Eleccion Binaria

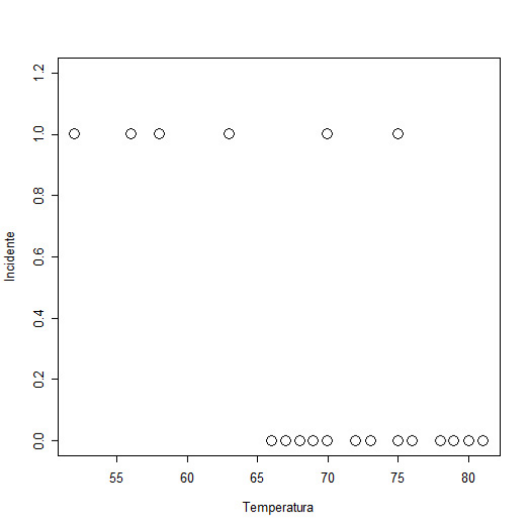
Walter Sosa-Escudero

El desastre del Challenger

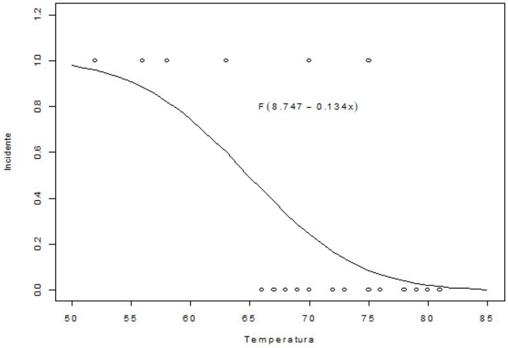
- 28/1/1986: taxi espacial Challenger explota luego de su lanzamiento, todos sus integrantes murieron.
- Causa: baja temperatura de lanzamiento, falla de una pieza llamada .º-ring”
- La temperatura de lanzamiento fue de 31 F.
- Previamente hubieron 23 lanzamientos.

Se podria haber previsto? Cual habria sido el efecto de esperar a que la temperatura suba?

Al menos una falla



Probit



	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	8.74749	3.76809	2.321	0.0203	*
tem	-0.13488	0.05481	-2.461	0.0139	*

Eleccion binaria

- $Y,\ y \in \{0,1\}$, ocurrencia de un evento, i.e., challenger falla, admision a un posgrado, asistir al secundario, etc.
- x vector de K variables explicativas.

Modelo de elección binaria: modelo para:

$$p \equiv Pr(Y = 1|x)$$

Ejemplo: admsion dadas las calificaciones.

Importante: y condicional en x tiene distribucion de *Bernoulli*.

Entonces:

$$E(y|x) = 1\ p + 0\ (1 - p) = p$$

y

$$V(y|x) = p\ (1 - p)$$

El modelo lineal de probabilidad

Un primer modelo podria ser:

$$y = x'\beta + u \qquad \text{con} \qquad E[u|x] = 0$$

- *Modelo lineal de probabilidad (LPM)*.
- β puede ser estimados consistentemente por MCO, regresando y en x en base a $(y_i, x_i), i = 1, \ldots, n$.

Problemas con MLP

- *1) Genera predicciones inconsistentes con una probabilidad:*

  Notar que:

  $$E(y|x) = x'\beta$$

  Problema: $E(y|x) = p$ de modo que $0 \leq E(y|x) = p \leq 1$. Pero en el modelo lineal general $-\infty < x'\hat{\beta} < \infty$.

- *2) u es heterocedastico.*

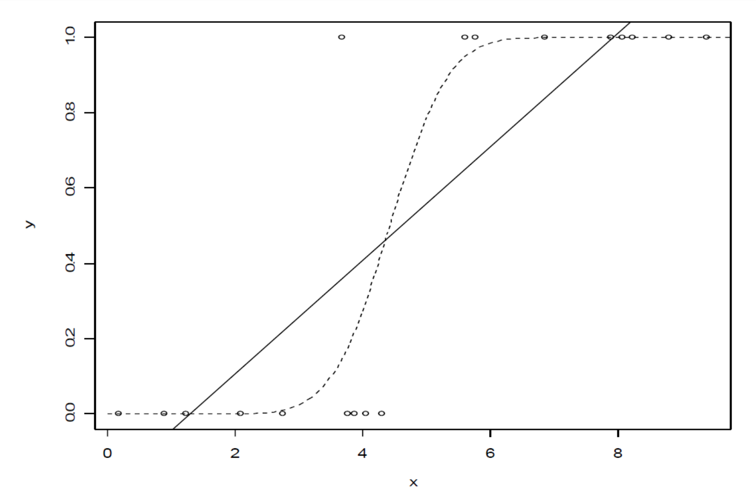
$$V(u_i|x_i) = V(y_i|x_i) = p_i(1 - p_i) = x_i'\beta(1 - x_i'\beta)$$

