

# REPORTE

## ANÁLISIS COMPARATIVO DE MODELOS DE SCORING PARA LA EVALUACIÓN DE UNA CARTERA DE CRÉDITO

*Angel Josué Mejía Nájera*

15 de marzo de 2024

# Índice

<b>1. Descripción de la base de datos</b>	<b>2</b>
1.1. Características específicas de la base de datos . . . . .	2
<b>2. Variables Aleatorias y Modelos</b>	<b>3</b>
2.1. Variables Aleatorias Utilizadas . . . . .	3
<b>3. Ajuste de Modelos (RLM, Logit y Probit)</b>	<b>3</b>
3.1. Resultados . . . . .	4
<b>4. Análisis de Odds Ratio (Coeficientes de momios)</b>	<b>6</b>
4.1. Interpretación de los coeficientes de momios (Modelo Logit . . . .	6
<b>5. Matriz de confusión</b>	<b>7</b>
<b>6. Curvas ROC y PR</b>	<b>8</b>
<b>7. Scoring</b>	<b>9</b>
7.1. Modelo RLM . . . . .	9
7.2. Modelo Logit . . . . .	10
7.3. Modelo Probit . . . . .	11
<b>8. Conclusiones</b>	<b>12</b>
<b>9. Referencias</b>	<b>12</b>

## 1. Descripción de la base de datos

La base de datos generada de manera aleatoria se compone por la información de 1,000 personas que solicitaron un crédito anteriormente y se les fue concedido, dándonos la información de si son cumplidos o incumplidos a la hora de pagar sus respectivos créditos. Para otorgar un crédito de manera responsable y adecuada, es necesario contar con información histórica relevante sobre los solicitantes para inferir si es adecuado autorizar un crédito o no, según el riesgo de crédito estimado.

La base de datos contiene la siguiente información de los solicitantes:

- Incumplimiento (estatus de pago con el crédito concedido). (**Incumplimiento**)
- Número de hijos. (**Hijos**)
- Número de trabajos formales en los que ha estado. (**Trabajos**)
- Edad del prestatario. (**Edad**)
- Años de educación. (**AniosEdu**)
- Salario. (**Salario**)
- Monto total de las deudas del solicitante. (**Deuda**)

Estos datos son vitales para evaluar la capacidad de pago del solicitante y para estimar el riesgo de crédito asociado. Al evaluar el riesgo de crédito se puede determinar el monto que se puede otorgar y establecer las condiciones del préstamo.

### 1.1. Características específicas de la base de datos

El crédito ofrecido está dirigido a un sector específico de la población, por lo que la base de datos incluye ciertas características específicas, tales como:

- Número de hijos: Personas con hasta 15 hijos.
- Número de empleos formales previos: Se consideran hasta 20 empleos anteriores, incluyendo el trabajo actual.
- Años de educación: Personas con hasta 16 años de estudios, incluyendo aquellas sin estudios o con una licenciatura.
- Edad: Personas mayores a 18 años y menores a 65.
- Salario: Se incluyen personas con salarios entre uno y seis salarios mínimos.
- Monto total de deudas actuales: Se incluyen personas con deudas entre 0MXN y 100,000MXN.

Estas características permiten identificar el perfil de los posibles solicitantes de crédito y ajustar las condiciones del préstamo de acuerdo con sus condiciones específicas.

## 2. Variables Aleatorias y Modelos

Para la generación de los datos, se usaron variables aleatorias discretas y continuas según la variable a simular. Es importante notar que los datos simulados tienen cierta dependencia ante la variable “Incumplimiento”, pues para hacer más realista la base de datos, se atribuyeron ciertas características a los solicitantes, modificando significativamente los parámetros de las variables aleatorias usadas según el incumplimiento.

