

Consideremos el ejercicio trabajado en la guía anterior:

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 cada cliente se le observó en distintos puntos de corte en el tiempo. Además considere la variable *gdp\_time* que corresponde al Crecimiento del producto interno bruto (PIB), donde para cada registro cambia en el tiempo:

```
mortgage[which(mortgage$id=="797"),c('time', 'default_time', 'gdp_time')]
# A tibble: 7 x 3
   time default_time gdp_time
  <dbl>         <dbl>    <dbl>
1    24             0     3.12
2    25             0     2.90
3    26             0     2.15
4    27             0     2.36
5    28             0     1.23
6    29             0     1.69
7    30             1     2.27
```

En la guía pasada aplicamos *slice* para trabajar con los últimos valores de las covariables. Sin embargo, resulta de interés plantear un modelo de Cox donde nuestra covariable *gdp\_time* sea incorporada en función del tiempo.

- ¿Qué tipo de variable es *gdp\_time*? ¿Por qué es importante determinar la naturaleza de los predictores en un modelo de análisis de sobrevivencia tiempo dependiente? ¿Qué podría ocurrir?
- Reestructure los datos de modo que sea aplicable un modelo de Cox tiempo dependiente siguiendo el ejemplo visto en clases.
- Plantee el modelo de Cox considerando que *gdp\_time* covariable que varía en el tiempo y plantee otro modelo asumiendo que fuera tiempo independiente. ¿Cuál es la diferencia estructural entre ambos modelos? ¿Qué sucede con el supuesto de riesgos proporcionales de Cox?