Las regresiones en las series temporales

Oscar Centeno Mora

Preámbulo

- ¿Por qué estudiamos la regresión en las series temporales?

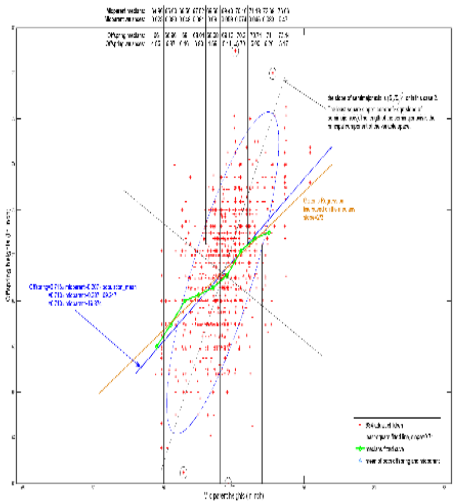
La regresión se origina en un contexto temporal…

Preámbulo

- Sir Francis Galton creo el modelo de regresión para predecir o pronosticar el tamaño de los hijos tomando las mediciones de sus padres.

- A partir de medidas en el tiempo, y comparando con los tamaños de los padres, le permitieron pronosticar, a nivel de la media, para un determinado tiempo Y_{t+h}, la estatura promedio que iba a adquirir el hijo(a).

- El termino regresión proviene del hecho de regresar, a partir de un punto de referencia (el padre), la estatura, y entonces ir conociendo el crecimiento medio del hijo al pasar del tiempo.



- Su base fue en un contexto temporal…

Preámbulo

- La regresión es parte de la familia de modelos de la primera generación: se adecuan según los componentes de una serie temporal.
- El presente tema explica las diferencias entre las regresiones aplicadas a datos transversales y longitudinales.
- Además, se abordan los métodos de estimación, los predictores de los modelos, el pronóstico y otras variantes no lineales en los modelos de regresión para las series temporales.
- El curso se enfoca en las series temporales univariadas. Sin embargo, ¿se puede utilizar la regresión múltiple en las series temporales? ¿Se debe hacer el mismo análisis de adecuación de los modelos (diagnóstico)?

Índice

Índice

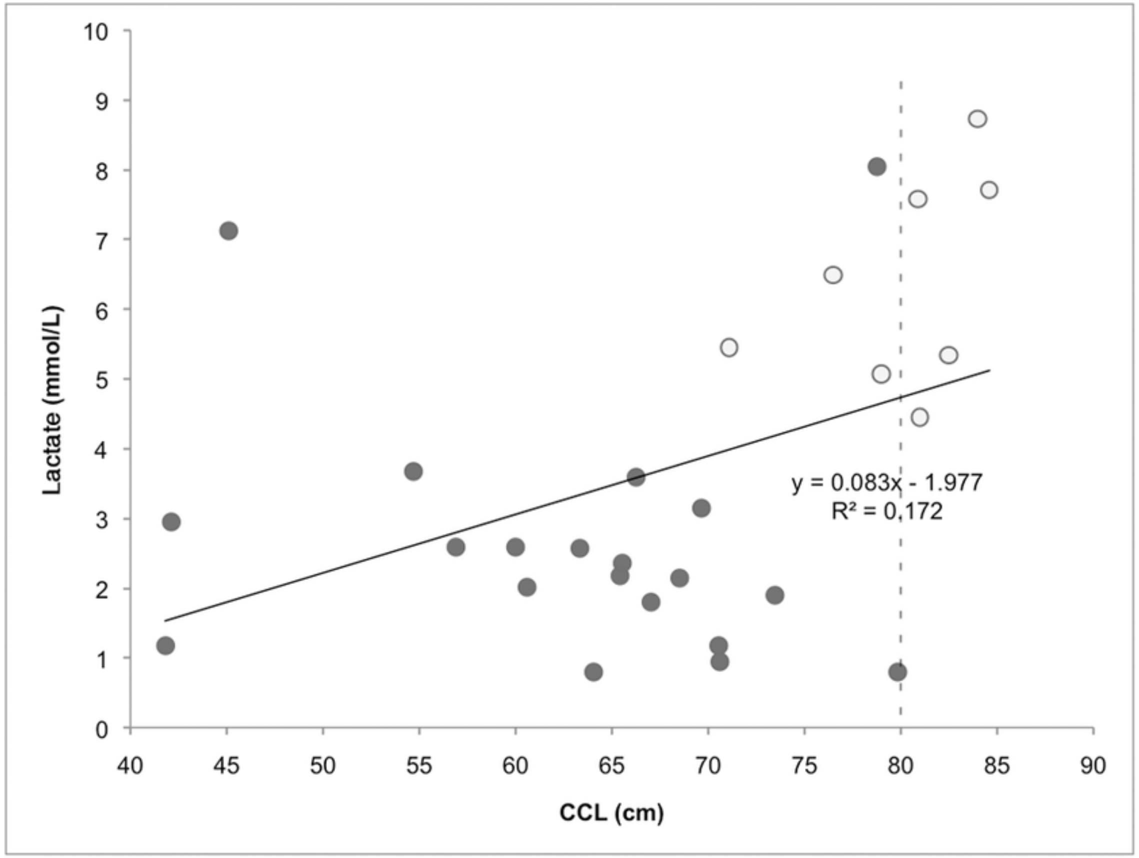
1

Regresión: transversal vs longitudinal

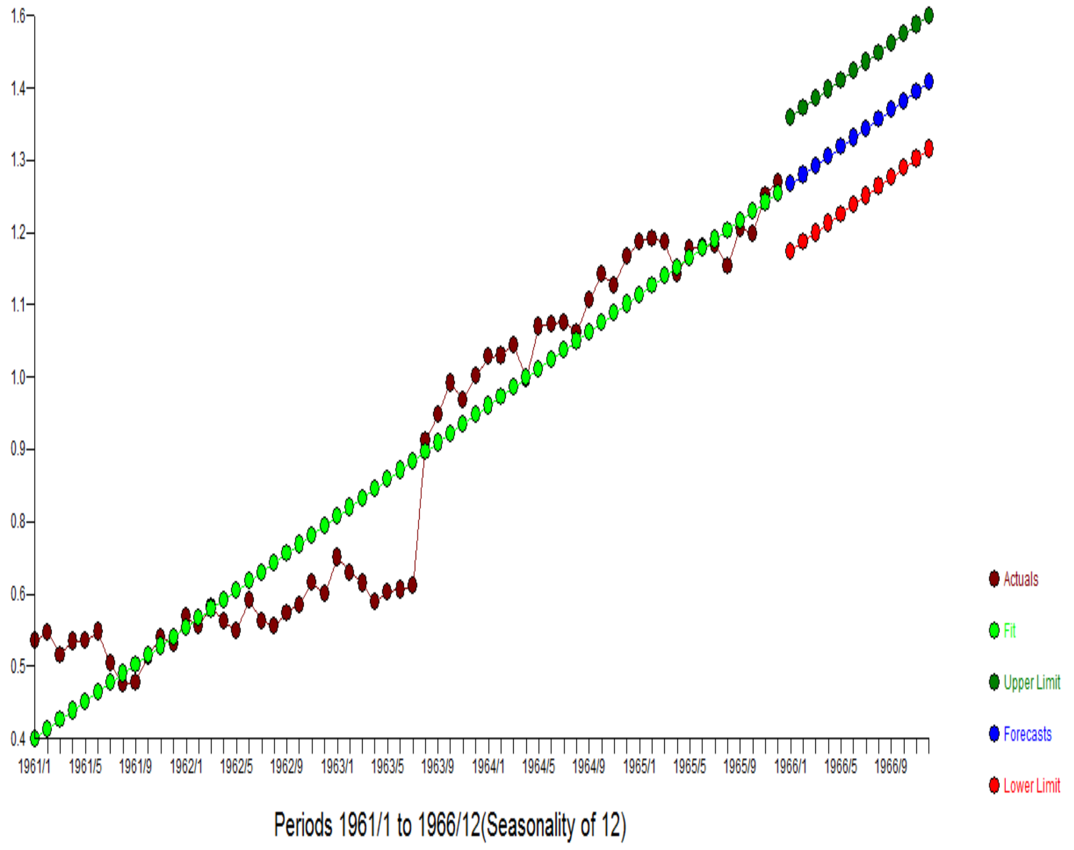
Regresión: transversal vs longitudinal

- ¿Qué diferencias existen entre la regresión transversal y la longitudinal?

Transversal

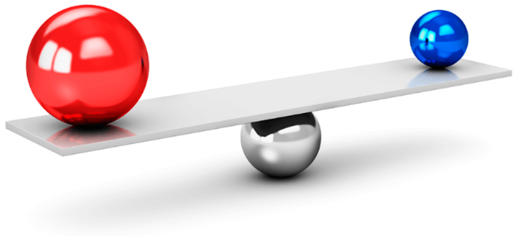


Longitudinal



Regresión: transversal vs longitudinal

- La regresión transversal (*RT*) trabaja en un determinado momento t, mientras que regresión en series de tiempo (*RST*) en T momentos.



- La *RT* utiliza sujetos o unidades de estudio, mientras que la regresión en series de tiempo trabaja con períodos.

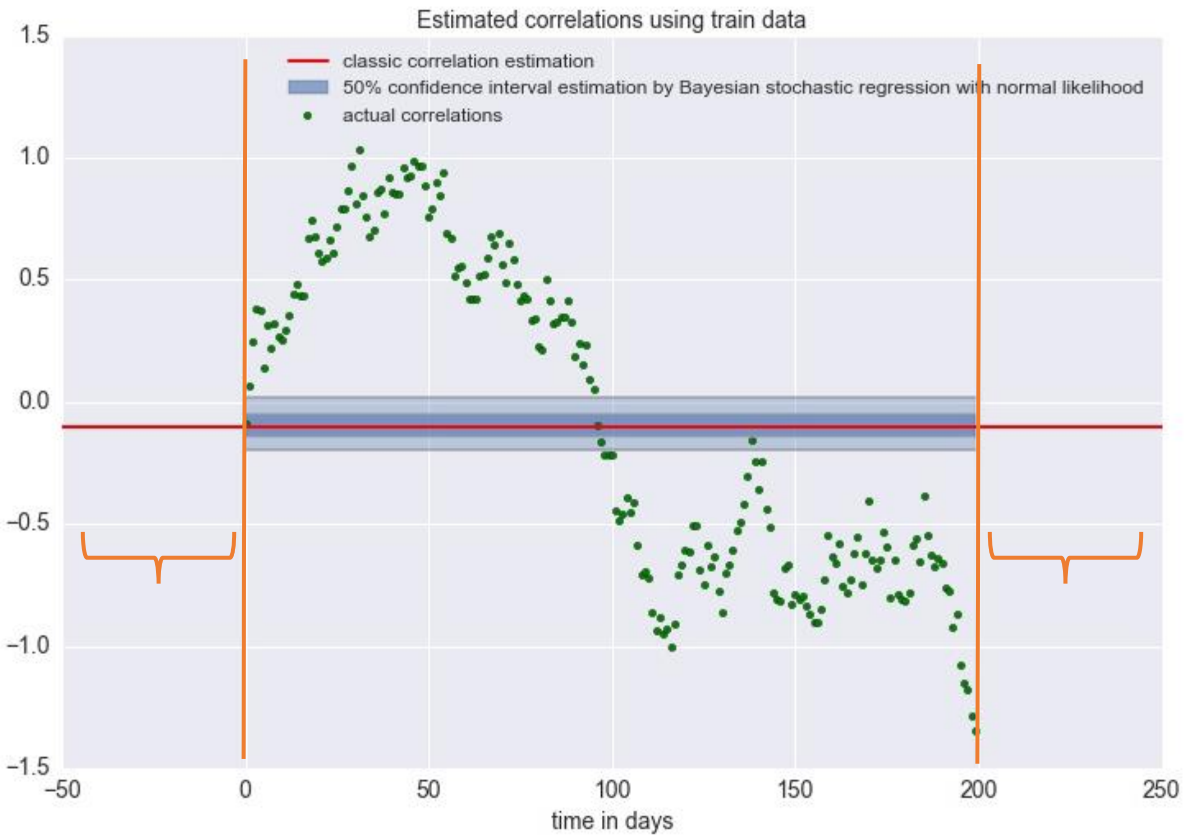
- Notamos para la *RT* a los sujetos los notamos como $i = 1,2, \ldots, n$. Por otra parte, notamos para la *RST* a los periodos como $t = 1,2, \ldots, T$.

- Sobre las predicciones, en las *RT* denominamos como estimación o extrapolación a los valores obtenidos a partir de la recta de regresión, en las *RST* llamamos estimación o pronóstico a los valores obtenidos de la recta de mejor ajuste.

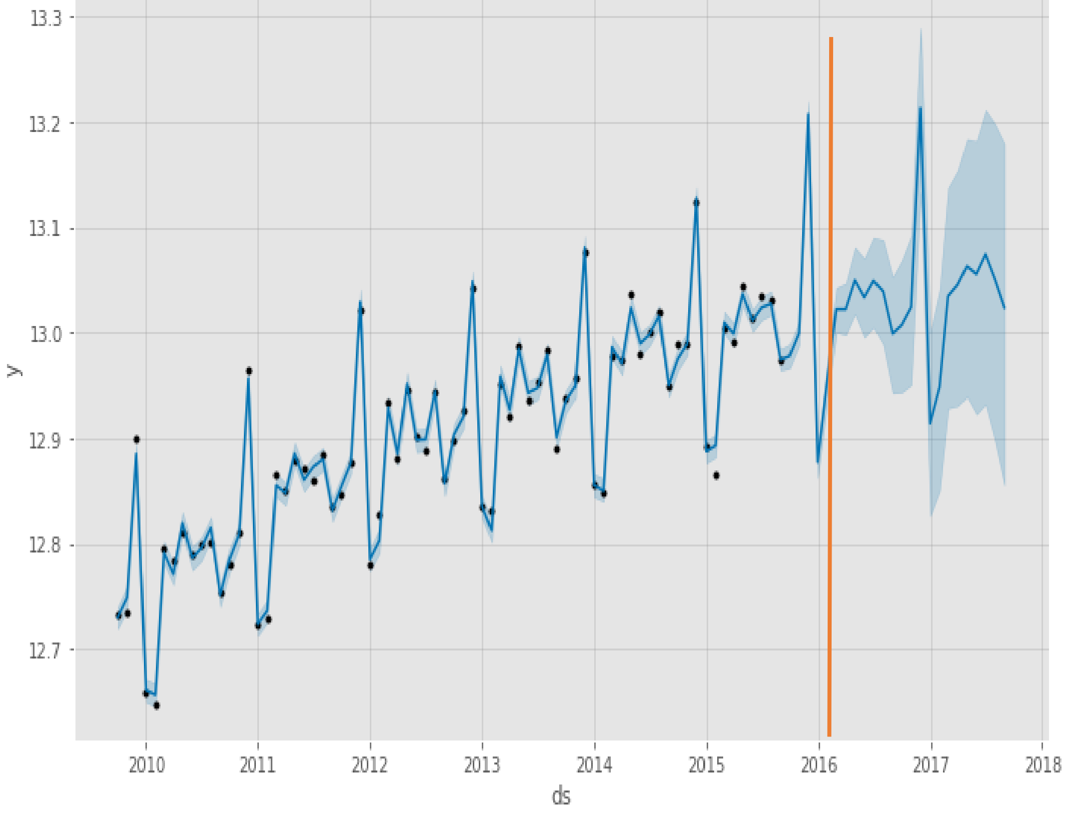
- Mientras que la *RT* no aconseja llevar a cabo extrapolaciones (Neter 2004), el objetivo mismo de la *RST* es hacer pronósticos.

