

Auxiliar Pre-Control

Daniel Olcay

8 de septiembre de 2014

Índice

- ▶ Test significancia individual
- ▶ Test múltiples restricciones
- ▶ Test de Chow

Test de hipótesis

- ▶ Estamos interesados en testear diferentes configuraciones en el modelo

$$Y = X\beta + \varepsilon$$

- ▶ Se tienen N observaciones y K parámetros (Algunas referencias usan K como el número de parámetros asociados a variables independientes)
- ▶ Valor p = mínimo nivel de significancia (α) al cual se rechazaría H_0 . Si $p < \alpha$, rechazamos H_0 , pues el nivel de significancia requerido está por debajo del mínimo que podríamos pedir

Significancia individual

- ▶ Estamos interesados en testear si el valor de cierto parámetro β_k es estadísticamente distinto de 0 (intervalo de confianza no contiene al 0)
- ▶ Se puede mostrar que

$$\frac{\hat{\beta}_k - \beta_k}{\hat{SD}(\hat{\beta}_k)} \sim t_{N-K}$$

- ▶ Bajo $H_0 : \beta_k = 0$, se tiene que

$$T = \frac{\hat{\beta}_k}{\hat{SD}(\hat{\beta}_k)} \sim t_{N-K}$$

- ▶ Se rechaza H_0 cuando $|T| > t_{\alpha/2, N-K}$

Test restricciones lineales múltiples

- ▶ Hay veces que es de interes testear un set de restricciones lineales de la forma $R_{q \times K} \beta_{K \times 1} = r_{q \times 1}$
- ▶ Se puede mostrar que el estadístico para realizar el test de restricciones múltiples es

$$\frac{1}{q\hat{\sigma}^2} (R\hat{\beta} - r)' (R(X'X)^{-1}R')^{-1} (R\hat{\beta} - r) \sim F_{q, N-K}$$

- ▶ Ejemplo: En el modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i$$

Testear la hipótesis nula múltiple $\beta_1 + \beta_2 = 3$ y $\beta_3 - 2\beta_4 = 0$

- ▶ $R?$, $r?$, $q?$, $K?$

Significancia global

- ▶ La hipótesis nula $H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_K = 0$ es posible de ser testeada mediante el estadístico F (notar que no es de interes testear la constante)
- ▶ En dicho caso, el estadístico es

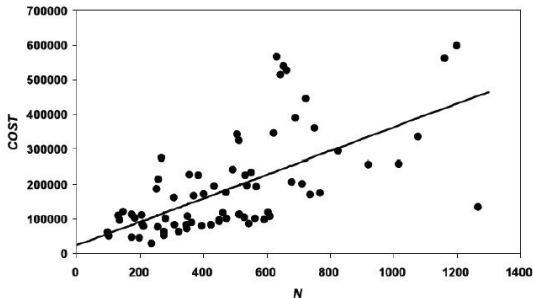
$$\frac{R^2(N - K)}{(1 - R^2)(K - 1)} \sim F_{K-1, N-K}$$

- ▶ En caso de tener que un modelo restringido (R) es un caso particular de un modelo no restringido (NR), el estadístico es

$$\frac{(R_{NR}^2 - R_R^2)(N - K_{NR})}{(1 - R_{NR}^2)q} \sim F_{q, N-K_{NR}}$$

Test de Chow

- Supongamos que tenemos un modelo que intenta explicar el costo en el que incurre un centro de enseñanza en función del número de estudiantes



Test de Chow (1)

► Resultados regresión

```
. reg COST N
```

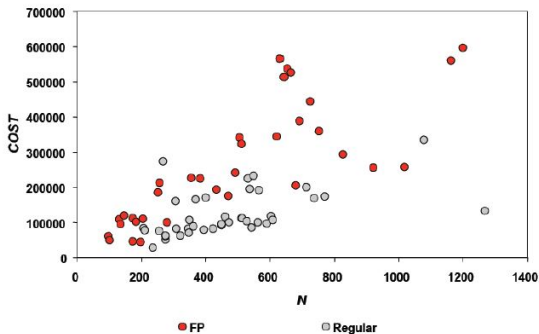
Source	SS	df	MS
Model	5.7974e+11	1	5.7974e+11
Residual	8.9160e+11	72	1.2383e+10
Total	1.4713e+12	73	2.0155e+10

