

Tema 4

Generalización: regresión lineal múltiple

(actualizadas el 27-10-2022)

Tema 4. Generalización: regresión lineal múltiple

4.1 El modelo de regresión lineal múltiple

4.2 Interpretación de coeficientes

4.3 Medidas de bondad de ajuste: R^2 ajustado

4.4 Contrastes de hipótesis sobre un conjunto de parámetros: estadístico F

4.5 Predicción puntual y por intervalos

Bibliografía

- Ezequiel Uriel (2013): Capítulo 3 y Capítulo 4 (epígrafes 4.3 y 4.5)
- Wooldridge (2015): Capítulo 3 (epígrafes 3.1 y .32) y Capítulo 4 (epígrafes 4.4 a 4.5)
- Stock y Watson (2012): Capítulo 6 y Capítulo 7 (epígrafes 7.1 a 7.3)

4.1 El modelo de regresión lineal múltiple

El MRLM es muy similar al MRLS

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_k x_{ki} + u_i$$

¿para qué necesitamos el MRLM? ¿no nos vale con el MRLS?

- ¿Qué ocurre si la variable endógena depende de más de una variable como explicativa de su comportamiento?
- En ese caso, debemos incluir esos otros factores explicativos explícitamente en el modelo (si no lo hacemos, tendríamos un problema de omisión de variables relevantes)

El MRLM

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_k x_{ki} + u_i$$

- Como veis, el MRLM es muy parecido al MRLS pero tiene $(k - 1)$ variables explicativas y k **regresores**

Ventajas de la regresión múltiple

- Es evidente que y puede ser influenciada o explicada por más de una variable, por lo tanto necesitamos ampliar el modelo, para así poder controlar explícitamente por otras variables y **estimar con mayor precisión** el efecto de una x en y .
- El MRLM, al introducir más regresores, puede explicar una mayor parte de la variación en y . De esta forma también puede proporcionar **mejores predicciones** de y .
- Otra ventaja adicional es que permite incorporar **relaciones funcionales** entre regresores y regresando **muy generales**. Por ejemplo:

$$\text{consumo} = \beta_1 + \beta_2 \text{renta} + \beta_3 \text{renta}^2 + u$$

Terminología en el MRLM

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_k x_{ki} + u_i$$

- Seguimos refiriéndonos a la variable y como variable dependiente, variable a explicar, **regresando** , ...
- Seguimos refiriéndonos a la variables x_k como variables independientes, variables explicativas o **regresores** sólo que tenemos más.
- Seguimos refiriéndonos a la variable u como término de error o **perturbación** aleatoria
- β_1 sigue siendo el término independiente
- $\beta_2, \beta_3, \dots, \beta_k$ siguen siendo los parámetros de "pendiente" y son **parámetros**, que no conocemos, y queremos estimar

4.2 Interpretación de coeficientes

Es similar a la del MRLS pero ...

los parámetros (β)

En el **Modelo teórico**: $y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k + u$

- β_1 **sigue siendo** el término independiente.
 - Gráficamente sería la ordenada en el origen; es decir sería el valor de y si el resto de variables (x_2, x_3, \dots, x_k , además de u) fuesen cero.
 - Generalmente no tiene una interpretación con sentido económico o teórico.
- β_j es el parámetro que acompaña a la variable x_j .
 - Gráficamente es la pendiente (en el espacio y, x_j)
 - Matemáticamente es la derivada parcial de y respecto de x_j
 - **Económicamente** representa, o es, el efecto marginal de x_j sobre y ; es decir, indica cuantas unidades aumentaría y si, ceteris paribus, x_j aumentase en 1 unidad.

y los estimadores y estimaciones ($\hat{\beta}$) ¿serán "iguales"?

Interpretación de los estimadores

$$\hat{y} = \hat{\beta}_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3 + \dots + \hat{\beta}_k x_k$$

- Los estimadores Siguen siendo las expresiones con las que obtener estimaciones de los β .
- Los estimadores siguen obteniéndose al minimizar la SCR

¿Serán iguales los estimadores en el MRLS y en el MRLM?