Regresión: transversal vs longitudinal

Estimación y extrapolación



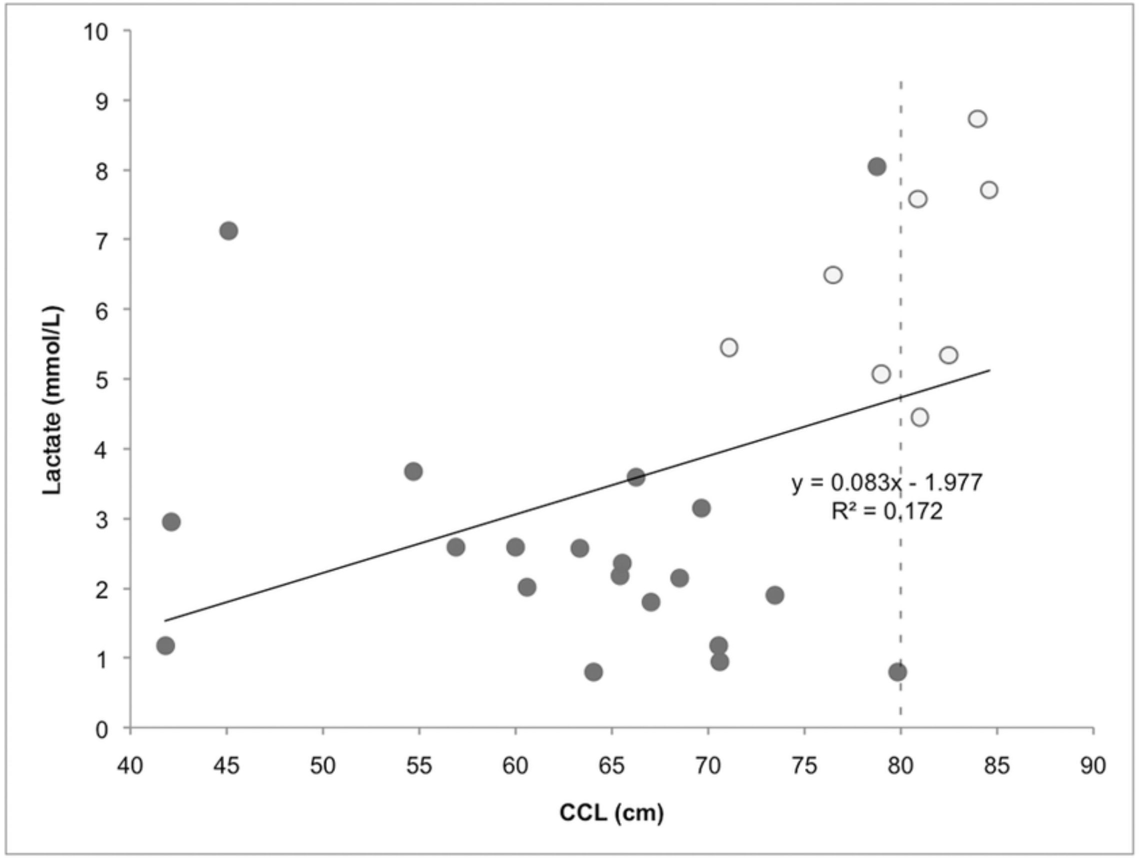
Estimación y pronóstico



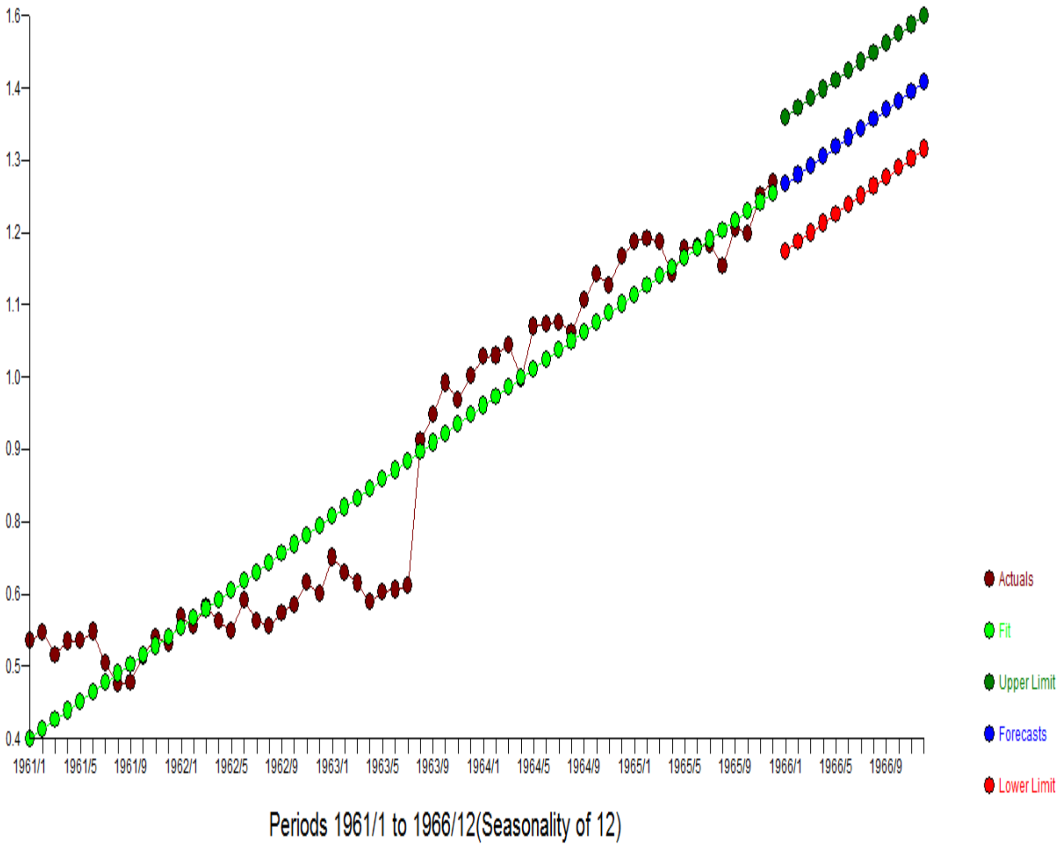
Regresión: transversal vs longitudinal

- ¿Qué similitudes existen entre la regresión transversal y la longitudinal?

Transversal



Longitudinal



Regresión: transversal vs longitudinal

- La regresión transversal (*RT*) y longitudinal persiguen en mismo objetivo: encontrar la función de mejor ajuste a los datos:

RT: $$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon_i$$

RST: $$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{k,t} + \varepsilon_i$$

(dentro de poco hablaremos de los predictores en las regresiones)

- Ambos son métodos estocásticos.

- Y sin duda, y más importante, poseen y comparten ***casi*** las mismas propiedades para los predictores: estimadores BLUE.

Regresión: transversal vs longitudinal (recordatorio)

- BLUE: Best Linear Unbiased Estimator.

- Es el teorema de Gauss-Markov que formula que cualquier modelo general cumple:

1. Correcta especificación. Modelo es una combinación de parámetros y variables
$$Y = X\beta + u$$
2. Correcta identificación: la matriz no es de rango completo: $rg(k) = k < N$
3. MIA: muestra de observaciones $(y_i, x_{1i}, \dots, x_{ki})$ es aleatoria simple, por lo que los vectores (y_i, X_i') y (y_i, X_j') son ortogonales (independientes).
4. Esperanza condicional: independencia entre la perturbación (error) y las variable, $\mathrm{E}(u_i, X_i') = 0$.
5. Homocedasticidad: variancia constante, $\mathrm{Var}(u|X) = \sigma^2 I$.

Regresión: transversal vs longitudinal

- Para el cumplimiento del teorema Gauss-Markov, se verifica:

1. Correcta especificación del modelo
2. Linealidad en los parámetros.
3. Media condicional es cero.
4. Homocedasticidad.
5. No hay correlación entre las perturbaciones (errores).
6. Covariancia nula entre u_i y x_i.
7. Mayor número de observaciones que de parámetros.
8. Presencia de variabilidad entre los x.
9. No hay multicolinealidad perfecta.
10. Las x no son estocásticas, es decir, son fijas en muestras repetidas.

Más sobre los estimadores BLUE:

http://willett.ece.wisc.edu/wp-uploads/2016/01/17-BLUE.pdf

http://www.ws.binghamton.edu/fowler/fowler%20personal%20page/EE522_files/EECE%20522%20Notes_11%20Ch_6.pdf

Regresión: transversal vs longitudinal

- La regresión transversal y longitudinal comparten muchas más propiedades: el elemento temporal es la base de sus diferencias…

- De igual forma, se pueden aplicar las mismas acciones a un análisis *RTS*: transformación, selección de variables, diagnósticos, medidas remediales, funciones de enlace, etc…

- Al utilizar el R-Studio, la regresión lineal utilizaba la función lm(). Para la RST se utiliza la función tslm(), y otra variación podría ser la stl(), sin embargo nos apropiaremos de la primera. Las funciones TS buscan incorporar el elemento temporal al análisis.

- También, para otras variación de los modelos por GLM, se utiliza la función glm(), su contra parte en las RTS también existen (tscount(), etc…).

Índice

Estimación: Mínimos Cuadrado Ordinarios (MCO)

- Veamos la gran similitud que existe entre los MCO para datos transversales y longitudinales.

Sea la recta de mejor ajuste:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{1,k} + \varepsilon_t$$

Para obtener el vector de coeficientes β' , se utiliza el principio de minimizar los errores:

$$\sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} (y_t - \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{1,k})^2$$

Estimación: Mínimos Cuadrado Ordinarios (MCO)

Sea $\boldsymbol{y} = (y_1, \dots, y_T)´$ el vector con los valores temporales, $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_T)´$ las perturbaciones, $\boldsymbol{\beta} = (\beta_0, \dots, \beta_k)´$ los parámetros y X la matriz con las variables independientes

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_{1,1} & x_{2,1} & \cdots & x_{k,1} \\ 1 & x_{1,2} & x_{2,2} & \cdots & x_{k,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1,T} & x_{2,T} & \cdots & x_{k,T} \end{bmatrix}$$

Se busca obtener los β a partir de la siguiente ecuación $\boldsymbol{y} = X\beta + \varepsilon$, sabiendo que el principio es el minimizar los residuos: $\boldsymbol{\varepsilon}´\boldsymbol{\varepsilon} = (y - X\beta)´(y - X\beta)$.

La ecuación obtenida es equivalente entre el mundo transversal y longitudinal:

$$\widehat{\boldsymbol{\beta}} = (X´X)^{-1}X'Y$$

Estimación: Mínimos Cuadrado Ordinarios (MCO)

- La estimación del error estándar de la regresión posee principios equivalentes:

$$\hat{\sigma}_e^2 = \frac{1}{T-k}\sum_{t=1}^{T} \varepsilon_t^2 \qquad , \qquad \hat{\sigma}_e = \sqrt{\frac{1}{T-k}\sum_{t=1}^{T} \varepsilon_t^2}$$

En términos matriciales, el resultado final sería:

$$\hat{\sigma}_e^2 = \frac{1}{T-k}\left(y - X\widehat{\boldsymbol{\beta}}\right)'\left(y - X\widehat{\boldsymbol{\beta}}\right)$$

Estimación: Máxima Verosimilitud (MVE)

Recordatorio

La función de verosimilitud toma una función de distribución $f(y)$, un vector de parámetros por estimar, $\Theta' = (\Theta_1, \Theta_2, \ldots, \Theta_k)$, y un total de datos o períodos T.

Se quiere que para cada observación $y_1, y_2, \ldots, y_T$, estén asociados a una función de distribución y así poder determinar los parámetros asociados. La función de verosimilitud se presenta como:

$$L = f(y_1, y_2, \ldots, y_T) = \prod f(y_t; \Theta)$$

Importante: en los métodos por MCO se ajustan los datos para probar la adecuación de estos a una función de distribución, y así obtener los parámetros. En la MLE es inverso: se ajusta una distribución para probar la adecuación conjunta (productora) de estas a los datos, y así entonces obtener los parámetros.

Estimación: Máxima Verosimilitud (MVE)

MLE en la regresión con series temporales

Sea la ecuación de mejor ajuste: $y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{1,k} + \varepsilon_t$

$$y_t = x_t{'}\beta + \varepsilon_t$$

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

Donde se asume que $E(\varepsilon_t) = 0$, $E({\varepsilon_t}^2) = \sigma^2$ y $E(\varepsilon_t \varepsilon_s) = 0 \quad s \neq t$

La verosimilitud está dado por:

$$L = (2\pi)^{-T/2}(\sigma^2)^{-T/2}\exp\,[-\left(\frac{1}{2\sigma^2}\right)\varepsilon{'}\varepsilon]$$

Estimación: Máxima Verosimilitud (MVE)

Aplicando el logaritmo natural:

$$\ln(L) = -\left(\frac{T}{2}\right)\ln(2\pi) - \left(\frac{T}{2}\right)\ln(\sigma^2) - \left(\frac{1}{2\sigma^2}\right)(Y - X\beta)'(Y - X\beta)$$

Al maximizar la ecuación de verosimilitud anterior respecto a al vector β, sería equivalente (dado el signo negativo) a minimizar $(Y - X\beta)'(Y - X\beta)$, y bajo esto sabemos que el MLE bajo la normalidad equivale a los resultados obtenidos por MCO.

La maximización del log verosimilitud se obtiene obteniendo las derivadas respecto a los parámetros $\Theta' = (\beta', \sigma^2)$ e igualando a cero,

$$\frac{\delta \ln(L)}{\delta \beta} = -\left(\frac{1}{\sigma^2}\right)(X'Y - X'X\beta) = 0$$

$$\frac{\delta \ln(L)}{\delta \sigma^2} = -\left(\frac{T}{2\sigma^2}\right) + \left(\frac{1}{2\sigma^4}\right)(Y - X\beta)'(Y - X\beta) = 0$$

Estimación: Máxima Verosimilitud (MVE)

Satisfaciendo la condición regular $\frac{\delta \ln(L)}{\delta\, \Theta} = 0$, si resolvemos las ecuaciones sobre β y $\hat{\sigma}^2$ se obtiene que:

$$\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'Y$$

$$\hat{\sigma}^2 = \frac{\left(Y - X\hat{\boldsymbol{\beta}}\right)'(Y - X\hat{\boldsymbol{\beta}})}{T} = \frac{e'e}{T}$$

Finalmente, aunque hay que volver a derivar de forma parcial para conocer los resultados asociados al segundo orden, estos no se presentarán, pero si los resultados de las distribuciones de cada parámetro para ver la casi concordancia con el método por MCO.

$$T^{\frac{1}{2}}(\hat{\beta} - \beta) \rightarrow N[0, \sigma^2(X'X)^{-1}]$$

$$T^{\frac{1}{2}}(\hat{\sigma}^2 - \sigma^2) \rightarrow N[0, 2\sigma^2]$$

Bajo el supuesto de normalidad, el método por MLE del β es equivalente al obtenido por los estimadores de MCO. No obstante para σ^2 si hay una pequeña diferencia: el divisor por MCO es de $T - k$, mientras que en MLE es por T. Esto produce que en pequeñas muestras el MLE será sesgado para σ^2, aunque para muestras grandes por la propiedad de la asintotisidad el parámetro permanece insesgado.

