

Cómputo Estadístico

Tarea 2

Marcelo Alberto Sanchez Zaragoza

25 de octubre de 2021

1. PROBLEMA 1

Si se ha extraído toda la información sistemática con un modelo de pronóstico, entonces lo que queda, el residual, debe ser ruido blanco. Con más precisión, las innovaciones verdaderas son ruido blanco, y si un modelo es buena aproximación de Wold, entonces sus errores de pronóstico a una etapa se deben aproximar al ruido blanco. Los residuales del modelo están en el análogo dentro de la muestra de los errores de pronóstico a una etapa fuera de la muestra. En consecuencia, vemos la utilidad de varias pruebas de la hipótesis que los residuales son ruido blanco. La de Durbin-Watson es la prueba más común. Recuérdese que la d-estadística de Durbin Watson, descrita en el apéndice del capítulo 1 es

$$DW = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}$$

Obsérvese que

$$\sum_{t=2}^T (e_t - e_{t-1})^2 \approx 2 \sum_{t=2}^T e_t^2 - 2 \sum_{t=2}^T e_t e_{t-1}$$

y así $DW = s(1 - \hat{p}(1))$

y entonces la prueba de Durbin-Watson se basa efectivamente solo en la correlación de la primera muestra, y en realidad solo prueba si la primera autocorrelación es cero. En consecuencia, se dice que la de Durbin-Watson es una prueba de correlación seriada de primer orden o correlación en serie de primer orden. Además, la prueba de Durbin-Watson no es válida en presencia de variables dependientes rezagadas. En ambos casos nos gustaría un marco más general y flexible para diagnosticar la correlación seriada. El correlograma de residuales, formado por las autocorrelaciones muestrales residuales, las autocorrelaciones parciales muestrales y los estadísticos Q asociados desempeñan este papel.

- a) Cuando describimos el correlograma en la prueba, nos enfocamos al caso de una serie temporal observada, para el que demostramos que los estadísticos Q se distribuyen como X_m^2 . Sin embargo, ahora deseamos evaluar si las perturbaciones no observadas del modelo son ruido blanco. Para hacerlo usaremos los residuales del modelo, que son estimados de las perturbaciones no observadas. Ya que un modelo se ajusta para obtener los residuales, necesitamos tomar en cuenta los grados de libertad usados. La consecuencia es que la distribución del estadístico Q con la hipótesis de ruido blanco se aproxima mejor como una variable aleatoria X_{m-k}^2 en la que k es la cantidad de parámetros que se estiman. Es la razón, por ejemplo, por la que no se mencionan los valores p (de hecho ni los programas estadísticos los calculan) para los estadísticos Q asociados con el correlograma

de residuales de nuestro modelo de pronóstico de empleo, sino hasta que m̃k.

- b) La prueba h de Durbin es una alternativa en la prueba de Durbin-Watson. Como en el caso de la prueba Durbin-Watson, el fin es detectar correlación en serie de primer orden, pero es valida en presencia de variables dependientes demoradas. Busque información acerca de las generalidades y de la prueba h de Durbin en la bibliografía y escriba lo encontrado.
- c) La prueba de Brewsch-Godfrey es otra alternativa a la de Durbin-Watson. Permite detectar correlación seriada de orden p , y también es valida en presencia de variables rezagadas. Investigue en la bibliografía acerca del procedimiento Brewsch-Godfrey y escriba lo que aprendió.
- d) ¿Cual de las pruebas es la más útil para evaluar las propiedades de residuales a partir de modelos de pronóstico, el correlograma de residuales, la prueba h de Durbin o la prueba de Brewsch-Godfrey? ¿porque?

Solución

Inciso b)

La prueba H de Durbin sigue siendo válida cuando se incluyen valores rezagados de la variable dependiente:

$$y_t = \alpha_0 + \alpha_1 y_t + \dots + \alpha_k y_{t-k} + \alpha_{k-1} y_t + u_t$$

La H de Durbin se define como:

$$H = p \sqrt{\frac{n}{[(1 - nV(\alpha_1))]]}$$

El estadístico d de Durbin-Watson puede no usarse para detectar la correlación serial (de primer orden) en modelos autorregresivos, porque el valor d calculado en tales modelos generalmente tiende hacia 2, que es el valor de d esperado en un verdadera secuencia aleatoria, es decir, al calcular el estadístico d para tales modelos, existe un sesgo incorporado contra el descubrimiento de la correlación serial (de primer orden).

Algunas generalidades:

- El principal inconveniente que tiene este contraste es que si el radicando es negativo, el test falla.
- Dado que la prueba es una prueba para muestras grandes, su aplicación en muestras pequeñas no está estrictamente justificada.
- No importa cuántas variables X o cuántos valores rezagados de Y se incluyan en el modelo de regresión. Para calcular h , necesitamos considerar solo la varianza del coeficiente Y_{t-1}
- La prueba no es aplicable si $nvar(\hat{\alpha}_2)$ excede 1. En la práctica, sin embargo, esto no suele suceder.

Inciso c)

El contraste de Breuch-Godfrey se especifica con la finalidad de analizar si existe o no autocorrelación de orden superior a uno. Es un test de autocorrelación en los errores y residuos estadísticos en un modelo de regresión. Hace uso de los errores generados en el modelo de regresión y un test de hipótesis derivado de éste. La hipótesis nula es que no exista correlación serial de cualquier orden sobre p .

El test es más general que el del estadístico de Durbin-Watson (o estadístico h de Durbin), el cual es solo válido para regresores no-estocásticos y para testear la posibilidad de un modelo autorregresivo de primer orden para los errores de regresión. El test BG no tiene restricciones, y es estadísticamente más poderoso que el estadístico de h de Durbin.

- Los regresores incluidos en el modelo de regresión pueden contener valores rezagados de la regresión Y , puede encontrarse variables explicativas. Comparada con la restricción de Durbin-Watson en la que no puede haber valores rezagados de la regresión entre los regresores.
- Una desventaja de la prueba Breuch-Godfrey es que el valor de p , la longitud de desfase, no se puede especificar a priori.

Inciso d)

La prueba más útil para evaluar las propiedades de residuales a partir de modelos de pronóstico es la prueba de Breusch-Godfrey ya que es estadísticamente más poderoso ya que analiza todas las autocorrelaciones hasta el retraso h y la prueba de Durbin-Watson solo analiza la autocorrelación en el retraso 1.

2. PROBLEMA 2

Demuestre paso a paso que:

$$\gamma(\tau) = E(\gamma_t \gamma_{t-\tau}) = E((\varepsilon_t + \theta \varepsilon_{t-1})(\varepsilon_{t-\tau} + \theta \varepsilon_{t-\tau-1}))$$

donde $\theta \sigma^2$, $\tau = 1$ y 0 en otro caso.

Completando los pasos que faltan evaluando en forma explícita la expectativa $E((\varepsilon_t + \theta \varepsilon_{t-1})(\varepsilon_{t-\tau} + \theta \varepsilon_{t-\tau-1}))$.

Solución

Comenzamos con la siguientes expresiones:

$$\begin{aligned} y_t &= \varepsilon_t + \theta \varepsilon_{t-1} \\ y_{t-\tau} &= \varepsilon_{t-\tau} + \theta \varepsilon_{t-\tau-1} \end{aligned}$$

las anteriores expresiones fueron proporcionadas en clase.

$$E(\gamma_t \gamma_{t-\tau}) = E(\varepsilon_t \varepsilon_{t-\tau} + \theta \varepsilon_t \varepsilon_{t-\tau-1} + \theta \varepsilon_{t-1} \varepsilon_{t-\tau-1} + \theta^2 \varepsilon_{t-1} \varepsilon_{t-\tau-1})$$

Ya que hemos desarrollado partimos de $\tau = 1$, así:

$$\begin{aligned} E(\varepsilon_t \varepsilon_{t-\tau} + \theta \varepsilon_t \varepsilon_{t-\tau-1} + \theta \varepsilon_{t-1} \varepsilon_{t-\tau-1} + \theta^2 \varepsilon_{t-1} \varepsilon_{t-\tau-1}) = \\ E(\varepsilon_t \varepsilon_{t-1}) + E(\theta \varepsilon_t \varepsilon_{t-2}) + E(\theta \varepsilon_{t-1} \varepsilon_{t-2}) + \theta \varepsilon_{t-1} \varepsilon_{t-1} + E(\theta^2 \varepsilon_{t-1} \varepsilon_{t-2}) \end{aligned}$$

Los valores esperados con epsilon de diferentes tiempo se vuelven igual a cero y los que coinciden tienen una varianza de error σ^2 . Así:

$$0 + 0 + 0 + \theta \sigma^2 + 0 = \theta \sigma^2$$

De esta forma hemos observado que pasa cuando $\tau = 1$, nos resta ver que sucede con $\tau > 1$.

Pero al tener $\tau > 1$ se puede ver que ningún ε coincidirá por lo que la tendremos que todo es igual a cero.

El último caso es cuando $\tau = 0$ vamos a tener que la varianza y el resultado es $\sigma^2 + \theta \sigma^2$.

3. PROBLEMA 3

Modelos de agregación y de desagregación pronóstico de arriba abajo y de abajo arriba. El asunto de la agregación se relaciona con el de los métodos y la complejidad. Con frecuencia se desea pronosticar un agregado, como por ejemplo las ventas totales de una empresa manufacturera, pero podemos emplear un método agregado o desagregado. Supongamos que las ventas totales están formadas por las de 35 productos cuya información se encuentra en Tarea02_Datos.txt. El método agregado, o de arriba abajo o

macro es simplemente modelar y pronosticar las ventas totales. El método desagregado, o de abajo arriba, o micro, es modelar y pronosticar por separado las ventas de los productos individuales, para después sumarlas. Quizá sea sorprendente, pero es imposible saber cual de los métodos es mejor, el agregado o desagregado. Todo depende de las circunstancias del caso la única forma de saberlo es probar ambos métodos y comparar los resultados del pronóstico. Argumente ventajas y desventajas de utilizar un método agregado o desagregado en función de la información proporcionada en el archivo Tarea02_Datos.txt.

Solución

```
library('readr')
library(forecast)

Jordan <- function(data, articulo, m1){
  if(articulo == 'Todos'){
    data_1 <- data.frame(data)
    data_orde <- data[order(data$Semana),]
    dat1 <- data_orde
    nd <- nrow(dat1)
    agr <- aggregate(dat1[,3:6],list(substring(dat1[,1],1,8)),sum)
    na <- nrow(agr)
    dat <- agr
    Fec <- dat[, "Group.1"]
    if(m1 == 1){
      dat <- dat[1:122,]
    }
    num_art <- 35
  }
  else{
    datos_ordenados <- data[order(data$Semana),] ## ordenamos datos
    data_articulo <- datos_ordenados[datos_ordenados$Articulo == articulo,]
    dat <- data_articulo
  }
}
```

```

#dat <- data_articulo[1:122,] #####
nd <- nrow(dat)
agr <- dat
na <- nrow(dat)
num_art <- 1
Fec <- dat[, "Semana"]
}
##### paso 1 #####
Y <- dat[, "Venta_Pesos"]
nY <- length(Y)

## Inter
X <- rep(1, nY)

RR <- lm(Y ~ -1 + X)
RES <- RR$res

X <- cbind(Inter = rep(1, nY), Tend = 1:nY)

RR <- lm(Y ~ -1 + X)
RES <- RR$res

##### paso 2 #####

Mes <- as.numeric(substring(Fec, 5, 6))
XMes <- matrix(0, nY, 12)
for(i in 1:nY){
  XMes[i, Mes[i]] <- 1
}

XMes <- XMes[, -1]

X <- cbind(Inter = rep(1, nY), Tend = 1:nY, XMes)