- NO, no son iguales.
- Piensa que ahora, aunque el problema teórico para obtener los estimadores conceptualmente es idéntico al que planteamos en el MRLS, en la práctica tendremos que derivar respecto a los k -estimadores, así que en lugar de 2, tendremos k derivadas, k ecuaciones normales y k - estimadores.
- Simplemente os pondré uno de los estimadores:

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x}_2 - \hat{\beta}_3 \bar{x}_3 - \dots - \hat{\beta}_k \bar{x}_k$$

¿Serán iguales las estimaciones en el MRLS y en el MRLM?

- Evidentemente NO.
- Imagina que hemos estimado por MCO los siguientes 2 modelos:

$$\hat{y} = 5.1 + 3.2x_2$$

$$\hat{y} = 7 + 2.5x_2 + 0.27x_3$$

- Como veis, tenemos 2 modelos estimados ¿Cuál es el efecto de x_2 sobre y ? ¿Es 3.2 o 2.5?

4.3 Medidas de bondad de ajuste: R^2 ajustado

$\overline{R^2}$ es similar al R^2 pero ... y ...

R^2 para comparar modelos

- En el tema 3 vimos el coeficiente de determinación R^2 como medida de bondad de ajuste, pero muchas veces se usa para **comparar modelos**
- Estrictamente, para poder comparar modelos en base al R^2 , los modelos han de tener el mismo **regresando**, haber sido estimados con el mismo tamaño muestral y el **mismo número de regresores**.

¿por qué el mismo número de regresores?

- R^2 **nunca disminuye si añadimos regresores** al modelo, sino que generalmente aumenta.
- La inclusión de una variable, aunque sea muy poco explicativa, incrementa el R^2 , pero reduce $(N - k)$ los grados de libertad.
- Si lo que se desea es comparar entre modelos diferentes, con igual regresando, pero diferente número de variables explicativas, **necesitamos otro estadístico**: el R^2 -corregido o $\overline{R^2}$

R^2 -corregido o $\overline{R^2}$

- $\overline{R^2}$ se construyó expresamente para penalizar a los modelos que añaden variables al modelo, de forma que $\overline{R^2}$ sólo aumenta si el nuevo regresor explica lo suficiente la varianza total.

$$\overline{R^2} = 1 - \frac{SCR/(N - k)}{SCT/(N - 1)}$$

- Por construcción, $\overline{R^2}$ penaliza a los modelos con elevado número de regresores (k)
- R^2 **no** puede disminuir al aumentar el número de regresores ($\uparrow k$), pero $\overline{R^2}$ sí, ya que ...
- ... cuando se introduce un regresor más al modelo, SCR generalmente cae, pero también lo hace los grados de libertad ($N - k$), de forma que el cambio en $\overline{R^2}$ no está determinado

R^2 y $\overline{R^2}$ están relacionados

$$\overline{R^2} = 1 - (1 - R^2) \frac{(N - 1)}{(N - k)}$$

- Por lo tanto:
 - si $k = 1$, entonces, $\overline{R^2} = R^2$
 - si $k > 1$, entonces, $\overline{R^2} < R^2$
 - si $R^2 = 1$, entonces $\overline{R^2} = 1$
 - la cota superior del $\overline{R^2}$ es 1, igual que la del R^2
 - $\overline{R^2}$ puede ser negativo; por ejemplo si $k = 2$ y $R^2 = 0$

Recuerda que ...

- $\overline{R^2}$ se puede utilizar para comparar modelos con distinto número de regresores, pero, para poder compararlos, los modelos **tienen que tener** el mismo regresando.
- La razón es sencilla: si el regresando fuese distinto la SCT también y, por tanto no se podrían hacer comparaciones

4.4 Contrastes de hipótesis sobre un conjunto de parámetros: estadístico F

el t -ratio solo puede contrastar hipótesis con una sola restricción

Recordando para qué servía el t-ratio

- El t-ratio, tal y como lo hemos visto, sirve para plantear contrastes de hipótesis en los que la H_0 tiene **una única restricción** y ... además en esa restricción aparece un sólo parámetro (β)
- Por ejemplo: $H_0 : \beta_2 = 0$ o $H_0 : \beta_2 = 3$ o,
- Es decir, sólo podemos utilizar el t-ratio para contrastes del tipo: $H_0 : \beta_i = valor$

El t-ratio no sirve para todas las H_0

- Repetimos: el **t-ratio** sólo sirve para realizar contrastes de hipótesis en los que la H_0 tiene **una única restricción** y ... además en esa restricción **aparece un sólo parámetro** (β)
- En otros casos tendremos que usar otro estadístico: el **estadístico F**, que veremos en breve.

