

Estadística y Técnicas de Machine Learning con R

Prueba 1

true

1 de agosto de 2018

Contents

1 Modelos de Probabilidad	1
2 Regresión de Cox	1

1 Modelos de Probabilidad

- ¿Cuál es la probabilidad de tener un ingreso superior al promedio en ecuador? Use un modelo *logit*
- ¿Los modelos son significativos? (test de Wald)
- Calcule los efectos marginales e interprete los resultados.

Nota 1: la fuente de datos es: `per12_2010.dta`, la variable ingreso es `ingr1`, sexo es `p02` edad es `p03`, area es `area`, tiempo libre en horas es `p62k`, ingreso recibido del exterior es: `p74b`. En la variable ingreso se debe tomar únicamente los valores entre 0 y 500000.

Nota 2: Crear una variable dicotómica que sea igual a 1 si el ingreso es superior al promedio y 0 caso contrario.

$$ingreso = \beta_0 + \beta_1 sexo$$

$$P(ingreso \geq \bar{ingreso}) = \beta_0 + \beta_1 sexo + \beta_2 edad + \beta_3 area$$

$$P(ingreso \geq \bar{ingreso}) = \beta_0 + \beta_1 sexo + \beta_2 edad + \beta_3 tiempolibre + \beta_4 area$$

2 Regresión de Cox

Se desea estimar un modelo de supervivencia con los datos de crédito de un banco alemán. Particularmente, la probabilidad de supervivencia entre los grupos de cuartiles formados por el monto asignado.

La variable de censura será `Default` y la de tiempo será `duration`. Crea una variable categórica en cuartiles del monto asignado (`amount`)

- Ajustar curvas de KM en forma descriptiva y comentar las diferencias entre Q2 y Q4.
- Ajustar una regresión de riesgos proporcionales de Cox y comentar los resultados, particularmente entre Q2 y Q4

Datos:

```
credit <- read.csv("http://www.biz.uiowa.edu/faculty/jledolter/DataMining/germancredit.csv")
```