no constante. Razon para abandonar MCO, es facil de arreglar.

- *3) LPM implica derivadas parciales constantes:*

$$\frac{\partial p}{\partial x_k} = \beta_k.$$

Intuitivamente, en la mayoria de los casos relevantes no queremos esto. Efectos 'debiles en las puntas'. Ejemplos: asistencia al secundario, admision a posgrado.



1.0
0.8
0.6
0.4
0.2
0.0
y
0
2
4
6
8
x

Logits y Probits

El modelo *no-lineal* que propondremos es:

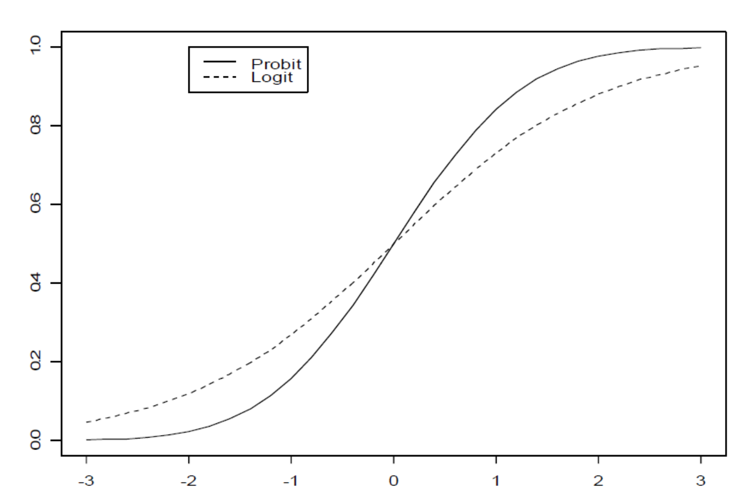
$$p = F(x'\beta)$$

$F(.)$ tiene las siguientes propiedades:

$$F(-\infty) = 0, \; F(\infty) = 1, \; f(z) = dF(z)/dz > 0$$

- Probit:
$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{s^2}{2}} ds$$
- Logit:
$$F(z) = \frac{e^z}{1 + e^z}$$

Notar que LMP corresponde a $F(z) = z$.


Probit
Logit
1.0
0.8
0.6
0.4
0.2
0.0
-3
-2
-1
0
1
2
3

Interpretaciones

$$\frac{\partial p}{\partial x_k} = \beta_k \ f(x_i'\beta),$$

no constante. Notar que:

$$\text{sgn}\left(\partial p/\partial x_k\right) = \text{sgn}\ \beta_k$$

de modo que el signo es interpretable, no su valor

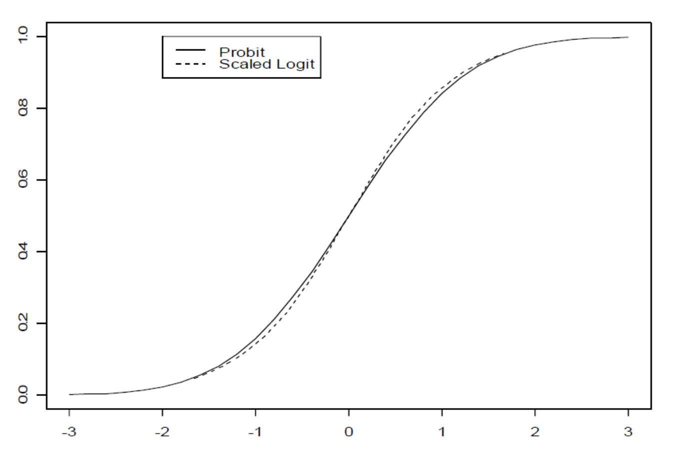
Alternativa: efecto marginal en las medias.

$$\frac{\partial p}{\partial x_k}_{x=\bar{x}} = \beta_k f(\bar{x}'\beta_k),$$

Podriamos hacerlo en cualquier punto. O graficar esta derivada moviendo solo una variable y fijando las restantes en algun punto interesante.

