

# Universidad Autónoma del Estado de México

## UN MODELO DE CALIFICACIÓN ESTADÍSTICA PARA UNA INSTITUCIÓN MEXICANA ESPECIALIZADA EN MICROCRÉDITOS

Rafael Morales Ibarra,  
Ma. Eugenia Valdez Pérez,  
Gandhi González Guerrero

Agosto, 2013



# Objetivo

- Emplear la regresión logística, para construir un modelo de *credit scoring* aplicada a una institución microfinanciera.
- Estimar la probabilidad de impago de futuros créditos potenciales.
- Caracterizar el funcionamiento y operatividad de las Instituciones microfinancieras.

# Antecedentes

- 1970. Orígenes de *la* evaluación del riesgo de portafolios financieros.
- 1990. Las técnicas del Credit Scoring, se hacen populares. Generalmente asociados a modelos de minería de datos y a técnicas estadístico-matemática.
- Actualmente. El *credit scoring* en microfinanzas no son comunes, fundamentalmente por la limitada información.
- Escasa literatura y/o trabajos empíricos en microfinancieras y más escaso aún la aplicación de técnicas a nivel regional.

# Justificación

- El problema de la morosidad tiene el mismo nivel de importancia que en cualquier institución bancaria. Las microfinanzas no son la excepción.
- Cualquier institución financiera requiere de mecanismos eficientes que reduzcan los índices del no pago.
- El *credit scoring* viene a coadyuvar al criterio humano en la difícil tarea de filtrar solicitudes de crédito.

# Marco Teórico

- Microcrédito. Son pequeños préstamos realizados a personas de menores ingresos o pobres.
- Importancia; radica en que personas sin recursos puedan acceder a financiamiento.
- El credit scoring (CS); procedimiento estadístico empleado para clasificar a los solicitantes de un crédito como usuario de riesgo 'bueno' y 'malo'. Hand and Henley (1997)
- El CS, se refiere al uso de conocimiento sobre desempeño y características de un préstamos en el pasado para pronosticar el desempeño de préstamos en el futuro. Schreiner (2011)

## ***El Credit Scoring (CS)***

Se aplican para obtener un conocimiento sobre:

- a) El historial financiero en los productos solicitados e índice de morosidad.
- b) Relación riesgo - rentabilidad.
- c) Coste de la operación

# Métodos y Técnicas

Técnicas comúnmente empleadas en el CS:

Paramétricas	No paramétricas
<ul style="list-style-type: none"><li>• Análisis Discriminante</li><li>• Modelos econométricos</li><li>• Inteligencia artificial</li><li>• Regresión logística</li></ul>	<ul style="list-style-type: none"><li>• Programación Lineal</li><li>• Redes neuronales artificiales</li></ul>

- Empleando estos modelos un cliente puede clasificarse en un grupo de liquidez, rentabilidad, apalancamiento y actividad. Orgler, (1971)

# El modelo

El modelo de regresión logística puede formularse como:

$$\text{Log} \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_{k1} \quad (1)$$

Dónde:

$P$  = probabilidad de que el cliente incurra en no pago.

Tal que:

$$p = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}} \quad (2)$$

Donde:

$$Z = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_{k1} \quad (3)$$



# Resultados

Con  $n=100$  microcréditos y el empleo de la Regresión Logística Binaria, se obtiene como resultado la siguiente ecuación:

$$z = 3.183 - 0.32Uso_{Cred} - 0.987Lab_{Usf} - 3.062Casa_i - 0.801Pib - 1.13Cta_{Loc} - 2.016Tip_{Garant} - 3.67Ti + 0.147Ad_{cred} \quad (4)$$

Empleando (4) y (5)

$$P(\text{probabilidad de incumplimiento}) = \frac{1}{1 + e^{-z}} \quad (5)$$

Se obtiene la probabilidad de no pago  $P$  para cada prestatario.

# VARIABLES ESTADÍSTICAMENTE SIGNIFICATIVAS

Definición	Variable	Parámetro	p-value ( $\alpha=0.05$ )
Cuál es el destino o uso del crédito	Uso_cred	-0.324	(0.000)
Actual situación laboral del cliente	Lab_Usf	+0.019	(0.025)
Número de atrasos anteriores	Atr_pas	-6.331	(0.055)
Número de cuotas atrasadas	Num_atrss	-2.795	(0.0670)
Comportamiento de la economía a nivel nacional	Pib	+1.891	(0.000)
Tasas de interés en el mercado	Ti_m	-0.298	(0.844)
El cliente cuenta con casa propia	Casa_i	+0.423	(0.005)
Estado civil del cliente	E_civil	+0.002	(0.060)
Número de préstamos anteriores	Num_cred	-4.893	(0.055)
Años funcionando del negocio o actividad económica actual	Num_exp	-0.078	(0.667)
Cuenta con local distinto a su casa	Cta_local	+0.689	(0.000)
Ingresos netos	Ynet_s	-0.069	(0.500)
Monto solicitado	Mon_s	+0.0226	(0.650)
Monto otorgado	Mon_o	-0.0240	(0.650)
Tipo de garantía	Tip_garant	-3.561	(0.035)
Analista de crédito	Ad_cred	-2.041	(0.000)
Nivel de estudios del analista	N_est	-0.0027	(0.500)

## El modelo final queda definido de la siguiente manera:

$$Z = 3.183 - 0.32Uso\_Cred - 0.987Lab\_Usf - 3.062Casa\_i - 0.801Pib - 1.13Cta\_Loc - 2.016Tip\_Garant - 3.67Ti + 0.147Ad\_cred \quad (6)$$

Las ocho variables en (6) influyen en la probabilidad de que un cliente sea clasificado como moroso o no. Una variable con signo positivo y siete con signo negativo por lo que dichas .

⇒ Con la variable **Ad\_cred** (signo positivo), al aumentar la variable independiente en una unidad, el logaritmo de las probabilidades de ser un cliente moroso aumenta en el valor del coeficiente respectivo, y viceversa para los estimadores con signo negativo (**Uso\_Cred; Lab\_Usf; Casa\_i; Pib; Cta\_Loc; Tip\_Garant y Ti**).

# Bondad de ajuste

- La prueba de razón de verosimilitud, cuyo valor es de 1469,492, el  $R^2$  de Cox y Snell y  $R^2$  de Nagelkerke, indican un ajuste aceptable.
- Hosmer y Lemeshow, no es satisfactorio por su nivel de significación, quizás por el tamaño de muestra.
- Los residuos, se asumen como aleatorios.
- La eficacia predictiva del modelo; el porcentaje correcto de clasificación obtenido para el modelo es de 83.1%.
- La mínima variabilidad en los  $\beta$ 's de las variables endógenas indica que los “*outliers*” no son significativos.

# Conclusiones

- Las variables predictoras del modelo son del tipo cuantitativo, de igual manera se agrupan en tres bloques: perfil del cliente, tipología de los créditos y variables macroeconómicas.
- Las variables macroeconómicas mejoran sustancialmente los resultados del modelo.
- Existe una relación inversa entre el riesgo y el puntaje del score; a mayor puntaje menor riesgo de no pago del cliente.

# Conclusiones....

- El modelo predice correctamente en más del 83% de los créditos de la cartera de MF.
- La técnica empleada tiene la ventaja del calcular la probabilidad de incumplimiento del cliente.
- En resumen. el modelo de CS solo es un indicativo, la determinación fundamental radica en el recurso humano de las instituciones.



**Gracias!!!**

[rmoralesi@uaemex.mx](mailto:rmoralesi@uaemex.mx)