Por ejemplo, el t-ratio no sirve para ...

Hipótesis nulas con **1 restricción** pero en la que **aparecen varios parámetros**:

- $H_0 : \beta_2 + \beta_3 = 1$
- $H_0 : \beta_2 = \beta_3$
- $H_0 : \beta_2 = \beta_3 + \beta_5$

Hipótesis nulas con **más de 1 restricción**

- $H_0 : \beta_2 = \beta_3 = 0$
- $H_0 : \beta_2 = 4 \quad \text{y} \quad \beta_3 = -7$
- $H_0 : \beta_2 = 4 \quad \text{y} \quad \beta_3 + \beta_5 = \beta_4$

Restricciones lineales múltiples

- A veces, es interesante o necesario contrastar **hipótesis compuestas de varias restricciones** ... y esto no podemos hacerlo con el t -ratio.
- Un caso típico (y el más sencillo) de restricciones múltiples son las “restricciones de exclusión”: queremos saber si un grupo de variables independientes no tienen efecto parcial en el regresando.
- Por ejemplo, si estamos trabajando con el modelo:

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

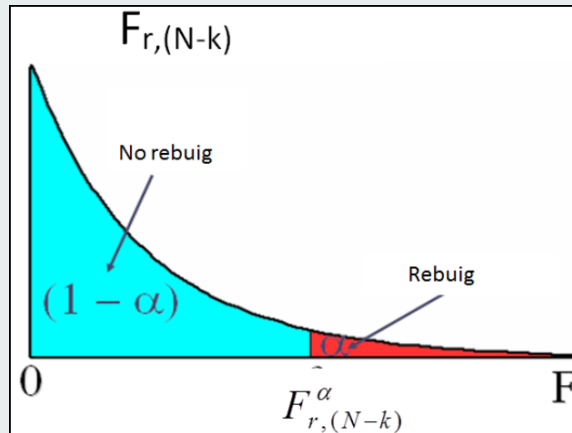
- Igual queremos contrastar la siguiente $H_0 : \beta_3 = \beta_4 = 0$
- La H_0 tiene 2 restricciones ($r = 2$) y estamos planteando si los parámetros β_3 y β_4 son **conjuntamente no significativos**, ya que sus efectos parciales serían nulos.

Restricciones conjuntas: ¿qué forma toma la alternativa?

- Para la siguiente $H_0 : \beta_3 = \beta_4 = 0$, la hipótesis alternativa será
- $H_1 : \exists \beta_j \neq 0$ para $j = 3, 4$
- Fíjate que **no es necesario** que todos ellos sean distintos de cero, con que uno de ellos sea no-nulo ya se incumpliría la H_0
- Es decir, la H_1 se define como la **negación de la nula** ($H_1 : H_0$ no es cierta); o sea, el contraste se construye de forma que detecte cualquier alejamiento de la nula.
- Puede ser tentador contrastar la anterior H_0 mediante una sucesión de contrastes individuales con el estadístico t , pero esta opción no es apropiada.
- Necesitamos un medio de **contrastar las restricciones conjuntamente**. Existe la posibilidad de que ninguna de las tres variables sean individualmente significativas pero, si lo sean conjuntamente

