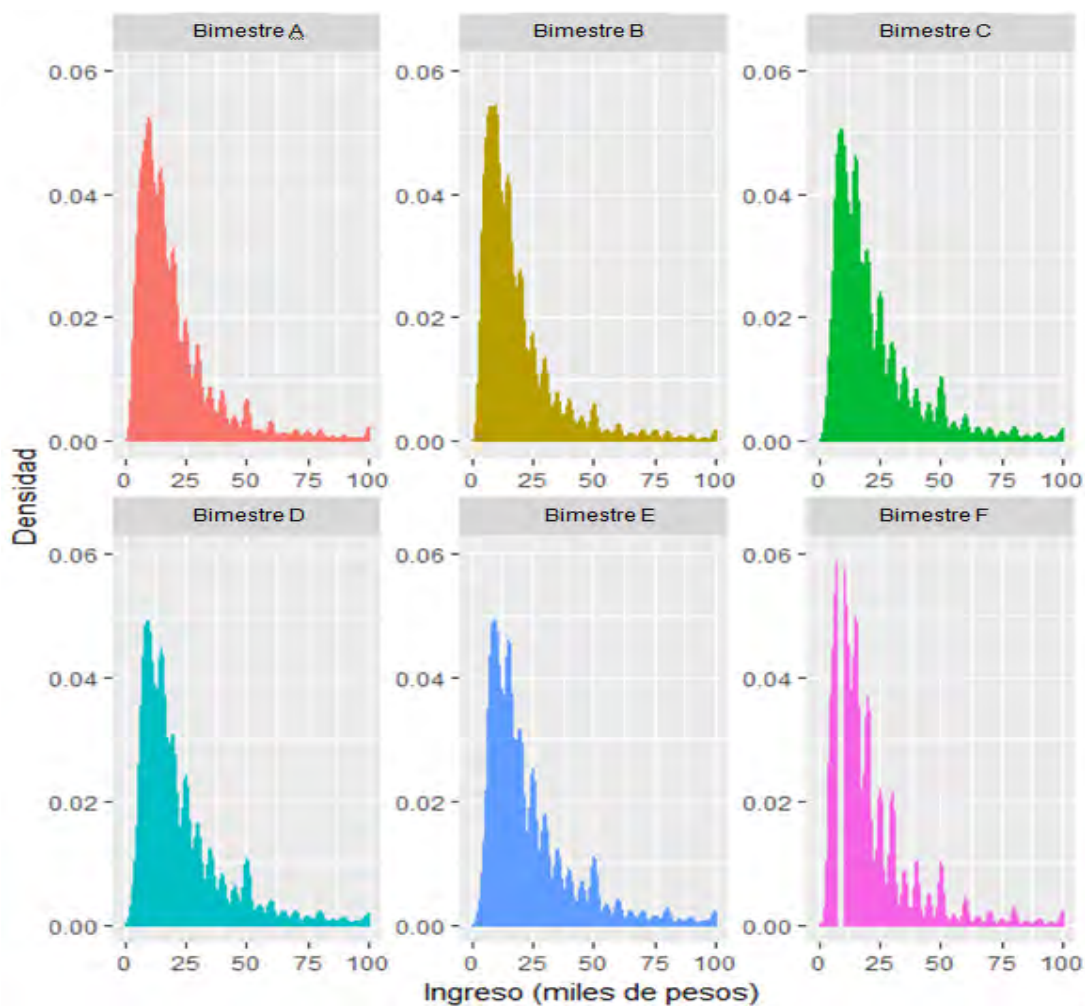
- julio, septiembre y noviembre 2016
- enero, marzo, mayo 2017

### **2.1.2. Ingreso por periodos**

El Estimador de Ingresos es un producto que se realizó con información que corresponde a un año con periodos bimestrales. En el análisis de esta información fue necesario estabilizar los ingresos reportados y minimizar la temporalidad que puede intervenir en los registros, ya que el ingreso puede verse influenciado por ingresos extras como bonos, aguinaldos y prestaciones laborales.

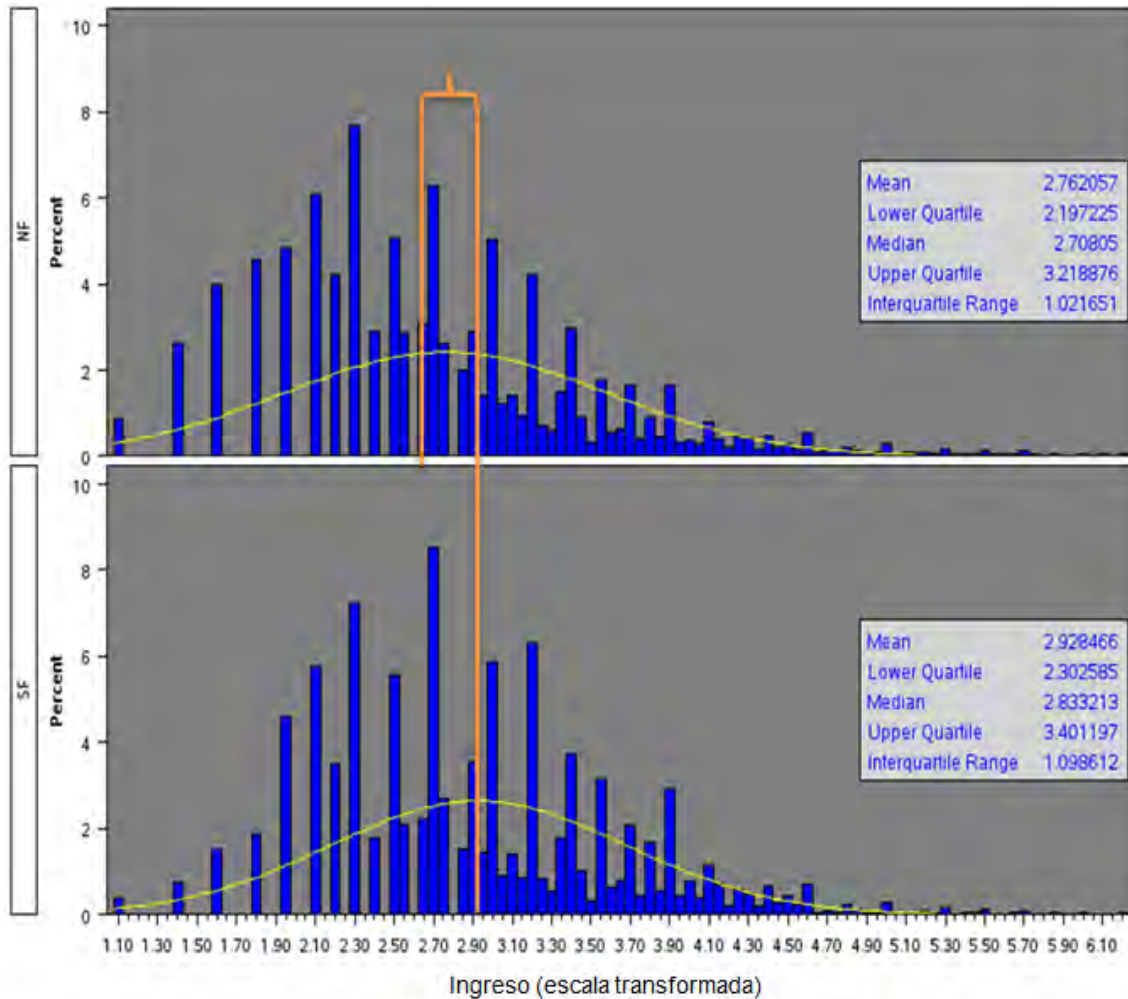
Entonces, el ingreso mensual fue analizado estadísticamente en cada uno de los periodos. .Por lo que se considera necesario realizar un análisis de series tiempo para conocer la tendencia temporal. Gráficamente la distribución por periodos se muestra a continuación:

Figura 3.1. Distribución del Ingreso Mensual en miles de pesos por periodo



Para tener una mejor perspectiva de los cambios en el ingreso se trabajó con una transformación, modificando la escala. Utilizando esta información se divide a la muestra en dos grupos de ingresos que dependen del periodo, como se muestra a continuación:

Figura 3.2. Distribución del Ingreso Mensual en miles de pesos por grupos con mayor y menor ingreso.



