UNAP
FINESI

Series de Tiempo

V I Semestre
Grupo: B

PRONÓSTICO CON MODELOS ARIMA

Mtr. Alcides Ramos Calcina
FINESI

Introducción

- Las tres primeras fases de la elaboración de un modelo ARIMA constituyen un proceso iterativo cuyo resultado final es la obtención de un modelo estimado que sea compatible con la estructura de los datos.

- La fase siguiente consiste en utilizar este modelo estimado en el pronóstico de valores futuros de la variable objeto de estudio.

Objetivo

- Obtener pronósticos óptimos de Yt en algún momento futuro basadas en un conjunto de información formado por el pasado disponible de la serie temporal.

$$I_T = \left(Y_T, Y_{T-1}, Y_{T-2}, Y_{T-3}, \cdots\right)$$

Supongamos que se observa una serie temporal denotada por *Yt*, para t = 0 hasta t = *T*, donde *T* es la última observación de que disponemos.

Introducción

- El pronóstico de series temporales, es decir valor que tomará en momentos futuros $\boldsymbol{T + \ell}$, donde ℓ representa el número de periodos en el futuro que estamos considerando.

Introducción

- Al pronóstico de $Y_{T+\ell}$ con información hasta el momento T la vamos a denotar por $Y_T(\ell)$.

  Si $\ell = 1$, $\Rightarrow$ se pronóstica el valor de Y_{T+1} y se calcula lo que se denomina predicción un periodo hacia adelante.

  Si $\ell = 3$, $\Rightarrow$ se pronóstica el valor de Y_{T+3} y se calcula la predicción tres periodos hacia adelante, etc.

- Una manera sensata de proceder a una evaluación de muestreo a posteriori es dividir los datos en dos conjuntos:

  1. El primer conjunto denominado: datos de estimación, se emplea para estimar los coeficientes del modelo.

  2. Los puntos de los datos restantes, denominados datos de predicción, se emplean para medir la exactitud del pronóstico del modelo.

1. Pronóstico de modelos Estacionarios (*ARMA*)

1.1. Pronóstico de modelos MA(*q*)

- Comencemos por un modelo de medias móviles sencillo, por ejemplo, el MA(2) de media cero:

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

- La función de pronóstico es:

$$Y_{T+1} = a_{T+1} - \theta_1 a_T - \theta_2 a_{T-1}$$

$$Y_T(1) = E_T\left[Y_{T+1}\right] = E_T\left[a_{T+1} - \theta_1 a_T - \theta_2 a_{T-1}\right] = -\theta_1 a_T - \theta_2 a_{T-1}$$

$$Y_{T+2} = a_{T+2} - \theta_1 a_{T+1} - \theta_2 a_T$$

$$Y_T(2) = E_T\left[Y_{T+2}\right] = E_T\left[a_{T+2} - \theta_1 a_{T+1} - \theta_2 a_T\right] = -\theta_2 a_{T-1}$$

$$Y_{T+3} = a_{T+3} - \theta_1 a_{T+2} - \theta_2 a_{T+1}$$

$$Y_T(3) = E_T\left[Y_{T+3}\right] = E_T\left[a_{T+3} - \theta_1 a_{T+2} - \theta_2 a_{T+1}\right] = 0$$

$$Y_T(\ell) = E_T\left[Y_{T+\ell}\right] = 0 \quad \forall \ell > 2$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

- Estos resultados se pueden generalizar fácilmente para el modelo MA(q):

$$Y_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \qquad t = 1, 2, \ldots$$

- La función de pronóstico es:

$$Y_T(\ell) = \begin{cases} Y_T(1) = -\theta_1 a_T - \theta_2 a_{T-1} - \ldots - \theta_q a_{T+1-q} \\ Y_T(2) = -\theta_2 a_T - \theta_3 a_{T-2} - \ldots - \theta_q a_{T+2-q} \\ \ldots \\ Y_T(q) = -\theta_q a_T \\ Y_T(\ell) = 0 \quad \forall \ell = q+1, q+2, \ldots \end{cases}$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

- Como el modelo MA(2) está escrito directamente en forma medias móviles, se obtiene la varianza del error de pronóstico:

$$V\left[e_T(\ell)\right] = \begin{cases} V\left[e_T(1)\right] = \sigma^2 \\ V\left[e_T(2)\right] = \left(1+\theta_1^2\right)\sigma^2 \\ \ldots \\ V\left[e_T(q)\right] = \left(1+\theta_1^2+\theta_2^2+\ldots+\theta_{q-1}^2\right)\sigma^2 \\ V\left[e_T(\ell)\right] = \left(1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2\right)\sigma^2 = V\left(Y_t\right) \quad \ell = q+1, q+2, \ldots \end{cases}$$

- Aunque la varianza del error de pronóstico es una función creciente de ℓ, el horizonte de pronóstico, tiene una cota máxima que viene dada por la varianza no condicionada del proceso y que se alcanza para $\ell = q$.

