Clases 7: MLE

Denise Laroze

7 de noviembre de 2018

CESS - Universidad de Santiago
denise.laroze@usach.cl

Resumen de contenidos

Más allá de las relaciones lineales

Cuando el OLS no es suficiente

Uno de los supuestos más importantes de las regresiones lineales es que la variable dependiente es continua. Pero a veces los fenómenos de interés son discretos

1. Una persona hizo un emprendimiento o no (dummy - $\{1,0\}$)
2. Cuántos emprendimientos tuvo una persona antes de lograr un éxito (ordinal, $1 < 2 < 3$)
3. Cuántos alumnos entraron a la carrera de Contador Público y Auditor como primera opción, por universidad (una cuenta $(1, 2, 3, 4, ...)$).

Relación no-lineal

Para analizar variables dependientes discretas se suele utilizar métodos que apliquen una formula del tipo:

$$y_i = h(x_i, \beta) + \epsilon_i.$$

Donde la función h varía de acuerdo a los supuestos que se están haciendo sobre cómo las variables inciden en Y, que termina por estar asociado a la distribución de Y.

Relación no-lineal binaria

En una relación no lineal binaria a la variable dependiente convencionalmente se le asigna uno de dos valores $\{0,1\}$, por lo que no se puede calcular el cambio esperado en Y por un cambio en una unidad de X. Lo que se calcula, en cambio, es:

$$Y = \left\{ \begin{array}{cc} 1 & \text{Si la respuesta es Sí} \\ 0 & \text{otra cosa} \end{array} \right\}$$

Entonces se estima:

$$p(Y = 1|X) = \Phi(x\beta).$$

Bayes rule

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} = \frac{P(Y,X)}{P(X)}.$$

$$P(Y,X) = P(Y|X)P(X).$$

Queremos saber

Consideremos la probabilidad de que un estudiante se transforme en emprendedor/a. Se denota $Y = 1$, si la persona A se transforma en emprendedor/a y $Y = 0$, si no ocurre. Se asume que la ocurrencia de $Y = 1$, depende de: su nivel socio-económico, experiencias familiares de emprendimiento, aversión al riesgo, etc. Todas estas covariables se identifican como X.

$$p(Y = 1|X) = F(X, \beta).$$
$$p(Y = 0|X) = 1 - p(Y = 1|X) = 1 - F(X, \beta).$$

Entonces, queremos saber cuál es el modelo correcto para $F(X, \beta)$

Queremos saber

La **Probabilidad** de observar $Y^* = 0, 1$, pero lo que realmente observamos es:

