

## Incumplimiento de pago de hipoteca

Cuando una persona solicita una hipoteca, los prestamistas (bancos) quieren saber el valor del riesgo que asumirían al prestar dinero. En este caso, el fallo es incumplimiento del cliente. La duración es el tiempo (en meses) entre el comienzo de la hipoteca y el incumplimiento de pago del cliente. Algunos clientes aún no han incumplido, por lo que se les etiquetará como censurados. Los datos se encuentran en el archivo *mortgage*. La variable *time* indica la duración de la observación y la variable *default\_time* indica con 1 si el cliente incumplió la fecha de pago de la hipoteca y con 0 si aún no ha incumplido. La información en detalle de las variables se encuentra en el siguiente sitio:

<http://www.creditriskanalytics.net/datasets-private2.html>

- a) A cada cliente se le observó en distintos puntos de corte en el tiempo, por ejemplo:

```
mortgage[which(mortgage$id=="797"),c('time', 'default_time')]
# A tibble: 7 x 2
  time default_time
  <dbl>         <dbl>
1    24             0
2    25             0
3    26             0
4    27             0
5    28             0
6    29             0
7    30             1
```

Indica que al cliente con ID 797 se le observó desde 24 a 30 meses (periodos acumulativos) y a los 30 meses registró incumplimiento. Sin embargo, hay otros clientes que en todo el periodo de observación no registraron incumplimiento (columna *default\_time* siempre cero). Obtenga registros únicos por cliente (ID) donde se indique, si no presentó incumplimiento, el máximo tiempo en meses observado, y si registró incumplimiento, el menor tiempo en el que presenta incumplimiento.

- b) ¿Cuántos clientes en el periodo de observación respectivo, incumplieron el pago de su hipoteca? ¿Cuántos no? ¿Cómo se distribuyen los tiempos de incumplimiento? Comente.
- c) Realice un gráfico de estimación de la función de sobrevivencia mediante un método no paramétrico, también muestre cómo cambia la cantidad de fallos en el tiempo. Comente.
- d) Respecto al modelo asumiendo riesgos proporcionales, consideraremos la siguiente información:
- *hpi\_time*: Índice de precios de la vivienda
  - *interest\_rate\_time*: Tasa de interés de la hipoteca
  - *uer\_time*: Tasa de desempleo
  - *gdp\_time*: Crecimiento del producto interno bruto (PIB)
  - *REtype\_SF\_orig\_time*: Indica 1 si es vivienda unifamiliar y 0 en otro caso
- d.1) Realice una inspección gráfica del supuesto de riesgos proporcionales respecto al factor de si es una vivienda unifamiliar o no. Comente.
- d.2) Plantee el modelo de Cox y evalúe por variable el supuesto de riesgos proporcionales con el test correspondiente. Interprete aquellos coeficientes tal que  $\exp(\hat{\beta}_i) < 1$ . Comente sobre la significancia de las variables al 5%. ¿Se cumplen los supuestos de riesgos proporcionales?