

UNIVERSIDAD DE COSTA RICA

ESCUELA DE MATEMÁTICA

DEPARTAMENTO DE MATEMÁTICA PURA Y CIENCIAS ACTUARIALES
DISTRIBUCIÓN DE PÉRDIDAS

Proyecto Distribución de Pérdidas

BITÁCORAS

Realizado por

Daniel Núñez

Andrés González



UNIVERSIDAD DE
COSTA RICA

EMat

Escuela de
Matemática

1. Bitácora 1

La idea inicial que se planteó va de la mano con calcular la pérdida esperada de un Banco a la hora de otorgar créditos basado en el perfil de la persona. Bajo este contexto se pueden encontrar dos tipos de riesgos:

- 1) El primer riesgo que se distingue es aquel en el que la persona presente una probabilidad alta de que no pague el crédito y que aún así el banco lo apruebe.
- 2) El segundo haciendo alusión al riesgo de que una persona tenga una probabilidad de pagar de vuelta el crédito pero el banco lo niegue, lo que resulta en una pérdida para la entidad bancaria.

Cuatro formas diferentes de reenplantarse la idea anterior son:

- 1) ¿Cuál sería la pérdida esperada de un banco a la hora de determinar la elegibilidad de una persona para un crédito?
- 2) ¿Por qué es importante conocer las probabilidades de pago o de impago de una persona solicitante?
- 3) ¿Cuáles son los factores que más importan para determinar la probabilidad de impago de un individuo?
- 4) ¿Cómo se relacionan las pérdidas con un buen sistema de elegibilidad para créditos?

Base de datos

Para efectos de la investigación se va a trabajar con una base de datos de un banco alemán que contiene información de personas solicitantes de un crédito y con base en esta información se determina su elegibilidad. Los datos fueron recuperados de kaggle y originalmente fueron obtenidos de Penn State Eberly College of Science. Los datos son públicos y son de libre acceso, sin embargo, por motivos de confidencialidad el nombre del banco nunca se menciona. Vale la pena mencionar que los datos no están delimitados temporalmente.

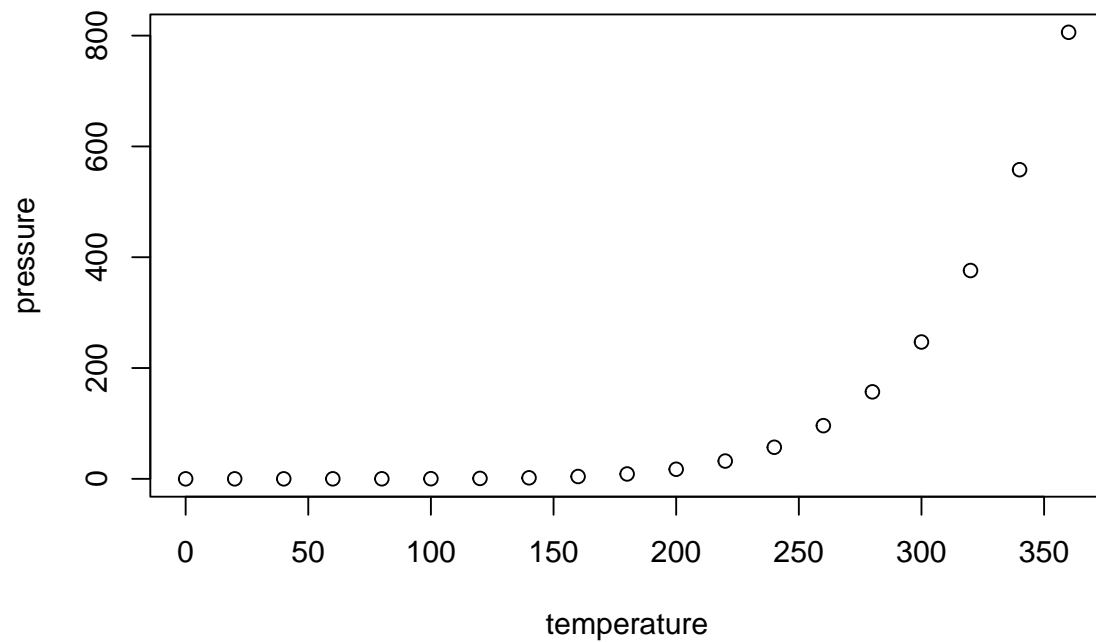
La población de estudio se define como las personas que solicitaron un crédito en esta entidad bancaria ubicada en Alemania mientras que la unidad estadística se define como la persona solicitante de un crédito en el banco alemán estudiado. La muestra para el desarrollo de dicha investigación consta de 1000 individuos. Asimismo, la base cuenta con 21 variables de interés donde en su columna matriz se encuentra la variable binaria de elegibilidad, que toma el valor de 1 si fue elegible y el valor de 0 si no lo fue.