### 2.1. Variables Aleatorias Utilizadas

- Incumplimiento: Bernoulli con una probabilidad de incumplimiento del 25
- Hijos: Binomial con parámetros  $n = 15$  y probabilidad  $p = 2/15$  para personas cumplidas y  $p = 3/15$  para personas incumplidas.
- Trabajos: Binomial con parámetros  $n = 19$  y probabilidad  $p = 4/15$  para personas cumplidas y  $p = 8/15$  para personas incumplidas. Se suma 1 al valor observado para tener en cuenta el trabajo actual, mientras que los trabajos anteriores al actual se generan mediante la variable aleatoria.
- Edad: Binomial con parámetros  $n = 46$  y probabilidad  $p = 12/46$ , considerando personas mayores de 18 años y sumando 18 al valor generado para que los solicitantes tengan al menos 18 años y hasta 64 años.
- Años de Educación: Binomial con parámetros  $n = 16$  y probabilidad  $p = 12/16$  para personas cumplidas y  $p = 9/16$  para personas incumplidas.
- Salario: El salario de las personas cumplidas se modela mediante una distribución beta con parámetros  $\alpha = 5$  y  $\beta = 3,5$ . Los salarios se ajustan al rango (6310, 31550) para que estén dentro del rango deseado, y se elige esta distribución para sesgar los datos hacia la izquierda (salarios altos dentro del rango considerado). Para las personas incumplidas, se utiliza una distribución beta con parámetros  $\alpha = 3,5$  y  $\beta = 3,5$ , generando una distribución más simétrica y salarios en promedio más bajos que los de las personas cumplidas.
- Deuda: Beta con parámetros  $\alpha = 1$  y  $\beta = 2$ , ajustando los datos al rango (0,75000) para personas cumplidas y (0,100000) para personas incumplidas. Se considera este rango para que las personas cumplidas tengan deudas menores que las incumplidas, y se eligen los parámetros para generar montos de deuda pequeños.

## 3. Ajuste de Modelos (RLM, Logit y Probit)

En el contexto de un problema de regresión lineal múltiple y de modelos lineales generalizados (Logit y Probit), donde la variable dependiente (cumplimiento) sigue una distribución Bernoulli, las betas positivas indican una mayor

probabilidad de incumplimiento (1), mientras que las betas negativas sugieren un mayor cumplimiento (0), siendo este último el nivel de referencia de la variable dependiente.

Esto pues, la función liga inversa de los modelos probit  $[\Phi(\eta_i) (CDF)]$  y logit  $\left[\frac{e^{\eta_i}}{1+e^{\eta_i}}\right]$  ( $\eta_i$  es el componente lineal) son crecientes, por lo que a valores positivos se espera un aumento en la esperanza de la variable dependiente y viceversa. En el caso de la RLM, el efecto es más directo sobre la esperanza de  $y_i$ , pues no existe una función liga de por medio.

Es importante destacar que las betas se obtuvieron mediante el método de mínimos cuadrados ordinarios en Python, y no mediante máxima verosimilitud. Por lo tanto, no se garantiza la normalidad de las estimaciones de las betas, lo que limita la inferencia estadística sobre estas estimaciones.

En el caso en el que se cumplan los supuestos, se puede realizar pruebas de hipótesis individuales de los parámetros mediante distribuciones asintóticas como la Ji-Cuadrada, o la distribución F para pruebas simultáneas. Las primeras pruebas consisten en comprobar si, dado que las demás covariables se encuentran en el modelo, la covariable a comprobar es significativa o no. De no ser el caso, es viable considerar la eliminación de la variable. Esto puede afectar los p-valores individuales cuando se valore de nuevo la significancia del modelo.

Con esto en mente, se ajustaron los modelos correspondientes considerando interceptor, a continuación, se presentan los resultados obtenidos.

### 3.1. Resultados

Los resultados de los modelos son los siguientes:

#### ■ Modelo RLM

OLS Regression Results						
Dep. Variable:	Incumplimiento	R-squared:	0.545			
Model:	OLS	Adj. R-squared:	0.543			
Method:	Least Squares	F-statistic:	238.2			
Date:	Fri, 15 Mar 2024	Prob (F-statistic):	3.47e-167			
Time:	07:49:22	Log-Likelihood:	-186.84			
No. Observations:	1000	AIC:	385.7			
Df Residuals:	994	BIC:	415.1			
Df Model:	5					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	-0.3025	0.078	-3.886	0.000	-0.455	-0.150
Hijos	0.0344	0.005	6.363	0.000	0.024	0.045
Trabajos	0.0154	0.004	4.020	0.000	0.008	0.023
AniosEdu	-0.0116	0.005	-2.431	0.015	-0.021	-0.002
Salario	1.047e-05	1.99e-06	5.266	0.000	6.57e-06	1.44e-05
Deuda	1.179e-05	4.17e-07	28.244	0.000	1.1e-05	1.26e-05
Omnibus:	212.572	Durbin-Watson:	1.898			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	382.457			
Skew:	1.294	Prob(JB):	8.92e-84			
Kurtosis:	4.575	Cond. No.	2.93e+05			
Notes:						
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.						
[2] The condition number is large, 2.93e+05. This might indicate that there are strong multicollinearity or other numerical problems.						