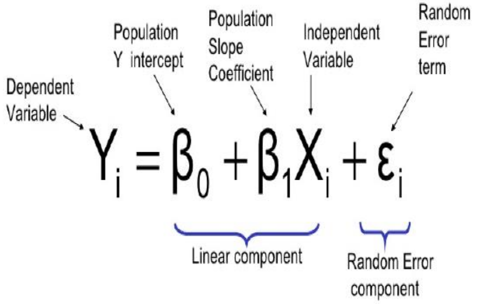
Índice

Los predictores: la tendencia

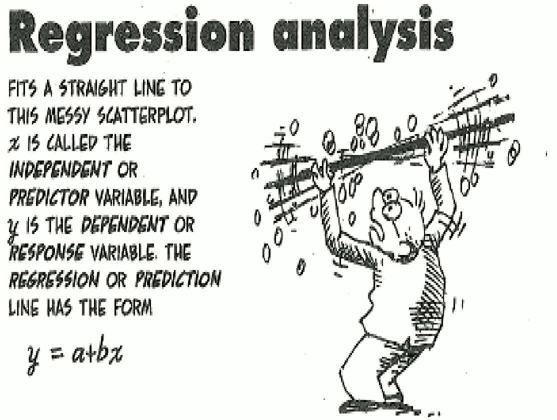
- La regresión en series de tiempo también trata de explicar los componentes de la serie de tiempo.

- Dentro de los componentes se puede tomar la tendencia como un predictor en la ecuación de mejor ajuste.

- Es común que los datos de las series temporales sean modelados como una tendencias. Una tendencia lineal se puede modelar simplemente utilizando $x_{1,t} = t$ como un predictor:



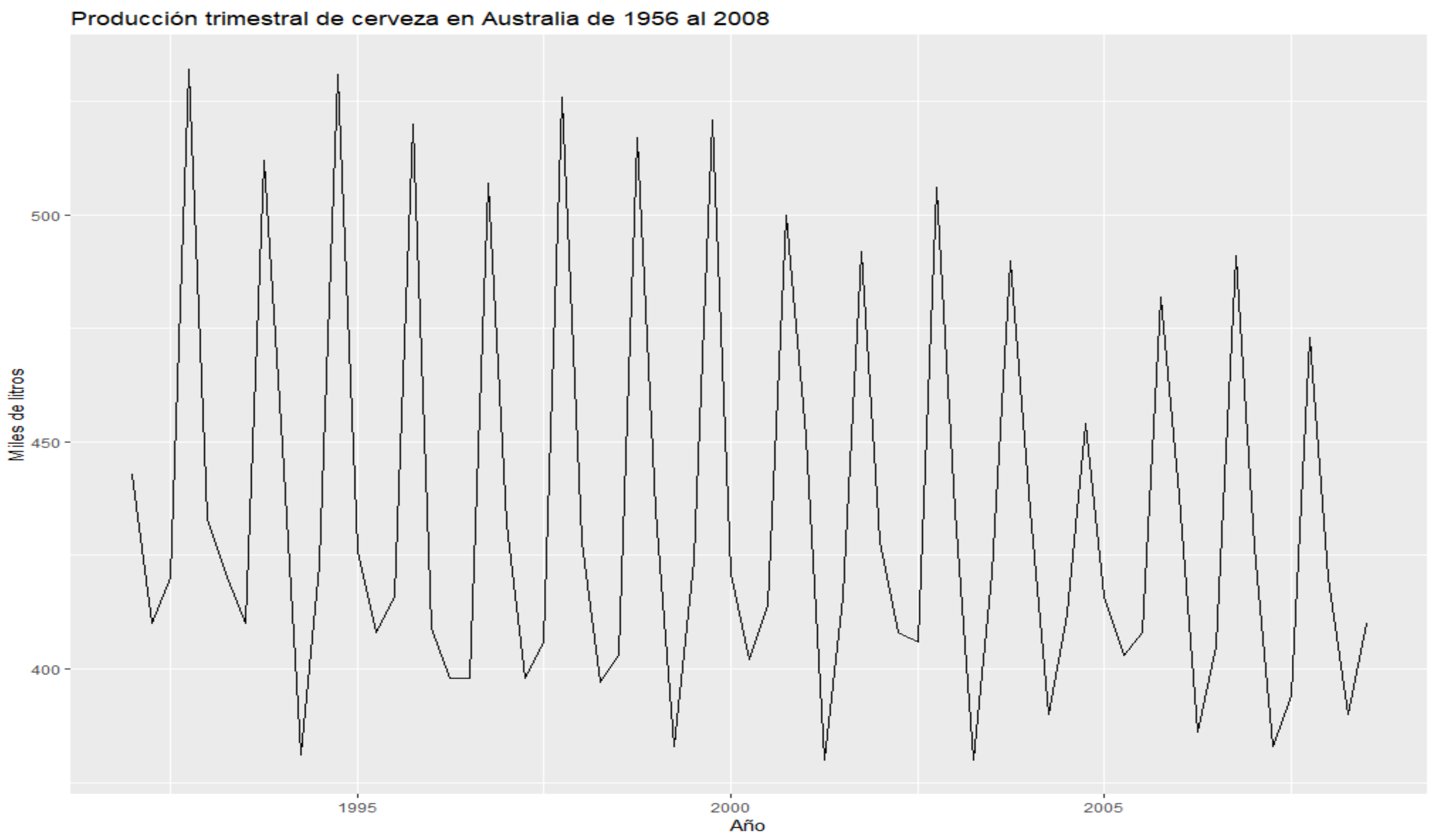
$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$



Donde $t = 1, \ldots, T$.

Los predictores: la tendencia

- Vamos a utilizar el archivo de datos ausbeer de la librería fpp. Estos son los miles de litros de producción de cerveza por trimestre de 1956 al 2008, en Australia.



Los predictores: la tendencia

- Estimemos un primer modelo utilizando únicamente la tendencia: $y_t = \beta_0 + \beta_1 t + \varepsilon_t$

- Para esto en R-Studio, vamos a utilizar la función tslm(), y vamos a escribir esto como:

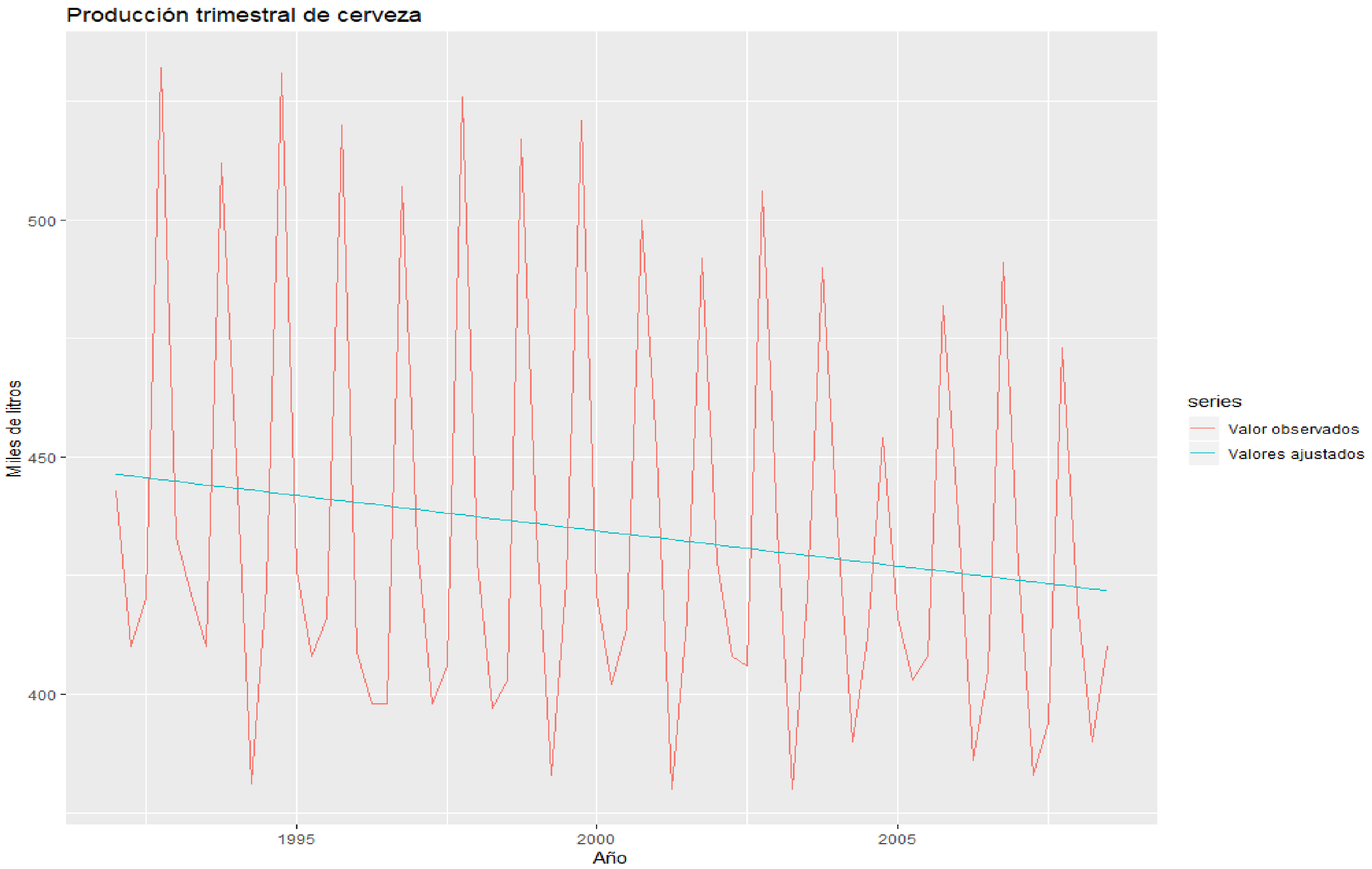
```
serie.1.tendencia <- tslm(serie.1 ~ trend)
summary(serie.1.tendencia)
```

- Obtenemos como resultado lo siguiente. ¿Qué podemos decir del modelo?

```
Call:
tslm(formula = serie.1 ~ trend)

Residuals:
   Min     1Q Median     3Q    Max 
-62.02 -32.13 -13.46  15.25  88.73 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 446.7354    10.6693  41.871   <2e-16 ***
trend        -0.3719     0.2728  -1.364    0.177    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43.18 on 65 degrees of freedom
Multiple R-squared:  0.02781,	Adjusted R-squared:  0.01285 
F-statistic: 1.859 on 1 and 65 DF,  p-value: 0.1774
```

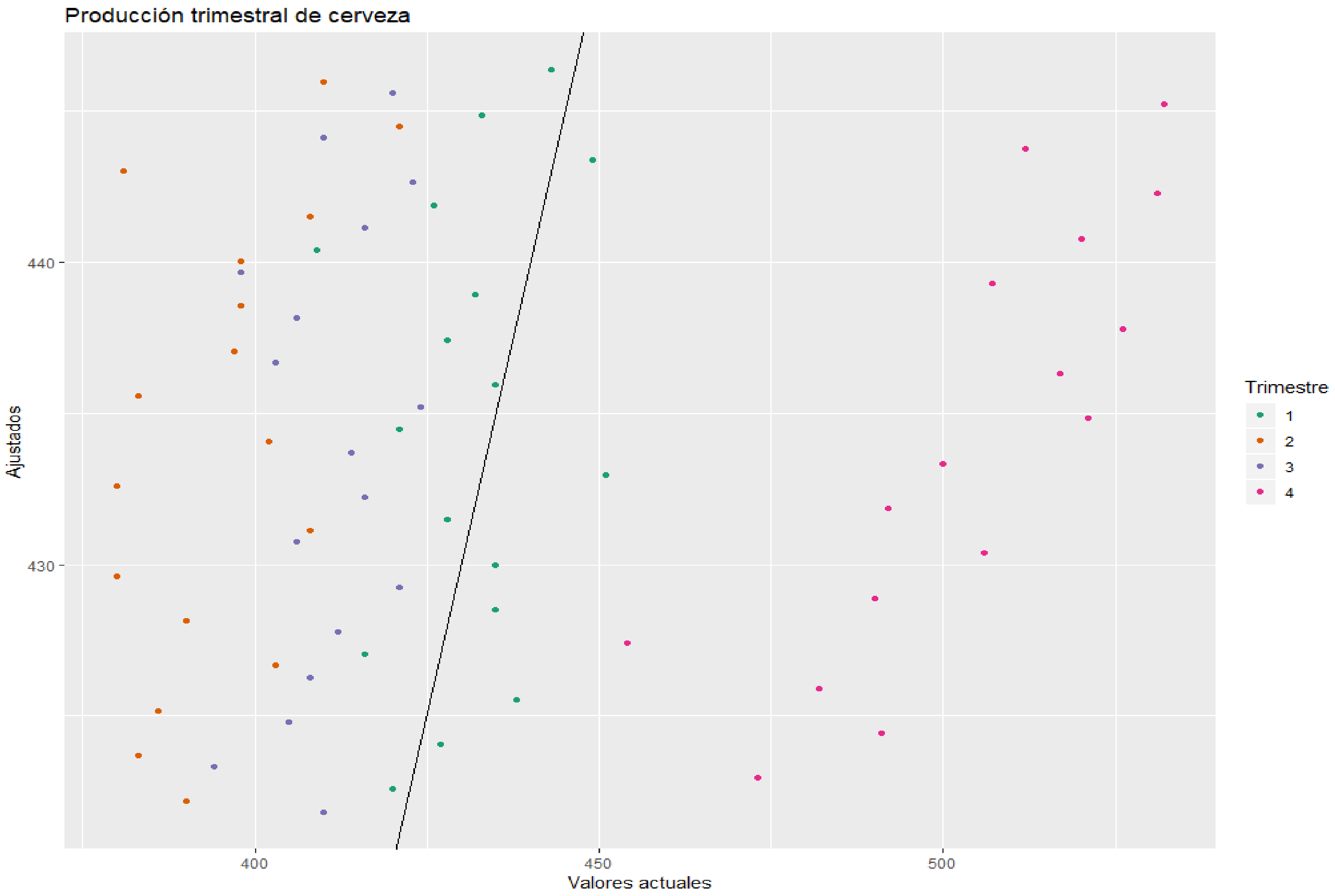
Los predictores: la tendencia

Visualización de los valores observador y ajustados.



Los predictores: la tendencia

Rendimiento interno de los valores ajustados contra los observados.



Los predictores: la tendencia

- El presente modelo no ajusta correctamente a la serie temporal en cuestión…

- Debemos de tomar en cuenta que estamos trabajando con una serie con estacionalidad: los trimestres.

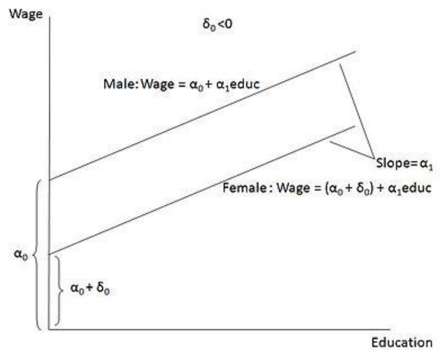
- Se debe ajustar un modelo con dichas características.

- A continuación introduciremos la estacionalidad como un componente adicional.

- Utilizaremos para esto las variables “dummy” para ajustar el componente estacional.

Los predictores: las variables control (dummy variables)

- Para poder captar la estacionalidad en los modelos de regresión de series de tiempo se suele recurrir a las variables control o "*dummy*".