Ahora se procede a dar una leve explicación de cada una de las variables que conforman esta base:

- Elegibilidad: toma el valor de 1 si fue elegible y el valor de 0 si no lo fue.
- Account Balance: variable categórica que toma el valor de 1 si la persona no cuenta con ninguna cuenta en el banco, el valor de 2 si no tiene un balance pendiente con el banco y el valor de 3 si sí tiene un balance pendiente.
- Duration: la duración en meses del crédito solicitado
- Payment Status of Previous Credit: variable categórica que toma el valor de 1 si el individuo presenta problema con el pago del crédito anterior, el valor de 2 si ya lo pagó y el valor de 3 si no tiene problemas con el crédito anterior

1. BITÁCORA 1

- Purpose: variable categórica que toma el valor de 1 si es para un auto nuevo, 2 si es para un auto de segunda mano, 3 si es para una casa/apartamento y 4 si es para cualquier otra cosa.
- Credit Amount: el monto del crédito solicitado en “Deutsche Mark” (DM), que es la unidad monetaria usada en la base.
- Saving/Stock value: toma el valor de 1 si no tiene nada de ahorros o de stock, de 2 si el valor es menor a los 100 DM, 3 si se está en el intervalo [100, 500[DM , 4 si está en [500, 1000[DM y 5 si está arriba de los 1000 DM.
- Length of current employment: variable categórica que toma el valor de 1 si es desempleado, de 2 si tiene menos de año, de 3 si es de 1 a 4 años, de 4 si es de 4 a 7 años y de 5 si es mayor a 7 años.
- Sex/marital status: toma el valor de 1 si es un hombre divorciado, el de 2 si es hombre soltero, el de 3 si es hombre casado/viudo y el de 4 si es mujer
- Guarantors: variable binaria que toma el valor de 1 si la persona tiene un fiador y de 0 si no lo tiene.
- Duration in current address: variable categórica que determina cuánto tiempo lleva la persona viviendo en la última dirección registrada. Toma el valor de 1 si es menos de un año, el de 2 si lleva entre uno y 4 años, el de 3 si lleva entre 4 y 7 años y el de 4 si lleva más de 7 años.
- Most valuable asset: toma el valor de 1 si no tiene ninguno, el de 2 si es un carro, el de 3 si es un seguro de vida y el de 4 si son bienes raíces.
- Edad: edad en años
- Tarjetas de créditos: toma el valor de 1 si el aplicante tiene tarjetas con otros bancos, el valor ed 2 si tiene tarjetas de créditos con empresas y el de 3 si no tiene nada.
- Type of department: toma el valor de 1 si no paga renta/hipoteca, el de 2 en caso de pague renta o hipoteca y el de 3 si es dueño de la vivienda/apartamento.
- No. of credits at this bank: toma el valor de 1 si solo tiene 1, el de 2 si tiene entre 2 y 3 créditos, el de 3 si tiene entre 4 y 5 créditos y el de 4 tiene más de 6 créditos.
- Occupation: 1 en caso de que sea desempleado o no calificado, 2 en caso de que sea un residente permanente no calificado, 3 en caso de que sea una persona calificada y 4 en caso de que sea un ejecutivo/a.
- No. of dependents: número de personas que mantiene. Toma el valor de 1 si son más de 3 y de 2 si son entre 0 y 2 personas.
- Foreign worker: variable binaria que toma el valor de 0 en caso de que sea un trabajador extranjero y 1 en caso de que no lo sea.



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.