Figura 1: Resultados de la Regresión Lineal Múltiple

## ■ Modelo Logit

```

Optimization terminated successfully.
Current function value: 0.251100
Iterations 8

Logit Regression Results
=====
Dep. Variable:      Incumplimiento    No. Observations:      1000
Model:              Logit             Df Residuals:          994
Method:              MLE              Df Model:              5
Date:               Fri, 15 Mar 2024   Pseudo R-squ.:         0.5526
Time:               07:49:23          Log-Likelihood:        -251.10
converged:           True              LL-Null:               -561.23
Covariance Type:     nonrobust         LLR p-value:           8.439e-132
=====
               coef    std err          z      P>|z|    [0.025    0.975]
-----
const         -8.1090      1.052     -7.711    0.000    -10.170    -6.048
Hijos           0.4007      0.067      5.964    0.000      0.269      0.532
Trabajos        0.1930      0.048      3.981    0.000      0.098      0.288
AniosEdu       -0.1096      0.059     -1.873    0.061     -0.224      0.005
Salario         0.0001     2.57e-05      5.043    0.000      7.91e-05      0.000
Deuda           0.0001     8.79e-06     12.357    0.000      9.14e-05      0.000
=====

```

Figura 2: Resultados del Modelo Logit

## ■ Modelo Probit

```

Optimization terminated successfully.
Current function value: 0.251027
Iterations 7

Probit Regression Results
=====
Dep. Variable:      Incumplimiento    No. Observations:      1000
Model:              Probit            Df Residuals:          994
Method:              MLE              Df Model:              5
Date:               Fri, 15 Mar 2024   Pseudo R-squ.:         0.5527
Time:               07:49:23          Log-Likelihood:        -251.03
converged:           True              LL-Null:               -561.23
Covariance Type:     nonrobust         LLR p-value:           7.848e-132
=====
               coef    std err          z      P>|z|    [0.025    0.975]
-----
const         -4.4033      0.543     -8.116    0.000     -5.467     -3.340
Hijos           0.2233      0.036      6.136    0.000      0.152      0.295
Trabajos        0.1071      0.026      4.076    0.000      0.056      0.159
AniosEdu       -0.0639      0.031     -2.034    0.042     -0.125     -0.002
Salario        6.887e-05   1.36e-05      5.067    0.000      4.22e-05      9.55e-05
Deuda          6.003e-05   4.55e-06     13.192    0.000      5.11e-05      6.89e-05
=====

```

Figura 3: Resultados del Modelo Probit

Al ejecutar el modelo completo (Con solo efectos principales y sin considerar transformaciones sobre las covariables), podemos notar en los tres modelos que todas las variables son significativas a un nivel de confianza del 95 %, pues el P-Value ( $P > [z]$ ) es menor que 0.05, a excepción de la variable “Edad”, cuya interpretación del resultado sería que, no se rechaza la hipótesis nula de que el coeficiente asociado a la covariable edad sea 0. Por lo tanto, podemos no contemplar esa variable en el modelo y, en consecuencia, volver a estimar los coeficientes sin considerar dicha variable.

## 4. Análisis de Odds Ratio (Coeficientes de momios)

Cuando se ajusta un modelo logit o probit, se obtienen coeficientes para cada variable independiente, que representan el cambio esperado en la probabilidad de que la variable dependiente tome el valor de 1, para un cambio unitario en la variable independiente correspondiente, manteniendo las demás covariables constantes y bajo el efecto de la correspondiente función liga.

El coeficiente de momios, también conocido como odds ratio en inglés, es una medida de la magnitud de la asociación entre la variable independiente y la variable dependiente en términos de probabilidades que es la relación entre la probabilidad de que ocurra un evento (comúnmente favorable) y la probabilidad de que no ocurra, y este se suele obtener comúnmente para el modelo logit debido al uso del logaritmo del coeficiente de probabilidades que facilita la interpretación.

La regresión logística es un modelo lineal para el logaritmo de las probabilidades de éxito. Entonces, un aumento de unidad en una variable explicativa resultará en un aumento o disminución de las probabilidades pronosticadas por un factor de  $\exp(B_i)$ , sin importar el valor de la variable explicativa o cuáles son los valores de las otras variables explicativas (suponiendo que no se haya incluido interacciones en el modelo).