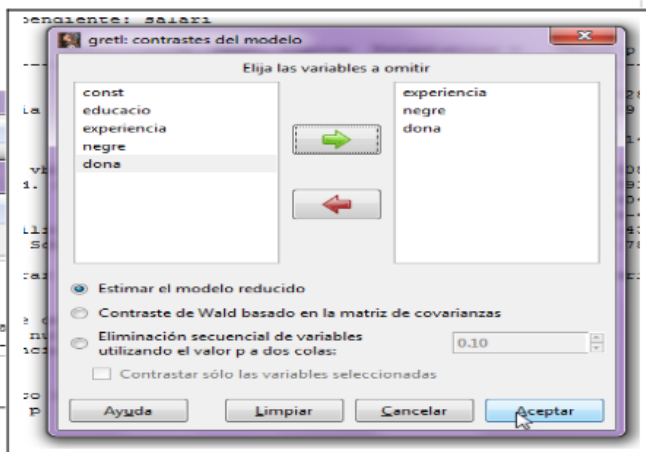
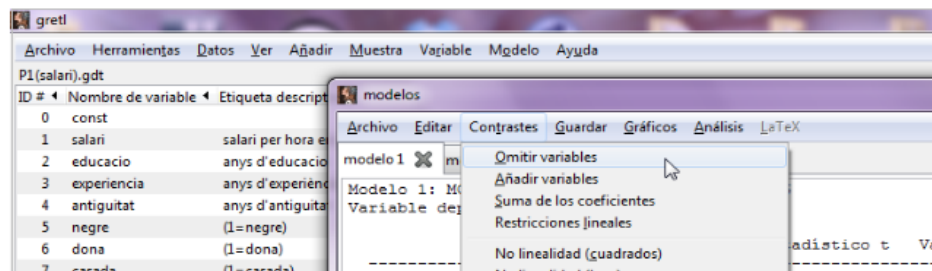
Restricciones conjuntas: ¿qué estadístico usamos?

- ¿Qué estadístico usamos? El **estadístico F** , del cual ni siquiera os voy a mostrar su expresión, ya lo calculará Gretl por nosotros.
- El estadístico F , bajo la H_0 , y recordando que se han de cumplir las h.e.b, se distribuye como: $F \rightarrow F_{r, (N-k)}$
- De forma que (para un α dado) se rechazará la H_0 si el valor muestral del estadístico F excede el valor crítico de las tablas (el correspondiente percentil de la distribución $F_{r, (N-k)}$)



- Si se rechaza la H_0 , se dirá que las variables son estadísticamente **significativas de forma conjunta** al nivel de significatividad α

Restricciones de exclusión con Gretl



Contraste sobre el Modelo 1:

Hipótesis nula: los parámetros de regresión son cero para las variables
experiencia, negre, dona

Estadístico de contraste: $F(3, 521) = 36.3586$, Valor p $2.39033e-021$

Al omitir variables mejoraron 0 de los 3 estadísticos de selección de
modelos considerados.

Modelo 2: MCO, usando las observaciones 1-526

Variable dependiente: salari

	Coefficiente	Desv. Típica	Estadístico t	Valor p
const	-0.904852	0.684968	-1.321	0.1871
educacio	0.541359	0.0532480	10.17	2.78e-022 ***

Media de la vble. dep.	5.896103	D.T. de la vble. dep.	3.693086
Suma de cuad. residuos	5980.682	D.T. de la regresión	3.378390
R-cuadrado	0.164758	R-cuadrado corregido	0.163164
F(1, 524)	103.3627	Valor p (de F)	2.78e-22
Log-verosimilitud	-1385.712	Criterio de Akaike	2775.423
Criterio de Schwarz	2783.954	Crit. de Hannan-Quinn	2778.764

Relación entre los estadísticos t y F

- ¿Qué ocurre si utilizamos el estadístico F para contrastar la significatividad estadística de una sola variable explicativa?
- Por ejemplo, utilizar el estadístico F para, por ejemplo, contrastar $H_0 : \beta_j = 0$.
- Ya sabemos que en este caso podemos utilizar el t -ratio. Entonces ¿hay dos formas de contrastar la misma hipótesis? La respuesta es NO.
- Se puede demostrar que cuando se contrasta $H_0 : \beta_j = 0$, el estadístico F es exactamente el cuadrado del correspondiente t -estadístico. Por lo tanto, **los dos nos conducirían al mismo resultado** (siempre que la alternativa sea a 2 colas).
- Como el **estadístico t es más flexible** (puede utilizarse para contrastar alternativas de una y de dos colas) y es más fácil de calcular, no hay ninguna razón para usar el estadístico F cuando se quiere contrastar hipótesis con una única restricción.