Logits o probits?

- Si X es logistica, $V(X) = \pi^2/3$, entonces $V(Z = X\sqrt{3}/\pi) = 1$. Se puede mostrar que la distribucion de Z (logistica re-escalada) es *muy* similar a la normal estandar.
- En general, los coeficientes del logit exceden a los de probit en $\pi/\sqrt{3}$.
- Utilizar cualquier especificacion.


Probit
Scaled Logit
1.0
0.8
0.6
0.4
0.2
0.0
-3
-2
-1
0
1
2
3

Estimacion e inferencia

$(y_i, x_i), i = 1, \ldots, n$, muestra iid. Y_i tiene distribucion de Bernoulli con $p_i = Pr(y_i = 1)$

La funcion de verosimilitud sera:

$$L(\beta) = \prod_{y_i=1} p_i \prod_{y_i=0} (1 - p_i) = \prod_{i=1}^{n} p_i^{y_i} (1 - p_i)^{1-y_i}$$

y su logaritmo:

$$\begin{aligned} l(\beta) &= \sum_{i=1}^{n} \left[y_i \ln p_i + (1 - y_i) \ln(1 - p_i) \right] \\ &= \sum_{i=1}^{n} \left[y_i \ln F(x_i'\beta) + (1 - y_i) \ln(1 - F(x_i'\beta)) \right] \end{aligned}$$

Las condiciones de primer orden son:

$$\sum_{i=1}^{n} \frac{(y_i - F_i) f_i \, x_{ki}}{F_i(1 - F_i)} = 0, \qquad\qquad k = 1, \ldots, K$$

con $F_i \equiv F(x_i' \hat{\beta}),\ f_i \equiv f(x_i' \hat{\beta})$.

- Es un sistema de K ecuaciones *no-lineales* con K incognitas.
- No satisface las condiciones del teorema de la funcion implicita: es imposible 'despejar' $\hat{\beta}$.
- Existe una solucion siempre que los regresores sean linealmente independientes y no haya un *clasificador perfecto*. Ver notas de clase.

Tests de especificacion

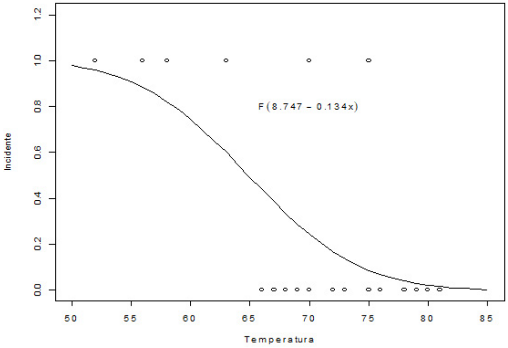
Habiendo estimado los parametros por maxima verosimilitud:

$$\sqrt{n}\left(\hat{\beta}_{MV} - \beta_0\right) \xrightarrow{d} N(0, V_{MV})$$

- *Significatividad individual:* $H_0 : \beta_k = \beta_{0k}$ vs. $H_A : \beta_k \neq \beta_{0k}$

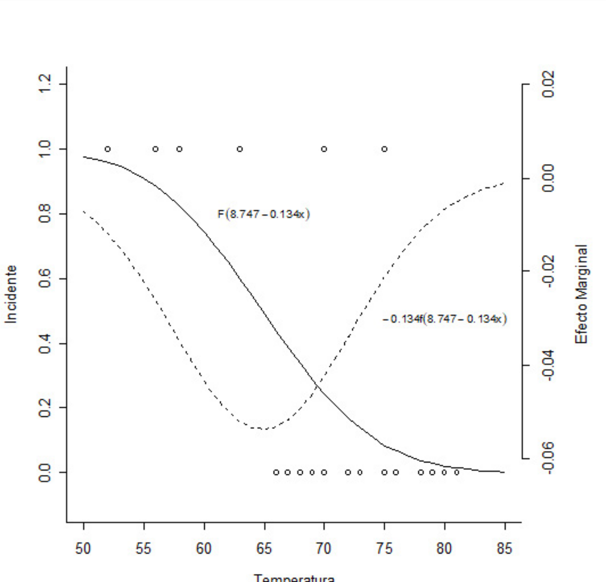
$$t_k = \frac{\hat{\beta}_k - \beta_{k0}}{\sqrt{\hat{V}(\hat{\beta}_k)/n}} \xrightarrow{d} N(0, 1)$$