En cambio, para el modelo probit, dado que la función liga es la inversa de la función de distribución acumulativa, el valor del ratio cambiará en función del valor de las variables explicativas, por lo que su interpretación o uso ya no es directamente comprensible.

### 4.1. Interpretación de los coeficientes de momios (Modelo Logit)

Para interpretar un coeficiente de momios, se debe considerar su valor y, como en cualquier proceso de inferencia estadística, su intervalo de confianza.

Un coeficiente de momios mayor que 1 indica que la variable independiente tiene una asociación positiva con la variable dependiente, es decir, que aumentar el valor de la variable independiente aumenta las probabilidades de que ocurra el evento de riesgo, en este caso incumplimiento.

Por otro lado, un coeficiente de momios menor que 1 indica que la variable independiente tiene una asociación negativa con la variable dependiente, es decir, que aumentar el valor de la variable independiente disminuye las probabilidades de que ocurra el evento de riesgo. Un coeficiente de momios igual a 1 indica que no hay asociación entre la variable independiente y la variable dependiente.

```

const      0.000301
Hijos      1.492802
Trabajos   1.212824
AniosEdu   0.896180
Salario    1.000129
Deuda      1.000109
dtype: float64

```

Figura 4: Coeficientes de Momios. Modelo Logit

A través de los resultados, las covariables Salario y Deuda son variables descartables, pues sus valores son casi 1, lo cual indica que no hay un efecto significativo ante variaciones de la variable independiente (Salario o Deuda) sobre la dependiente (Cumplimiento).

- Mayor a 1
  - **Hijos:** Se puede afirmar con un 95 % de confianza que, si esta variable aumenta, la persona en cuestión tendrá un 49 % más de probabilidades de incumplir.
  - **Trabajos:** Se puede afirmar con un 95 % de confianza que, si esta variable aumenta, la persona en cuestión tendrá un 21 % más de probabilidades de incumplir.
- Menor a 1
  - **AniosEdu:** Se puede afirmar con un 95 % de confianza que, si esta variable aumenta, la persona en cuestión tendrá un 11.584 % ( $1/0.896 - 1$ ) más de probabilidades de cumplir con el pago del crédito.

## 5. Matriz de confusión

Se presenta la matriz de confusión para evaluar el rendimiento de los modelos:

- Modelo Logit

```

Matriz de confusión para el modelo logit:
[[668  83]
 [ 48 201]]

```

Figura 5: Matriz de confusión para el modelo Logit

A partir de la matriz, se observa lo siguiente:

- El modelo estimó de manera correcta 668 personas como cumplidas.



- Estimó a 83 personas como cumplidas cuando en realidad son incumplidas.
- Estimó a 48 personas como incumplidas cuando en realidad son cumplidas.
- Estimó correctamente a 201 personas como incumplidas.

Es decir, obtuvo 131 estimaciones incorrectas y 869 correctas.

#### ■ Modelo Probit

```
Matriz de confusión para el modelo probit:
[[651 100]
 [ 45 204]]
```

Figura 6: Matriz de confusión para el modelo Probit

A partir de la matriz, se observa lo siguiente:

- El modelo estimó de manera correcta 651 personas como cumplidas.
- Estimó a 100 personas como cumplidas cuando en realidad son incumplidas.
- Estimó a 45 personas como incumplidas cuando en realidad son cumplidas.
- Estimó correctamente a 204 personas como incumplidas.

Es decir, obtuvo 145 estimaciones incorrectas y 855 correctas.

## 6. Curvas ROC y PR

Se muestran las curvas ROC y PR para evaluar el rendimiento de los modelos Logit y Probit.

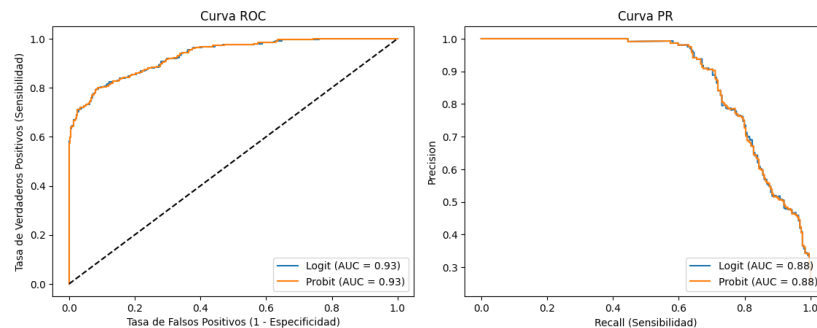


Figura 7: Curvas ROC y PR para el modelo Logit y Probit

El área bajo la curva azul (Logit) es de 0.93 al igual que en la naranja (Probit), lo que indica que el modelo está discriminando de manera satisfactoria, por lo cual ambos modelos están bien calibrados.