```



```

RR <- lm(Y~-1+X)
RES <- RR$res

##### paso 3 #####
colnames(XMes) <- paste("Mes_",2:12,sep="")

X <- cbind(Inter=rep(1,nY),Tend=1:nY,XMes,
           Trafico=dat[, "Cant_Tickets"]/num_art)#### ojo
#X[1:4,]

RR <- lm(Y~-1+X)
RES <- RR$res

##### paso 4 #####
YMT <- Y/(dat[, "Num_Tiendas"]/num_art)

X <- cbind(Inter=rep(1,nY),Tend=1:nY,XMes,
           Trafico=dat[, "Cant_Tickets"]/num_art)
RR <- lm(YMT~-1+X)
RES <- RR$res

##### paso 5 #####
SS <- scale(RES)
OutPos <- ifelse(SS > 2,1,0)
OutNeg <- ifelse(SS < -2,1,0)
X <- cbind(Inter=rep(1,nY),Tend=1:nY,XMes,
           Trafico=dat[, "Cant_Tickets"]/num_art,
           OutPos,OutNeg)
RR <- lm(YMT~-1+X)
RES <- RR$res

##### paso 6 #####
Precio <- dat[, "Venta_Pesos"]/dat[, "Unidades"]

```

```

X <- cbind(Inter=rep(1,nY),Precio,
           Tend=1:nY,XMes,
           Trafico=dat[, "Cant_Tickets"]/num_art,
           OutPos,OutNeg)
RR <- lm(YMT~1+X)
RES <- RR$res

##### paso 7 #####
LYMT <- log(YMT)

X <- cbind(Inter=rep(1,nY),Precio=log(Precio),Tend=log(1:nY),
           XMes,Trafico=log(dat[, "Cant_Tickets"]/num_art),
           OutPos,OutNeg)
RR <- lm(LYMT~1+X)
RES <- RR$res
#plot.ts(RES,ylab="Residuales",
#         main="Modelo: Log (Inter + Tend + Traf + Out + Precio)",
#         xlab="",xaxt="n")
#axis(1,1:na,agr[,1],las=2,cex.axis=0.5)
lista = c(LYMT, X)

return( list(X = X,LYMT = LYMT) )
}

data_prueba1 <- read.delim("C:/Users/Marcelo Sanchez/Downloads/Tarea02_Datos.txt",se
a <- 'Todos'
l <- c('Art_01', 'Art_02', 'Art_03', 'Art_04', 'Art_05', 'Art_06',
      'Art_07', 'Art_06', 'Art_07', 'Art_08', 'Art_09', 'Art_10',
      'Art_11', 'Art_12', 'Art_13', 'Art_14', 'Art_15', 'Art_16',
      'Art_17', 'Art_18', 'Art_19', 'Art_20', 'Art_21', 'Art_21',
      'Art_22', 'Art_22', 'Art_23', 'Art_24', 'Art_25', 'Art_26',
      'Art_27', 'Art_28', 'Art_29', 'Art_30', 'Art_31', 'Art_32',
      'Art_33', 'Art_34', 'Art_35')

```

```
#####
##### tomando cada uno de los articulos #####
#####

matriz_datos <- matrix(, nrow = 11, ncol = 35)
n <- 122
for (i in 1:35) {
  cat(" ")
  cat(" ")
  cat("Articulo",i)
  prueba1 <- Jordan(data_prueba1, l[i])
  LYMT <- data.frame( prueba1[2] )
  X <- data.frame( prueba1[1] )
  model = auto.arima(LYMT[1:n,], max.p = 12,
                    max.q = 12, d = 0, stepwise = FALSE,
                    xreg = cbind(X$X.Precio[1:n],
                                X$X.Tend[1:n],X$X.Mes_2[1:n],
                                X$X.Mes_3[1:n],X$X.Mes_4[1:n],
                                X$X.Mes_5[1:n],X$X.Mes_6[1:n],
                                X$X.Mes_7[1:n],X$X.Mes_8[1:n],
                                X$X.Mes_9[1:n],X$X.Mes_10[1:n],
                                X$X.Mes_11[1:n],X$X.Mes_12[1:n],
                                X$X.Trafico[1:n]) )

  prediccion <- forecast(model, xreg = cbind(X$X.Precio[n:132],
                                             X$X.Tend[n:132], X$X.Mes_2[n:132],
                                             X$X.Mes_3[n:132],X$X.Mes_4[n:132],
                                             X$X.Mes_5[n:132], X$X.Mes_6[n:132],
                                             X$X.Mes_7[n:132], X$X.Mes_8[n:132],
                                             X$X.Mes_9[n:132],X$X.Mes_10[n:132],
                                             X$X.Mes_11[n:132], X$X.Mes_12[n:132],
                                             X$X.Trafico[n:132]) ,h=10)
```

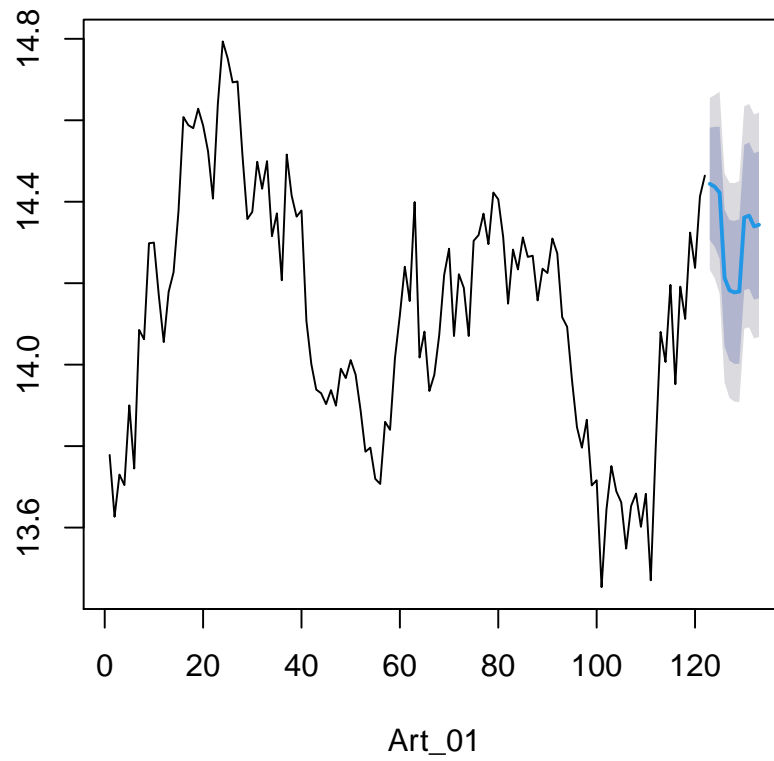
```

##### agregamos un grafico para observar la predicci
m4 <- prediccion
# verificando el ajuste del m
plot(m4, main="Grafica de un articulo",
      xlab=l[i] )
autoplot(m4) + autolayer(fitted(m4), series="Ajuste")
#####
s_1 <- data.frame(prediccion)
l22 <- c(s_1$Point.Forecast)
matriz_datos[,i] <- l22
}

Articulo 1

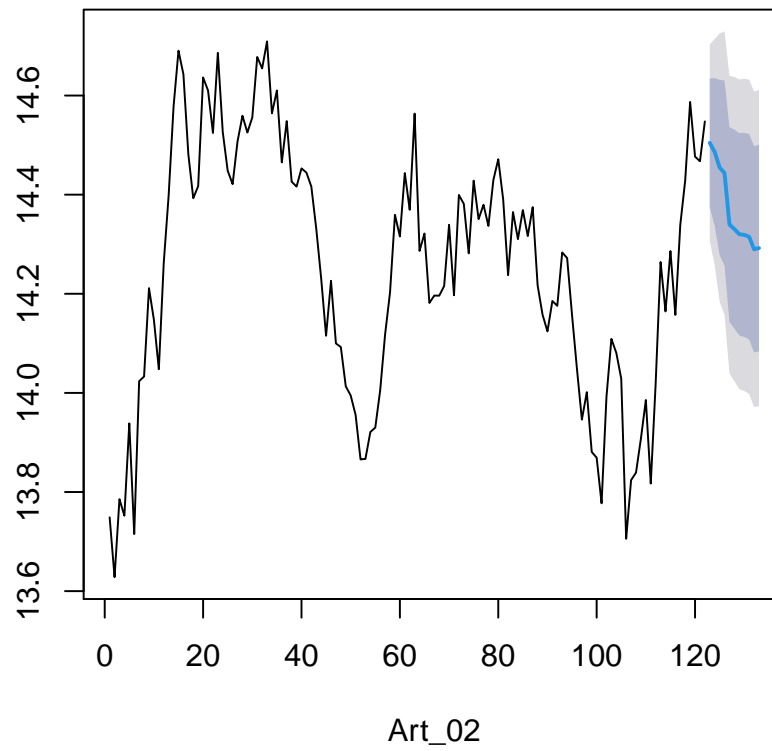
```

