Fundamentos de analítica 2: Modelos de pronósticos basados en series de tiempo.

Diego Agudelo
UNIVERSIDAD ICESI

Temas del día

- Introducción
- Estimación de la tendencia
- Pronóstico de la estacionalidad
- Consideraciones finales

Introducción a las series de tiempo

- Ya hemos visto como encontrar un pronóstico por medio de:
  - Media móvil
  - Suavización Exponencial
  - Simple
  - Holt
  - Holt-Winters

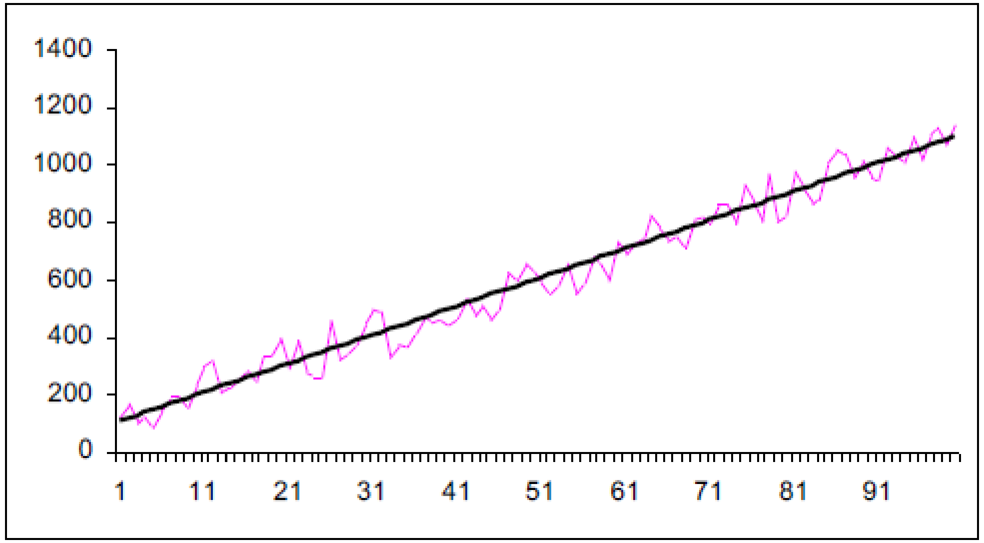
- Ahora veremos otra manera de explicar el comportamiento de una serie.

Diego Agudelo
UNIVERSIDAD ICESI

Estimación de tendencia –lineal

- Tendencia: “Evolución lenta y a largo plazo”