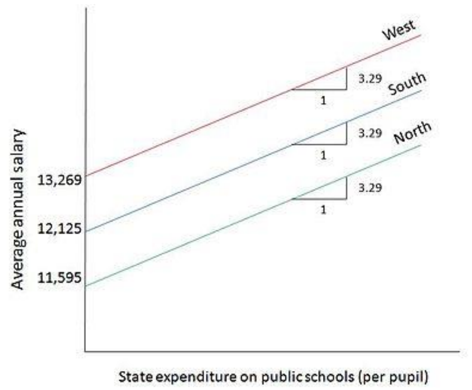
- Sea s el número de períodos que conforman la estacionalidad. Paara diferentes niveles de estacionalidad, s tendrá como valores:

Semanal: $s = 7$.

Mensual: $s = 12$.

Trimestral: $s = 4$.

Semestral: $s = 2$.



Las variables *dummy* trabajan con el concepto de la no redundancia entre las variables.

Los predictores: las variables control (dummy variables)

¿Qué establecía en principio de la no redundancia entre variables?

Los predictores: las variables control (dummy variables)

- El principio de la NO redundancia es que para las variables dummy, siempre se toma $k-1$ (en este caso $s-1$) períodos para controlar cierto evento.

- Sea el caso donde se quiere controlar una semana por los días de la semana. ¿Qué efecto posee el domingo? ¿Cómo se asignaría el domingo en la ecuación de mejor ajuste?

	$d_{1,t}$	$d_{2,t}$	$d_{3,t}$	$d_{4,t}$	$d_{5,t}$	$d_{16,t}$
Lunes	1	0	0	0	0	0
Martes	0	1	0	0	0	0
Miércoles	0	0	1	0	0	0
Jueves	0	0	0	1	0	0
Viernes	0	0	0	0	1	0
Sábado	0	0	0	0	0	1
Domingo	0	0	0	0	0	0
Lunes	1	0	0	0	0	0
Martes	0	1	0	0	0	0
…	…	…	…	…	…	…

Los predictores: las variables control (dummy variables)

- Solo se necesitan seis variables dummy para codificar siete categorías. Esto se debe a que la séptima categoría (en este caso, el domingo) es capturada por la intersección, y se especifica cuando las variables dummy están todas especificadas como cero.

- Se podría intentar agregar una séptima variable dummy. Esto se conoce como la "trampa de variable dummy", porque hará que la regresión falle por la redundancia (saturación de parámetros). La regla general es usar menos variables dummy que categorías. Entonces, como regla general utilizamos $s-1$ variables dummy. El modelo como tal estaría dado por la ecuación

$$y_t = \beta_0 + \beta_1 t + \beta_2 d_{2,t} + \cdots + \beta_{s+1} d_{s-1,t} \varepsilon_t$$

- **Recordar**: la interpretación de cada coeficiente se hace respecto a la base seleccionada (en este caso el domingo). Si dice el efecto visto al lunes, en comparación a lo reportado al domingo.

Los predictores: las variables control (dummy variables)

- Entonces, volviendo al ejemplo de la producción de cerveza en Australia, ajustemos un modelo con tendencia y estacionalidad. Este estaría dado por la ecuación:

$$y_t = \beta_0 + \beta_1 t + \beta_2 d_{2,t} + \beta_3 d_{3,t} + \beta_4 d_{4,t} \varepsilon_t$$

- En el R-Studio debemos poner:

```
serie.2.estacionalidad <- tslm(serie.1 ~ trend+ season)
summary(serie.2.estacionalidad)
```

- Obtenemos como resultado lo siguiente (siguiente diapositiva). ¿Qué podemos decir del modelo? ¿Y los coeficientes? ¿Otras medidas que les llame la atención?

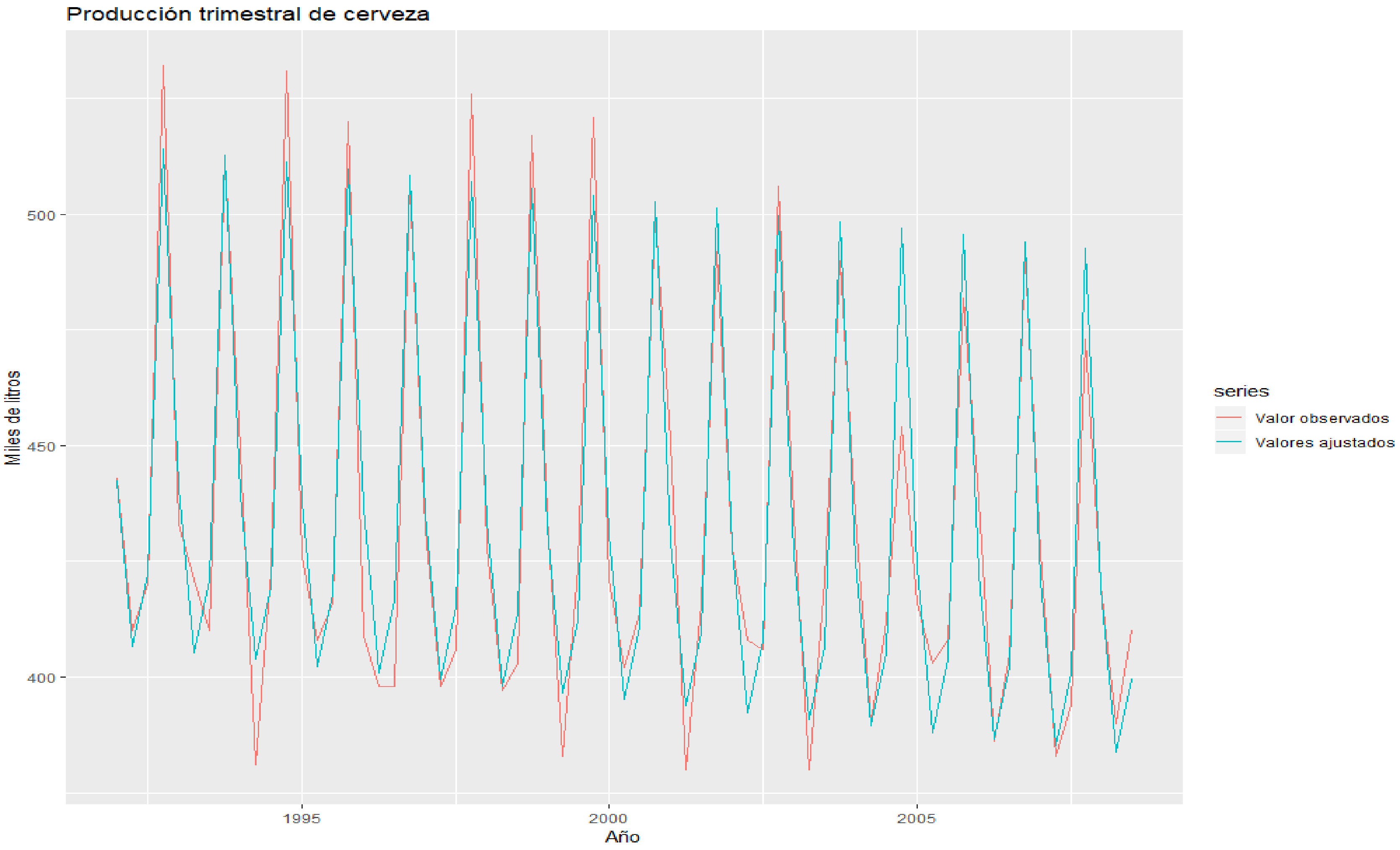
Los predictores: las variables control (dummy variables)

```
Call:
tslm(formula = serie.1 ~ trend + season)

Residuals:
    Min      1Q  Median      3Q     Max 
-42.916  -7.877  -0.070   7.594  21.494 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 442.78341    3.98067 111.233  < 2e-16 ***
trend        -0.35886    0.07866  -4.562 2.45e-05 ***
season2     -35.40585    4.26869  -8.294 1.22e-11 ***
season3     -19.28229    4.27086  -4.515 2.89e-05 ***
season4      72.79268    4.33485  16.792  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 12.44 on 62 degrees of freedom
Multiple R-squared:  0.923,	Adjusted R-squared:  0.918 
F-statistic: 185.8 on 4 and 62 DF,  p-value: < 2.2e-16
```

- Nótese que ahora todos los coeficientes son significativos.
- Además las medidas de ajuste son mejores que para el modelo con tendencia.

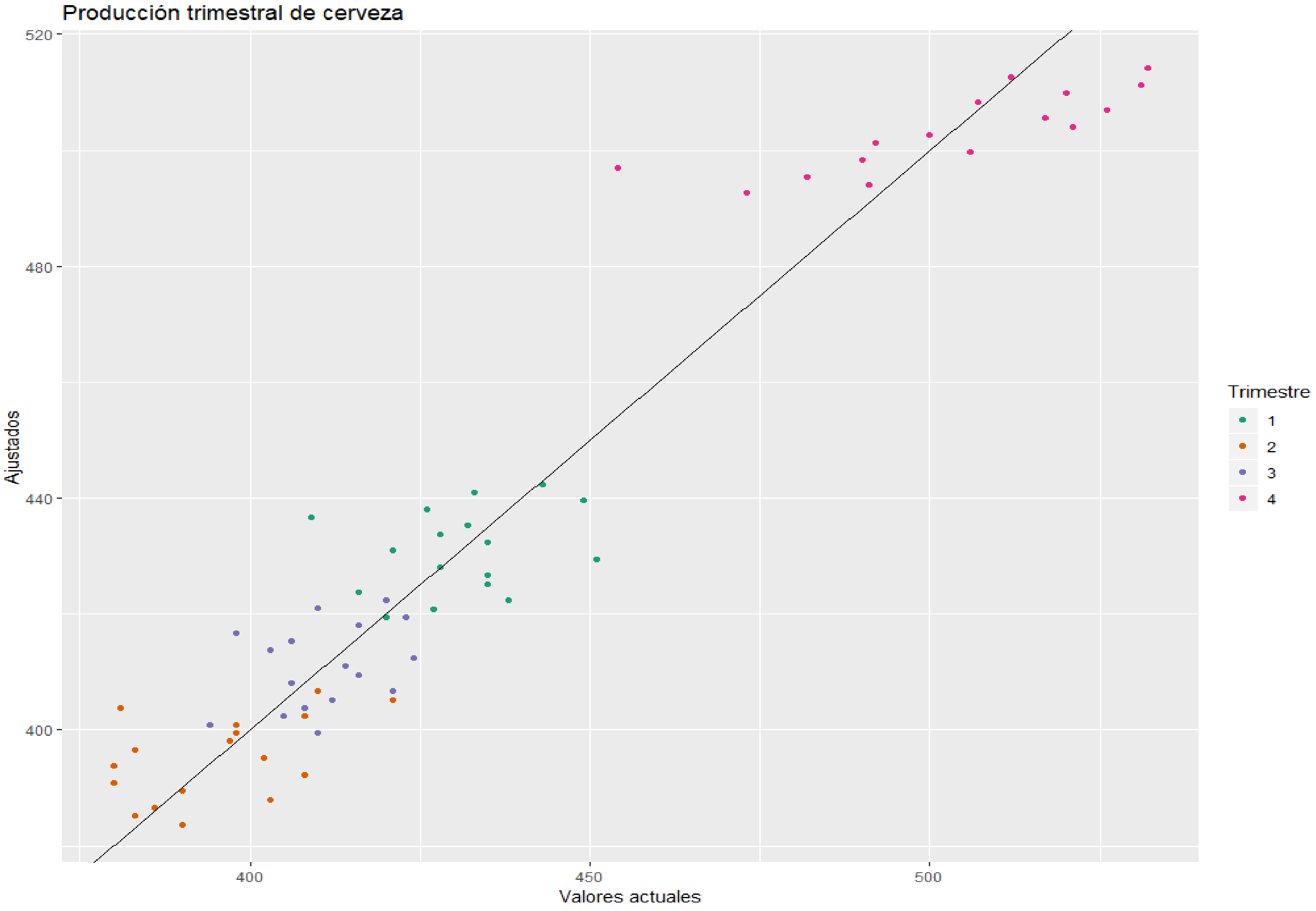
Los predictores: las variables control (dummy variables)

Visualización de los valores observador y ajustados



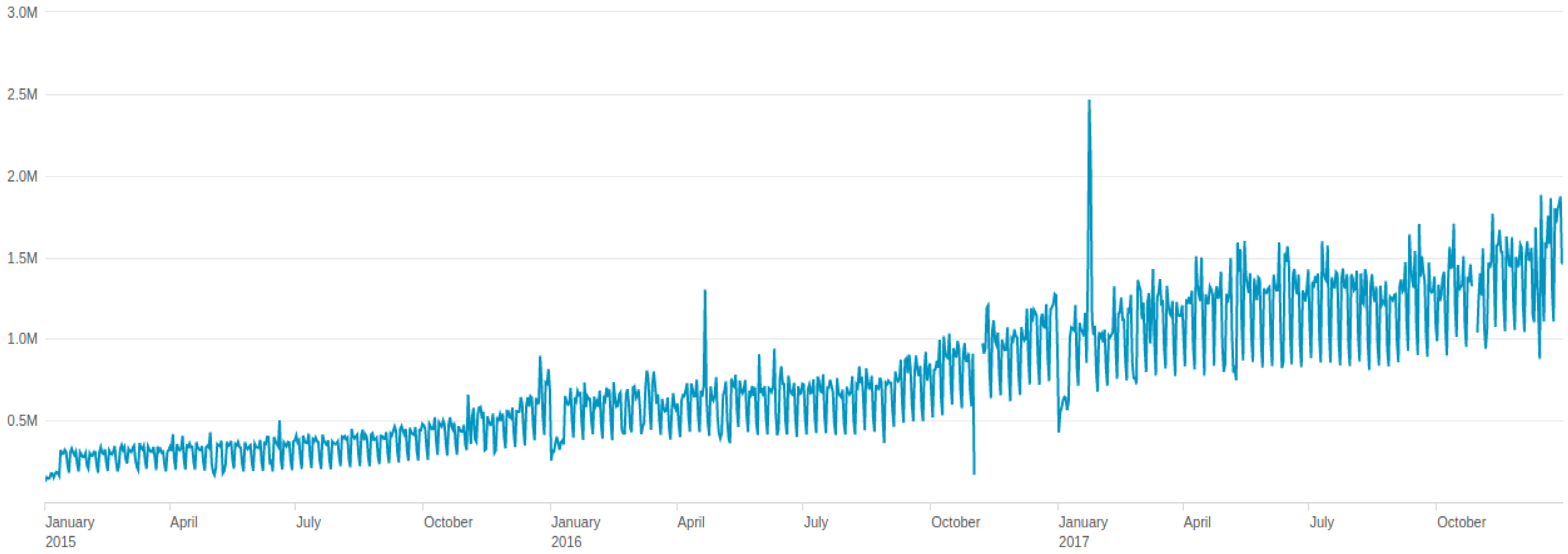
Los predictores: las variables control (dummy variables)