Conclusión: para contrastar una única restricción es mejor utilizar el t -ratio

Contraste de significatividad global (del modelo)

- Un caso particular de restricciones de exclusión es:

$$H_0 : \forall \beta_j = 0 \quad j = 2, \dots, k$$

- Es decir, ninguna de las variables explicativas afecta al regresando.
- La alternativa consistirá en que al menos uno de los regresores es significativo

$$H_1 : \exists \text{ algun } \beta_j \neq 0 \text{ para } j = 2, \dots, k$$

- Para este caso particular, el estadístico F se simplifica mucho y queda como:

$$\frac{R^2/(k-1)}{(1-R^2)/(N-k)} \longrightarrow \mathcal{F}_{(k-1), (N-k)}$$

Contraste de significatividad global (output de regresión)

- El contraste de significatividad global, suele ser ofrecido automáticamente por todos los paquetes econométricos cuando se efectúa una regresión por MCO.

Modelo 1: MCO, usando las observaciones 1-706

Variable dependiente: SLEEP

	coeficiente	Desv. típica	Estadístico t	valor p	
-----	-----	-----	-----	-----	
const	3638,25	112,275	32,40	1,47e-141	***
TOTWRK	-0,148373	0,0166935	-8,888	5,19e-18	***
EDUC	-11,1338	5,88457	-1,892	0,0589	*
AGE	2,19988	1,44572	1,522	0,1285	
Media de la vble. dep.	3266,356	D.T. de la vble. dep.	444,4134		
Suma de cuad. residuos	1,23e+08	D.T. de la regresión	419,3589		
R-cuadrado	0,113364	R-cuadrado corregido	0,109575		
F(3, 702)	29,91889	Valor p (de F)	3,28e-18		
Log-verosimilitud	-5263,106	Criterio de Akaike	10534,21		
Criterio de Schwarz	10552,45	Crit. de Hannan-Quinn	10541,26		

Contrastar restricciones conjuntas mediante SCR

- Necesitaremos estimar dos modelos el **modelo general** (sin las restricciones incluidas) y el **modelo restringido** (modelo que incorpora las restricciones)
- Por ejemplo, si estamos trabajando con el modelo:

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

- y queremos contrastar: $H_0 : \beta_3 = \beta_4 = 0$
 - El modelo general (MG) será: $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + u_i$
 - El modelo restringido (MR) será: $y_i = \beta_1 + \beta_2 x_{2i} + u_i$
- Es evidente que siempre ocurrirá que $SCR_R \geq SCR_G$. ¿Por qué?
- Se trata de evaluar si el cambio en SCR es lo suficientemente grande como para no excluir las variables x_3 y x_4 del modelo.
- Si al incluir las restricciones (modelo restringido) se produce un gran incremento en SCR, es **evidencia en contra de la H0**.
- Si por contra, SCR_R próximo a SCR_G , esto indicará que las restricciones son **aproximadamente ciertas**.

Comparando los SCR de los modelos general y restringido

- Por lo tanto, incluso aunque las SCR , por sí mismas, no nos dicen nada acerca de la validez de H_0 , el incremento de SCR cuando se imponen las restricciones puede ayudarnos a decidir sobre la validez de H_0 .
- Hay que valorar si el incremento observado en SCR (al imponer las restricciones) es lo suficientemente grande respecto a la SCR del modelo general para decidir no rechazar la H_0
- Como en cualquier contraste de hipótesis la respuesta depende del nivel de significación (α).
- Pero también necesitaremos un estadístico con distribución conocida bajo la nula (si la nula fuese cierta).

El estadístico F (en términos de SCR)

- El estadístico F toma la forma

$$\frac{(SCR_R - SCR_G)/r}{SCR_G/(N - k)} \longrightarrow \mathcal{F}_{r, (N-k)}$$

siendo:

- SCR_G : Suma de cuadros residual del modelo general
 - SCR_R : Suma de cuadros residual del modelo restringido
 - $(N - k)$: grados de libertad del modelo general
 - r : número de restricciones en la hipótesis nula
-
- Fíjate que el denominador del estadístico F es precisamente el estimador de σ^2 en el modelo general.