$$X_t = \beta_0 + \beta_1 t$$



Estimación de tendencia –lineal

- Entonces el pronóstico será

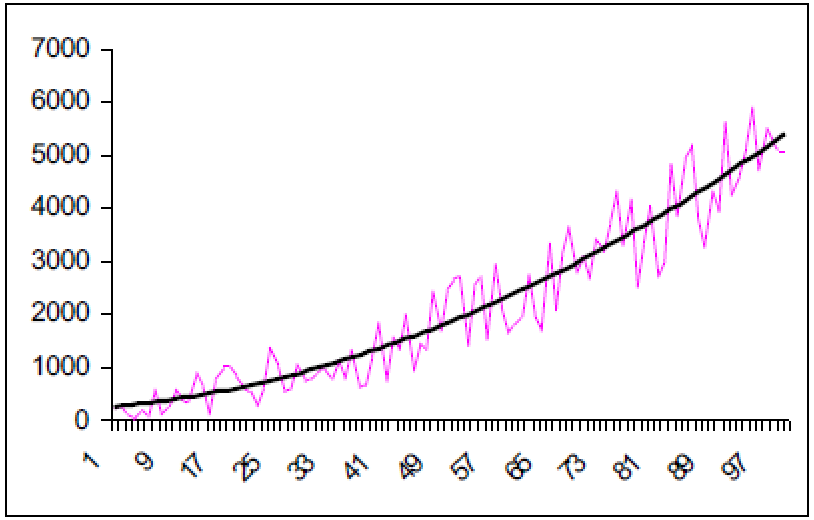
$$F_{t+m} = \beta_0 + \beta_1 (t + m)$$

$$F_{t+m} = pronóstico\ (m\ adelante)$$

Estimación de tendencia – Cuadrática

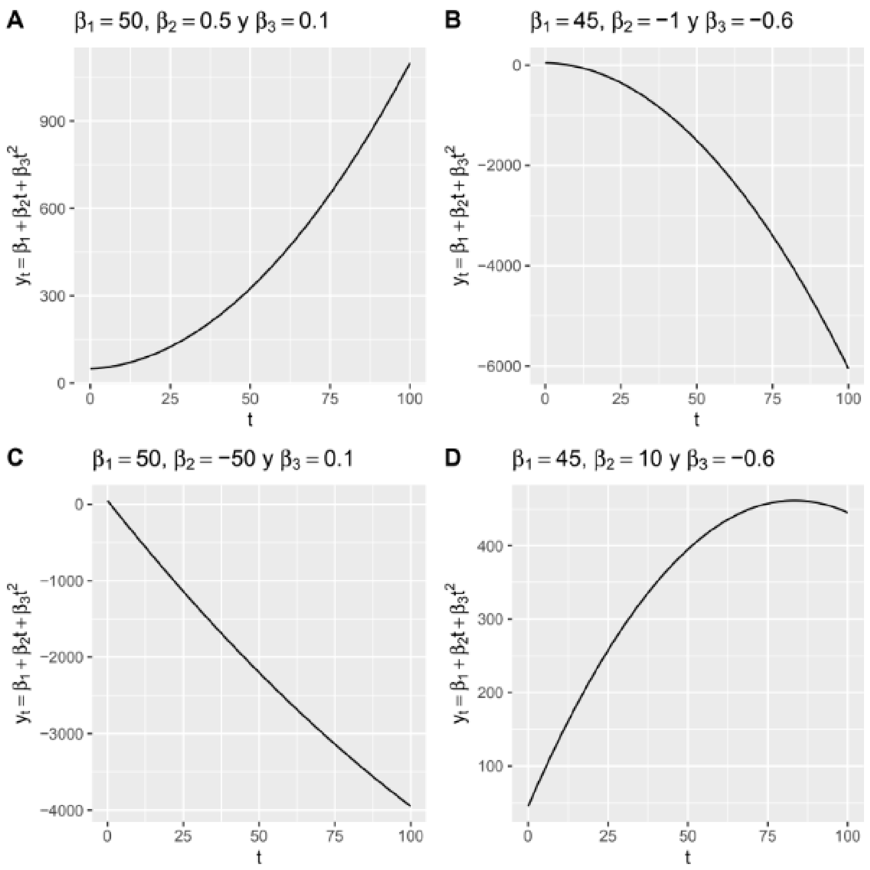
- No siempre las series tienen una tendencia lineal:

$$X_t = \beta_0 + \beta_1 t + \beta_2 t^2$$



Estimación de tendencia – Cuadrática

- Permite mayor flexibilidad



Diego Agudelo

Estimación de tendencia – Cuadrática

- En este caso:

$$F_{t+m} = \beta_0 + \beta_1 (t+m) + \beta_2 (t+m)^2$$

$$F_{t+m} = pronóstico\ (t+m)$$

Estimación de la estacionalidad

- Variables dummies:

gender	gender_m	gender_f
male	1	0
female	0	1
male	1	0
male	1	0
female	0	1
male	1	0
female	0	1
male	1	0
female	0	1

Estimación de la tendencia + estacionalidad

- Consideremos que tenemos datos trimestrales:

$$D_{1t} = \begin{cases} 1 & si \text{ 1er trimestre} \\ 0 & \text{caso contrario} \end{cases} \qquad D_{3t} = \begin{cases} 1 & si \text{ 3er trimestre} \\ 0 & \text{caso contrario} \end{cases}$$

$$D_{2t} = \begin{cases} 1 & si \text{ 2do trimestre} \\ 0 & \text{caso contrario} \end{cases} \qquad D_{4t} = \begin{cases} 1 & si \text{ 4to trimestre} \\ 0 & \text{caso contrario} \end{cases}$$

$$X_t = \beta_1 t + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t}$$

$$F_{t+m} = X_t = \beta_1 (t+m) + \alpha_1 D_{1(t+m)} + \alpha_2 D_{2(t+m)} + \alpha_3 D_{3(t+m)} + \alpha_4 D_{4(t+m)}$$

Consideraciones finales

- Ya conocemos varias formas de hacer pronósticos.
- ¿Cómo comparar los diferentes modelos?

  - La raíz cuadrada media del error de pronóstico, (RMSE Root Mean Square.)