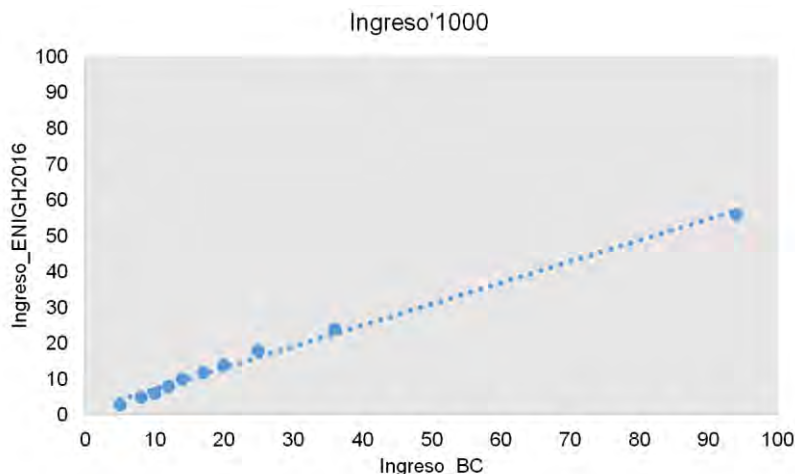
Las distribuciones con la transformación favorecen que el ingreso sea independiente del tiempo además de incrementar la cercanía entre periodos. De tal forma que la transformación reduce la variabilidad entre periodos, logrando homogeneizar los periodos. Ya que el objetivo es encontrar una estimador del ingreso de cualquier individuo debemos buscar que la muestra tenga representatividad en Buró de Crédito y a nivel nacional para México.

Existen diversas encuestas que han estimado el ingreso de la población mexicana, la más reciente con estimación a nivel nacional en 2016 la realizó el INEGI: **ENIGH2016**, misma que es posible utilizar como punto de referencia y así evaluar si la base de ingresos de Buró de Crédito es representativa de la población mexicana.

La ENIGH2016 presenta el ingreso corriente promedio por trimestre, donde se detalla el ingreso por deciles y sus correspondientes intervalos de confianza de un hogar. La ENIGH estimó un promedio de 3.67 integrantes por hogar y de ellos 1.69 integrantes ocupados, lo que implica que para cada hogar sólo 1.69 de los integrantes aportan ingresos al hogar. Para efectos de nuestro análisis consideraremos que el Ingreso trimestral reportado por la ENIGH corresponde al ingreso trimestral de sólo un individuo pues es bien sabido que existe subregistro por diversas cuestiones socioculturales.

Considerando los registros en la ENIGH, se encontró una transformación de los datos del ingreso en Buró de crédito para que su distribución fuera muy similar a la reportada en la ENIGH2016. Para los datos del ingreso comprobado desde las bases de Buró de Crédito calculamos los deciles de la totalidad de la muestra, en la Figura 3.3 la distribución de los percentiles del ingreso a nivel nacional para la base de Buró de Crédito se compara con el símil de la ENIGH2016.

Figura 3.3. Distribución del Ingreso en miles de pesos



Se aplica una transformación de tal forma que se alinean los Ingresos reportados en Buró de Crédito con lo de la ENIGH2016 de la Figura 3, haciendo que la base de Buró de Crédito sea también representativa a nivel nacional. Se sigue, que la ecuación que dará origen al ingreso alineado o transformado incluye todo este trabajo de alineación para hacerla representativa y estable para predecir.

### 2.1.3 Fraccionamiento de la población

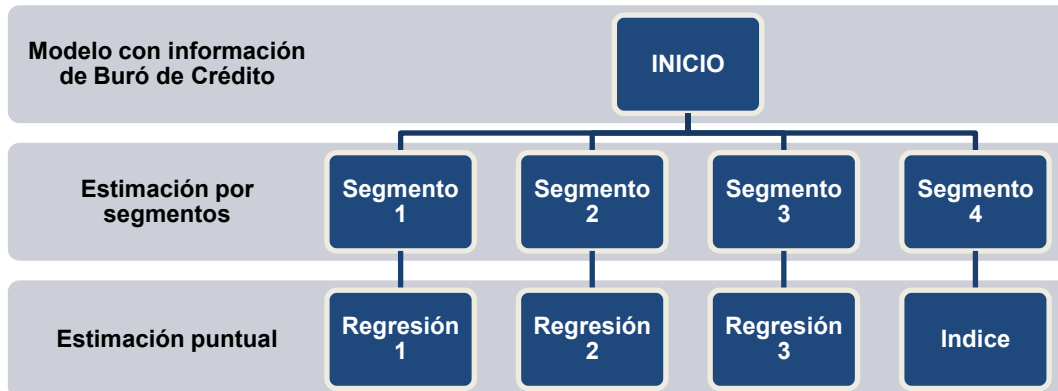
Se analizaron los datos con el ingreso alineado identificando diferentes grupos de población homogéneos al interior y que requirieron sus propios análisis con el fin de obtener una estimación óptima. El resultado de este análisis, se denomina fraccionamiento, mismo que nos dirige a una estimación puntual que incluye diferentes modelos.

El uso de modelos dirigidos a alguna población en particular dentro de un grupo poblacional produce un modelo más robusto, es por ello, que el Estimador de Ingresos ofrecerá un valor puntual de cada prospecto que considera el grupo poblacional donde se encuentra ubicado.

### 2.1.4. Escalamiento de los rangos de análisis

El análisis del Ingreso transformado tiene un enfoque en dos etapas, en la etapa primer se buscan grupos de individuos que tengan características similares y para la muestra de desarrollo se han generado cuatro rangos de análisis. La segunda etapa es buscar una regresión al interior de cada uno de los grupos. Enseguida se muestra el flujo de análisis de la modelación por rangos es la siguiente:

Figura 3.4. Flujo del Modelo para clientes que tienen historial crediticio



Para cada uno de los rangos se tiene internamente un modelo de regresión, entonces tendremos una estimación para cada individuo, pero esa estimación puede calcularse desde regresiones diferentes.

## 2.2. Modelo de estimación por rangos puntual

El comportamiento del ingreso mensual de un individuo de la población mexicana debe considerar la variabilidad presente entre los individuos dentro del rango de ingreso mensual al que pertenecen. Se evaluaron variables calculadas con información de los segmentos de créditos y consultas. El tipo de variables analizadas corresponde principalmente al domicilio, tipos de crédito de los clientes, historial crediticio y consultas.

El modelo propuesto se usará primordialmente para describir las relaciones entre el ingreso y las características de los datos que están agrupados que aporten más al modelo, según uno o más factores de clasificación. El mecanismo para calcular el ingreso para cualquier expediente será:

1. Se ejecutan los cuatro modelos de rangos y en cada uno de ellos se tendrá una probabilidad.
2. Se elige aquel rango que resulte más adecuado para el Estimador de Ingresos al calcularse la regresión puntual del ingreso.
3. Finalmente se ejecuta la regresión correspondiente al rango elegido en el punto 2.