El estadístico F en términos del R^2

- Hay una formulación alternativa del estadístico F muy fácil de calcular y, por tanto, es conveniente conocerla.
- Debido a que $SCR = SCT(1 - R^2)$, podemos reformular el estadístico F como:

$$\frac{(R_G^2 - R_R^2)/r}{(1 - R_G^2)/(N - k)} \longrightarrow \mathcal{F}_{r, (N-k)}$$

- Este estadístico es muy conveniente para contrastar restricciones conjuntas, pero no puede aplicarse a todos los tipos de restricciones lineales.
- Recordad que el estadístico F mide el aumento relativo en SCR o en R^2 al movernos del modelo general al modelo restringido.

Ejemplo: contraste de significatividad global ... mediante *SCR*

- Hagamos el contraste de significatividad global PERO comparando MG y MR

Modelo 1: MCO, usando las observaciones 1-526

Variable dependiente: salari

	Coficiente	Desv. Típica	Estadístico t	Valor p	
const	-1.71453	0.761686	-2.251	0.0248	**
educacio	0.601749	0.0513538	11.72	2.61e-028	***
experiencia	0.0642164	0.0104108	6.168	1.39e-09	***
negre	-0.0838851	0.444298	-0.1888	0.8503	
dona	-2.15649	0.270605	-7.969	1.01e-014	***

Media de la vble. dep.	5.896103	D.T. de la vble. dep.	3.693086
Suma de cuad. residuos	4945.334	D.T. de la regresión	3.080910
R-cuadrado	0.309351	R-cuadrado corregido	0.304048
F(4, 521)	58.34070	Valor p (de F)	1.09e-40
Log-verosimilitud	-1335.718	Criterio de Akaike	2681.436
Criterio de Schwarz	2702.762	Crit. de Hannan-Quinn	2689.786

Modelo 1: MCO, usando las observaciones 1-526

Variable dependiente: salari

	Coficiente	Desv. Típica	Estadístico t	Valor p
const	5.89610	0.161026	36.62	1.14e-146 ***
Media de la vble. dep.	5.896103	D.T. de la vble. dep.	3.693086	
Suma de cuad. residuos	7160.414	D.T. de la regresión	3.693086	
R-cuadrado	0.000000	R-cuadrado corregido	0.000000	
Log-verosimilitud	-1433.060	Criterio de Akaike	2868.121	
Criterio de Schwarz	2872.386	Crit. de Hannan-Quinn	2869.791	

¿Usamos el estadístico F en función de los SCR o de los R^2 ?

- La forma básica del estadístico F en términos de SCR es posible usarla para cualquier conjunto de restricciones lineales.
- Los cálculos (a mano) son más fáciles si usas el estadístico F en función de los R^2 , pero no siempre podremos usarlo

ejemplo: no podemos usar F en función de los R^2

- MG: $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + u_i$
- $H_0 : \beta_3 = 1, \beta_4 = 0$
- El MR queda como: $(y_i - x_{3i}) = \beta_1 + \beta_2 x_{2i} + u_i$
- En este caso, como el regresando del MG y MR no son iguales, no podremos usar el estadístico F expresado en función de los R^2
- La razón estriba en que, al tener distintos regresandos, la SCT de los dos modelos es diferente.

4.5 Predicción puntual y por intervalos

Uno de los principales usos de los modelos estadísticos

Predicción: otro de los usos de los modelos estadísticos

- Hasta ahora nos hemos centrado en obtener y contrastar hipótesis acerca de los parámetros del MRL, pero ...
- En Economía y también en el mundo de la empresa interesa, y mucho, disponer de técnicas para predecir el valor futuro de una variable (y). Hay múltiples métodos de predicción (cualitativos y cuantitativos).
- La predicción es la **anticipación del comportamiento de la variable endógena** en el futuro, a partir de la información de la que se dispone hasta el momento presente
- ¿Podemos usar nuestro MRL para predecir? **SÍ**. De hecho uno de los principales métodos de predicción consiste en el uso de modelos estadísticos
- Hay que tener en cuenta que la **predicción puntual** consistirá es una "simple" extrapolación; mientras que la **predicción por intervalos** recoge, en cierta forma, una medida de la precisión de la predicción (fiabilidad de la predicción).
- Por otro lado también, a posteriori, interesará evaluar cuál ha sido la **capacidad predictiva** del modelo. ¿Es el modelo adecuado para predecir?