1. Pronóstico de modelos Estacionarios (*ARMA*)

- El pronóstico por intervalo viene dada por::

$$\ell = 1 \quad \left[-\theta_1 a_T - \theta_2 a_{T-1} - \cdots - \theta_q a_{T+1-q} \pm N_{\alpha/2}\sqrt{\sigma^2}\right]$$

$$\ell = 2 \quad \left[-\theta_2 a_T - \theta_3 a_{T-2} - \cdots - \theta_q a_{T+2-q} \pm N_{\alpha/2}\sqrt{\sigma^2\left(1+\theta_1^2\right)}\right]$$

$$\ell = q \quad \left[-\theta_q a_T \pm N_{\alpha/2}\sqrt{\sigma^2\left(1+\theta_1^2+\theta_2^2+\cdots+\theta_{q-1}^2\right)}\right]$$

$$\ell > q \quad \left[0 \pm N_{\alpha/2}\sqrt{\sigma^2\left(1+\theta_1^2+\theta_2^2+\cdots+\theta_{q-1}^2+\theta_q^2\right)}\right]$$

- La amplitud de los intervalos de pronóstico va creciendo con ℓ, con el límite impuesto por

$$\pm N_{\alpha/2}\sqrt{\sigma^2\left(1+\theta_1^2+\theta_2^2+\cdots+\theta_{q-1}^2+\theta_q^2\right)} = \pm N_{\alpha/2}\sqrt{V\left(Y_t\right)}$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

1.2. Pronóstico de modelos AR(p)

- Consideremos el modelo autorregresivo más sencillo, el AR(1).

$$Y_t = \phi_1 Y_{t-1} + a_t \qquad t = 1, 2, \ldots$$

- La función de pronóstico es:

$$Y_{T+1} = \phi Y_T - a_{T+1}$$

$$Y_T(1) = E_T\left[Y_{T+1}\right] = E_T\left[\phi Y_T + a_{T+1}\right] = \phi Y_T$$

$$Y_{T+2} = \phi Y_{T+1} - a_{T+2}$$

$$Y_T(2) = E_T\left[Y_{T+2}\right] = E_T\left[\phi Y_{T+1} + a_{T+2}\right] = \phi E_T\left[Y_{T+1}\right] = \phi Y_T(1)$$

$$Y_{T+3} = \phi Y_{T+2} - a_{T+3}$$

$$Y_T(3) = E_T\left[Y_{T+3}\right] = E_T\left[\phi Y_{T+2} + a_{T+3}\right] = \phi E_T\left[Y_{T+2}\right] = \phi Y_T(2)$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

- La trayectoria de la función de pronóstico depende de la estructura de la parte autorregresiva:

$$
\begin{aligned}
&Y_T(1) = \phi Y_T \\
&Y_T(2) = \phi Y_T(1) = \phi\phi Y_T = \phi^2 Y_T \\
&Y_T(3) = \phi Y_T(2) = \phi\phi^2 Y_T = \phi^3 Y_T \\
&Y_T(4) = \phi Y_T(3) = \phi\phi^3 Y_T = \phi^4 Y_T \\
&\ldots \\
&Y_T(\ell) = \phi^\ell Y_T \qquad \forall \ell = 1, 2, 3, \ldots
\end{aligned}
$$

- Como el proceso autorregresivo es estacionario, $|\phi| < 1$, y por lo tanto cuando nos alejamos en el futuro la función de pronóstico tiende hacia la media no condicionada del proceso:

$$\lim_{\ell \to \infty} Y_T(\ell) = 0 = E(Y_t)$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

- Por lo que la varianza del error de pronóstico se obtiene:

$$V\left[e_T(\ell)\right]=\begin{cases} V\left[e_T(1)\right]=\sigma^2 \\ V\left[e_T(2)\right]=\left(1+\phi^2\right)\sigma^2 \\ V\left[e_T(3)\right]=\left(1+\phi^2+\left(\phi^2\right)^2\right)\sigma^2 \\ V\left[e_T(4)\right]=\left(1+\phi^2+\left(\phi^2\right)^2+\left(\phi^3\right)^2\right)\sigma^2 \\ \ldots \\ V\left[e_T(\ell)\right]=\left(1+\phi^2+\left(\phi^2\right)^2+\cdots+\left(\phi^{\ell-1}\right)^2\right)\sigma^2 \end{cases}$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