## 7. Scoring

Después de aplicar los modelos a la base de datos utilizando Python y obtener las respectivas betas junto con las probabilidades estimadas para la variable de incumplimiento (Y), se procedió a realizar el proceso de Scoring para los tres modelos. Esto implicó dividir la base de datos en un conjunto de entrenamiento y otro de prueba, utilizando un muestreo aleatorio estratificado con proporción 80/20, con el fin de evaluar el desempeño de los modelos. Con las estimaciones obtenidas y los datos reales de incumplimiento, se determinaron los umbrales de riesgo. Al registrar los resúmenes para obtener estos umbrales, se asignaron calificaciones crediticias de manera conservadora en relación con el nivel de riesgo.

### 7.1. Modelo RLM

Intervalos	Niveles	Probabilidad Estimada	% Cumplidos	% Incumplidos
37	AAA	-9.83	6.16	0.0
74	AA+	-6.04	12.31	0.0
111	AA	-3.08	18.47	0.0
148	AA-	-0.24	24.63	0.0
185	A+	2.27	30.45	1.01
222	A	4.75	36.44	1.51
259	A-	7.23	42.43	2.01
296	BBB+	9.75	48.42	2.51
333	BBB	11.88	54.58	2.51
370	BBB-	14.02	60.07	4.52
407	BB+	16.99	65.06	8.04
444	BB	20.0	70.72	9.55
481	BB-	22.05	76.04	12.06
518	B+	25.5	81.53	14.07
555	B	29.09	86.69	17.09
592	B-	34.85	91.85	20.1
629	CCC+	39.8	96.17	25.63
666	CCC	49.88	99.17	35.18
703	CCC-	67.63	100.0	51.26
740	CC	90.16	100.0	69.85
777	C	111.96	100.0	88.44
800	D	146.74	100.0	100.0

Figura 8: Scoring - Modelo RLM

Al evaluar las diferentes opciones, se optó por designar el umbral de color verde como el umbral conservador. Este umbral indica que, con una probabilidad del 34.85 %, se aceptará al menos al 91.85 % de las personas cumplidas, mientras que, al mismo tiempo, se aceptará como máximo al 20.10 % de los incumplidos. En contraste, el umbral amarillo se definió como el umbral moderado, ya que garantiza prácticamente el 100 % de las personas cumplidas, aunque acepta una cantidad no tan considerable de incumplidos. Por último, el umbral naranja se

identificó como el umbral de alto riesgo, dado que, aunque se acepta casi el 100 % de los cumplidos, también se acepta más de la mitad de los incumplidos.

Considerando la postura adversa al riesgo adoptada en este trabajo, se determinó que el umbral conservador representaba el nivel de riesgo ideal. Por lo tanto, se comparó este umbral con los resultados obtenidos en la base de prueba, donde se observó lo siguiente:

	Cumplidos	Incumplidos
<b>Referencia</b>	91.85%	20.10%
<b>Test</b>	95.33%	34.00%

Figura 9: Comparación del Scoring con el umbral conservador - Modelo RLM

Se confirma que el modelo funciona adecuadamente, al igual que el umbral, ya que en esta prueba se registra que, con una probabilidad de hasta 34.71 %, se acepta al 95.33 % de las personas cumplidas, lo cual supera el umbral establecido del 91.85 %. No obstante, se observa que se está aceptando al 34 % de las personas incumplidas, una cifra por encima del límite del 20.10 % definido en nuestro umbral conservador. Esta discrepancia representa una desventaja del modelo en este aspecto.

## 7.2. Modelo Logit

Intervalos	Niveles	Probabilidad Estimada	% Cumplidos	% Incumplidos
37	AAA	0.55	6.16	0.0
74	AA+	0.76	12.31	0.0
111	AA	1.12	18.47	0.0
148	AA-	1.42	24.63	0.0
185	A+	1.81	30.78	0.0
222	A	2.34	36.77	0.5
259	A-	3.14	42.76	1.01
296	BBB+	3.81	48.75	1.51
333	BBB	5.03	54.74	2.01
370	BBB-	6.08	60.57	3.02
407	BB+	7.52	65.89	5.53
444	BB	9.5	71.21	8.04
481	BB-	12.23	76.04	12.06
518	B+	16.44	81.2	15.08
555	B	22.05	86.69	17.09
592	B-	31.13	91.85	20.1
629	CCC+	42.88	95.51	27.64
666	CCC	65.48	99.0	35.68
703	CCC-	88.57	99.83	51.76
740	CC	98.44	100.0	69.85
777	C	99.79	100.0	88.44
800	D	100.0	100.0	100.0

Figura 10: Scoring - Modelo Logit

Misma situación que con el modelo anterior, continuamos asignando los mismos umbrales en las calificaciones B-, CCC y CCC-. Asimismo, escogemos el umbral conservador como nuestro nivel de riesgo idóneo.