Grafica de un articulo



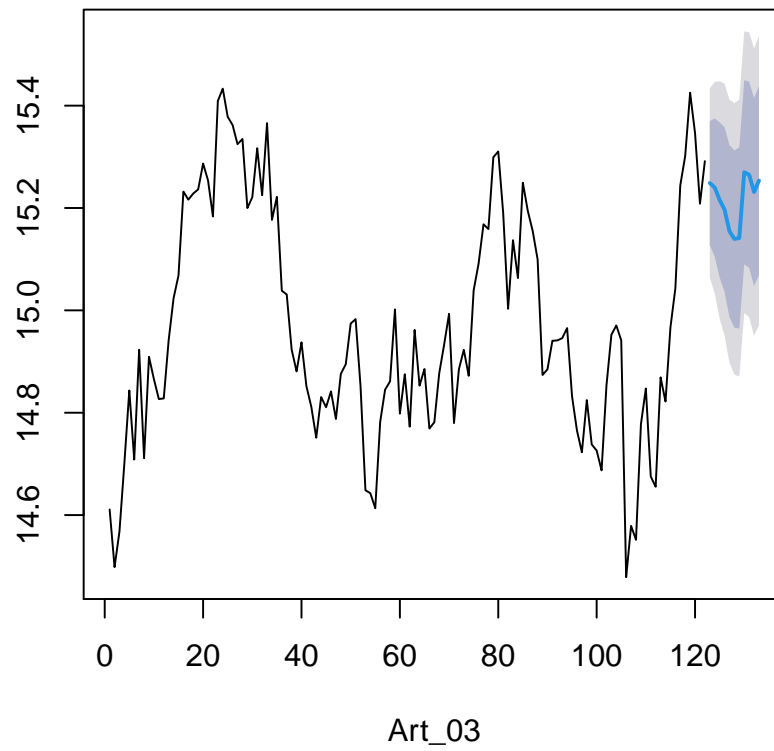
Articulo 2

Grafica de un articulo



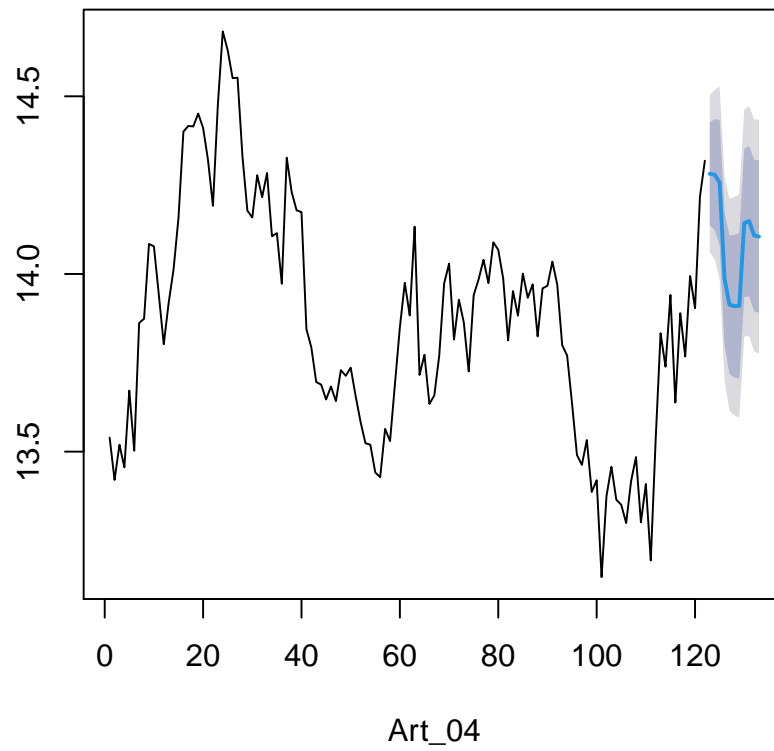
Articulo 3

Grafica de un articulo



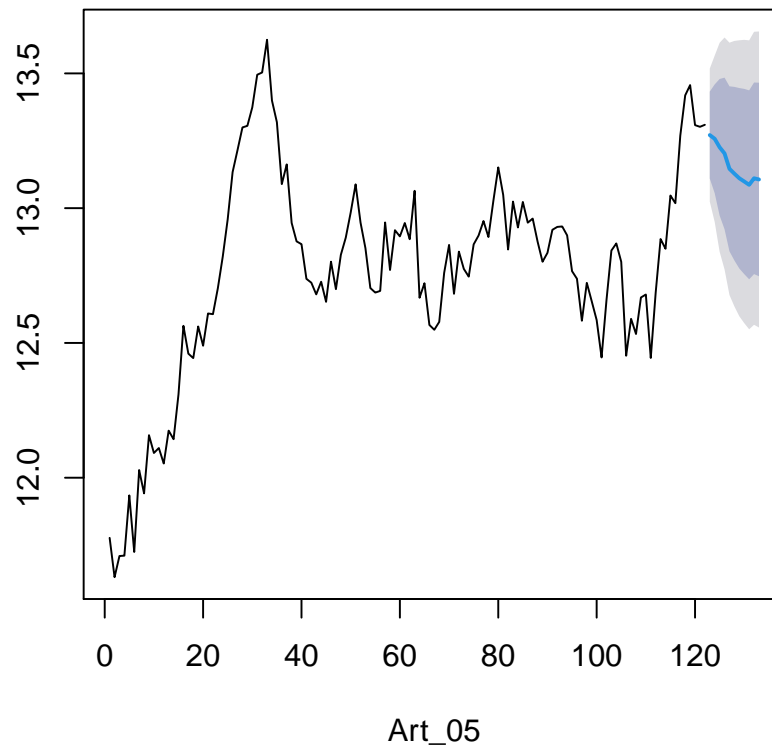
Articulo 4

Grafica de un articulo



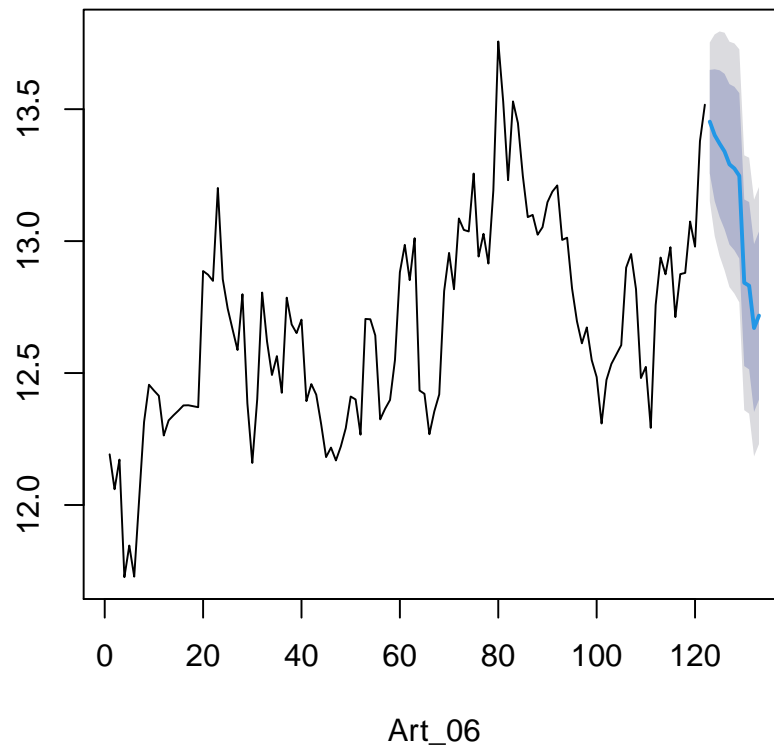
Articulo 5

Grafica de un articulo



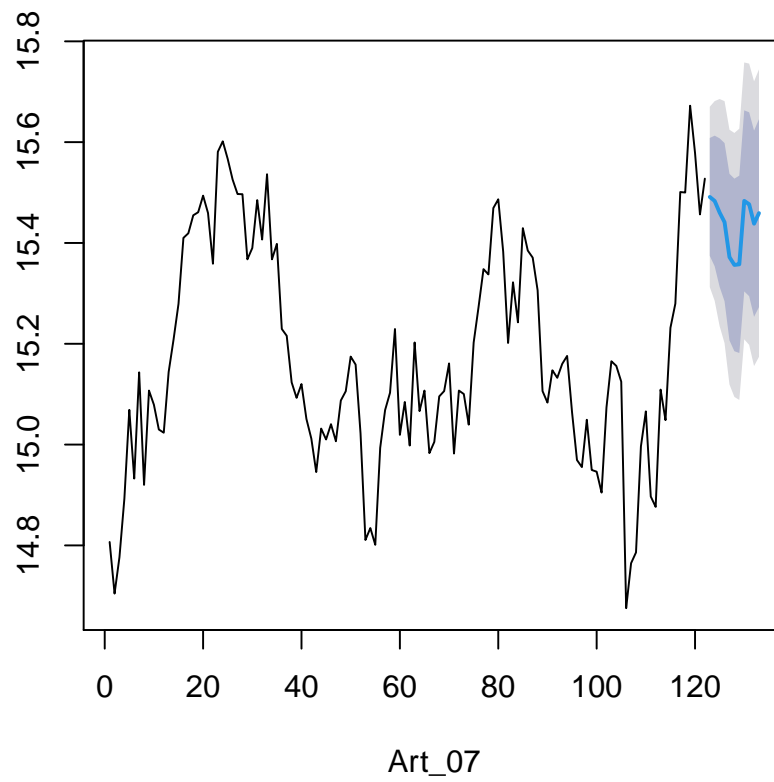
Articulo 6

Grafica de un articulo



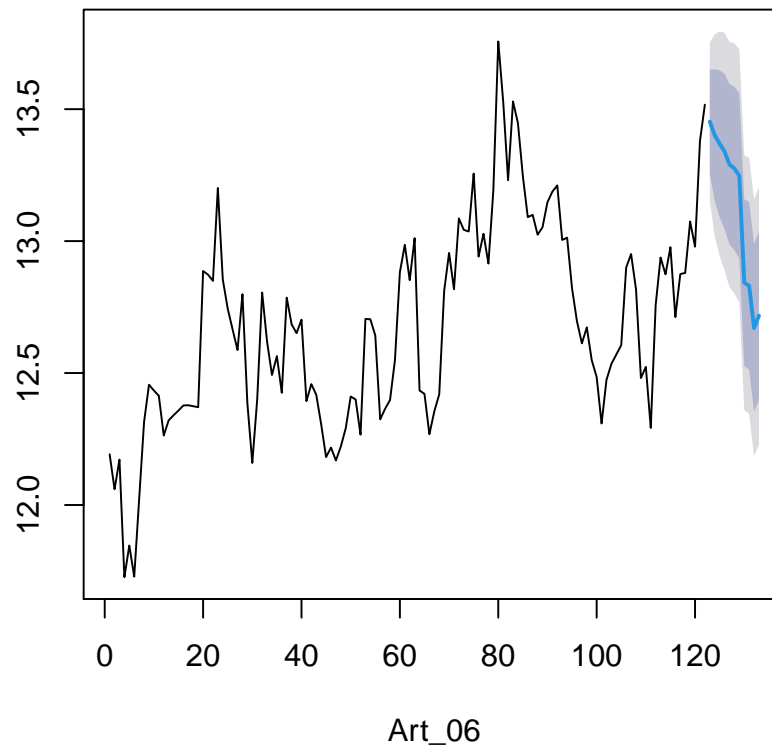
Articulo 7

Grafica de un articulo



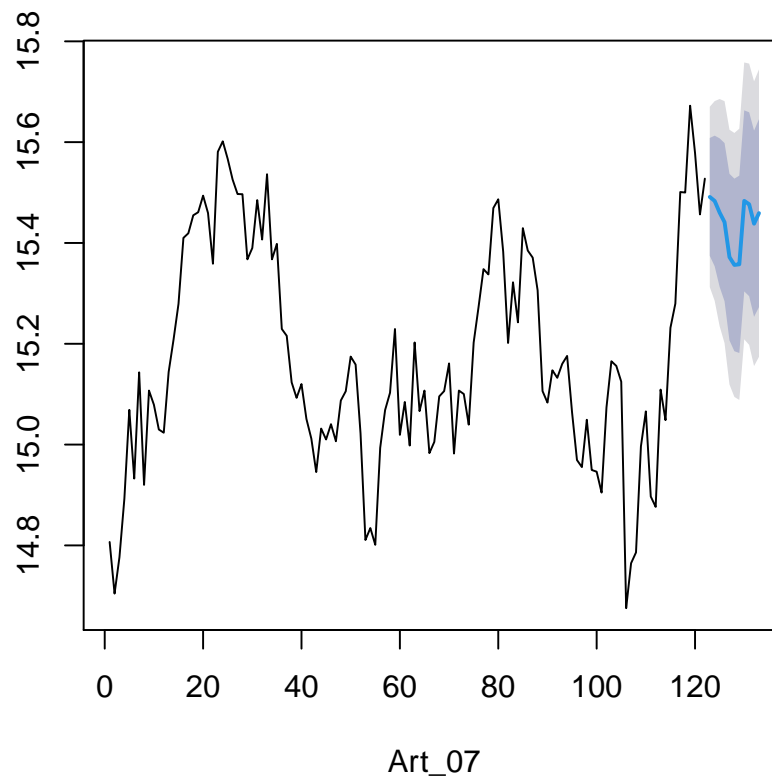
Articulo 8

Grafica de un articulo



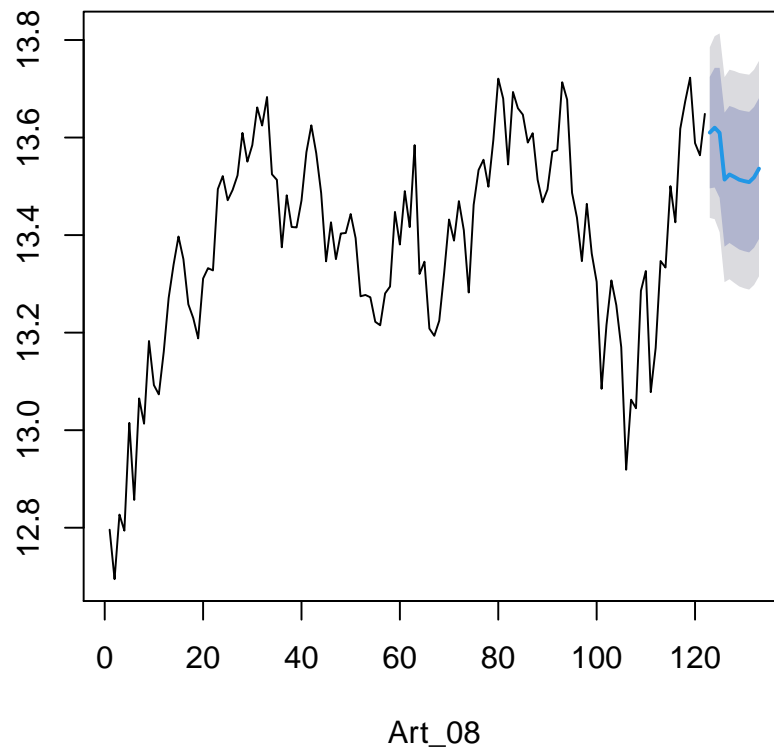
Articulo 9

Grafica de un articulo



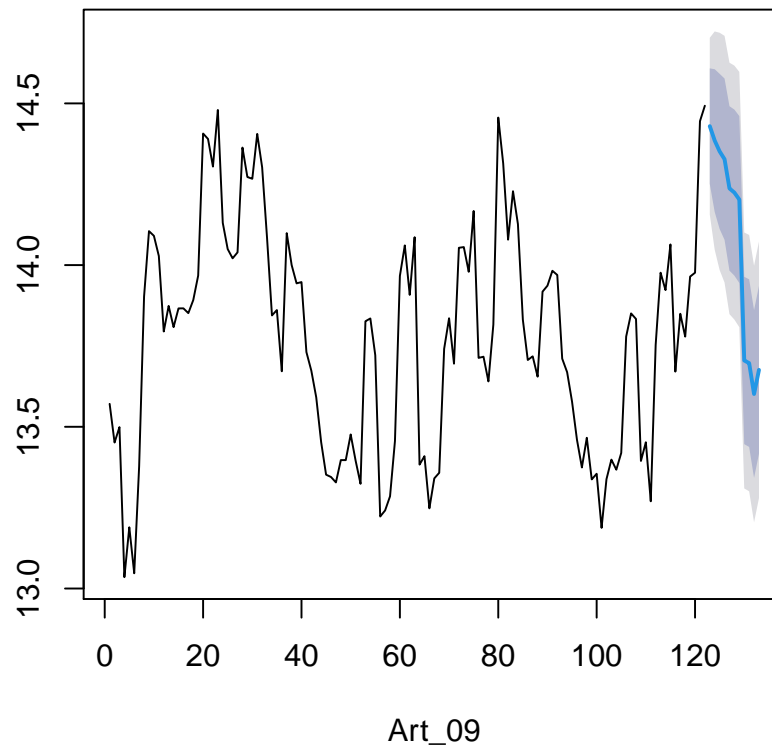
Articulo 10

Grafica de un artículo



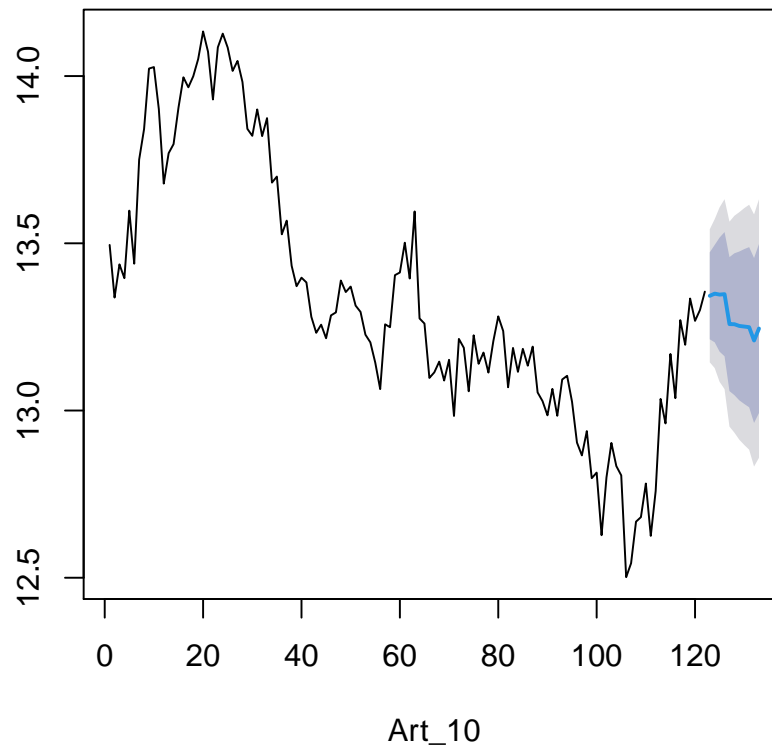
Artículo 11

Grafica de un articulo



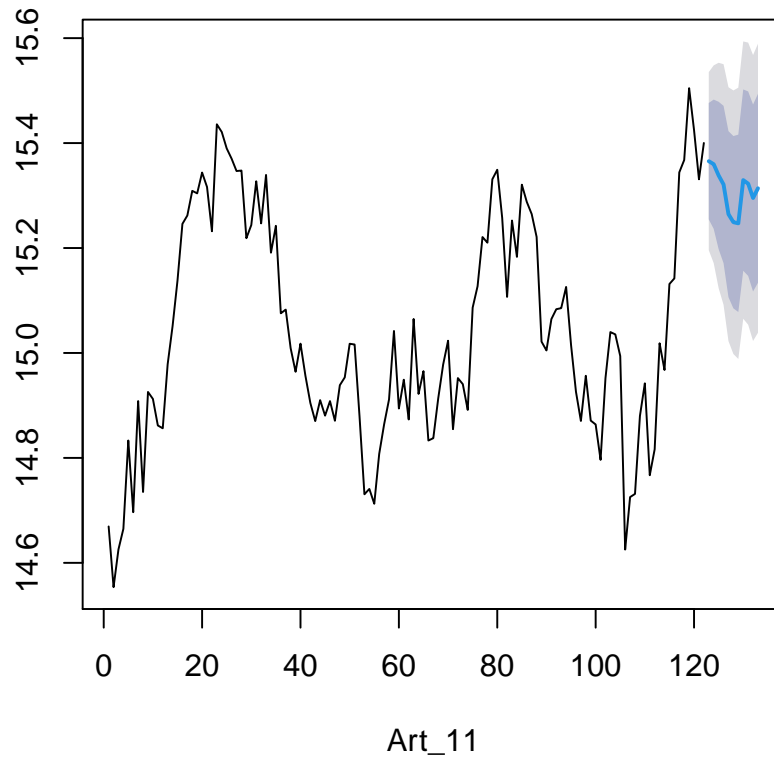
Articulo 12

