

Bootstrap en el modelo de regresión logística

Contents

1 Bootstrap en regresión logística	1
------------------------------------	---

1 Bootstrap en regresión logística

Vamos a calcular, utilizando bootstrap, el standard error y los intervalos de confianza para los parámetros del modelo:

```
d = read.csv("datos/MichelinNY.csv")
str(d)
```

```
## 'data.frame': 164 obs. of 6 variables:
## $ InMichelin : int 0 0 0 1 0 0 1 1 1 0 ...
## $ Restaurant.Name: chr "14 Wall Street" "212" "26 Seats" "44" ...
## $ Food : int 19 17 23 19 23 18 24 23 27 20 ...
## $ Decor : int 20 17 17 23 12 17 21 22 27 17 ...
## $ Service : int 19 16 21 16 19 17 22 21 27 19 ...
## $ Price : int 50 43 35 52 24 36 51 61 179 42 ...
```

Bootstrap:

```
set.seed(99)
B = 500
n = nrow(d)
beta_e = matrix(0, nrow = B, ncol = 5)
for (b in 1:B){
  pos_b = sample(1:n, n, replace = T)
  d_b = d[pos_b,]
  m_b = glm(InMichelin ~ Food + Decor + Service + Price, data = d_b, family = binomial)
  beta_e[b,] = coef(m_b)
}
```

- Standard errors calculados con bootstrap:

```
apply(beta_e, 2, sd)
```

```
## [1] 3.20370046 0.14221078 0.10042618 0.13469161 0.04590046
```

- Intervalos de confianza calculados con bootstrap:

```
alfa = 0.05
apply(beta_e, 2, quantile, probs = c(alfa/2, 1-alfa/2))
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]
## 2.5% -19.322348 0.1867948 -0.08561463 -0.50060459 0.02105848
## 97.5% -6.269294 0.7423821 0.30084908 0.04150634 0.20071627
```

Se puede comprobar que los resultados de bootstrap concuerdan con los obtenidos mediante la distribuciones asintóticas.