Con este proceso de estimación se minimiza el error de estimación puntual.

### 2.2.1. Intervalos de predicción y códigos de razón

El modelo ofrecerá una estimación del ingreso transformado y además se generará un intervalo de predicción donde se encuentra el verdadero valor del ingreso transformado. En nuestro caso se considera un intervalo que oscila entre el 80% y el 90% dependiendo del monto de ingreso estimado. Es decir, a mayor ingreso a estimar será menor la confianza del intervalo. Pero este será suficientemente informativo para cualquier estimación del ingreso puntual y además tendrá un sentido de ordenamiento que se cumplirá en la mayoría de los casos reales.



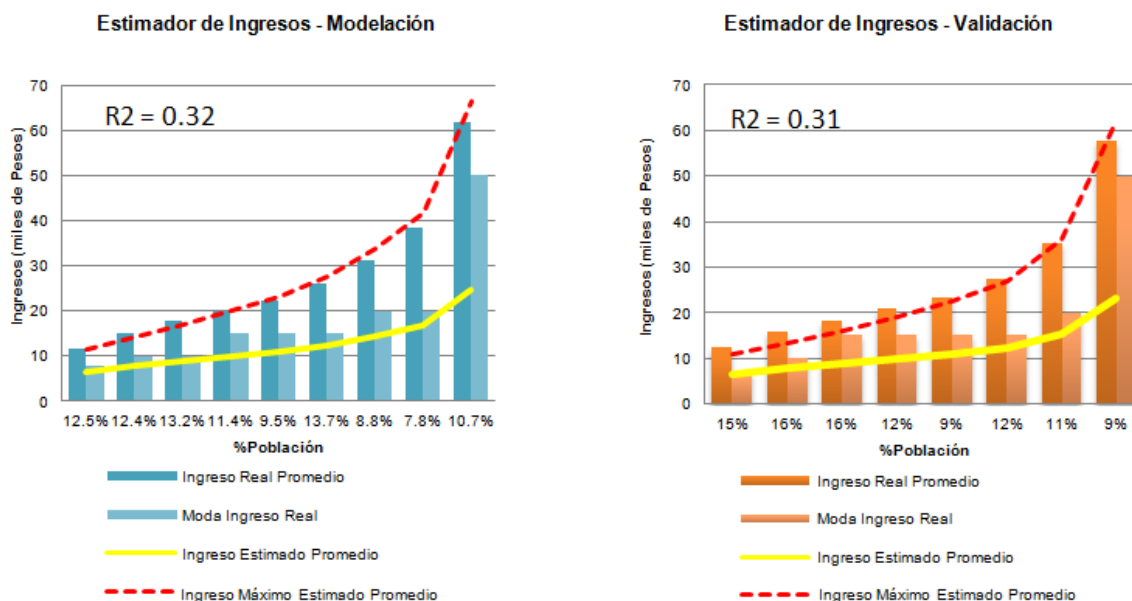
Finalmente, el otorgante de crédito recibirá el estimador del ingreso para cada uno de sus prospectos de análisis, además de tres códigos de razón. Los dos primeros códigos de razón corresponden a los límites de un intervalo de predicción donde oscila el verdadero valor de ingreso y el tercer código es un indicador de la fuente de datos utilizada para el cálculo del estimador; esto es, si el estimador del ingreso fue construido con información crediticia o a través de las fuentes de información pública. Los códigos de razón del estimador del ingreso se describen en el Apéndice 3.

## 2.2.2 Validación del modelo

Este modelo con información de crédito se evaluó con la información de validación y con una base diferente para corroborar el poder estadístico del modelo. El Coeficiente  $R^2$  es un elemento que determina la calidad del modelo para replicar los resultados, y la proporción de variación de los resultados que puede explicarse por el modelo.

El coeficiente  $R^2$  para el modelo fue de 0.32 y se validó con una muestra independiente de 2 millones de expedientes con un coeficiente  $R^2$  de 0.31, el coeficiente  $R^2$  es una medida de desempeño del modelo y valores superiores a 0.25 en variables económicas indican resultados muy robustos y consistentes.

Figura 3.5. Desempeño del modelo para clientes que tienen historial crediticio con muestras de Modelación y de Validación.



## 2.3. Modelo para clientes que no tienen historial crediticio (NOREF)

La encuestas nacionales tienen como principal objetivo mostrar la representatividad poblacional, considerando que los individuos dentro de una localidad, municipio o estado comparten características similares y por tanto, al extraer una muestra aleatoria de la población se cumple con el objetivo de representatividad, además de optimizar el tiempo y recursos.



El promedio del ingreso a nivel nacional por individuo se estimó a través de múltiples encuestas nacionales, con representatividad a nivel estatal por sexo: Estas estimaciones son de ayuda para la implementación de estrategias de desarrollo poblacional, sin embargo, actualmente se requiere de una estimación del ingreso de un individuo considerando un enfoque más detallado. Una estimación del ingreso de un individuo para un área de menor densidad geográfica sirve en la implementación y direccionamiento de estrategias socioeconómicas y de salud.

Buro de Crédito trabajo con la información pública disponible para generar un estimador del ingreso asociado al código postal (CÓDIGO POSTAL) del domicilio de un mexicano. Usando diversas fuentes de información se obtienen indicadores poblacionales a nivel CÓDIGO POSTAL del entorno socio-cultural y de salud que caracteriza a un individuo de cierto CÓDIGO POSTAL, con esta información se diseñó un modelo estadístico para la estimación del ingreso a nivel CÓDIGO POSTAL.

Se trabajó con las fuentes de información públicas que tuvieran información directa o indirectamente con el gasto y el ingreso de la población mexicana:

- INEGI 2010. Censo poblacional
- CONAPO 1995-2010. Índice y Grado de marginación. Localidad, Municipio, Estatal y Nacional
- SECRETARÍA DE SALUD. ENSANUT 2012. Nivel socioeconómico, inseguridad alimentaria, escolaridad y sexo del jefe de familia, programa oportunidades.
- PROSPERA. Acceso al programa de apoyo prospera
- SEDESOL. Acceso a algún beneficio de programas de crecimiento familiar
- SEPOMEX. Códigos postales identificados (Descarga en noviembre 2017)

Las fuentes de información poseen diferentes atributos y se diseñaron para cubrir objetivos diferentes, en el presente desarrollo se trabajó con la información y todos los elementos del diseño de la encuesta o registro de información, enseguida se presentan los enfoques de las fuentes de información:

- Censos poblacionales (INEGI)
- Padrones de acceso a algún beneficio (SEDESOL y PROSPERA)
- Encuestas representativas con diseño complejo (ENSANUT 2012)
- Proyecciones a 30 años de la población mexicana con el uso de herramientas estadísticas (CONAPO)

La ENSANUT 2012 tiene un diseño complejo y en particular posee información sobre la salud y nutrición de la población mexicana, por tanto, los análisis deben considerar el diseño estadístico de su construcción. Los beneficios de usar el diseño de la ENSANUT 2012 es que los resultados se pueden extrapolar hasta nivel estatal urbano y rural incluyendo en los términos de error la información del diseño muestral: conglomerado, estrato y ponderador.

### **2.3.1. Estimador del ingreso mensual promedio por Código Postal**

La definición antropológica y geográfica de una localidad o código postal indica que los individuos que viven en un área geográfica o espacio determinado de menor densidad constituyen un paisaje humano particular. Entonces es posible considerar que los individuos de cierto código postal comparten características y tienen

comportamientos similares. Los análisis serán para estimar el ingreso promedio mensual de un individuo que pertenece a cierto código postal.

