

Respuesta de la modelización

La respuesta del proceso de modelización, aún en el caso más sencillo, es algo compleja:

RESPUESTA o SOLUCIÓN

MODELO + INCERTIDUMBRE

- Por ejemplo,
 - Objetivo: Predecir el valor de Y que se espera observar
 - MODELO: "valor esperado de Y"= 5,95
 - RESPUESTA:

El valor de Y que se espera observar es 5,95 más un error del modelo o incertidumbre.





Añadir información al modelo

- Modelos menos sencillos se basan en suposiciones sobre el comportamiento de los datos.
 - En la variable objetivo y/o en otras
 - Relaciones con otras variables
- Los casos más habituales son:
 - Suponer una familia de leyes de probabilidad como candidatas a describir la variable respuesta u objetivo.
 - Suponer Y sigue una ley normal
 - Suponer que el valor esperado de Y sigue una ley normal, etc.
 - Fijar una relación lineal con otras variables
 - Suponer que el valor de Y es una proporcional al de X
 - Suponer que el valor esperado de Y es una combinación lineal de X, etc.



Ejemplo de añadir información al modelo

Objetivo: Predecir el valor de Y

Supongamos que Y sigue una distribución Normal
 RESPUESTA = N (5.95, 2.6²)
 RESPUESTA = 5.95 + N(0, 2.6²)
 RESPUESTA = MODELO + INCERTIDUMBRE

• Supongamos que el valor de Y depende de X

MODELO = 0.5 X + 4.3

**RESUESTA = 0.5 X + 4.3 + N(0, 2.42)*

RESPUESTA = MODELO + INCERTIDUMBRE

Х	Υ	Yc
0,5	2,9	1
0,2	0,6	1
1,8	5,3	1
0,5	6,1	1
3	8,4	0
2,1	8,6	0
4,6	3,2	1
4,2	6,7	0
3,9	9,1	0
5,1	5,9	1
6,6	8,2	0
6,8	6,4	1



Modelo de regresión básico: modelo lineal

RESPUESTA = MODELO + INCERTIDUMBRE

"pronóstico sobre Y" = $a \cdot X + b + error$

error tiene ley $N(0, \sigma^2)$

- Objetivos que contempla:
 - Predecir/explicar el valor de Y bajo el supuesto de normalidad
 - Predecir/explicar el valor que se espera observar de Y
- Calibración:
 - Parámetros: a, b, σ
 - Método habitual: OLS



Modelo de regresión básico: modelo lineal

- RESPUESTA = MODELO + INCERTIDUMBRE
- Valores predichos
- Residuos del modelo
- MSE

		Modelo 1		
Х	Υ	predicho	residuo	MSE
0,5	2,9	4,5	-1,6	5,0
0,2	0,6	4,4	-3,8	
1,8	5,3	5,2	0,1	
0,5	6,1	4,5	1,6	
3	8,4	5,8	2,6	
2,1	8,6	5,4	3,2	
4,6	3,2	6,6	-3,4	
4,2	6,7	6,4	0,3	
3,9	9,1	6,3	2,8	
5,1	5,9	6,9	-1,0	
6,6	8,2	7,6	0,6	
6,8	6,4	7,7	-1,3	



Modelo de clasificaión básico: modelo logístico

RESPUESTA = MODELO + INCERTIDUMBRE
$$log(p/(1-p)) = a \cdot X + b + error$$
 error tiene ley $N(0, \sigma^2)$

- Predecir/explicar las probabilidades de cada uno de los valores de la variable Y
 - Si solo toma dos valores: p y (1-p)
- Calibración:
 - Parámetros: a, b, σ
 - Método habitual: OLS
- La respuesta del modelo son los logit: log(p/(1-p)), por lo tanto

$$p = \exp(a.X+b) / (\exp(a.X+b) + 1)$$





Calibración del modelo