Number of obs = 74
F(1, 72) = 46.82
Prob > F = 0.0000
R-squared = 0.3940
Adj R-squared = 0.3856
Root MSE = 1.1e+05

COST	Coef.	Std. Err.	t	F> t	[95% Conf. Interval]	
N	339.0432	49.55144	6.842	0.000	240.2642	437.8222
_cons	23953.3	27167.96	0.882	0.381	-30205.04	78111.65

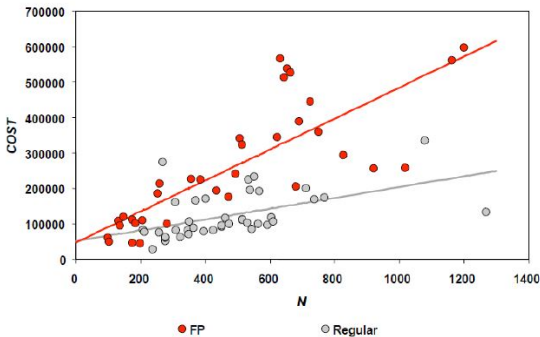
Test de Chow (2)

- ▶ Sin embargo, si desagregamos por tipo de institución (regular y privada) se tiene...



Test de Chow (3)

- Si ajustamos una curva para cada tipo de institución?



Test de Chow (4)

```
. reg COST N if PP==1
```

Source	SS	df	MS
Model	6.0538e+11	1	6.0538e+11
Residual	3.4895e+11	32	1.0905e+10
Total	9.5433e+11	33	2.8919e+10

```
Number of obs = 34
F( 1, 32) = 55.52
Prob > F = 0.0000
R-squared = 0.6344
Adj R-squared = 0.6229
Root MSE = 1.0e+05
```

COST	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
N	436.7769	58.62085	7.451	0.000	317.3701 556.1836
_cons	47974.07	33879.03	1.416	0.166	-21035.26 116983.4

```
. reg COST N if PP==0
```

Source	SS	df	MS
Model	4.3273e+10	1	4.3273e+10
Residual	1.2150e+11	38	3.1973e+09
Total	1.6477e+11	39	4.2249e+09

```
Number of obs = 40
F( 1, 38) = 13.53
Prob > F = 0.0007
R-squared = 0.2626
Adj R-squared = 0.2432
Root MSE = 56545
```

COST	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
N	152.2982	41.39782	3.679	0.001	68.49275 236.1037
_cons	51475.25	21599.14	2.383	0.022	7750.064 95200.43

Se justifica estadísticamente la inclusión de un modelo diferenciado para instituciones regulares y privadas?

Test de Chow (5)

- ▶ Básicamente, el test de Chow verifica si se justifica estadísticamente el uso de una regresión conjunta, en contraste con el modelo separado por submuestras
- ▶ Se basa en el cálculo de SCR del modelo conjunto (pooled) y los modelos separados por submuestras (1 y 2)
 1. Cálculo de SCR modelo conjunto (SCR_p)
 2. Cálculo de SCR de modelos restringidos (SCR_1 y SCR_2)
 3. Cálculo del estadístico

$$\frac{(SCR_p - (SCR_1 + SCR_2))(N - 2K)}{(SCR_1 + SCR_2)K} \sim F_{K, N-2K}$$

- ▶ H_0 : No hay cambio estructural, es decir, los coeficientes asociados al modelo pooled son iguales a ambos modelos por submuestras

Test de Chow (6)

- ▶ Metiendo todo a la fórmula (omitiendo los 10^{11})...

$$\begin{aligned}F_{2,70} &= \frac{(SCR_p - (SCR_1 + SCR_2)) (N - 2K)}{(SCR_1 + SCR_2)K} \\&= \frac{(8,91 - (3,49 + 1,22)) (74 - 2 \cdot 2)}{(3,49 + 1,22) \cdot 2} \\&= 31,2\end{aligned}$$

- ▶ Con un nivel de significancia de un 10 %, el valor crítico de $F_{2,70}$ es de 7,8
- ▶ Se rechaza la hipótesis nula, por lo que se justifica estadísticamente un cambio estructural

Eso es todo...ahora la auxiliar

Suerte en el Control! 😬😬 😊😊