Ya que los análisis se centrarán a nivel código postal, se utilizó la base de SEPOMEX que incluye 32,291 códigos postales, en donde para cada uno de los código postal se cuenta con el número de comunidades urbanas y rurales identificando el municipio y estado al que pertenecen.

Generalmente las encuestas nacionales no son diseñadas para ser representativas a nivel código postal, sino se tiene información a nivel localidad, municipio y estatal. Cuando no se dispone de información a nivel código postal, se utilizaran los valores estimados a partir de la metodología de estimación de área pequeña, usando la información del Municipio y el Estado que contiene al código postal. Para cada una de las fuentes de información se incluyó a las variables relacionadas con el ingreso mensual de un individuo en los código postal. Además, se construyen variables de apoyo por código postal con el número de comunidades urbanas, rurales y el tipo de localidades como rancherías, barrios, ejidos, etc.

El modelo estadístico considera un modelo de regresión multivariada, donde el objetivo principal de estimar el ingreso mensual promedio de un individuo por código postal –en miles de pesos- a partir de las estimaciones obtenidas en los código postal, utilizando variables de apoyo sobre las características de las localidades urbanas y rurales del código postal mediante SEPOMEX y desde las fuentes públicas todas aquellas variables que fueron estadísticamente significativas. . Se segmentó a la población objetivo para optimizar la estimación del ingreso mensual y se utilizaron las variables de sexo, edad y el ingreso mensual desde las bases de Buró de Crédito.

Básicamente el Modelo de clientes que no tienen información de crédito se divide en tres fases:

1. Análisis descriptivo de la edad y sexo de la base de Buró de Crédito para valorar la representatividad a nivel nacional de la muestra utilizada en los análisis.
2. Modelo multivariado del ingreso a nivel municipal
3. Modelo de estimación de área pequeña a nivel código postal usando las estimaciones a nivel Municipio

Considerando que la ENSANUT2012 posee representatividad nacional se compararon las muestras de la ENSANUT2012 y de desarrollo de Buró de Crédito. La densidad de la población por sexo y edad, se observa en la Figura 6, se puede observar que las distribuciones de la Buró de Crédito son cercanas a las reportadas por la ENSANUT2012.

Figura 3.6. Distribución del Ingreso transformado por tipo de fuente: Buró de Crédito y ENSANUT

Rango de edad	Buro de Crédito (%)		ENSANUT2012 (%)	
	Masculino (MAS)	Femenino (FEM)	Masculino (MAS)	Femenino (FEM)
Grupo 1	12.48	14.62	20.65	19.62
Grupo 2	22.16	23.34	19.14	19.8
Grupo 3	21.21	21.5	20.07	21.39
Grupo 4	21.61	20.15	20.57	22.28
Grupo 5	22.54	20.39	19.57	16.91

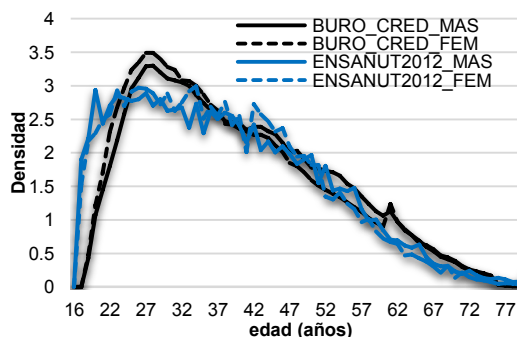


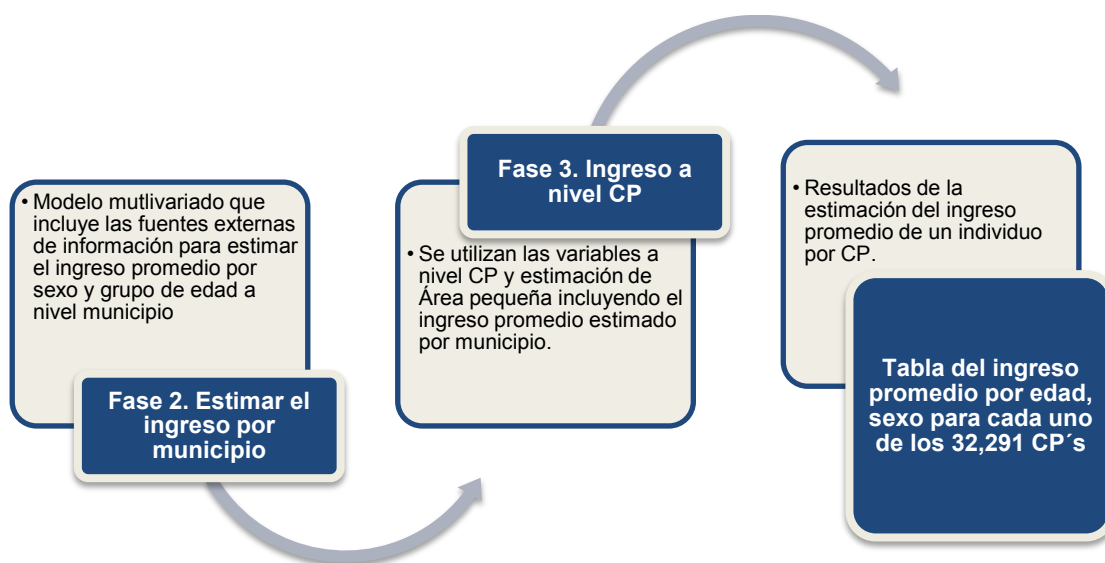
Tabla 3.1 Estadísticos descriptivos por edad y sexo

Ya que la distribución de la población en Buró de Crédito es representativa a nivel nacional, segmentaremos a la población en los grupos homogéneos. El ingreso transformado por código postal se analizó por cada uno de los subgrupos de edad para cada sexo. Teniendo 5 subgrupos de edad para el sexo masculino y otros 5 subgrupos de edad para el sexo femenino indicados en la tabla anterior.

La segunda fase utiliza información del municipio como el número de hogares, la posesión bienes en los hogar como radio, computadora, el índice de condiciones de bienestar, además del número de individuos dentro de los padrones de beneficiarios de algún programa de apoyo de desarrollo o salud, ésta información se incluye en el modelo multivariado para estimar el ingreso promedio por municipio para cada uno de los 10 subgrupos antes mencionados.

Finalmente en la tercera fase se usó la técnica de estimación de área pequeña para estimar el ingreso promedio por código postal, utilizando las variables que caracterizan al código postal y el ingreso por municipio estimado en la fase 2. El flujo de las fases 2 y 3 se presenta enseguida:

Figura 3.7. Flujo del Modelo NOREF

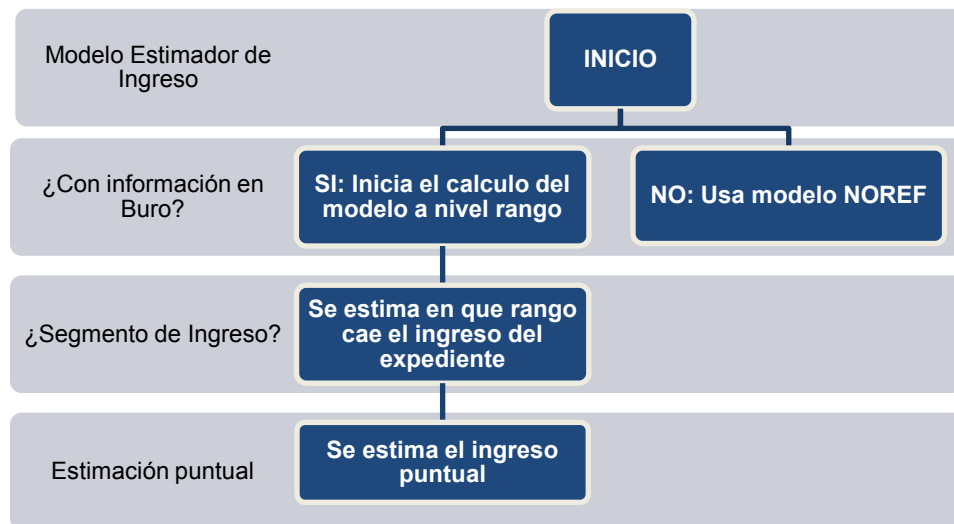


El modelo propuesto para la estimación del ingreso promedio para cada CÓDIGO POSTAL tuvo un buen desempeño, obteniendo un valor del coeficiente  $R^2$  general de 0.3, cabe mencionar que al centrarnos en un subgrupo en particular el coeficiente  $R^2$  es menor, descendiendo hasta 0.11.