Rendimiento interno de los valores ajustados contra los observados



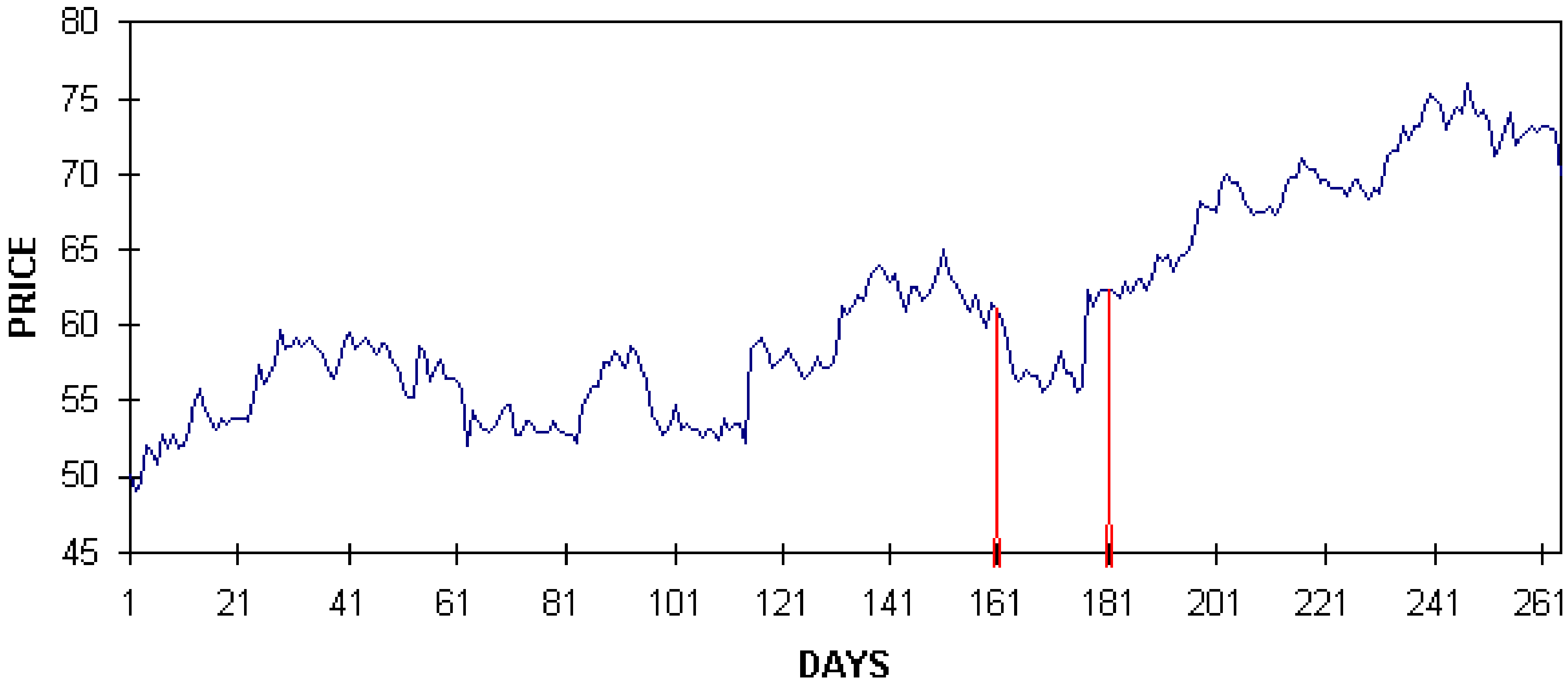
Los predictores: otras variaciones

- **Las variables de intervención**: pueden ser efecto coyunturales o fortuitos que producen un cambio brusco en la serie. Se puede utilizar una variable dummy o una denominada como variable “spike” Variable. En R-stdio se utiliza la función spikeslab().



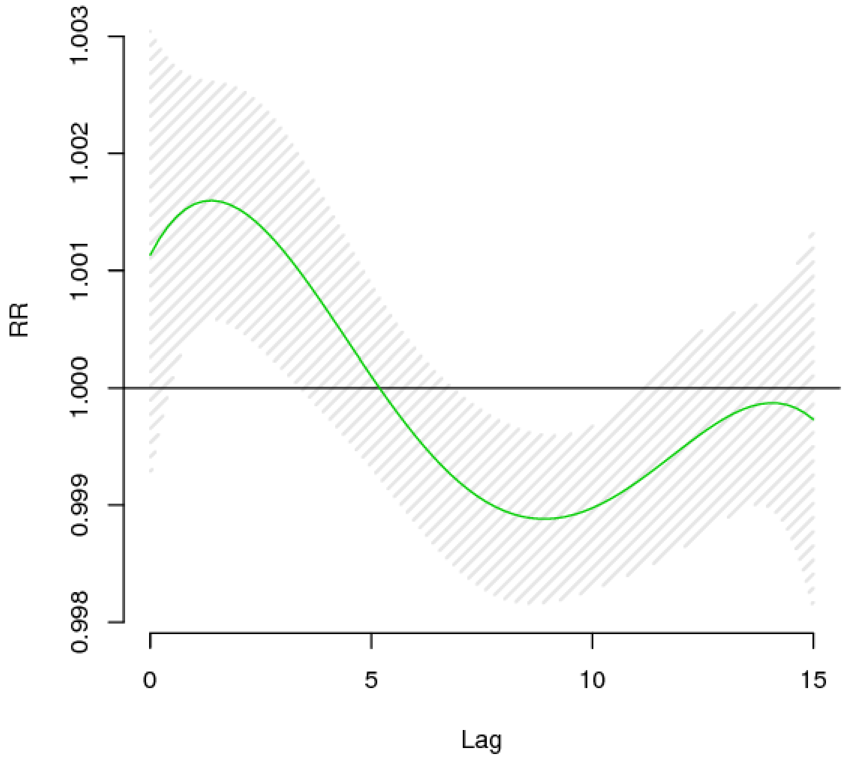
Los predictores: otras variaciones

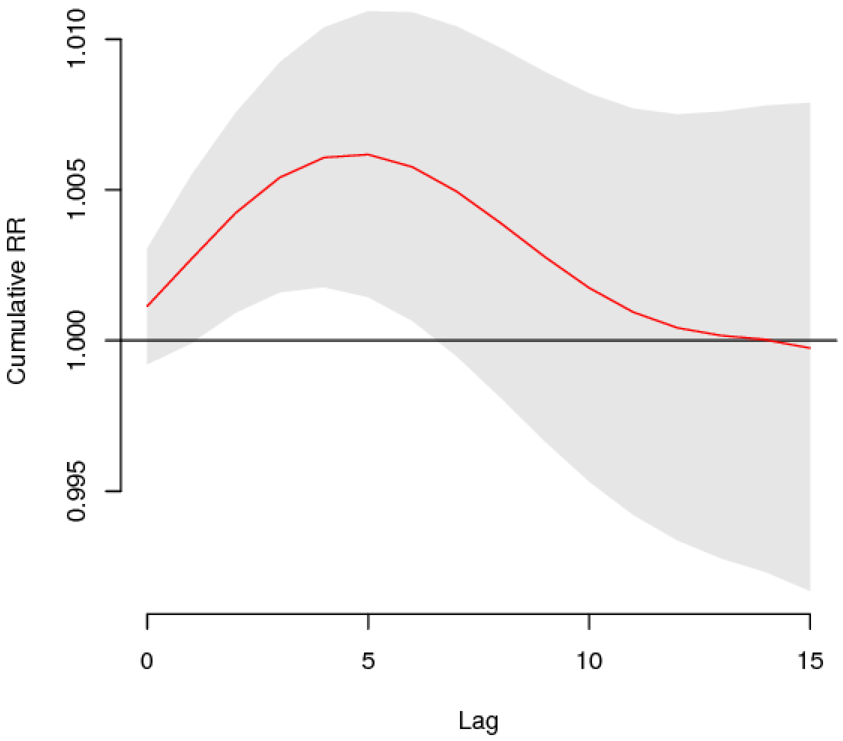
- **Trading days**: El número de días de negociación en un mes puede variar considerablemente y puede tener un efecto sustancial en los datos de ventas. Para permitir esto, se puede incluir el número de días de negociación en cada mes como predictor. En R-stdio se utiliza la función bizdays().



Los predictores: otras variaciones

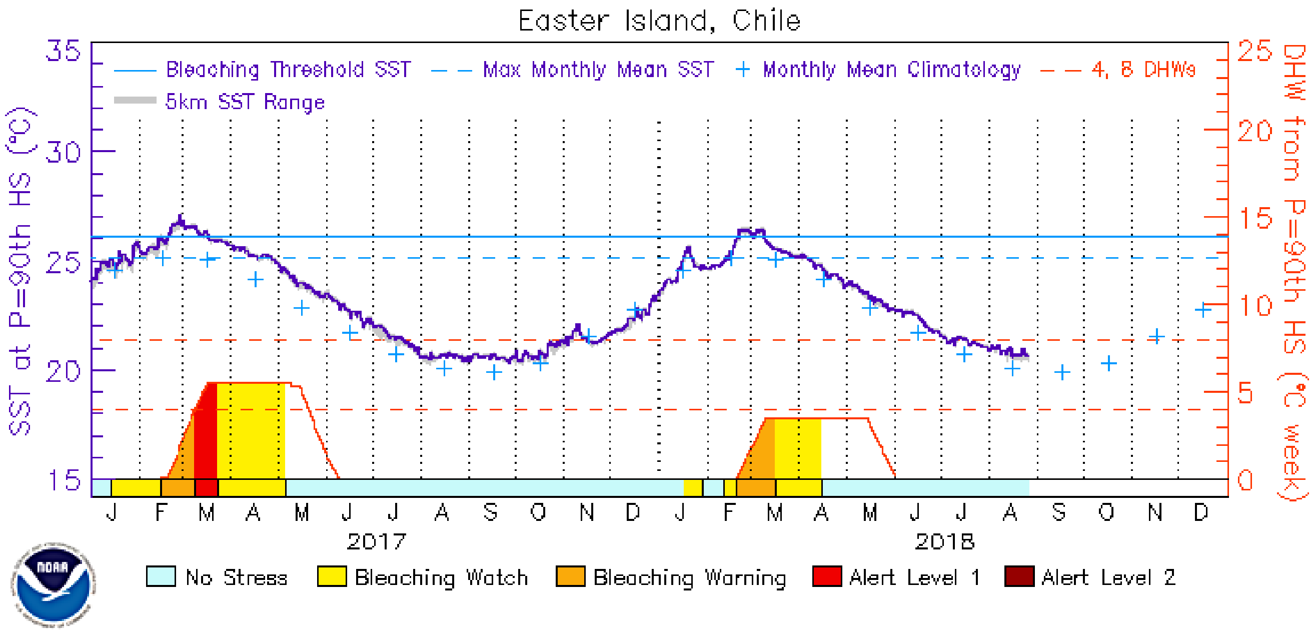
- **Retrasos distribuidos**: A menudo es útil incluir, a forma de ejemplo, los gastos de publicidad como un predictor. Sin embargo, dado que el efecto de la publicidad puede durar más allá de la campaña real, debemos incluir los valores rezagados del gasto publicitario. Sería incluir el tiempo hasta que ocurra el evento. En R-Studio los paquetes dlnm y dlsem se crearon para esto.





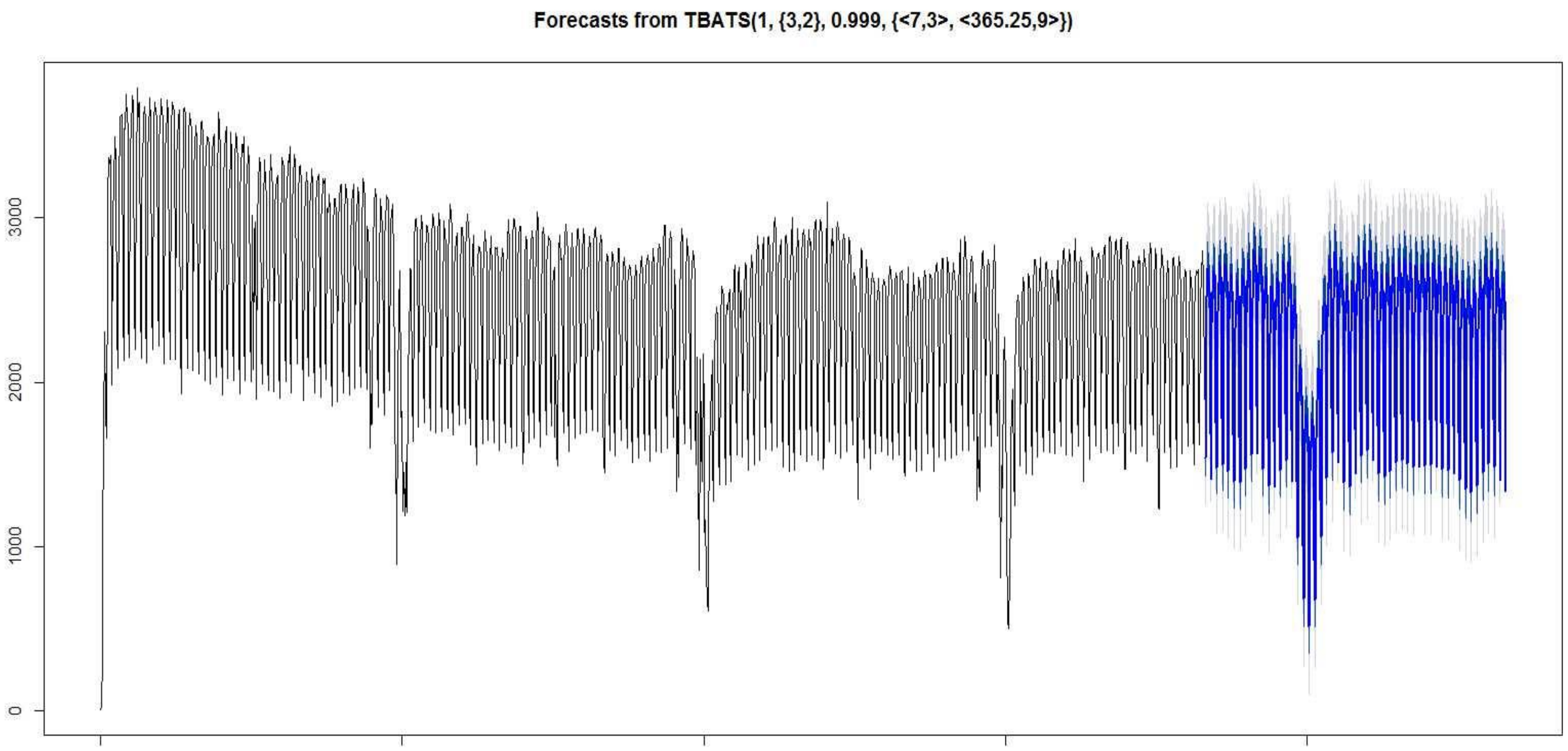
Los predictores: otras variaciones

- **Pascua (Easter)**: la Pascua difiere de la mayoría de las vacaciones porque no se celebra en la misma fecha cada año y su efecto puede durar varios días. En este caso, una variable dummy se puede usar con valor uno donde las vacaciones caen en el período de tiempo particular y cero de lo contrario. Con los datos mensuales, si la Semana Santa cae en marzo, la variable ficticia toma el valor 1 en marzo, y si cae en abril, la variable ficticia toma el valor 1 en abril. Cuando la Semana Santa comienza en marzo y finaliza en abril, la variable ficticia se divide proporcionalmente entre meses. En R-Studio se utiliza la función easter().



Los predictores: otras variaciones

- **Series de Fourier**: una alternativa al uso de variables dummy estacionales, especialmente para períodos estacionales largos, es usar términos de Fourier. Una serie temporal en términos del seno y del coseno de las frecuencias correctas puede aproximar cualquier función periódica. Podemos usarlos para patrones estacionales. En R-Studio se utiliza la función fourier().



$$x_{1,t} = \sin\left(\frac{2\pi t}{m}\right), x_{2,t} = \cos\left(\frac{2\pi t}{m}\right), x_{3,t} = \sin\left(\frac{4\pi t}{m}\right),$$

$$x_{4,t} = \cos\left(\frac{4\pi t}{m}\right), x_{5,t} = \sin\left(\frac{6\pi t}{m}\right), x_{6,t} = \cos\left(\frac{6\pi t}{m}\right),$$

Índice

Medidas de rendimiento

- Entonces recordemos que se poseen las medidas basadas en diferencias y las que se sirven de los criterios de información. A continuación se presentan ambos casos.