Ejemplo: Challenger (Probit)



	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	8.74749	3.76809	2.321	0.0203	*
tem	-0.13488	0.05481	-2.461	0.0139	*

Efectos marginales



Probits via variables latentes

$$y^* = x'\beta^* + u, \qquad u \sim N(0, \sigma^2)$$

- Si se observa $(y_i^*, x_i),\ i = 1, \ldots, n$ es posible estimar β y σ^2 consistentemente.
- Que es posible estimar si se observa $(y, x_i),\ i = 1, \ldots, n$, con $y_i = 1[y_i^* > 0]$?

Si $y^* = x'\beta^* + u$:

$$\begin{aligned} P(y=1|x) &= P(y^* > 0|x) \\ &= P(u > -x'\beta^*|x) \\ &= P(u/\sigma < x'\beta^*/\sigma \mid x) \qquad \text{(Simetria)} \\ &= \Phi(x'\beta) \end{aligned}$$

con, $\beta \equiv \beta^*/\sigma$ y $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2}\, dz$.

- $P(y=1|x) = \Phi(x'\beta)$ es el modelo probit.
- Es posible estimar β por MV en base a $(y_i, x_i), i = 1, \ldots, n$.

$$y^* = x'\beta^* + u, \qquad u \sim N(0, \sigma^2)$$

- Con (y_i^*, x_i) es posible recuperar β^* y σ^2.
- Con (y_i, x_i) solo $\beta = \beta^*/\sigma^2$.
- σ^2 y β^* no estan identificados en una muestra (y_i, x_i).
- *Ejemplo:* $\beta^* = 10$ y $\sigma^2 = 2$ producen los mismos (y_i, x_i) que $\beta^* = 5$ y $\sigma^2 = 1$.

MLP vs Probit (Logit)

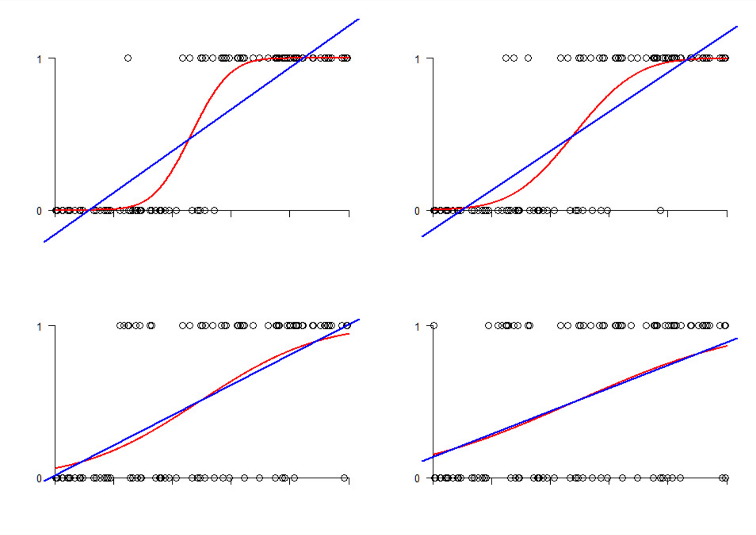
Variables latentes. $y^* = x'\beta^* + u$:

$$P(y = 1|x) = \Phi(x'\beta)$$

con, $\beta \equiv \beta^*/\sigma$. Recordar

$$\frac{\partial p}{\partial x_k} = \beta_k \, f(x_i'\beta),$$

Entonces, si $\sigma^2 \to \infty$, $\partial p/\partial x_k \to 0$. Idea: cuando σ^2 es alta, MPL y Probit son muy dificiles de distinguir


1
0
1
0
1
0
1
0