## 2.4. Modelo de ensamble

Finalmente se realiza la conexión entre los modelos interno y externo a través de un ensamble condicionado a la disponibilidad de información como se indica a continuación.

Figura 3.8. Flujo del Modelo NOREF



Por tanto, se tendrá una estimación para un individuo que puede ser sólo a través de su domicilio o un modelo más robusto par cuando se cuente con información del historial crediticio de un individuo.

El reporte tendrá la siguiente estructura:

**SCORE** = XXXX (Valores de 001 a 061)

**R1** = Limite inferior de SCORE redondeado a Miles (Valores de 001 a 300)

**R2** = Limite superior de SCORE redondeado a Miles (Valores de 001 a 300)

**R3** = 501: HIT en Buró (Información de BURÓ DE CRÉDITO) o 502 = NOREF en Buró (Fuentes externas)

Siguiendo con el compromiso de continuidad de la calidad en términos de modelación estadística, se realizarán actividades de monitoreo de los modelos, en particular para el Estimador de Ingresos, el monitoreo será desde el enfoque externo de forma periódica cuando se encuentren disponibles datos públicos de la fuentes que se utilizaron en los cálculos. Además, para el modelo interno se realizará el mantenimiento necesario considerando las normativas de la ley de valoraciones numéricas que rigen a todas las SICs.

Para cada punto el elemento de análisis será el desempeño del modelo y los intervalos de confianza, con el fin de buscar la mejora continua del modelo. Esto también se presentaría cuando incremente el número de datos disponibles.

## 2.4.1 Ficha técnica

Estimador de Ingresos		
Fuentes de Información Desarrollo	Fecha Desarrollo	2017
	Fecha Observación	Julio, Septiembre, Noviembre 2016 y Enero, Marzo y Mayo 2017
	Tamaño Muestra (individuos con ingreso comprobado)	971,710
	Tamaño Muestra de evaluación de ajuste	2 millones de individuos con ingreso real
	Fuentes externas	INEGI, CONAPO, PROSPERA, SEDESOL y ENSANUT
Resultados	# Subgrupos analizados	4
	Rango de ingreso estimado	1 – 60 Mil pesos mensuales
	Rango de confianza de estimación	1 – 300 mil pesos mensuales
	Nivel de confianza del intervalo (%)	80-90
	Número de variables	50
Ajuste del Modelo ( $R^2$ )	Modelación	0.32
	Validación	0.31

## Sección 3. EJEMPLOS DE USO

### 1.1. Utilización del Estimador de Ingresos

En esta sección se muestran algunos ejemplos de uso del Estimador de Ingresos, con un enfoque común según el sector analizado. Los ejemplos muestran algunas formas de cómo utilizar el modelo a lo largo de la vida del crédito. Es necesario puntualizar que el uso del Estimador de Ingresos depende de las políticas de cada otorgante, sus políticas de evaluación y la experiencia en el uso de los scores de crédito.

Sector	Oportunidades
<b>Nuevos emisores de productos crediticios</b>	Colocar productos en clientes de otros productos de crédito Asignación de límites de crédito Términos del crédito
<b>Emisores de productos crediticios</b>	Asignación de límites de crédito Términos del crédito Estrategias Up-grade

Los otorgantes pueden apoyarse en el Estimador de Ingresos en la búsqueda del equilibrio entre el riesgo y la colocación de algún producto crediticio, ya que la estimación se puede utilizar para evaluar la capacidad crediticia de un prospecto.

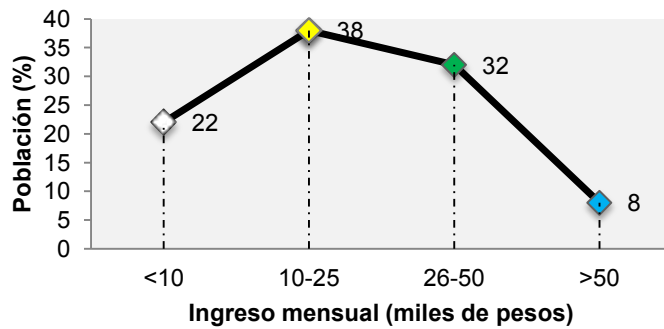
Con el uso del Estimador de Ingresos como insumo en los análisis de la institución se incrementarían las posibilidades para determinar la capacidad de endeudamiento de un prospecto y los costos del envío de las ofertas de colocación de un producto nuevo u ofertas según el perfilamiento de la población de cada institución.

### 1.2. Colocación de productos crediticios para otorgantes

Ejemplo teórico: Un otorgante habitualmente envía a todos sus clientes las ofertas de los nuevos productos crediticios, lo que implica grandes costos y busca una estrategia que le permita identificar a los clientes con mayor potencial para el nuevo producto crediticio y minimizar costos.

Solución: Usando la segmentación de habitual de los clientes del otorgante y el **Estimador de ingresos** se tiene la distribución del ingreso de sus clientes.

Figura. Distribución de la Estimación del ingreso mensual de una institución. (Ejemplo teórico)



El otorgante decide enviar la oferta del nuevo producto sólo a los clientes con un ingreso mensual mayor a 10 mil pesos, quienes representan el 78%. Además el otorgante considera enviar diferentes ofertas alineadas con el comportamiento crediticio de sus clientes a través del cálculo de BC Score. Aplicando información adicional del BC Score es posible diferenciar entre una oferta Estándar, Media o Premium..

Tabla 4.2-1 Ejemplo de estrategia de cálculo de límite de crédito inicial

Ingreso	Estrategia de colocación tradicional	BC Score	Estrategia con Estimador de ingresos
<10	Oferta	<650	No hay Oferta
		>650	Oferta Estándar
10-25	Oferta	650-700	Oferta Estándar
		>700	Oferta Media
26-50	Oferta	620-680	Oferta Estándar
		>680	Oferta Media
>50	Oferta	650-700	Oferta Media
		>700	Oferta Premium

Aplicando información adicional del Reporte de Crédito / Adviser, podemos diferenciar entre una oferta Estándar de una Premium, dependiendo de si el cliente ya tienen compromiso con otros productos crediticios.