Predicción puntual y por intervalos

- La **predicción puntual** es una extrapolación del valor de la variable endógena para una nueva observación.
- La **predicción por intervalos** se efectúa cuando nos interesa conocer la precisión de la estimación y el rango de valores donde se puede situar la variable para la nueva observación.

Predicción puntual

- Si tenemos el modelo estimado:

$$\hat{\text{salario}} = -1.73 + 0.60\text{educacion} + 0.06\text{experiencia} - 2.15\text{mujer}$$

- y conocemos los niveles de educación (15 años), experiencia (0 años) y el genero (mujer), de una persona (Marta) podemos predecir su salario
- La predicción puntual será una simple extrapolación: $\hat{\text{salario}}_{\text{Marta}} = -1.73 + 0.60 * 15 + 0.06 * 0 - 2.15 * 1 = 5.12$
- El modelo estimado predice que "una persona como Marta" tendrá un salario de 5.12
- ¿Nos equivocaremos? Seguro. Hay **múltiples fuentes de error**; así que, a pesar de que, bajo las h.e.b, el método o predictor que hemos usado es un predictor óptimo e insesgado, seguro que nos equivocamos.

3 fuentes de error

- Verdadero valor de y_0 : $y_0 = \beta_1 + \beta_2 x_{20} + u_0$

- Predicción: $\hat{y}_0 = \hat{\beta}_1 + \hat{\beta}_2 x_{20}$

1) Se ha utilizado el modelo estimado, no el modelo teórico; es decir, no se han utilizado los verdaderos valores de los parámetros

2) Perturbación aleatoria. En el valor de y_0 hay una parte aleatoria u_0

3) Los valores de las variables explicativas no se conocen exactamente y estos valores bajo los cuales se ha condicionado la predicción, puede ser erróneos

- Podemos dar una medida del **margen de error**, o de la fiabilidad de la predicción presentando una estimación por intervalos

Predicción por intervalos

- Intentamos dar un intervalo en el que con determinada confianza, generalmente al 95%, esté el valor verdadero valor de y_0 (salario de Marta)
- La amplitud de este intervalo difiere según se desee dar una **predicción individual** o una **predicción media**
- Es más difícil predecir el comportamiento de un sujeto que el comportamiento medio de una población. Por ejemplo:
 - Precio de una casa concreta en la calle tal, frente al precio medio de las casas de similares características de una ciudad
 - Nota media en econometría de Juan frente a la nota media del grupo
 - Salario de Marta frente al salario medio de las personas de perfil similar (personas con características similares a Marta: mujer, con 15 años de educación y sin experiencia)
- La amplitud del intervalo de predicción media es **menor** que la del intervalo de predicción individual

Predicción por intervalos

- Un intervalo de predicción al $(1 - \alpha)\%$ se calcula como:

$$\hat{y}_0 \pm t_{N-k}^{\alpha/2} ee$$

donde ee es el error estándar de la predicción

- Por ejemplo si, $\hat{y}_0 = 7$, y $ee = 0.6$, entonces un intervalo de confianza al 95% quedaría como:

$$7 \pm t_{N-k}^{0.05/2} 0.6$$

- De forma que si el modelo se ha estimado con $N = 124$ y $k = 4$, entonces el intervalo al 95% quedaría como:

$$7 \pm t_{124-4}^{0.025} 0.6$$

- Que finalmente quedaría como:

$$[7 - (1.98 * 0.6) ; 7 + (1.98 * 0.6)] = [5.812 , 8.188]$$

Predicción por intervalos

La amplitud del intervalo depende de:

- del nivel de confianza del intervalo: un intervalo al 100% siempre dará un intervalo muy amplio ($-\inf, +\inf$)
- La calidad del modelo de regresión: a mayor calidad menor amplitud del intervalo.
- El tamaño de la muestra: a mayor tamaño muestral menor amplitud del intervalo
- La dispersión en los datos de las variables exógenas: a mayor dispersión en los datos menor amplitud del intervalo
- Parecido entre el perfil del sujeto a predecir y los sujetos de la muestra: a mayor parecido menor amplitud del intervalo