Sin embargo, en contraste al modelo previo, se acepta un porcentaje menor de cumplidos e incumplidos que el indicado en el umbral. Esto sugiere que el modelo Logit proporciona un peor ajuste que el modelo RLM.

	Cumplidos	Incumplidos
<b>Referencia</b>	91.85%	20.10%
<b>Test</b>	90.67%	22.00%

Figura 11: Comparación del Scoring con el umbral conservador - Modelo Logit

### 7.3. Modelo Probit

Intervalos	Niveles	Probabilidad Estimada	% Cumplidos	% Incumplidos
37	AAA	0.19	6.16	0.0
74	AA+	0.34	12.31	0.0
111	AA	0.62	18.47	0.0
148	AA-	0.88	24.63	0.0
185	A+	1.2	30.62	0.5
222	A	1.86	36.77	0.5
259	A-	2.63	42.43	2.01
296	BBB+	3.6	48.25	3.02
333	BBB	4.91	54.24	3.52
370	BBB-	6.23	60.23	4.02
407	BB+	8.28	65.56	6.53
444	BB	11.62	70.38	10.55
481	BB-	14.05	75.54	13.57
518	B+	18.39	81.2	15.08
555	B	23.98	86.19	18.59
592	B-	32.71	91.51	21.11
629	CCC+	45.36	95.01	29.15
666	CCC	65.11	99.17	35.18
703	CCC-	90.15	99.83	51.76
740	CC	98.73	100.0	69.85
777	C	99.98	100.0	88.44
800	D	100.0	100.0	100.0

Figura 12: Scoring - Modelo Probit

	Cumplidos	Incumplidos
<b>Referencia</b>	91.51%	21.11%
<b>Test</b>	92.00%	20.00%

Figura 13: Comparación del Scoring con el umbral conservador - Modelo Probit

En cuanto a la selección de umbrales de riesgo, este modelo sigue la misma línea que los dos anteriores. Se observa que el porcentaje de cumplidos aceptados es mayor al del umbral y el de incumplidos aceptados es menor.

Por tanto, el modelo Probit ofrece las mejores especificaciones al obtener un rendimiento mejor al esperado. Por lo cual es el que ofrece mejor ajuste.

## 8. Conclusiones

Basándonos en lo expuesto, se concluye que el modelo Probit se presenta como la elección más adecuada para el análisis de riesgo crediticio. Además de estar en consonancia con los resultados obtenidos en el análisis realizado en Python, este modelo exhibe una serie de ventajas que respaldan su selección.

En primer lugar, es importante destacar que el modelo Probit es el único de los considerados cuyo porcentaje de aceptación de cumplidos e incumplidos se sitúa por encima y por debajo de los niveles de referencia respectivamente. Aunque el porcentaje de aceptación de cumplidos no sea tan elevado como en los otros modelos, es fundamental priorizar el menor nivel de incumplidos aceptados para salvaguardar la integridad de la cartera de crédito.

Esta mejora marginal adquiere especial relevancia en entornos donde incluso las diferencias mínimas en la precisión del modelo pueden tener consecuencias significativas en la toma de decisiones financieras.

Es crucial subrayar que la selección del modelo Probit como el más idóneo para nuestra base de datos se fundamenta en una evaluación exhaustiva de diversas métricas de desempeño. Esto incluye el análisis de la matriz de confusión, las curvas ROC y PR, así como la comparación de las probabilidades estimadas con los umbrales de riesgo establecidos. Este enfoque integral nos proporciona la confianza necesaria para afirmar que el modelo Probit es la mejor opción para nuestras necesidades específicas de evaluación de riesgo crediticio.

## 9. Referencias

- <https://colab.research.google.com/drive/1v16Z6fPtG4orJZKMsq3JTL6270zSsGdt?hl=es#scrollTo=0x2LswInTFq6>