$$Y = 1 \text{ si, } Y^* > 0$$

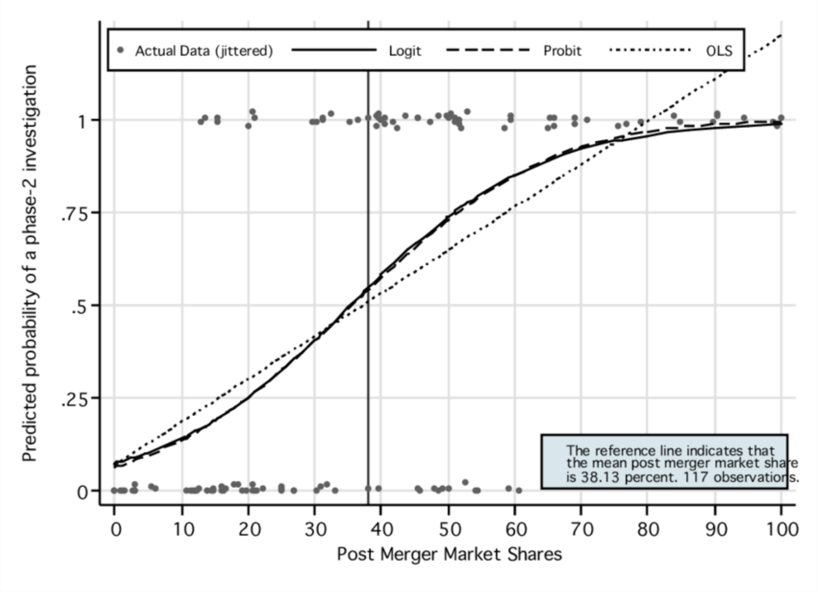
$$Y = 0 \text{ si, } Y^* \leq 0$$

$$p(Y^* > 0|X) = p(X\beta + \epsilon > 0|X) = p(\epsilon > -X\beta|X) :$$

Con distribuciones simétricas:

$$p(\epsilon > X\beta|X) = F(X; \beta)$$

Probabilidades lineales vs logit o probit



Problema - Probabilidades lineales

- Las PL no están limitadas a $\{0, 1\}$
- los errores ya no tienen, ni aproximadamente, una distribución normal y, por tanto, se viola en principio de homoskedasticidad.

Maximum Likelihood Estimation - MLE

Tipos de Distribuciones comunes

Lo que se necesita para estimar $F(X;\beta)$ es una función distribución acumulativa apropiada. Distintas alternativas comunmente usadas son:

Logit

$$p(Y=1|X) = \frac{e^{X\beta}}{1+e^{X\beta}} = \Lambda(X\beta).$$

Probit

$$p(Y=1|X) = \int_{-\infty}^{X\beta} \phi = \Phi(X\beta).$$

donde ϕ es la una función distribución acumulativa de la distribución normal.

Interpretación datos

En lógit la estimación que se hace es un logaritmo de una **proporción de probabilidades** (Odds Ratio - OR)

$$logit = log(OR) = log(\frac{\pi}{1-\pi})$$

donde $\pi = p(Y = 1)$

- Donde $logit = 0$ indica que ambos eventos tienen las mismas probabilidades.
- si, $logit > 0$ entonces la probabilidad de que el evento ocurra es mayor a la probabilidad que no ocurra ($\pi > 0{,}5$)
- si, $logit < 0$ entonces la probabilidad de que el evento ocurra es menor a la probabilidad que no ocurra ($\pi < 0{,}5$)

Interpretación datos II

Como pocas personas piensan en términos del logaritmo de una proporción de probabilidades, el logit se puede transformar en una simple proporción de probabilidades mediante su exponencial.

$$\pi = \frac{exp(logit)}{1 + exp(logit)} = \frac{Odds}{1 + Odds}$$

En ese caso, si

- si $OR = 1$, la probabilidad de una respuesta positiva es *igual* a la probabilidad de una respuesta negativa.
- si $OR > 1$, la probabilidad de una respuesta positiva es **mayor** a la probabilidad de una respuesta negativa.
- si $OR < 1$, la probabilidad de una respuesta positiva es **menor** a la probabilidad de una respuesta negativa.

¿Cuándo se rechaza la hipótesis nula entonces?

Interpretación datos III

Como tb pocas personas piensan en términos proporciones de probabilidades, lo que se hace es:

Presentar ejemplos de cambios en los valores esperados (predicted probabilities) Y^* cuando cambian los valores de las variables independientes de interés.

Miremos este ejemplo:
`https://stats.idre.ucla.edu/r/dae/logit-regression/`

en R

El código para ejecutar regressiones no lineales en R tiene la siguiente estructura:

En general:

`glm(formula, family=familytype(link=linkfunction), data=)`

Para logit (default):

`mylogit<-glm(d x1+x2+x3,data=mydata,family=binomial()`

Exp(mylogit): `exp(coef(mylogit))`

Exp(mylogit) + IC: `exp(cbind(OR = coef(mylogit), confint(mylogit)))`

Para Probit: `glm(d x1+x2+x3,data=mydata,family=binomial(link = "probit")`

Probit

Ejemplo en **Probit:**

`https://stats.idre.ucla.edu/r/dae/probit-regression/`

Fin