- Criterios clásicos de diferencias: utilizamos la función accuracy().

```
accuracy(serie.2.estacionalidad)
```

Estadístico	Valor
RMSE	11.96985
MAE	9.205978
MPE	-0.07013125
MAPE	2.112422
MASE	0.266796

Medidas de rendimiento

- Medidas según los criterios de información : utilizamos la función CV().

```
CV(serie.2.estacionalidad)
```

Estadístico	Valor
CV	167.2714550
AIC	344.6403984
AICc	346.0403984
BIC	357.8685541
AdjR2	0.9180188

Los pronósticos en RTS

- Para obtener los pronósticos a partir de una regresión, necesitamos dos cosas. Si pensamos en lo que necesitábamos para estimar en una regresión para datos transversales, ¿qué son esos dos elementos?

Los pronósticos en RTS

Los valores ajustados

Sea la ecuación de los valores ajustados: $\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \hat{\beta}_2 x_{2,t} + \cdots + \hat{\beta}_k x_{1,k} + \varepsilon_t$

$$\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'y = Hy$$

Utilizamos la matriz sombrero definida como: $H = X(X'X)^{-1}X'$

En la obtención del valor $\hat{y}_t$, nos interesa el valor de la matriz H en un determinado momento t, definido entonces por sus diagonales: $H_t = (h_1, h_2, \dots, h_T)$

Los valores $(h_1, h_2, \dots, h_T)$ se obtiene por validación cruzada a partir de la ecuación:

$$CV = \frac{1}{T}\sum_{t=1}^{T}\left[\frac{e_t}{1-h_t}\right]^2$$

Los pronósticos en RTS

Los pronósticos

Sea $\mathbf{x}^*$ el vector sobre el horizonte de pronóstico *H.* Los valores pronósticos se obtiene a partir de la ecuación:

$$\widehat{\boldsymbol{y}} = \mathbf{x}^*\widehat{\boldsymbol{\beta}} = \mathbf{x}^*(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{y}$$

La variancia de los pronósticos

La variancia está dada por la ecuación: $\hat{\sigma}_e^2\,[1 + \mathbf{x}^*(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}(\mathbf{x}^*)']$

Intervalos de confianza

Finalmente, el intervalo de confianza de los pronósticos se define como:

$$\widehat{\boldsymbol{y}} \pm \boldsymbol{t}_{\mathbf{1}-\frac{\boldsymbol{\alpha}}{\mathbf{2}};\mathbf{1}-\boldsymbol{\alpha}}\hat{\sigma}_e^2\sqrt{1 + \mathbf{x}^*(\boldsymbol{X}'\boldsymbol{X})^{-1}(\mathbf{x}^*)'}$$

$$\widehat{\boldsymbol{y}} \pm \boldsymbol{t}_{\mathbf{1}-\frac{\boldsymbol{\alpha}}{\mathbf{2}};\mathbf{1}-\boldsymbol{\alpha}}\hat{\sigma}_e^2\sqrt{1 + \frac{1}{T} + \frac{(x-\bar{x})^2}{(T-1)s_x^2}}$$

Los pronósticos en RTS

- En los pronósticos de series temporales utilizando la regresión, existen dos tipos de pronósticos: el ex-ante y el ex-post. Dependiendo de lo que se supone que es conocido cuando se calculan los pronósticos…

Pronóstico ex-ante: los pronósticos ex-ante son aquellos que se hacen usando solo la información que está disponible en el momento t y la pasada.

Pronóstico ex-post: los pronósticos ex-post son aquellos que se hacen usando información posterior sobre los predictores. Por lo tanto, para generar pronósticos ex-ante, el modelo requiere pronósticos también de otros predictores para un cierto horizonte ***h.***

Una evaluación comparativa de los pronósticos **ex-ante** y **ex-post** puede ayudar a separar las fuentes de la incertidumbre del pronóstico. Esto mostrará si los errores del pronóstico han surgido debido a pronósticos deficientes del predictor o debido a un modelo de pronóstico deficiente.

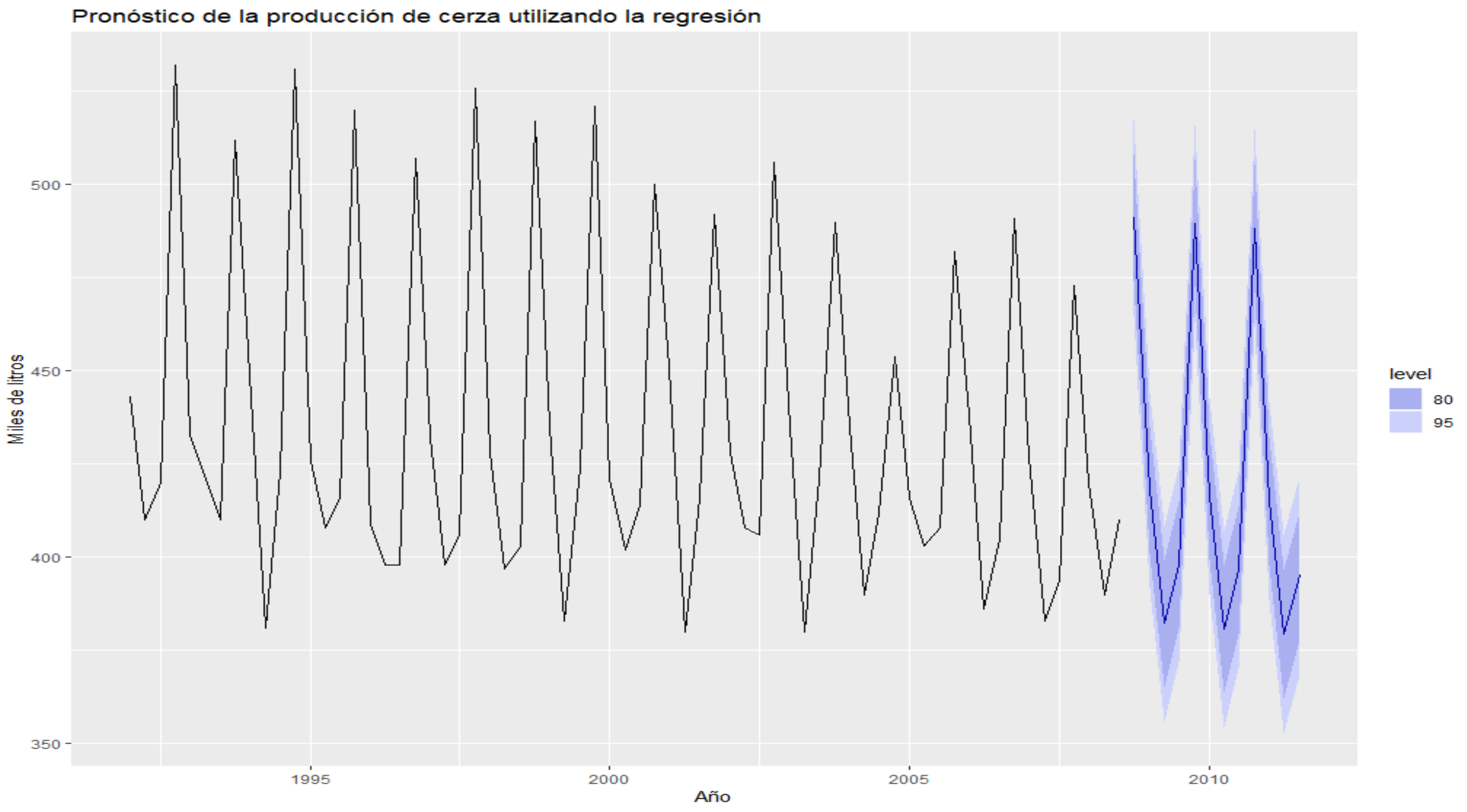
Los pronósticos en RTS

- ¿Nosotros cuál es la que utilizamos? ¿Podrían dar otro ejemplo de **ex-post**?

Los pronósticos en RTS

- Siguiendo el ejemplo de la estimación de litros de cerveza en Australia, vamos a pronosticar para los siguientes 3 años (h=12). Las siguientes figuras muestra los resultados del pronóstico.

Periodo	Pronóstico
2008 Q4	491.1739
2009 Q1	418.0224
2009 Q2	382.2577
2009 Q3	398.0224
2009 Q4	489.7385
2010 Q1	416.5870
2010 Q2	380.8223
2010 Q3	396.5870
2010 Q4	488.3031
2011 Q1	415.1515
2011 Q2	379.3868
2011 Q3	395.1515



Los pronósticos en RTS

- ¿Qué podemos decir de la producción de cerveza en los próximos 3 años?

¿Se relaciona con algún otra variable temporal para explicar dicho comportamiento?

Índice

1. Regresión: transversal vs longitudinal
2. Estimación de los modelos RST
3. Los predictores en la regresión
4. Medidas de rendimiento y el pronóstico
5. Otras variantes no lineales

Otras variantes no lineales de la *RST*

- Aunque la relación lineal asumida hasta ahora en este capítulo es a menudo adecuada, hay muchos casos en los que una forma funcional no lineal es mejor. Para mantener las cosas simples en esta sección, asumimos que solo tenemos un predictor.

- Existen varias formas no lineales conocidas: logaritmo, funciones cuadráticas, exponenciales, cambios de nivel, etc.

- Para optar por una función no línea, sea que cambiamos el valor de la variable observada, el valor de la variable predictora, o cambiamos la funciones de enlace que está ligada a la variable predictora. Estas dos ecuaciones presentan todos los casos:

$$\ln(y) = \beta_0 + \beta_1 \ln(\mathrm{x}) + \varepsilon$$

$$y = f(x) + \varepsilon$$

Otras variantes no lineales de la *RST*

¿Bajo la función $y = f(x) + \varepsilon$, qué modelos conocemos donde se aplica esta lógica?

Otras variantes no lineales de la *RST*

- Otra opción muy común donde la linealidad no se aplica es en los cambios de nivel a partir de un determinado periodo. Esto se presente como sigue:

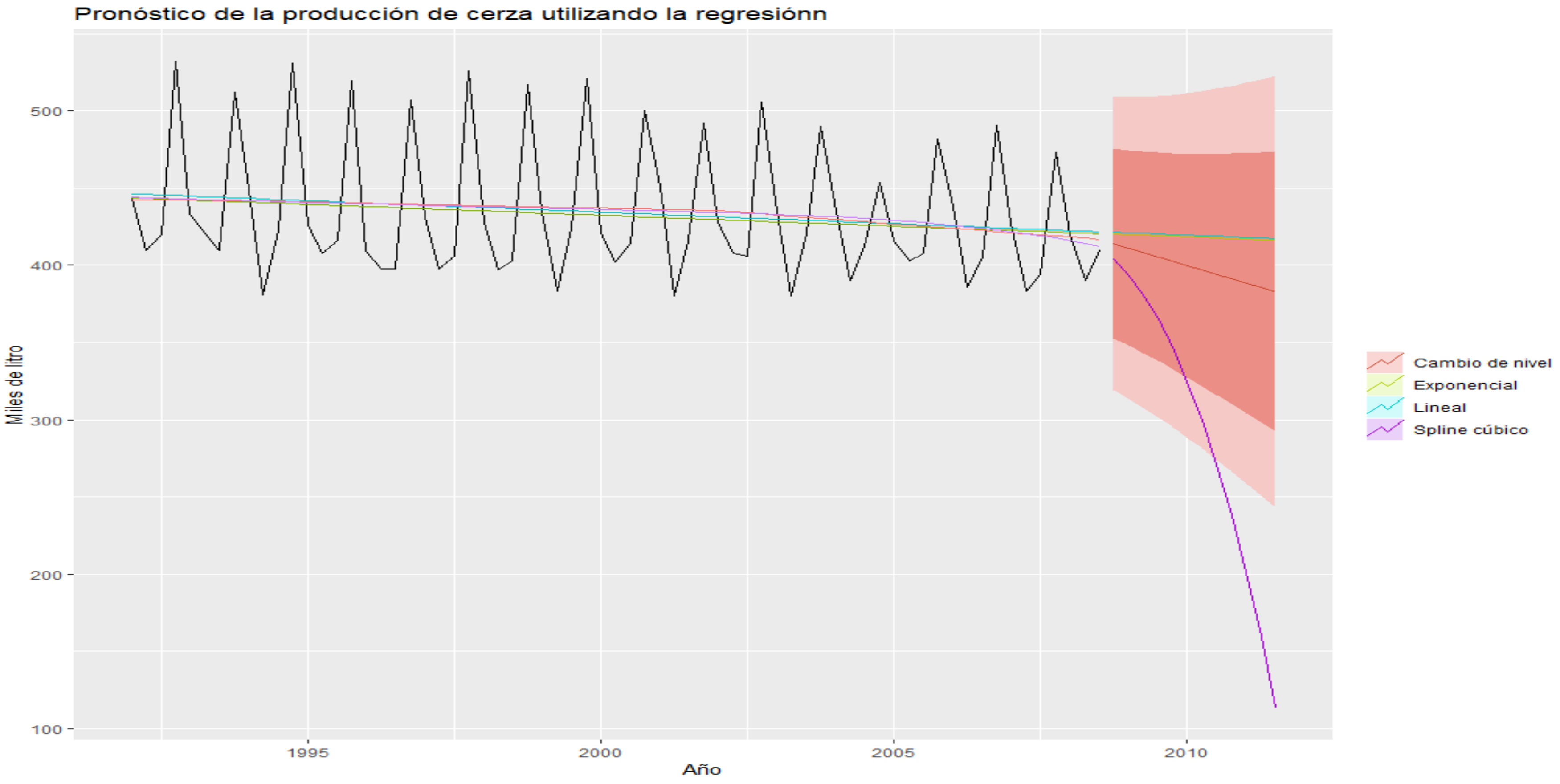
$$x_{j,t} = (x - t)_+ = \begin{cases} 0 & x < t \\ (x - t) & x \geq t \end{cases}$$

- Finalmente, la opción más famosa de no linealidad son las regresiones ***splines***. Estas se definen ajustando el valor de los variables según ciertas constantes o transformaciones para un determinado momento o punto:

$$x_1 = x \quad x_2 = x^2 \quad x_3 = \ x^3 \quad x_4 = (x - t_1) + \cdots + x_k = (x\_t_{k-3})$$