Grafica de un articulo



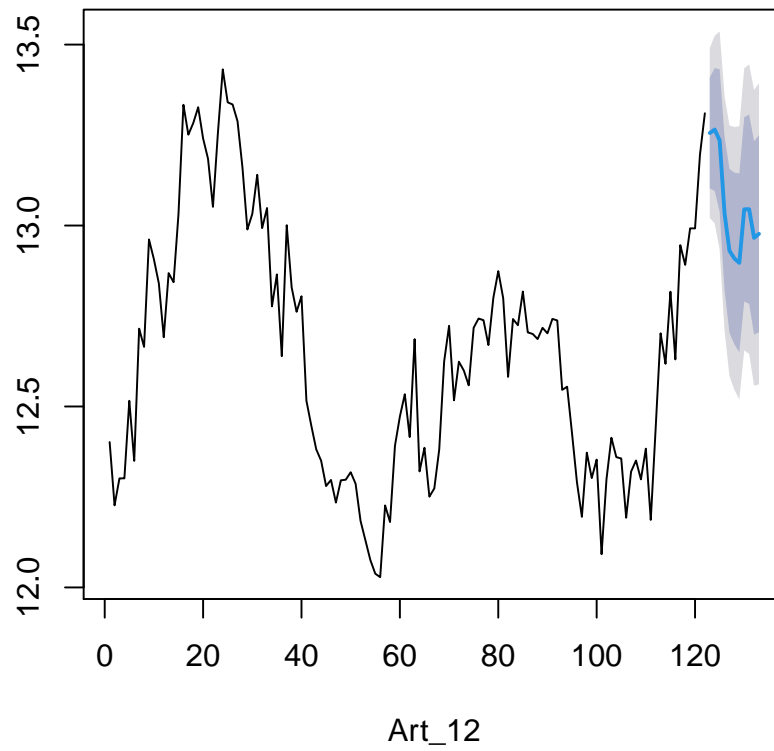
Articulo 13

Grafica de un articulo



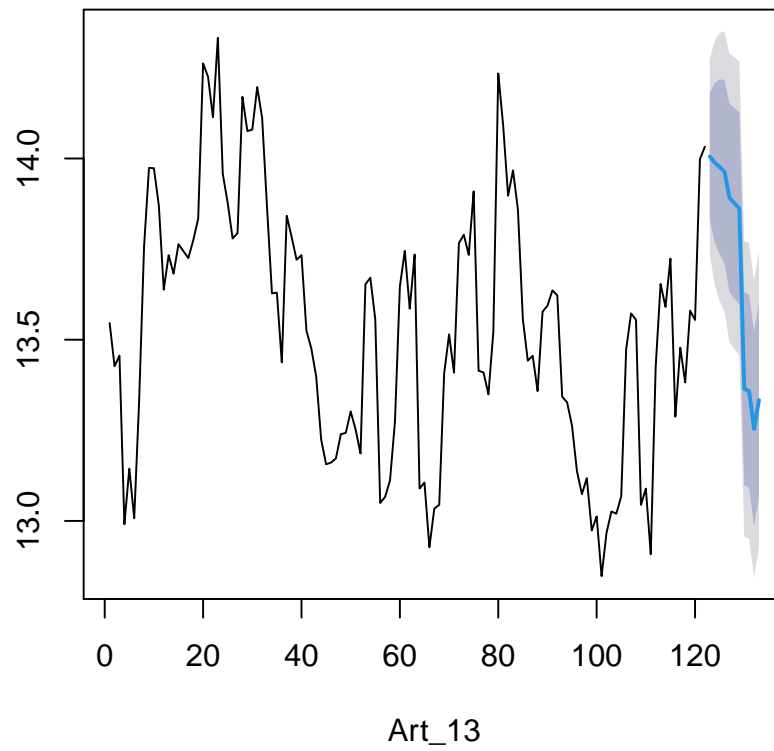
Articulo 14

Grafica de un articulo



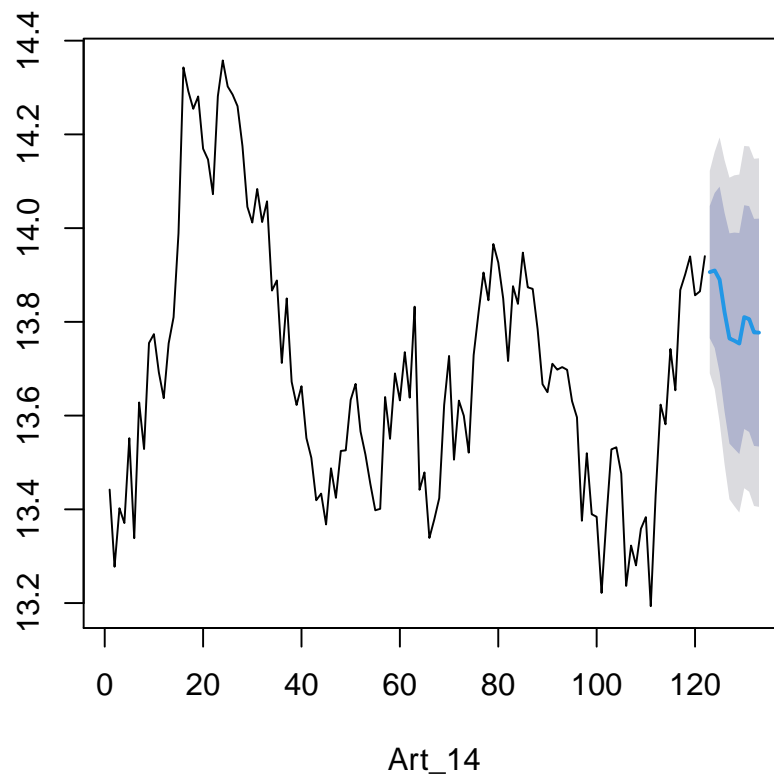
Articulo 15

Grafica de un articulo



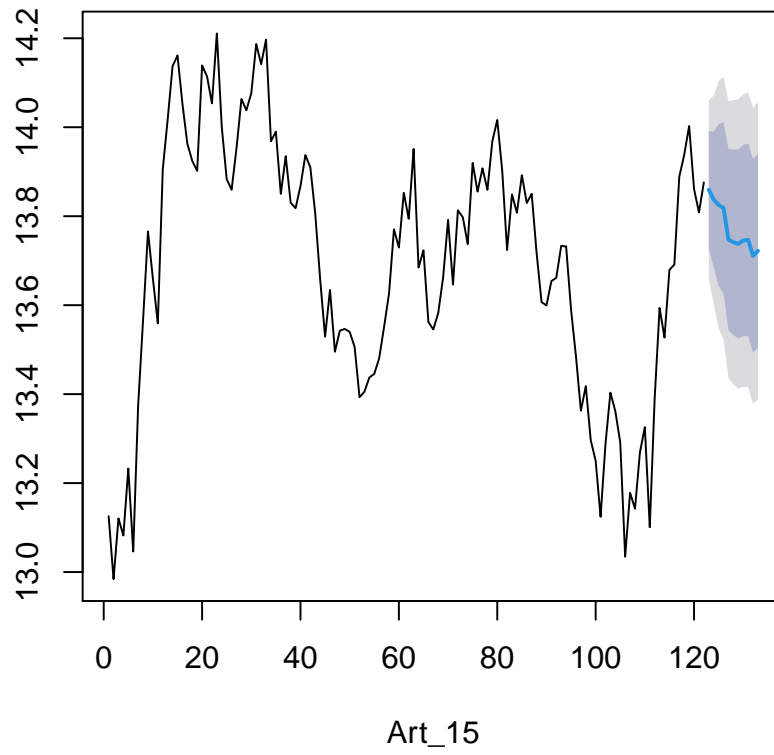
Articulo 16

Grafica de un articulo



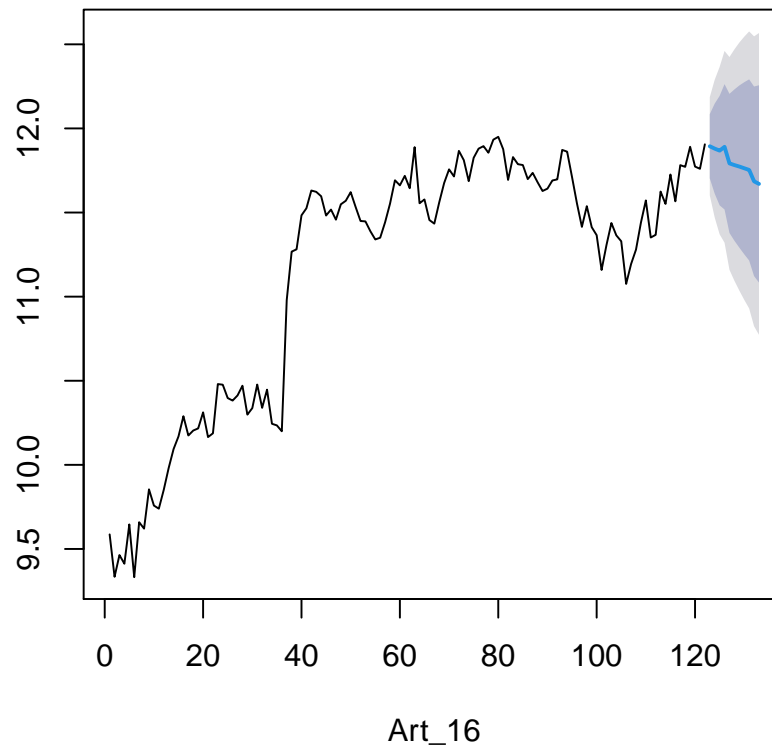
Articulo 17

Grafica de un articulo



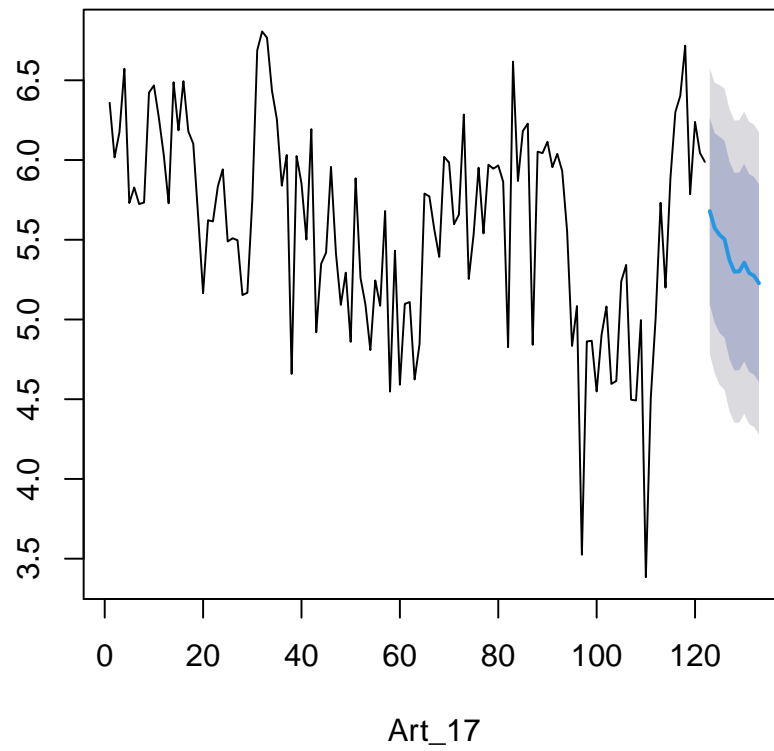
Articulo 18

Grafica de un artículo



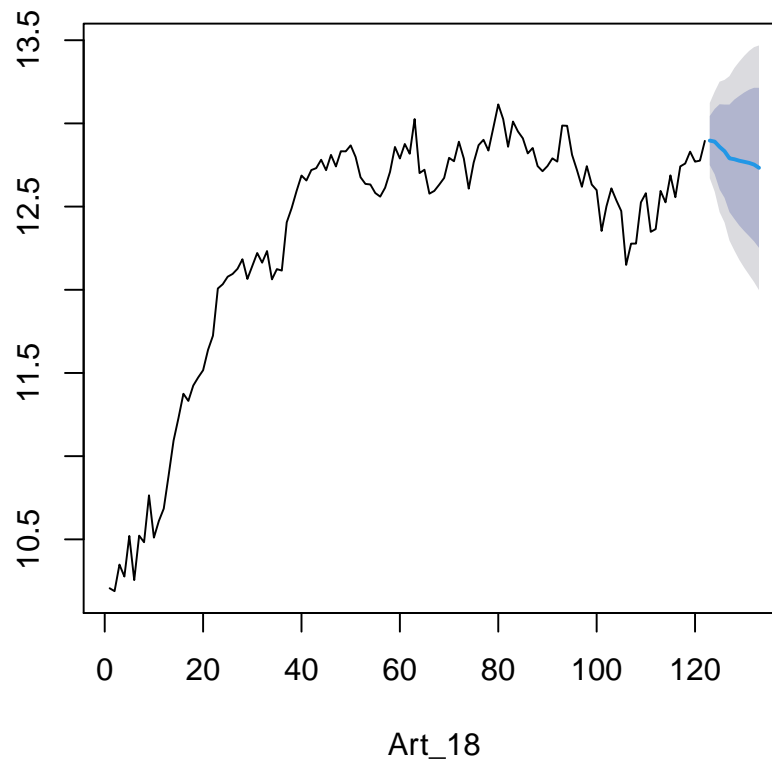
Artículo 19

Grafica de un articulo



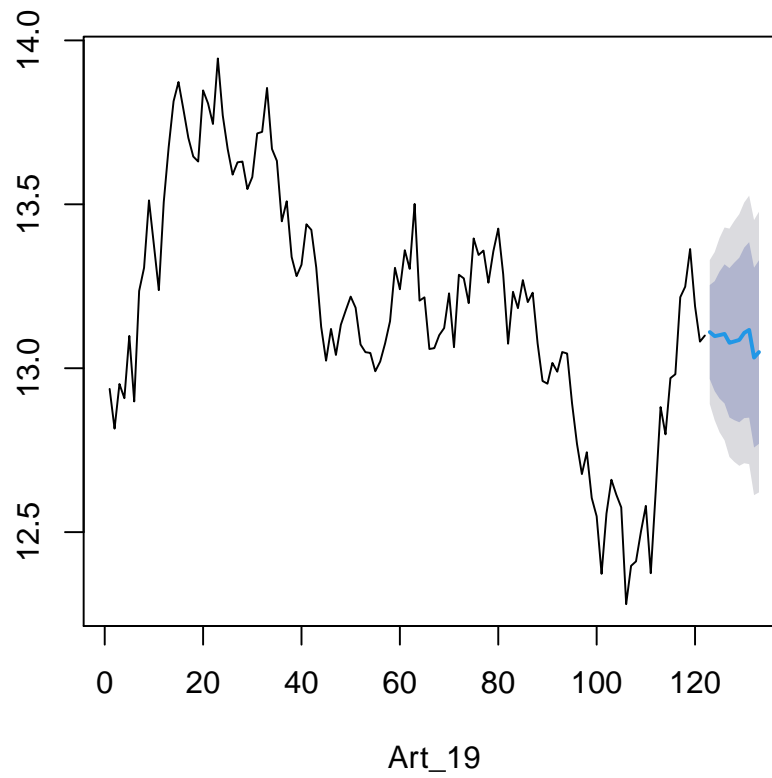
Articulo 20

Grafica de un articulo



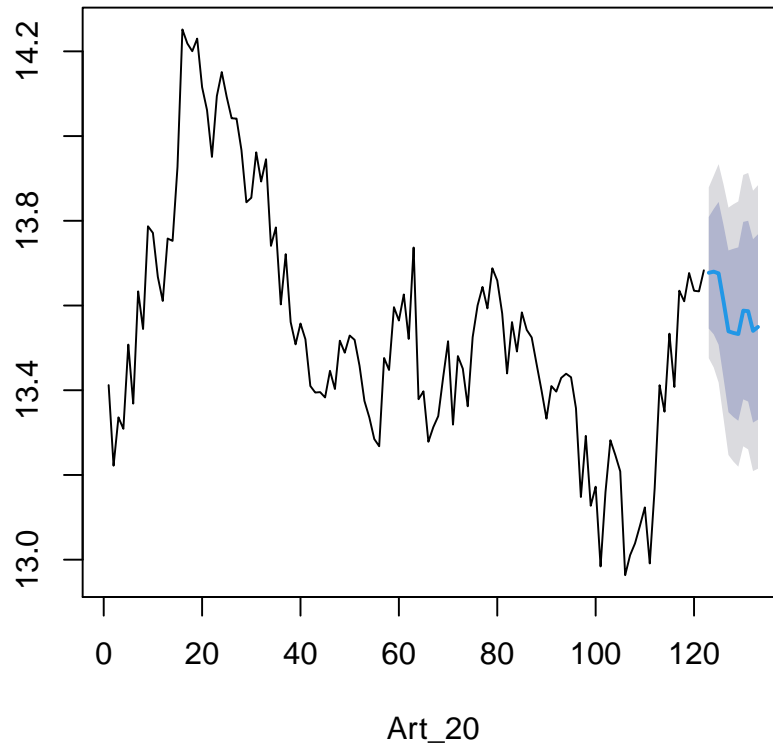
Articulo 21

Grafica de un articulo



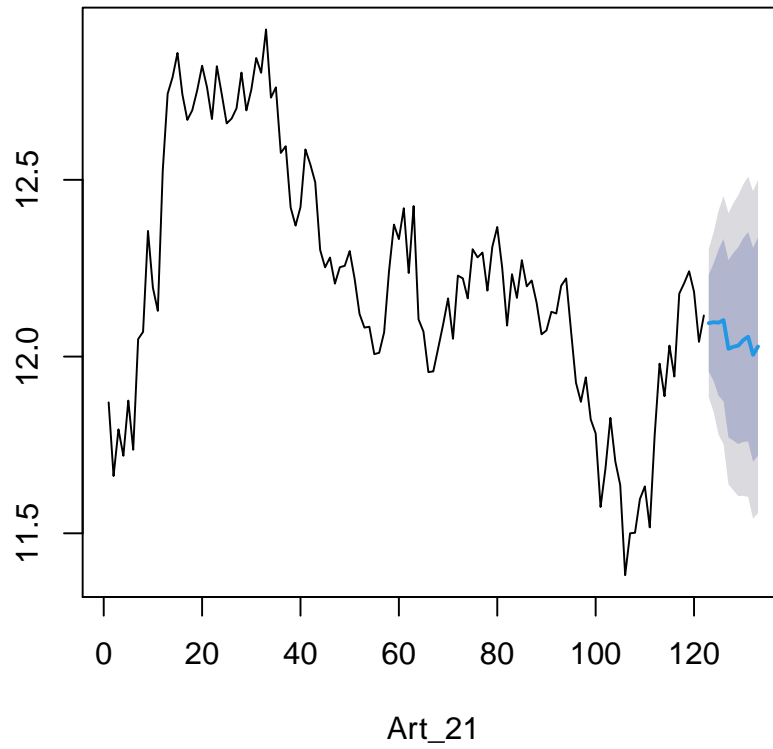
Articulo 22

Grafica de un articulo



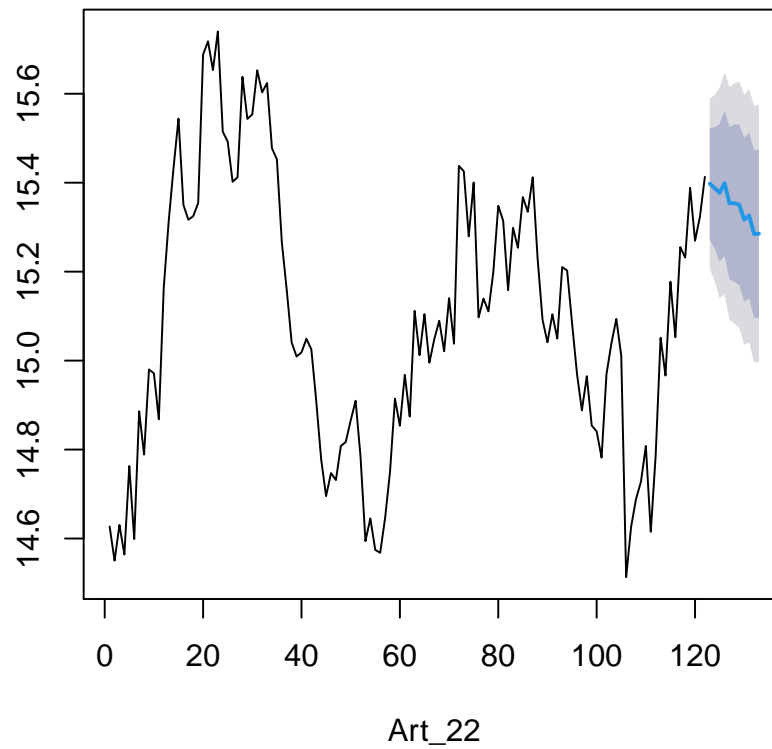