- El pronóstico por intervalo es:

$$\ell = 1 \quad \left[\phi Y_T \pm N_{\alpha/2}\sqrt{\sigma^2}\right]$$

$$\ell = 2 \quad \left[\phi Y_T(1) \pm N_{\alpha/2}\sqrt{\sigma^2\left(1+\phi^2\right)}\right]$$

$$\ell = 3 \quad \left[\phi Y_T(2) \pm N_{\alpha/2}\sqrt{\sigma^2\left(1+\phi^2+\left(\phi^2\right)^2\right)}\right]$$

$$\ell \;=\; \left[\phi Y_T(\ell-1) \pm N_{\alpha/2}\sqrt{\sigma^2\left(1+\phi^2+\left(\phi^2\right)^2+\cdots+\left(\phi^{\ell-1}\right)^2\right)}\right]$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

- La amplitud de los intervalos de pronóstico va creciendo con ℓ, con el límite impuesto por:

$$\pm N_{\alpha/2}\sqrt{\frac{\sigma^2}{1-\phi^2}} = \pm N_{\alpha/2}\sqrt{V\left(Y_t\right)}$$

- Los resultados obtenidos para el modelo AR(1) se pueden extender para el modelo AR(p). En general, las fun ciones de pronóstico de procesos autorregresivos puros, se obtendrán a partir de reglas de cadena:

$$Y_T\left(\ell\right) = \phi_1 Y_T\left(\ell-1\right) + \phi_2 Y_T\left(\ell-2\right) + \phi_3 Y_T\left(\ell-3\right) + \cdots + \phi_p Y_T\left(\ell-p\right)$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

1.3. Pronóstico de modelos ARMA(*p, q*)

- Consideremos un modelo ARMA(p,q) sencillo, el ARMA(1,2):

$$Y_t = \delta + \phi Y_{t-1} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \qquad t = 1, 2, \ldots$$

- La media de este proceso no es cero si δ = 0:

$$E(Y_t) = \frac{\delta}{1-\phi}$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

- Los pronósticos por punto son:

$$Y_{T+1} = \delta + \phi Y_T + a_{T+1} - \theta_1 a_T - \theta_2 a_{T-1}$$

$$Y_T(1) = E_T\left[Y_{T+1}\right] = E_T\left[\delta + \phi Y_T + a_{T+1} - \theta_1 a_T - \theta_2 a_{T-1}\right] = \delta + \phi Y_T - -\theta_1 a_T - \theta_2 a_{T-1}$$

$$Y_{T+2} = \delta + \phi Y_{T+1} + a_{T+2} - \theta_1 a_{T+1} - \theta_2 a_T$$

$$Y_T(2) = E_T\left[Y_{T+2}\right] = E_T\left[\delta + \phi Y_{T+1} + a_{T+2} - \theta_1 a_{T+1} - \theta_2 a_T\right] = \delta + \phi Y_T(1) - \theta_2 a_T$$

$$Y_{T+3} = \delta + \phi Y_{T+2} + a_{T+3} - \theta_1 a_{T+2} - \theta_2 a_{T+1}$$

$$Y_T(2) = E_T\left[Y_{T+2}\right] = E_T\left[\delta + \phi Y_{T+2} + a_{T+3} - \theta_1 a_{T+2} - \theta_2 a_{T+1}\right] = \delta + \phi Y_T(2)$$

$$\rightarrow \qquad Y_T(\ell) = E_T\left[Y_{T+\ell}\right] = \delta + \phi Y_T(\ell - 1) \quad \forall \ell > 2$$

1. Pronóstico de modelos Estacionarios (*ARMA*)

- La estructura de la función de pronóstico es la siguiente. Los dos primeros pronósticos dependen de la última observación Y_T (parte autorregresiva) y de los últimos errores de pronóstico un periodo hacia adelante a_T y a_{T-1} (parte medias móviles).

$$Y_T(3) = \delta + \phi Y_T(2)$$

$$Y_T(4) = \delta + \phi Y_T(3) = \delta + \phi\left(\delta + \phi Y_T(2)\right) = \delta(1+\phi) + \phi^2 Y_T(2)$$

$$Y_T(5) = \delta + \phi Y_T(4) = \delta + \phi\left(\delta(1+\phi) + \phi^2 Y_T(2)\right) = \delta\left(1+\phi+\phi^2\right) + \phi^3 Y_T(2)$$

$$\ldots$$

$$Y_T(\ell) = \delta + \phi Y_T(\ell-1) = \delta\left(1+\phi+\phi^2+\ldots+\phi^{\ell-3}\right) + \phi^{\ell-2} Y_T(2)$$

Ejemplo 1

Retomamos el Ejemplo 1 de una serie clásica en la literatura con datos anuales de la cantidad de Pieles de lince (Yt) captadas cada año, en el periodo 1821 a 1930, del capítulo anterior, para lo cual se pide realizar el pronóstico con los modelos planteados

Modelo 1: ARIMA(3, 0, 0)
Modelo 2: ARIMA(0, 0, 2)
Modelo 2: ARIMA(2, 0, 1)

Solución:

Para el análisis consideraremos los modelos anteriores, para los cuales se hará un pronóstico de 10 años, es decir, $\ell = 10$.