Otras variantes no lineales de la *RST*

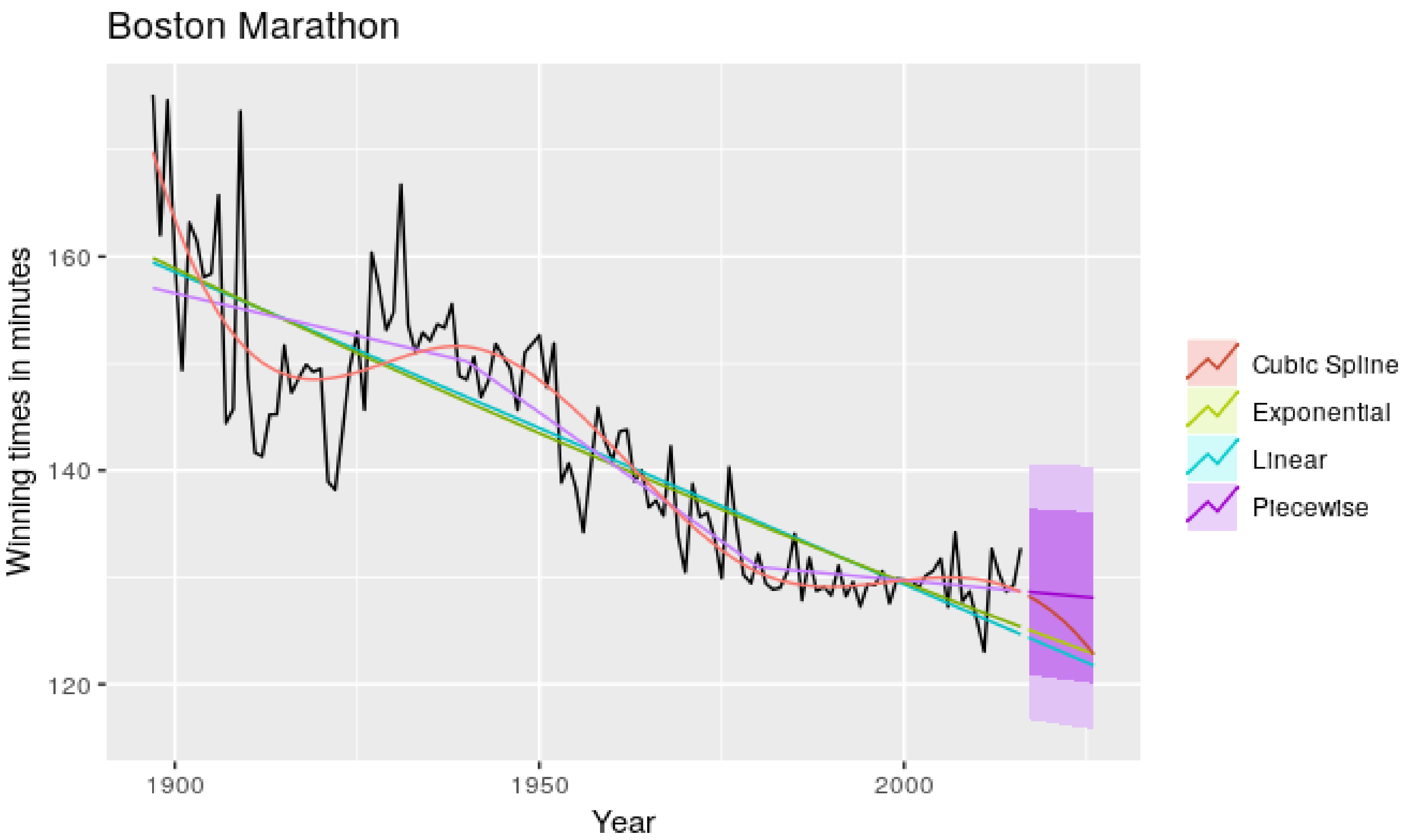
- Apliquemos como forma no lineales los cambios de nivel, una función con cambio exponencial ($\lambda = 0$), una forma lineal y un Spline Cúbico. El resultado es el siguiente:



¿Qué sucedió?

Otras variantes no lineales de la *RST*

- Otro ejemplo donde es más pertinente aplicar las mismas opciones no lineales.



¿Qué piensan al respecto?

Índice

1 Regresión: transversal vs longitudinal

2 Estimación de los modelos RST

3 Los predictores en la regresión

4 Medidas de rendimiento y el pronóstico

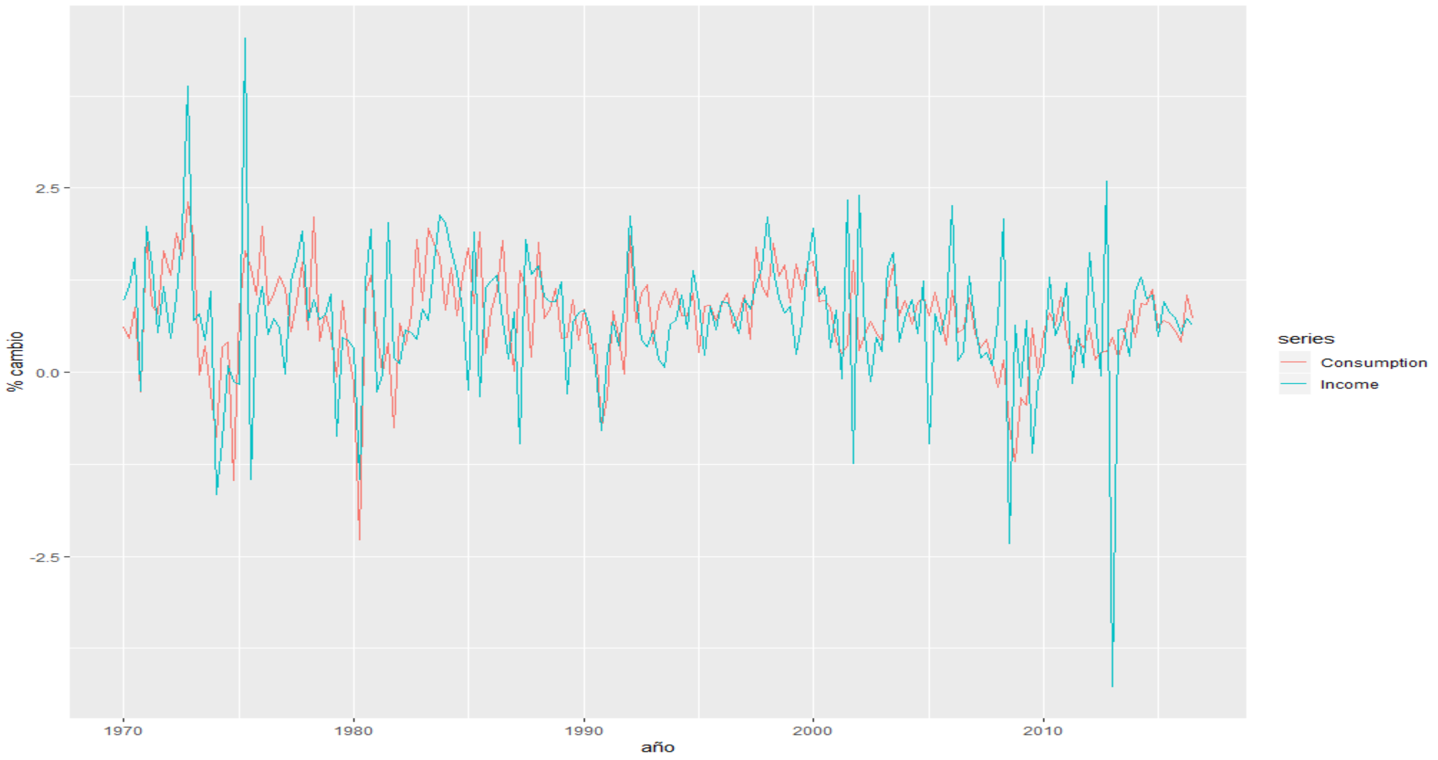
5 Otras variantes no lineales

6 Regresión múltiple

La regresión multivariada en *RST*

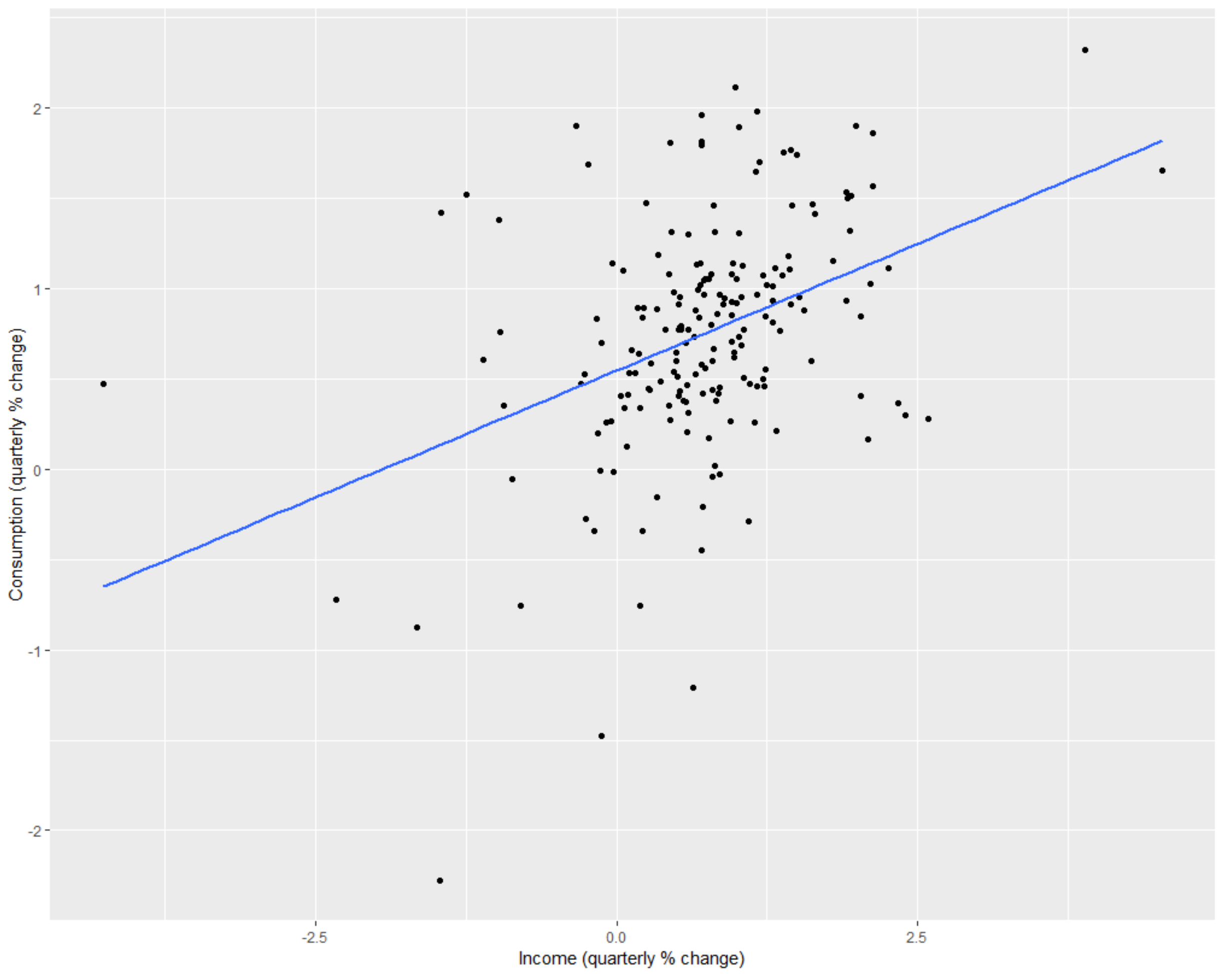
- Aunque el curso se centra en las series univariadas, se presenta brevemente la utilización de las regresiones múltiples en las series temporales.

- Empecemos por estimar una serie por algún otro predictor o variable también temporal. Utilizaremos el archivo uschange de la librería fpp2.



La regresión multivariada en *RST*

- Relación entre el consumo y el ingreso. ¿Sería pertinente pronosticar el consumo por el ingreso medio?

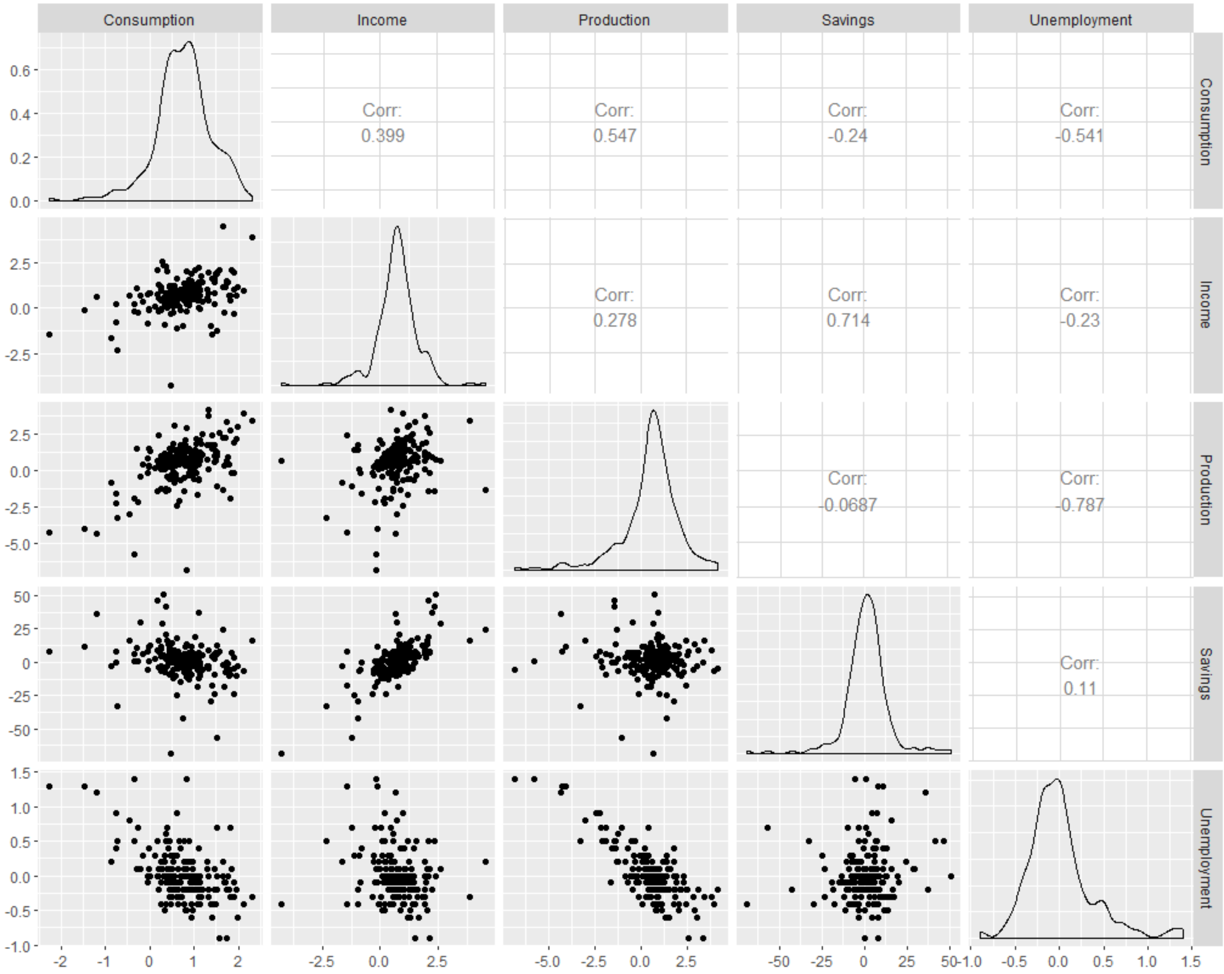


Ecuación de regresión

```
> summary(regre.multi.1)

Call:
tslm(formula = Consumption ~ Income, data = uschange)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.40845 -0.31816  0.02558  0.29978  1.45157 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.54510    0.05569   9.789  < 2e-16 ***
Income       0.28060    0.04744   5.915 1.58e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6026 on 185 degrees of freedom
Multiple R-squared:  0.159,	Adjusted R-squared:  0.1545 
F-statistic: 34.98 on 1 and 185 DF,  p-value: 1.577e-08
```

La regresión multivariada en *RST*

- ¿Sería pertinente pronosticar el tipo del consumo según las otras variables?