Articulo 23
Articulo 24

Grafica de un articulo



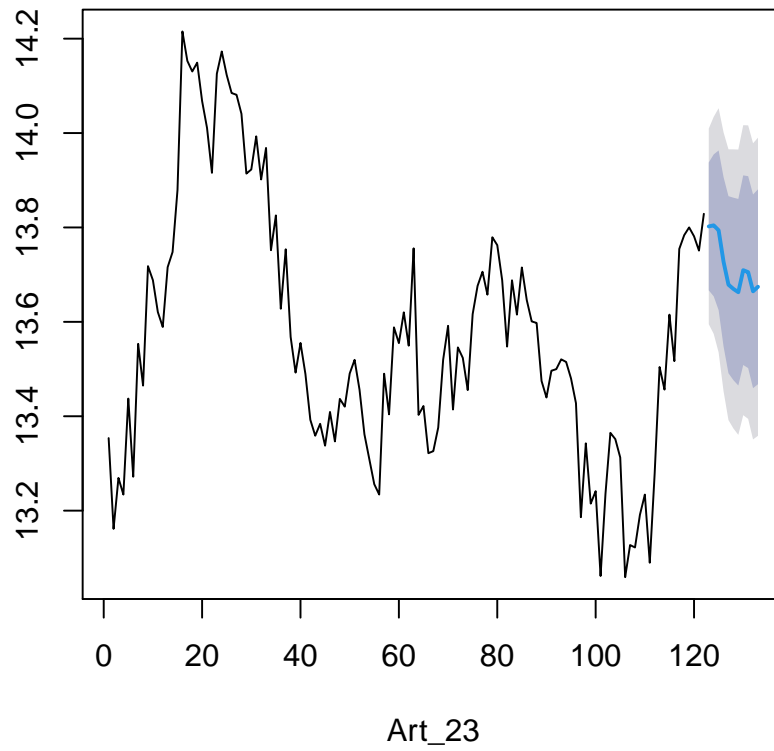
Articulo 25
Articulo 26

Grafica de un articulo



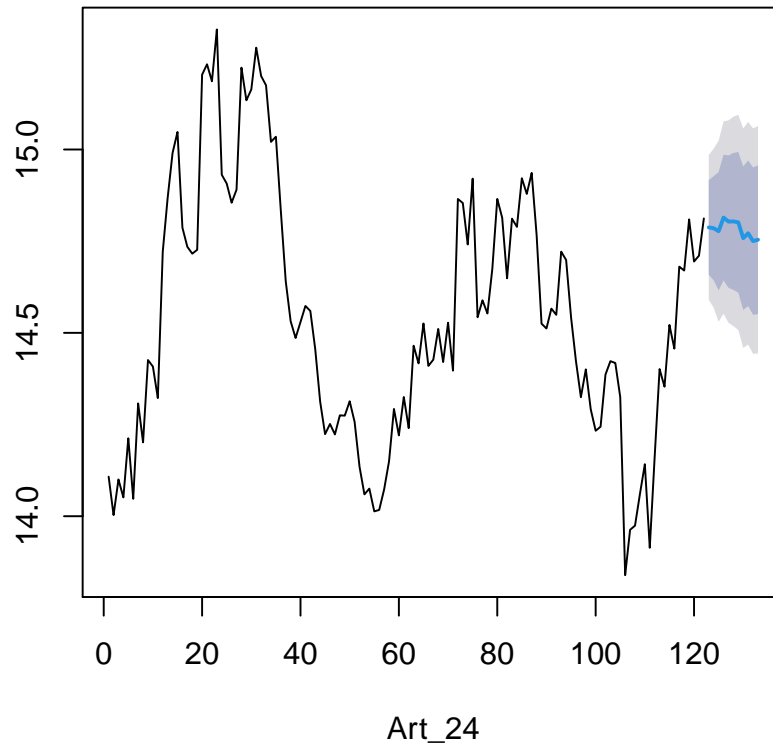
Articulo 27

Grafica de un articulo



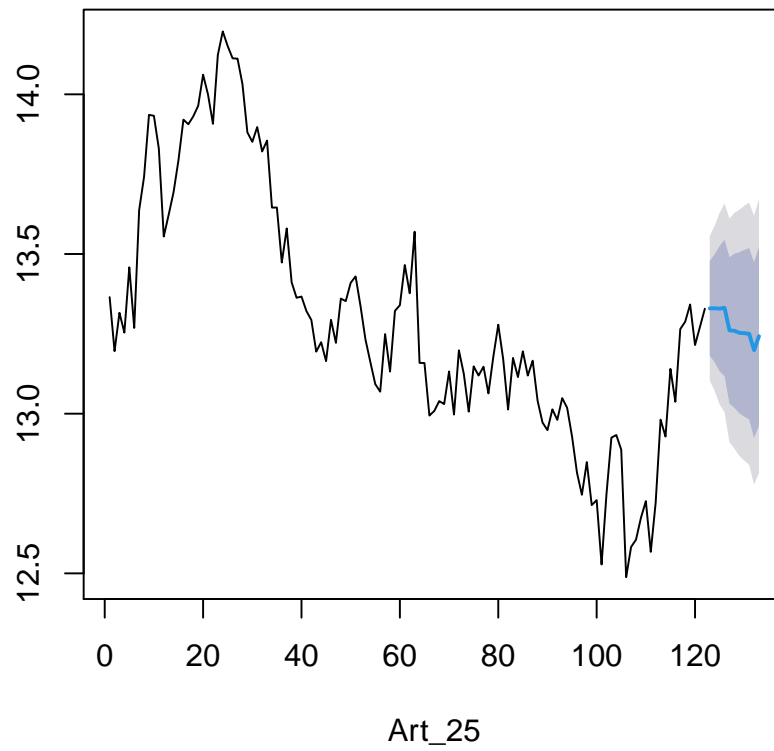
Articulo 28

Grafica de un articulo



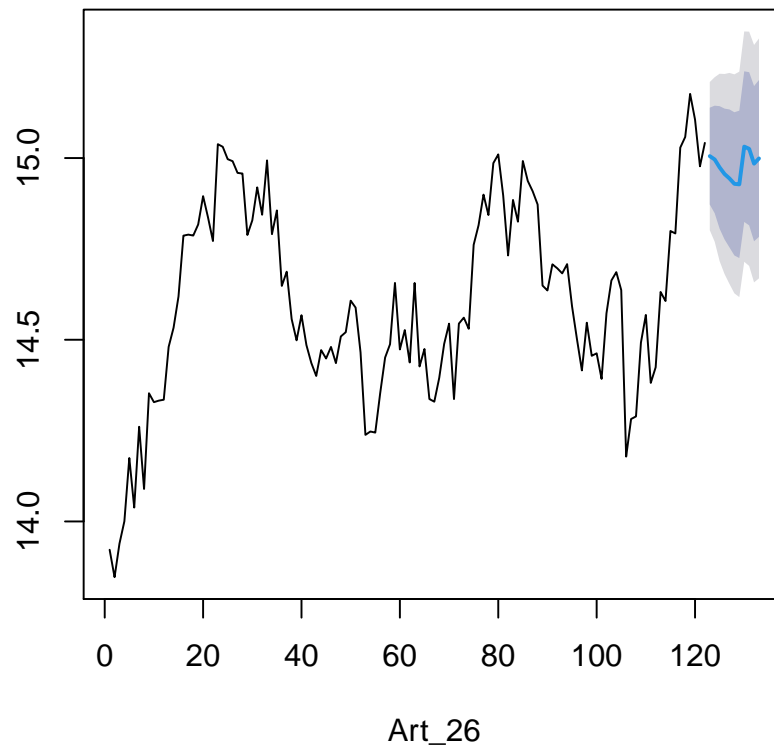
Articulo 29

Grafica de un articulo



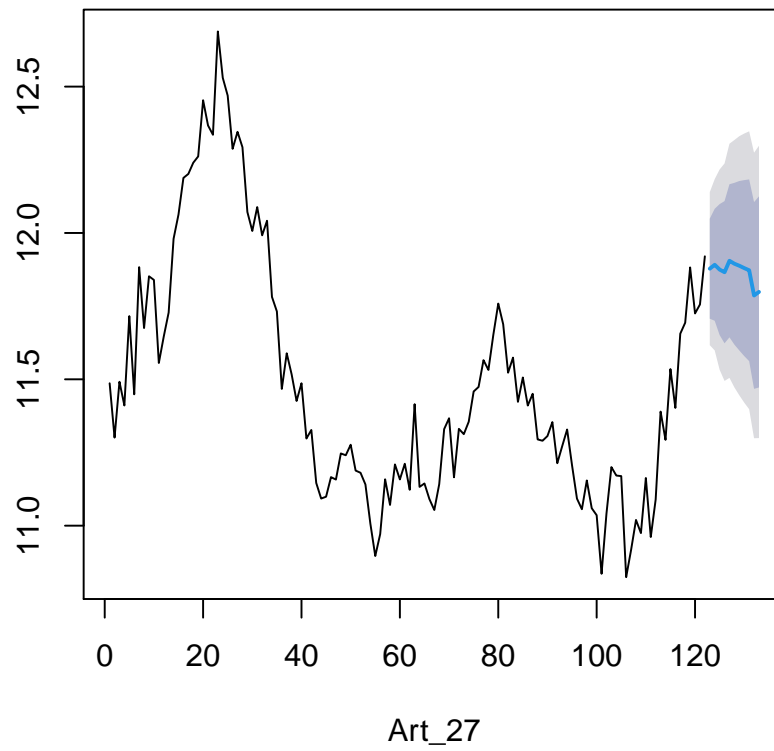
Articulo 30

Grafica de un articulo



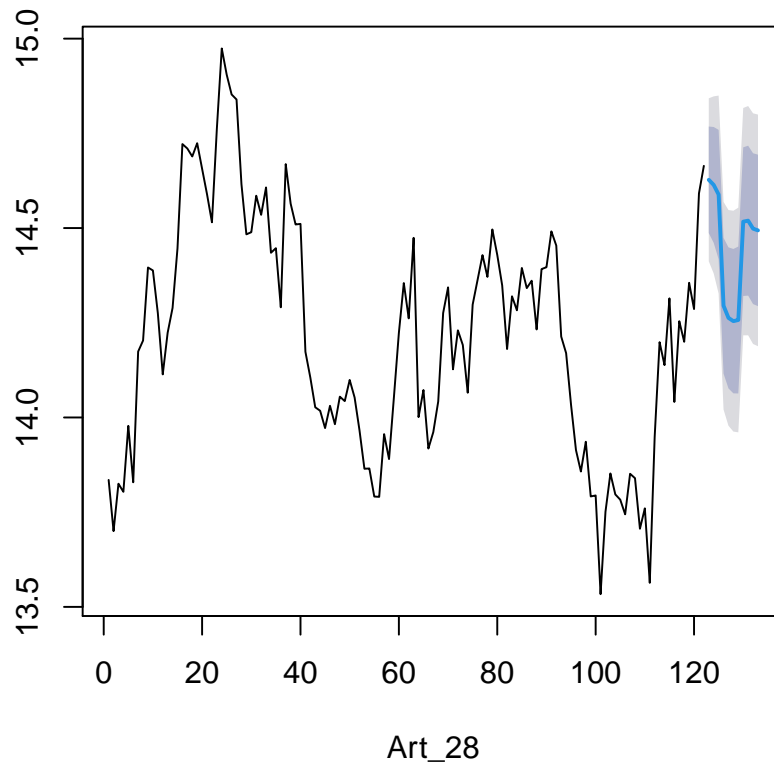
Articulo 31

Grafica de un articulo



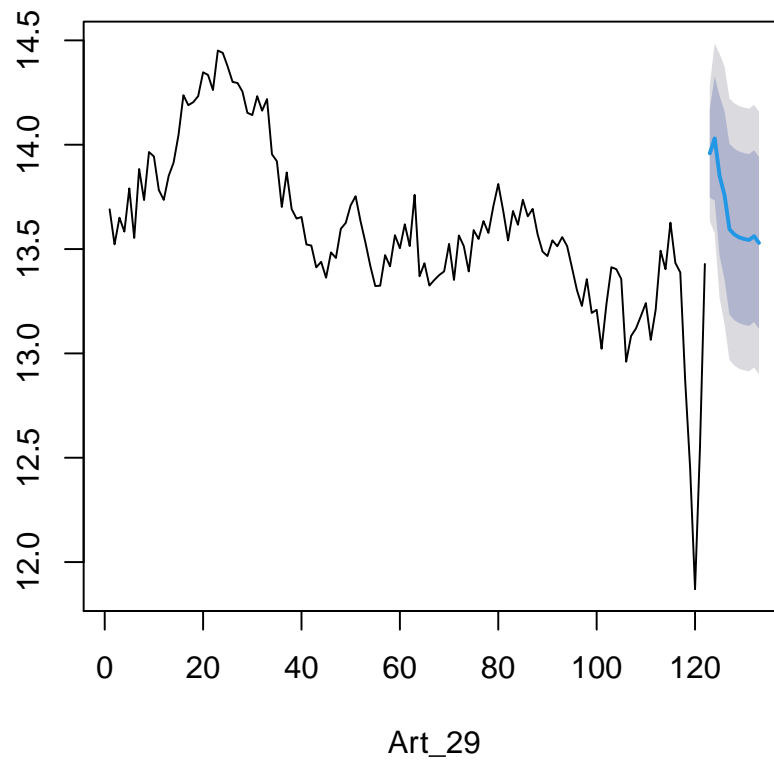
Articulo 32

Grafica de un articulo



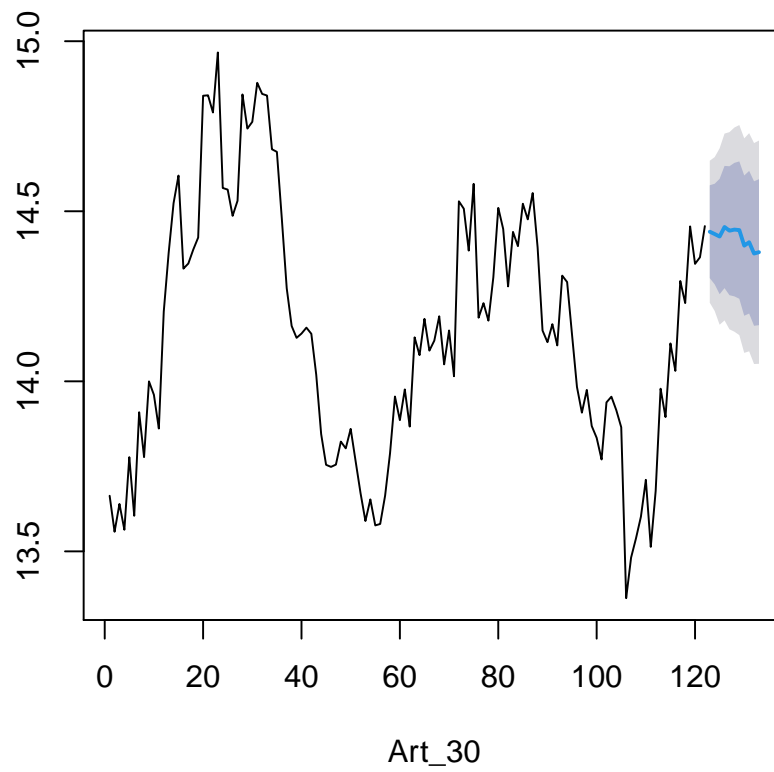
Articulo 33

Grafica de un articulo



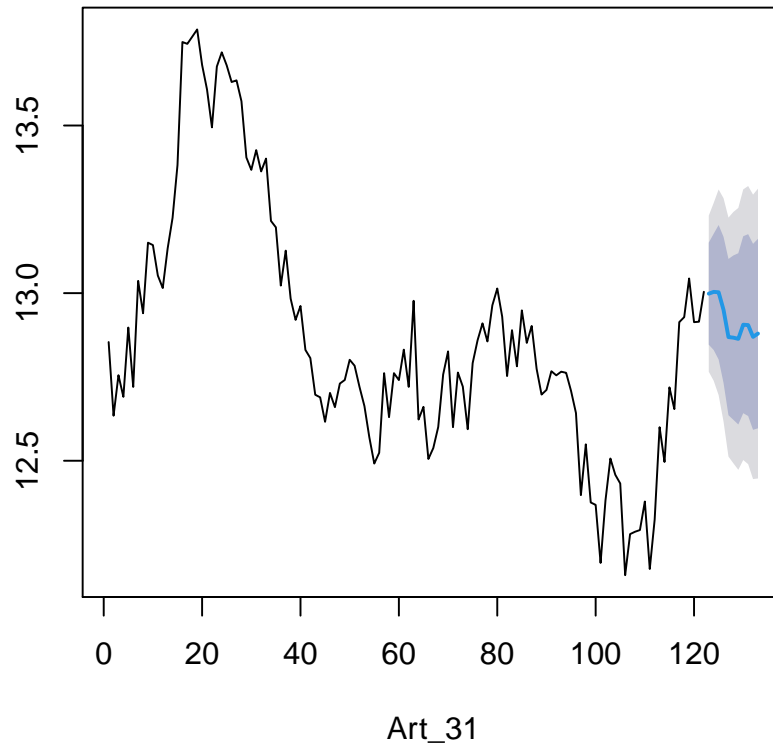
Articulo 34

Grafica de un articulo



Articulo 35

Grafica de un artículo



```