### 1.3. Asignación de límites de crédito

El Estimador de Ingresos en conjunto con un Score Buró de riesgo puede utilizarse para establecer lineamientos para la asignación de límites de crédito iniciales. Al establecer los límites iniciales, junto con el riesgo debe considerarse la capacidad del solicitante para asumir deudas adicionales. Existen diferentes medidas de capacidad, entre las que se encuentran: ingreso y/o ingreso disponible;

Tabla 4.3-1 Ejemplo de estrategia de cálculo de límite de crédito inicial

Score de riesgo	Ingreso nivel I	Ingreso nivel II	Ingreso nivel III
Bajo (Alto riesgo)	No hay oferta		
Medio inferior (Alto riesgo)	No hay oferta	Oferta Estándar-Línea baja	Oferta Estándar-Línea baja

<b>Medio (Riesgo marginal)</b>	Oferta Estándar-Línea baja	Oferta Estándar-Línea media	Oferta Estándar-Línea Alta
<b>Alto inferior (Bajo Riesgo)</b>	Oferta Estándar-Línea Media	Oferta Estándar-Línea Alta	Oferta Estándar-Línea Alta
<b>Alto (Bajo riesgo)</b>	Oferta Estándar-Línea Media	Oferta Premium-Línea Alta	Oferta Premium-Línea Alta

La tabla anterior muestra un ejemplo de una estrategia de límite de crédito que utiliza el Estimador de Ingresos y algún Score de riesgo para los clientes de un segmento Privado. Desde una perspectiva metodológica, la forma correcta de establecer el valor límite de crédito final se basaría en la estimación de dos dimensiones completamente separadas: la "disposición a pagar" que proporciona la indicación correcta del riesgo de crédito y la "capacidad de pago". Eso permite definir la accesibilidad financiera del cliente. Hay diferentes medidas de capacidad; ingresos, renta disponible, potencial financiero residual son ejemplos de medidas de capacidad.

Los límites de crédito pueden variar dentro de cada rango de riesgo (alto, medio y bajo) de acuerdo con la capacidad del solicitante para la deuda adicional. De los clientes que tenían un Score Buró de riesgo, aquellos con altos ingresos se les asignaron líneas de crédito más altas. A otros clientes que recibieron scores altos en Score Buró de riesgo, pero que tenían bajos ingresos, se les asignaron límites de crédito medios.

Las asignaciones de límites de crédito de esta forma maximizan los ingresos de buenos clientes, mientras que limitan el potencial de pérdidas de crédito entre los clientes más riesgosos. La matriz de límite de crédito se puede automatizar fácilmente, lo que permite una asignación de límites rápida y consistente.

Otros enfoques se centrarían en la utilización de un vector de coeficientes para aplicar a la disponibilidad financiera asociada al ingreso mensual más que en un límite de crédito predefinido, donde la estimación se define sobre la base del nivel de riesgo esperado.

## 1.4. Cobranza

La rentabilidad de las operaciones de cobranza de un otorgante son favorecidas con el uso de del Estimador de Ingresos, ya que es un indicador de la capacidad de pago de los clientes que se encuentran en mora.

Con el objetivo de sanar las cuentas morosas, será posible desarrollar análisis costo-beneficio basados en estrategias de cobranza, usando segmentos generados a partir del Estimador de Ingresos.

## 1.5. Resumen

Buró de Crédito ofrece un modelo de estimación del ingreso mensual en miles de pesos que utiliza metodología innovadora e información pública, ofreciendo específicamente para el mercado crediticio mexicano una herramienta robusta auxiliar que permitirá a los otorgantes de crédito evaluar con mayor precisión el riesgo crediticio de la población cuando se utilice en conjunto con un score de riesgo. El Estimador de Ingresos es un modelo único que puede ser útil como herramienta para rankear a los clientes de algún portafolio en particular o en la valoración de capacidad de pago.



El modelo se construyó a partir de los datos de aproximadamente 1 millón de expedientes de crédito de consumo tomados de todas las regiones de México y presenta un enfoque de rangos para segmentar mejor a la población. El Estimador de Ingresos es un modelo estadístico construido con información crediticia de la población mexicana que se utiliza para conocer el ingreso mensual de una persona. Es un modelo con características únicas, ya que, además de proveer una estimación precisa del ingreso, proporciona también dos elementos de apoyo para su comprensión:

- Un intervalo de confianza para el ingreso real y,
- Un identificador de las fuentes utilizadas para la construcción de la estimación, Buró de Crédito o información pública externa como INEGI, CONEVAL, CONAPO y ENSANUT.

El Estimador de Ingresos entrega la estimación del ingreso mensual en miles de pesos, desde \$1,000 hasta \$61,000 pesos, y un intervalo de predicción que garantiza con 90% de probabilidad que el ingreso real de la persona estará dentro de este intervalo.

## **Sección 4: Apéndices**

### **Apéndice 1: Preguntas frecuentes**

- ¿Qué es el Estimador de Ingresos?
  - Es un modelo estadístico construido con información crediticia y socio-demográfica de la población mexicana para estimar el ingreso mensual de una persona. Posee dos elementos de apoyo para su comprensión:
    - Un intervalo de predicción para el ingreso real y,
    - Un identificador de las fuentes utilizadas para la construcción de la estimación, Buró de Crédito o información externa.
- ¿Cómo empezar a usar Estimador de Ingresos?
  - Los usuarios que por primera vez utilizan Estimador de Ingresos tendrán como motivación la gran bandeja de opción para su uso, ya que pueden usarlo para rankear a sus clientes o en conjunto con un score de riesgo para potenciar su utilidad. Los otorgantes pueden tener una gran variedad de estrategias de decisión a través de una herramienta robusta para la validación, creación y mantenimiento de su cartera, considerando que se posee cierta comprensión de cómo funcionan las estimaciones y scores en su entorno crediticio.
- ¿Qué es el BC Score?
  - Una calificación de Buró de Crédito representa las probabilidades o posibilidad de que un cliente se vuelva moroso por lo menos en una cuenta en un período de 12 meses.
- ¿Qué es un modelo?
  - Un modelo generalmente está compuesto de diferentes segmentos o scorecards y es una serie de algoritmos matemáticos que calculan de manera sistemática el número de puntos (o calificación) de cada cuenta que se procesa a través de la clasificación, los modelos están diseñados para utilizar características que se consideran predictivas en el pronóstico de posibilidades de que un cliente actúe de una manera dada, como por ejemplo, volverse moroso en 90 días o más a partir de un período específico.
- ¿A quiénes se les puede calcular el Estimador de Ingresos?
  - A todos aquellos que cuenten con historial crediticio o bien, cuyo domicilio se encuentre dentro del listado de códigos postales activos al momento del análisis. Por ejemplo, no será posible calcular el Estimador de Ingresos de los códigos postales de reciente creación por SEPOMEX, sin embargo, el modelo será calibrado conforme lo dictamina la ley que regula a CIC's para considerar la actualización más reciente SEPOMEX.

- ¿Cuándo puede usarse Estimador de Ingresos?
  - La institución podrá estimar el ingreso a aquellos prospectos para un nuevo producto crediticio, o a quienes desea enviar una propuesta de mejores condiciones y evaluar el mantenimiento y/o extensión de productos actuales. Seleccionando el Estimador de Ingresos en combinación con un score de riesgo, podrá tener una herramienta poderosa para implementar acciones que evalúen de mejor manera al cliente.
- ¿Cómo puedo determinar si el Estimador de Ingresos es mejor que un comprobante de ingresos?
  - El Estimador de Ingresos no compite con un documento fidedigno, sino es un auxiliar en el diseño de estrategias grupales, permitiendo ubicar a un prospecto en un grupo determinado a quien es posible ofrecer un producto crediticio.
  - La utilidad del Estimador de Ingresos está muy presente en la revisión de límites de crédito o renovaciones de crédito, donde no se solicita un comprobante de ingresos.
  - Con el Estimador de Ingresos se tendrá mayor información del prospecto cuando las actualizaciones de información no son reciente o se tienen irregularidades en los documentos presentados.
  - En el área de cobranza cuando se desconoce el ingreso con el Estimador de Ingresos podríamos diseñar una estrategia de cobranza.
- ¿Cuál es el valor del Estimador de Ingresos recomendado para ofrecer un crédito?
  - Es imposible que se recomiende que para cierto Ingreso estimado se ofrezca un crédito para cualquier otorgante de crédito debido a que el entorno de cada otorgante es diferente. Buró de Crédito podría trabajar de forma más cercana con los otorgantes de crédito para ayudarles a establecer herramientas de evaluación usando el Estimador de Ingresos. En la Sección 3 se señalan algunos escenarios que podrían determinar las ocasiones donde es preferible usar algún Score de Buró de Crédito y el Estimador de Ingresos, para dirigir las estrategias de cierto producto crediticio. Finalmente cada otorgante de crédito debe determinar cuál es la combinación más apropiada del Ingreso estimado y el Score para sus objetivos.
- ¿Cómo puede determinar un otorgante de crédito el punto de corte más adecuado de un score de riesgo en conjunto con el Estimador de Ingresos?
  - El Estimador de Ingresos fue calculado con información crediticia de un individuo pero no es necesariamente la misma de un score de riesgo, es por ello que se sugiere usar el Estimador de Ingresos como apoyo de un score de riesgo. Ya con las tablas de probabilidad del score de riesgo es posible colocar los ingresos asociados a cierto corte dentro de la distribución del score.