La regresión multivariada en *RST*

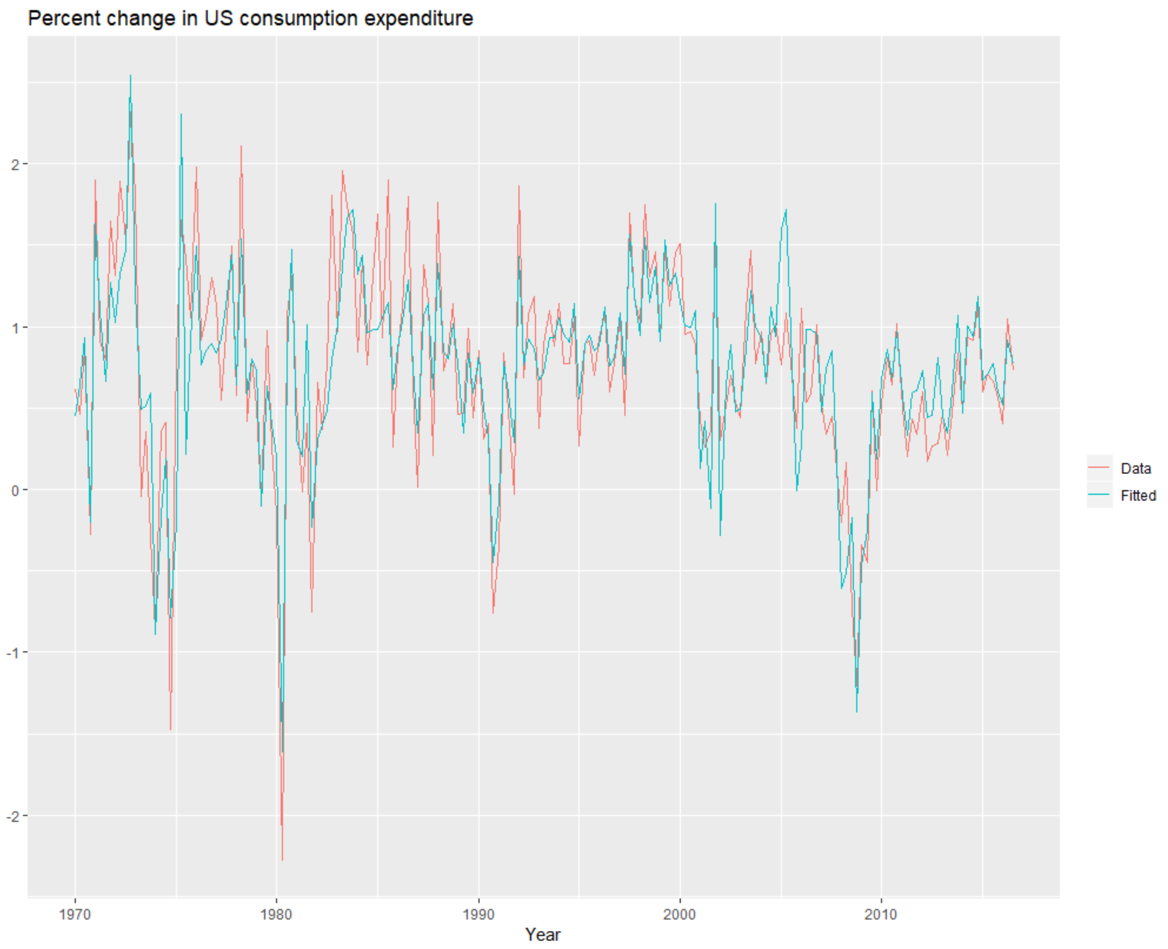
- Estimemos el modelo según las otras variables predictoras. Veamos el resultado:

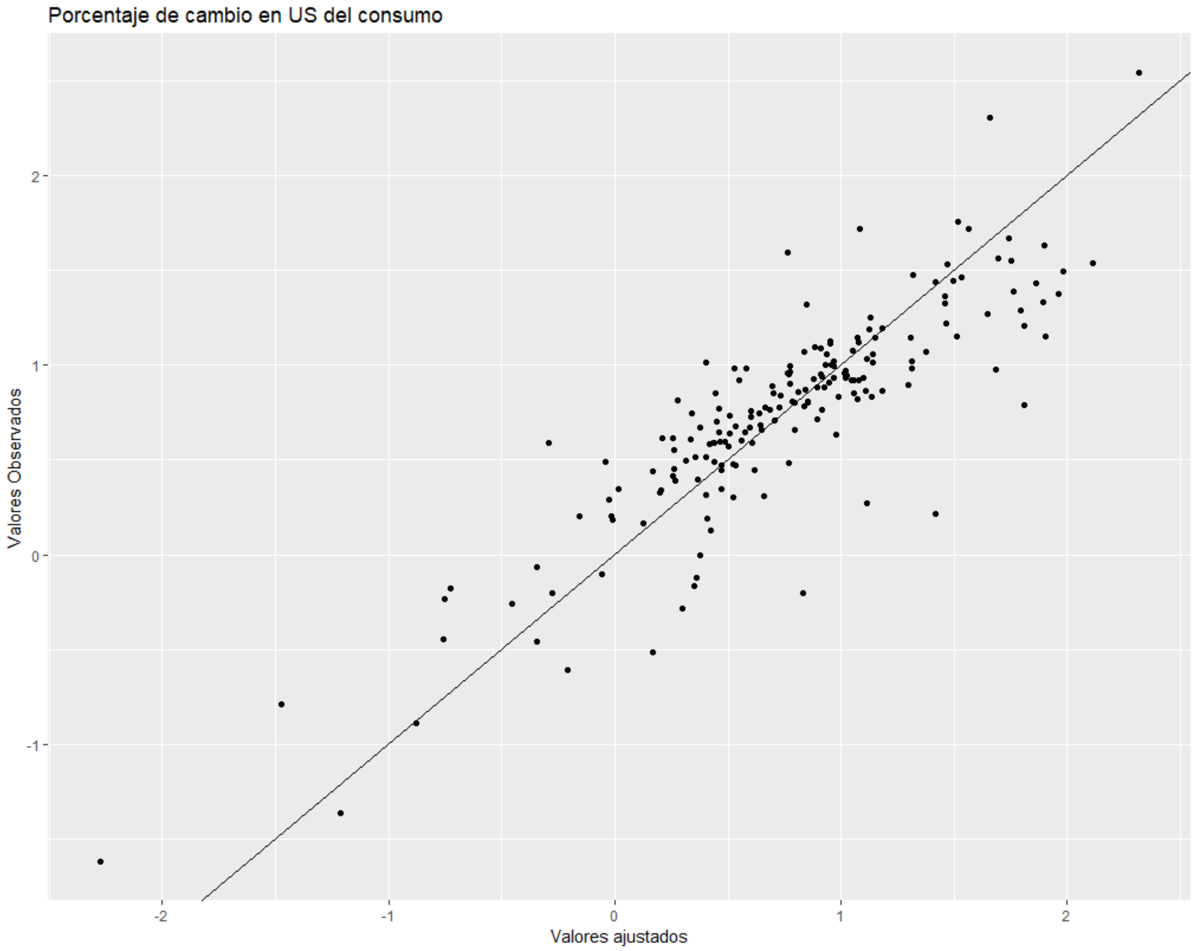
```
> summary(regre.multi.2)

Call:
tslm(formula = Consumption ~ Income + Production + Unemployment +
    Savings, data = uschange)

Residuals:
     Min       1Q   Median       3Q      Max
-0.88296 -0.17638 -0.03679  0.15251  1.20553

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.26729    0.03721   7.184 1.68e-11 ***
Income        0.71449    0.04219  16.934  < 2e-16 ***
Production    0.04589    0.02588   1.773   0.0778 .
Unemployment -0.20477    0.10550  -1.941   0.0538 .
Savings      -0.04527    0.00278 -16.287  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3286 on 182 degrees of freedom
Multiple R-squared:  0.754,	Adjusted R-squared:  0.7486
F-statistic: 139.5 on 4 and 182 DF,  p-value: < 2.2e-16
```

¿Qué concluimos?

La regresión multivariada en *RST*

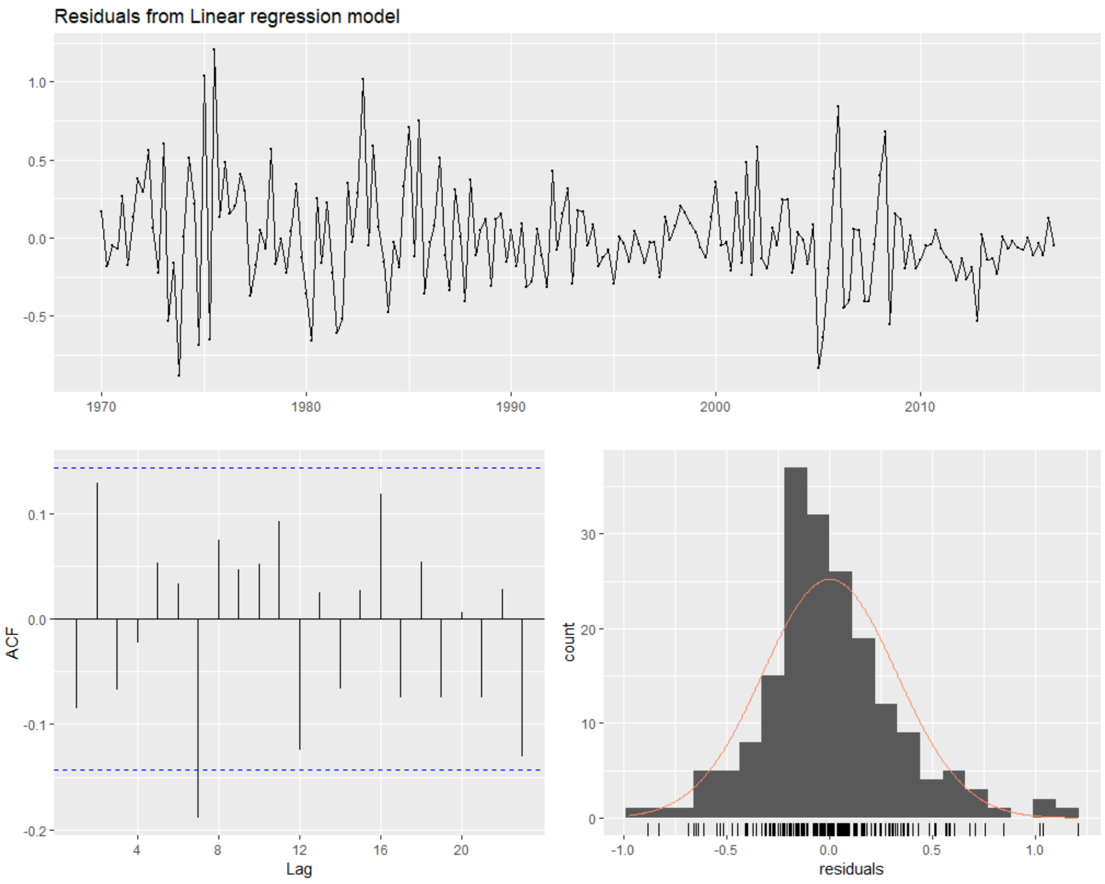
- Evaluemos los resultados de la regresión. ¿Parece ser una buena opción?

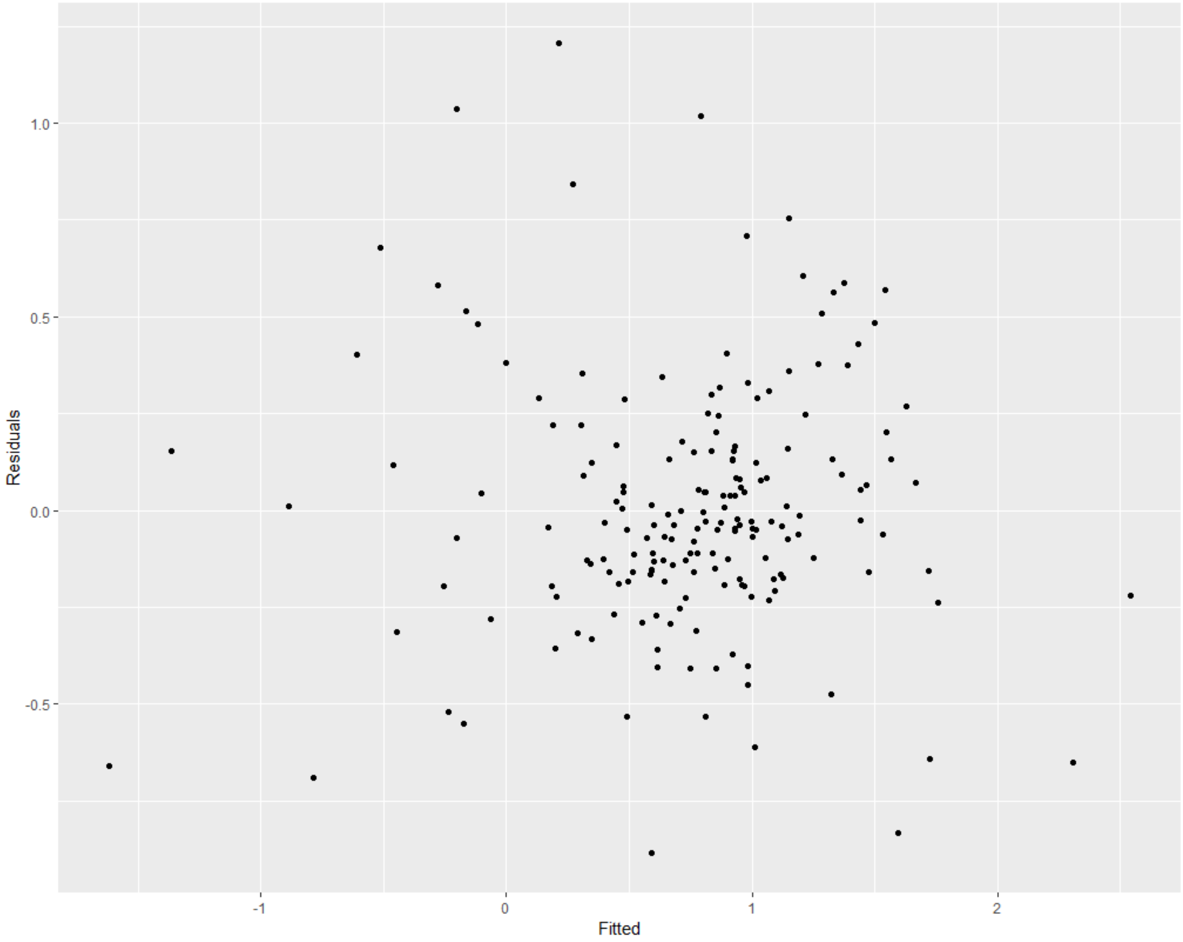




La regresión multivariada en *RST*

- ¿Debemos evaluar la bondad y ajuste del modelo? Sí…





Conclusión y discusión

- La clase presentó la regresión en las series temporales. Se estableció el contexto de la regresión en datos transversales como longitudinales.

- Se analizaron las diferentes secciones de una regresión con datos temporales: estimación, selección de los predictores, evaluación de las estimación y los pronósticos.

- También se repasaron otras opciones no lineales de la RTS, y la regresión múltiple en las series temporales

- Aunque la *RTS* constituye una herramienta de análisis, las series se suelen ajustar mejor mediante los modelos paramétricos exponenciales.

¿Dudas o preguntas?

The End