### Realizamos unos ajustes a los valores para encontrar  
## los valores verdaderos.  
matriz_exp <- exp(matriz_datos)  
suma_t <- rowSums(matriz_exp)  
log(suma_t)
```

```
[1] 17.90354 17.89554 17.87412 17.84024 17.78731 17.77853 17.77506 17.80976
```

```

[9] 17.80924 17.77167 17.78448

#####
##### con todos los datos #####
#####

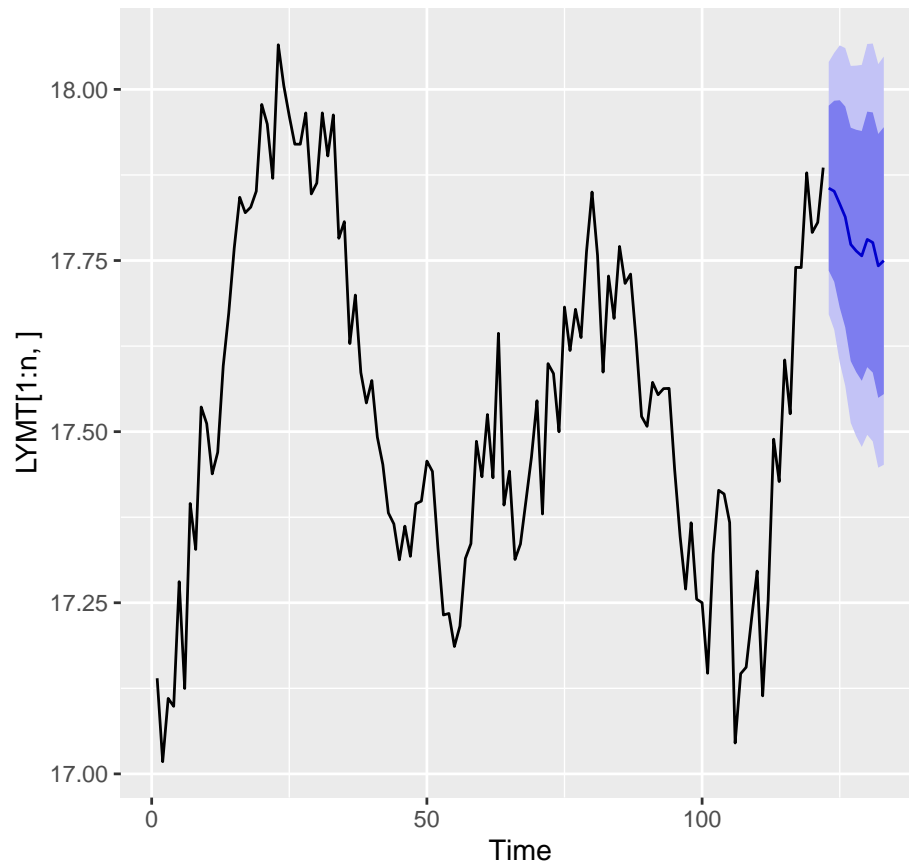
prueba1 <- Jordan(data_prueba1, a, 0)
LYMT <- data.frame( prueba1[2] )
X <- data.frame( prueba1[1] )
n <- 122
##### realizamos el pronostico #####
model = auto.arima(LYMT[1:n,], max.p = 12,
                  max.q = 12, d = 0, stepwise = FALSE,
                  xreg = cbind(X$X.Precio[1:n],
                              X$X.Tend[1:n], X$X.Mes_2[1:n],
                              X$X.Mes_3[1:n], X$X.Mes_4[1:n],
                              X$X.Mes_5[1:n], X$X.Mes_6[1:n],
                              X$X.Mes_7[1:n], X$X.Mes_8[1:n],