- ¿Cuáles son las ventajas de utilizar el Estimador de Ingresos y los scores de Buró de Crédito?
  - Poseer mayor información acerca de la capacidad de pago de un prospecto permite que un otorgante de crédito pueda disminuir sus índices de pérdida, ajustando su cartera a cierto ingreso permisible o bien diseñar herramientas ajustadas al ingreso de ciertos perfiles, tramitar más solicitudes en menos tiempo, reducir costos de operación y aumentar la precisión y consistencia de sus decisiones de otorgamiento de crédito.
  - Los Scores de Riesgo proporcionan un resumen de fácil interpretación sobre el perfil de riesgo del cliente y en conjunto con el Estimador de Ingresos se tendrá una potente herramienta de predicción de conducta futura en pagos y responsabilidades crediticia, lo que implica para los otorgantes de crédito un control rápido y flexible en el manejo de sus carteras.

- ¿Por qué el coeficiente de determinación  $R^2$  se puede usar para validar un modelo?

En la literatura se suele encontrar que los criterios para determinar la bondad de un modelo se basa en el examen de los valores del coeficiente de correlación y, principalmente, del coeficiente de determinación  $R^2$  (también denominado coeficiente de correlación múltiple al cuadrado o coeficiente de determinación múltiple). El  $R^2$  es un criterio de valoración de la capacidad de **explicación** de los modelos de regresión, y representa el porcentaje de la varianza justificado por la variable independiente. Si todas las observaciones están en la línea de regresión, el valor de  $R^2$  es 1, y si no hay relación lineal entre las variables dependiente e independiente, el valor de  $R^2$  es 0. El coeficiente  $R^2$  es una medida de la relación lineal entre dos variables. A medida que su valor es mayor, el acercamiento de la recta a los datos es mejor, puesto que la variación explicada es mayor; así, el desajuste provocado por la sustitución de los valores observados por los predichos es menor. Sin embargo, en Economía y Finanzas los modelos de estimación son considerados adecuados cuando el  $R^2$  es cercana a 0.3.

En definitiva, lo deseable es realizar predicciones lo más acertadas posibles, sin importarnos el método; siempre que sea un método científico que pueda ser reproducible. Y, para medir la capacidad de predicción se deben utilizar variables **predictivas** principalmente, no sólo **explicativas** que aporten validez al modelo. En general, siempre existirán otros elementos cuestionables cuando se está modelando una variable que pueden emerger durante la validez para un modelo predictivo, y que muchas veces no son considerados. Sólo una visión global de todas las variables, las covariables que integran el modelo y las puramente expresivas de capacidad de predicción nos pueden mostrar verdaderamente las posibilidades de nuestro método y de nuestros datos. Es por ello, que el Estimador de Ingresos de Buró de Crédito fue construido que forma su género integrando información pública al modelo final para ofrecer un modelo integral que considera los fuentes de información disponible para la población mexicana.

## **Apéndice 2: Salida del Reporte del Estimador de Ingresos**

La especificación técnica y descripción de los campos de salida relacionados con el Estimador de Ingresos, se muestran enseguida:

Monto en miles de pesos dentro del rango (0-61) y códigos de razón:

- Número de 3 dígitos para el código de exclusión, o
- Número de 3 dígitos para el rango: 001 a 061.

Cuando se desee estimar el ingreso mensual en miles de pesos para un individuo que posea información crediticia, el modelo interno del Estimador de Ingresos podrá ser calculado y además se tendrán tres códigos de razón se proporcionan junto con el valor, que corresponden al intervalo de confianza del 90% del valor real del ingreso del individuo.

**Códigos de razón:**

- El primer código de razón corresponde al límite inferior del intervalo de 90% de confianza.
- El segundo código de razón corresponde al límite superior del intervalo de 90% de confianza.
- El tercer código de razón es un identificador de las fuentes utilizadas para la construcción de la estimación, Buró de Crédito o información externa (INGEI, CONAPO, CONEVAL, ENSANUT).

## **Apéndice 3: Exclusiones del Score**

El Estimador de Ingresos podrá calcularse en la gran mayoría de los casos, con una estimación a través del modelo interno si cuenta con historial crediticio o bien, con un código postal válido, además de estar vivo. De lo contrario el Estimador de Ingresos no se calculará y se asignará el código de la exclusión, el cual explica la carencia de información crediticia o que el domicilio se encuentra en un código postal de reciente creación y no considerado en la base de SEPOMEX descargada en la fecha de estimación.

Tabla 5.3-1 Códigos de exclusión Estimador de Ingresos

Código de exclusión	Descripción
-001	Difunto
-009	No es posible calcular el Estimador de Ingresos (Sólo aplica para los nuevos códigos postales).

## Apéndice 4: Códigos de Razón

Los posibles códigos de razón son:

Tabla 5.4-1 Códigos de razón de Estimador de Ingresos

Código de razón	Descripción
001-300	Límite inferior en miles de pesos
001-300	Límite superior en miles de pesos
501 ó 502	501: HIT. Uso de información de : Buró de Crédito

## Apéndice 5: Validación del Estimador de Ingresos

Para evaluar que el Estimador de Ingresos es robusto, se usaron algunas medidas muy comunes, tales como el coeficiente de correlación y el coeficiente de determinación  $R^2$ , dando una idea equivocada sobre las capacidades predictivas del modelo de estimación en cuestión. Además de las medidas habituales la sugerencia en la literatura es examinar también medidas, claramente predictivas, como son el Nivel de Predicción y la Magnitud Media del Error Relativo.

En modelos de riesgo los que se han obtenido para el coeficiente  $R^2$  en los diferentes estudios publicados, por ejemplo, sobre los puntos de función varían desde 0,14 hasta 0,87. Apoyándose en estos valores, algunos autores afirman la validez de la técnica de los puntos de función. Sin embargo, es una conclusión que no se desprende directamente de esos datos. Fijémonos que son valores explicativos, no necesariamente predictivos. Tanto el  $R^2$  como el coeficiente de correlación son las medidas frecuentemente utilizadas para evaluar la **predicción** de un modelo; en el mejor de los casos se trata de medidas de desempeño de la ecuación a los datos e indirectamente de la capacidad predictiva del modelo. Es necesario aclarar que existen algunos casos donde los que nos transmite el  $R^2$  puede coincidir con la de las variables que teóricamente son mayormente utilizadas, sin embargo, existen otros casos donde la  $R^2$  no coincide con los estadísticos: *magnitud media del error relativo y nivel de predicción*.