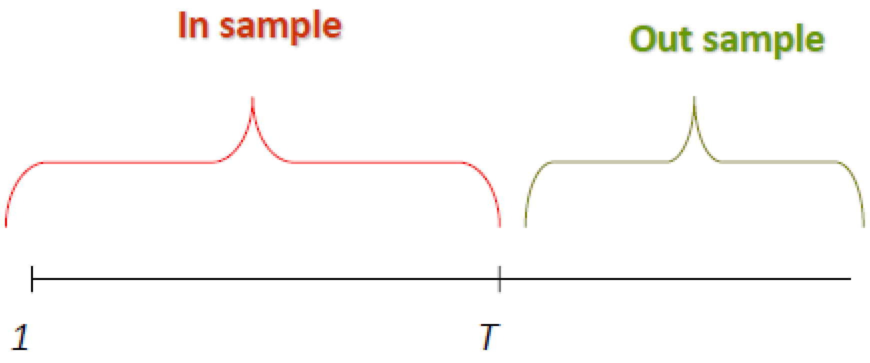
$$\mathrm{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{y}_t - y_t)^2}$$

El mejor modelo es el que tenga un menor Error

Consideraciones finales

- La evaluación se realiza normalmente fuera de la muestra:

  - In sample
  - Out of sample



Consideraciones finales

Nombre	Medida	Breve descripción
Error medio (EM)	$EM = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)$	Mide el sesgo promedio del pronóstico
Error absoluto promedio (EAP) MAE	$EAP = \frac{1}{T}\sum_{t=1}^{T}\left\lvert Y_t - \hat{Y}_t\right\rvert$	No tiene en cuenta el sesgo como si lo hace el EM.
Varianza de error (VE)	$VE = \frac{1}{T}\sum_{t=1}^{T}\left[\left(Y_t - \hat{Y}_t\right) - EM\right]^2$	Mide la dispersión de los errores del pronóstico
Error cuadrático promedio (ECP)	$ECP = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)^2$	penaliza los errores mayores
Error absoluto medio porcentual (EAMP) MAPE	$EAMP = \frac{1}{T}\sum_{t=1}^{T}\left\lvert\frac{Y_t - \hat{Y}_t}{Y_t}\right\rvert$	indica que tan grande son los errores pronosticados comparado con los datos reales en terminos porcentuales

Diego Agudelo

UNIVERSIDAD ICESI

Consideraciones finales

Nombre	*Medida*	*Breve descripción*
Error cuadrático promedio porcentual (ECPP)	$ECPP = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{Y_t - \hat{Y}_t}{Y_t}\right)^2$	penaliza los errores porcentuales mayores
Raíz cuadrada del error cuadrático promedio (RCECP)	$RCECP = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)^2}$	conserva las unidades de la variable
Raíz cuadrada del error cuadrático promedio porcentual (RCECPP)	$RCECPP = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{Y_t - \hat{Y}_t}{Y_t}\right)^2}$	Igualmente penaliza los errores porcentuales mayores
Coeficiente de desigualdad de Theil (U) Theil's U	$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}Y_t^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{Y}_t^2}}$	mide la RCECP en términos relativos

Consideraciones finales

- Guía para la selección de método de pronóstico apropiado:

Método de Pronóstico	*Patrón de datos*	*Cantidad de Datos*	*Horizonte de Pronóstico*
Métodos discrecionales			
Intuición (Ojímetro)	cualquiera	ninguna	cualquiera
Encuestas a clientes	ninguno	ninguna	Mediano a largo plazo
Jurado de opinión ejecutiva	cualquiera	pocos	cualquiera
Delphi	cualquiera	pocos	largo

Diego Agudelo
UNIVERSIDAD ICESI

Consideraciones finales

- Guía para la selección de método de pronóstico apropiado:

Método de Pronóstico	*Patrón de datos*	*Cantidad de Datos*	*Horizonte de Pronóstico*
Métodos Estadísticos			
Promedio Móvil	cualquiera	Mínimo = a la ventana	Muy Corto
Suavización Exponencial - simple	Sin tendencia ni estacionalidad	5 a 10 datos	Corto plazo
Suavización Exponencial - Holt	Con tendencia	10 a 15 datos	Corto a Mediano plazo
Suavización Exponencial – Holt-Winters	Con tendencia y estacionalidad	Por lo menos 3 o 4 por estación	Corto a Mediano plazo

Consideraciones finales

- Guía para la selección de método de pronóstico apropiado:

Método de Pronóstico	*Patrón de datos*	*Cantidad de Datos*	*Horizonte de Pronóstico*
Métodos Estadísticos			
Estimación de la tendencia	Con tendencia sin estacionalidad	Mínimo 10 datos	Corto a Mediano plazo
Estimación de la tendencia + estacionalidad	Con tendencia y estacionalidad	Mínimo 10 datos con 4 o5 por estacionalidad	Corto a Mediano plazo
Regresión múltiple	No importa	Mínimo 10 por variable explicativa	Corto, Mediano y Largo plazo