                              X$X.Mes_9[1:n], X$X.Mes_10[1:n],
                              X$X.Mes_11[1:n], X$X.Mes_12[1:n],
                              X$X.Trafico[1:n]) )

prediccion <- forecast(model, xreg = cbind(X$X.Precio[n:132],
                                           X$X.Tend[n:132], X$X.Mes_2[n:132],
                                           X$X.Mes_3[n:132], X$X.Mes_4[n:132],
                                           X$X.Mes_5[n:132], X$X.Mes_6[n:132],
                                           X$X.Mes_7[n:132], X$X.Mes_8[n:132],
                                           X$X.Mes_9[n:132], X$X.Mes_10[n:132],
                                           X$X.Mes_11[n:132], X$X.Mes_12[n:132],
                                           X$X.Trafico[n:132]) ,h=10)

#####
m4 <- prediccion
autoplot(m4)

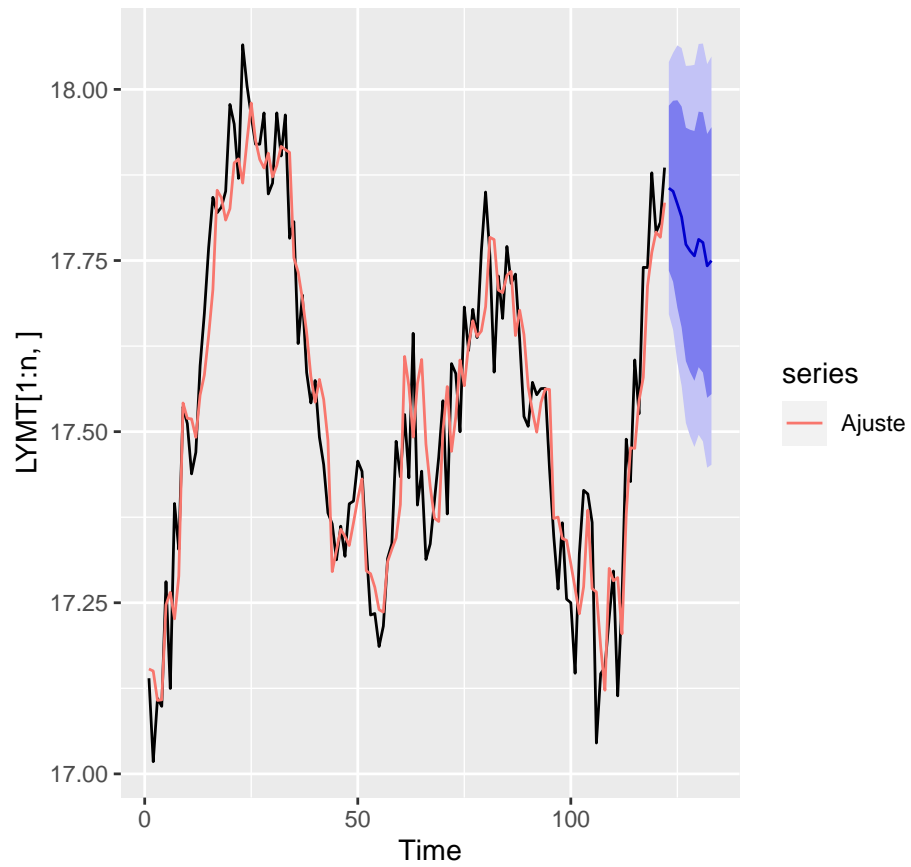
```