## Apéndice 6: Glosario

### Validación del modelo

Es la herramienta utilizada para evaluar un modelo que permite apoyar en la decisión si las predicciones son suficientes o se deben buscar modelos alternativos.

### Aleatorio(a)

Fundado sobre la intervención del azar o independientemente de otros acontecimientos.

### Análisis de regresión simple

Procedimiento para derivar una relación matemática, en forma de ecuación, entre una variable dependiente métrica y una variable independiente métrica.

Análisis de regresión:

Procedimiento estadístico para analizar las relaciones de asociación entre una variable dependiente métrica y una o más variables independientes.

Atributos:

Variables cualitativas que sólo poseen categorías.

BC Score:

Un BC Score representa los odds o probabilidad que un consumidor se vuelva moroso por 90 días o más en al menos una cuenta para un periodo de 12 meses.

Censo:

Conteo completo de los elementos de una población u objetos de estudio.

Códigos de exclusión

Los códigos de exclusión son las reglas y criterios con los cuales se excluye a un prospecto desde el cálculo estadístico del score.

Coeficiente de correlación parcial

Medida de la asociación entre dos variables después de controlar o ajustar los efectos de una o más variables adicionales.

Coeficiente de determinación  $R^2$

Coeficiente de determinación múltiple,  $R^2$ , y  $R^2$  ajustado, son algunas medidas habituales en el análisis de regresión, denotando el porcentaje de varianza justificado por las variables independientes. El  $R^2$  ajustado tiene en cuenta el tamaño del conjunto de datos, y su valor es ligeramente inferior al de su correspondiente  $R^2$ .

Datos externos:

Datos que se obtienen de una fuente diferente de la organización para la cual se está llevando a cabo la investigación.

Decil:

Percentil cuyo valor que indica su proporción es un múltiplo de diez. Percentil 10 es el primer decil, percentil el segundo decil, etc.

Desviación estándar:

Raíz cuadrada positiva de la varianza; medida de dispersión con las mismas unidades que los datos originales, más bien que en las unidades al cuadrado en que está la varianza.

Dispersión:

La extensión o variabilidad de un conjunto de datos.

Distribución asimétrica:

Se presenta cuando la distribución de un conjunto de datos resulta con un promedio, una mediana y una moda con valores diferentes; también se considera como una distribución "sesgada".

Distribución de la muestra:

La distribución de los valores de la estadística de una muestra (calculada para cada muestra posible), que pueda tomarse de la población meta de acuerdo con un plan de muestreo específico.

Distribución de probabilidad:

Lista de los resultados de un experimento con las probabilidades que se esperarían ver asociadas con cada resultado.

Distribución normal:

Distribución de una variable aleatoria continua que una curva de un solo pico y con forma de campana. La media cae en el centro de la distribución y la curva es simétrica con respecto a una línea vertical que pase por la media. Los dos extremos se extienden indefinidamente, sin tocar nunca el eje horizontal. Base para la inferencia estadística clásica que tiene forma de campana y apariencia simétrica. Todas sus medias de tendencia central son idénticas.

Estadística inferencial:

Es el tipo de estadística que interpreta la información de tal manera forma que nos apoye en obtener conclusiones válidas, a partir del estudio de una muestra.

Estadístico:

Descripción resumida de una medida en la muestra seleccionada.

Eventos independientes:

Dos eventos evento son independientes si el conocimiento de que uno ocurrirá o ya ha ocurrido no afecta la probabilidad del otro; más precisamente, si la probabilidad condicional de cada uno dada por el otro es la misma que la probabilidad incondicional.

Indicador:

Se trata de un número o de un índice (un valor sobre una escalera de medida) derivado de la observación de un conjunto de fenómenos. Variable que permite evaluar ciertos cambios en el curso del tiempo.

Intervalo de confianza:

Intervalo de valores que tiene designada una probabilidad de que incluya el valor real del parámetro de la población.

Intervalo muestral:

Tamaño de la distancia ente los elementos seleccionados en un muestreo sistemático; el recíproco de la fracción muestral.



Límites de confianza:

Límites inferior y superior de un intervalo de confianza.

Línea de regresión:

Una línea ajustada a un grupo de puntos para estimar la relación entre dos variables.

Media:

El promedio; valor que se obtiene al sumar todos los elementos en un conjunto y dividirlos entre el número de elementos.

Medidas de dispersión:

Estadísticas que expresan criterios para describir la ubicación relativa de los datos.

Medidas de variabilidad:

Estadística que indica la dispersión de la distribución.

Moda:

Medida de tendencia central que se da como el valor que ocurre con mayor frecuencia en la distribución de una muestra.

Modelo

Un modelo generalmente está compuesto de varias tablas de score o scorecards y es una serie de algoritmos que calcula de forma sistemática el número de puntos, o calificación, para cada cuenta que se pasa a través de la tabla de score.

Muestra:

Es una parte representativa que refleja las similitudes y diferencias de la población y que son importantes para la investigación, se podría decir que es el subconjunto seleccionado de la población; por eso se suele seleccionar un subgrupo que sea suficientemente representativo, pero tiene que tener datos que puedan servir para conclusiones generalizadas.

Muestra aleatoria / muestra de probabilidad:

Tipo de muestra caracterizada por una selección de sujetos basada en la ley de las probabilidades; un procedimiento de preparación de muestras es aleatorio, o probabilista, cuando todos los elementos de la población tienen una posibilidad de ser recuperados en la muestra: la probabilidad de elección de cada elemento de la población debe ser conocida por progreso. Se trata del único método general capaz de atribuir un valor numérico preciso a la estimación.

Multicolinealidad:

Problema estadístico que se presenta en el análisis de regresión múltiple, en el que la confiabilidad de los coeficientes de regresión se ve reducida debido a un alto nivel de correlación entre las variables independientes.

Nivel de confianza:

Probabilidad que los estadísticos asocian con una estimación de intervalo de un parámetro de población. Ésta indica qué tan seguros están de que la estimación de intervalo incluirá al parámetro de la población.

Parámetro:

Elemento variable en función del cual se explicitan las características esenciales de un fenómeno. Se trata de una unidad de medida desconocida y cuantitativa (tal como la renta total, la renta media, la producción total, el número de desempleados) utilizada por los investigadores para estudiar a una población entera u otros ámbitos de interés. Valores que describen las características de una población.

Ponderación:

Desempeño estadístico a los datos en el cual a cada caso o entrevistado en la base de datos se asigna un valor relativo a fin de reflejar su importancia relativa para otros casos o entrevistados.

Score o Calificación

Un score es una medida numérica que se utiliza para representar la probabilidad de un evento futuro.

Segmentación mediante Scorecards

La segmentación es el proceso de realizar un análisis para determinar si acaso diferentes sub-poblaciones dentro de la muestra de desarrollo del modelo podría alcanzar mayor capacidad predictiva utilizando clasificaciones claramente segmentadas.

## Contactos de Buró de Crédito

ÁREA	TELÉFONO
Dirección Comercial	5449 4945 5449 4917
Gerencia de Ventas	5449 4930 5449 4948 5449 4900 ext. 5976 5449 4900 ext. 4841 5449 4900 ext. 5914
Dirección de Operaciones	5449 4988
Dirección de Sistemas	5449 4973
Centro de Atención a Usuarios	5449 4949
Adquisición de Bases de Datos	5449 4923
Soporte Técnico	5449 4982 mesadeayuda@burodecredito.com.mx

## Propiedad y Confidencialidad

Este Manual es propiedad de **TransUnion de México, S.A., SIC**

Se prohíbe su reproducción total o parcial así como su venta y distribución a personas y/o empresas que no tengan acordado un contrato de confidencialidad con **TransUnion de México, S.A., SIC**