Forecasts from Regression with ARIMA(2,0,0) errors



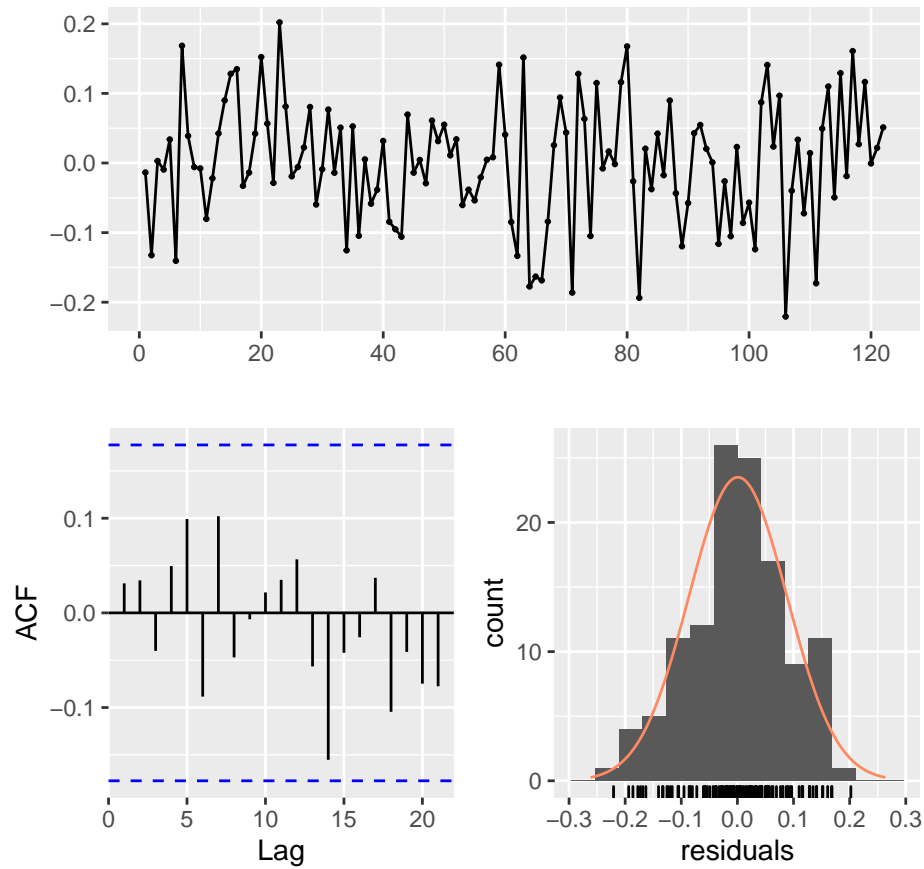
```
# verificando el ajuste del m<U+653C><U+3E39>todo  
autoplot(m4)+autolayer(fitted(m4), series="Ajuste")
```

Forecasts from Regression with ARIMA(2,0,0) errors



```
# verificando los residuales  
checkresiduals(m4)
```


Residuals from Regression with ARIMA(2,0,0) errors



Ljung-Box test

data: Residuals from Regression with ARIMA(2,0,0) errors
Q* = 12.455, df = 3, p-value = 0.005976

Model df: 17. Total lags used: 20

```

s_1 <- data.frame(prediccion)
l22 <- c(s_1$Point.Forecast)

#####
##### RESULTADOS #####
#####

##### sin exponencial
l22 #agregado

[1] 17.85577 17.85120 17.83272 17.81369 17.77356 17.76398 17.75678 17.78088
[9] 17.77633 17.74219 17.74981

LYMT[122:132,]

[1] 17.88577 17.86859 17.83217 17.87052 17.97583 17.96134 17.98131 18.00316
[9] 17.84440 17.90185 17.90907

log(suma_t) #desagregado

[1] 17.90354 17.89554 17.87412 17.84024 17.78731 17.77853 17.77506 17.80976
[9] 17.80924 17.77167 17.78448

##### ocupando exponencial
exp(l22)

[1] 56841091 56581883 55546056 54499029 52354921 51855809 51484072 52739764
[9] 52500160 50738218 51126446

exp(LYMT[122:132,])

[1] 58572011 57574670 55515146 57685762 64092055 63169864 64443917 65867498
[9] 56198634 59521711 59952669

```

```

exp(log(suma_t))

[1] 59622012 59147141 57893815 55965118 53079942 52615728 52433760 54285284
[9] 54256994 52256086 52929826

#####
##### MEJOR MODELO #####
#####
#####

ERROR_1 <- sum( (LYMT[122:132,] - 122)^2 )/10
ERROR_1 ## agregado

[1] 0.02396029

ERROR_2 <- sum( (LYMT[122:132,] - log(suma_t))^2 )/10
ERROR_2 ##desagregado

[1] 0.01863235

```

El procedimiento que se realizo fue el siguiente:

- 1.- Observar cuantos articulos contiene nuestra base de datos y ordenar por semanas los datos. Una vez realizada dicha tarea se realizo una suma por semana, observe que dicha suma solo realizo en el método agregado ya que como tenemos 35 articulos nos resultaron 35 datos con la misma semana, por lo que se opto por sumat todas estas semanas. Al final para ambos métodos obtuvimos un total de 132 semanas.
- 2.- Encontrar un modelo ARIMA adecuado para cada uno de los 35 articulos que contiene la base de datos y también encontrar un modelo tomando los 35 articulos, es decir, un modelo para el método agregado.

3.- Una vez encontrado dicho resultado realizamos una predicción de las últimas 10 semanas para ambos métodos. Al realizar dicha predicción seguimos las indicaciones que nos proporciona el problema, realizar la suma de los resultados que encontramos en el método desagregado. Cabe recalcar que para realizar dicha suma en el método desagregado hubo que aplicar logaritmo a cada uno de los 10 pronósticos y luego sumarlo para finalmente aplicarles el logaritmo, si no se hubiera hecho de esa forma nuestro resultado estaría mal ya que al sumarlas directamente estaríamos realizando un producto y esa no es la intención. Adicional a lo anteriormente comentado se realizaron algunas gráficas ilustrativas por medio de unos comandos para ambos métodos, las gráficas correspondientes al método desagregado tienen la etiqueta al artículo que corresponde y la gráfica correspondiente al agregado solo muestra los resultados de los pronósticos y no lleva etiqueta como en las anteriores.

4.- Finalmente buscamos el error en la predicción para cada uno de los métodos, en este caso el que mejor nos arrojó fue el método desagregado, los resultados de error fueron: *agregado* = 0,0239602 y *desagregado* = 0,0186323. Para encontrar dichos errores los valores que tomamos como base fueron los verdaderos valores de las 10 semanas tomando los 35 artículos.

Comentarios:

Observamos que el mejor método fue el desagregado ya que nos arrojó un error menor comparado con el método agregado. Podemos afirmar que al menos para este caso el método que mejor funciona es el desagregado pero como desventaja es el tiempo de cómputo ya que al ser 35 artículos había que dividir los datos en los 35 grupos, encontrar el mejor modelo para cada uno, predecir para cada uno de ellos y finalmente realizar la gráfica(opcional), mientras que el agregado solo se corrió una vez y el resultado fue más veloz.

4. PROBLEMA 4

A continuación presentaremos los modelos ARCH y GARCH, muy útiles para modelar y pronosticar fluctuaciones por volatilidad. Para conocer una descripción detallada, véase Diebold y Lopez 1995 y lo que sigue se basa en ese trabajo.

a) El proceso ARCH propuesto por Engle(1982) se define así:

$$\begin{aligned}\varepsilon_t | \mu_{t-1} &\sim N(0, h_t) \\ h_t &= w + \gamma(L)\varepsilon_t^2 \\ w > 0, \quad \gamma(L) &= \sum_{i=1}^p \gamma_i L^i \quad \gamma \neq 0, \text{ para toda } i, \gamma(1) < 1.\end{aligned}$$

El proceso se parametriza en función de la densidad condicional de $\varepsilon_t | \mu_{t-1}$, que se supone tiene distribución normal con promedio condicional igual a cero y varianza condicional que depende en forma lineal de innovaciones pasadas, elevadas al cuadrado. Así, aunque las ε_t son seriamente no correlacionadas, no son independientes (a menos que $\gamma(L)$ sea cero, en cuyo caso ε_t es simplemente iid al ruido con varianza w . En especial. La varianza condicional, que es una medida común de la volatilidad, fluctúa y es pronosticable ¿Cómo esperaría usted que se vea la correlograma de ε_t^2 ? ¿Por qué?

Solución

Se observa que es totalmente razonable hacer uso de las funciones de autocorrelación simple y parcial de los residuos $\hat{\varepsilon}_t^2$ para obtener el orden del proceso ARCH, dado que $\hat{\varepsilon}_t^2$ es un estimador insesgado de h_t e igual es buena opción usar la función de autocorrelación simple de los residuos al cuadrado $\hat{\varepsilon}_t^2$ para poder decir si hay heterocedasticidad e identificar el orden del proceso ARCH.

Donde los correlogramas obtenidos serían muy parecidos a los obtenidos al identificar un proceso ARIMA.

Para obtener el orden del modelo GARCH(r,m) se recomienda intentar con modelos de menor orden y posteriormente elegir el modelo más adecuado tomando como criterios algunos indicadores como el AIC y bayesiano BIC. El modelo GARCH contempla un parámetro extra, directamente relacionado con la media condicional del proceso generando un efecto suavizado. Además el modelo GARCH es más parcimonioso y el modelo ARCH sólo considera la varianza condicional y el GARCH suma la media de manera que reduce los cambios bruscos en los intervalos de tiempo.