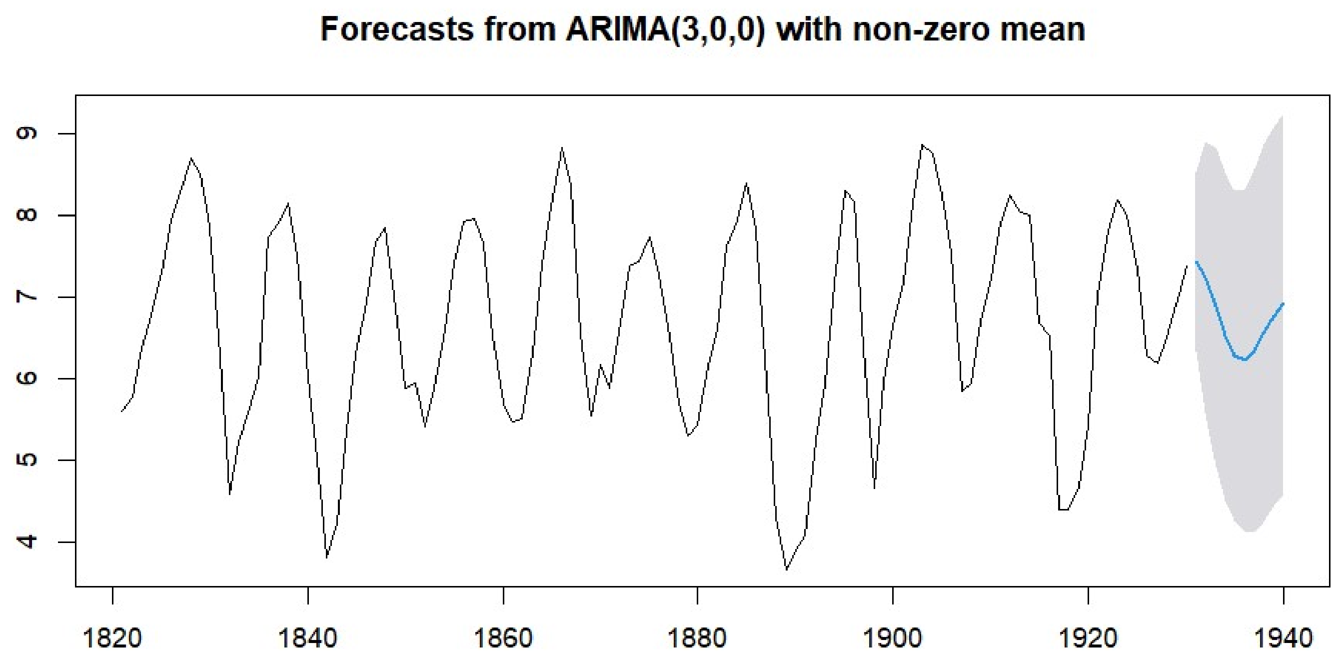
Nota: Previamente estimar el modelo haciendo uso de la función **Arima()**:

Ejemplo 1

- **Modelo 1:** ARIMA(3, 0, 0)

```
mod1 <- Arima(T.Yt, order = c(3,0,0),include.constant = T)
Pron1=forecast(mod1,level= c(95), h=10)
plot(Pron1)
```

En la figura se puede observar el comportamiento de la función de predicción por punto y por intervalo.



Ejemplo 1

Los datos proyectados para los siguientes 10 años son:

```
summary(Pron1)

Forecast method: ARIMA(3,0,0) with non-zero mean

Model Information:
Series: T.Yt
ARIMA(3,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3    mean
      1.1294  -0.3298  -0.2678  6.6754
s.e.  0.0911   0.1376   0.0909  0.1125

sigma^2 = 0.3123:  log likelihood = -91.25
AIC=192.49   AICc=193.07   BIC=205.99

Error measures:
                        ME      RMSE      MAE        MPE     MAPE      MASE
Training set -0.0002035144 0.5486022 0.424295 -0.8392937 6.913166 0.6166635
                    ACF1
Training set -0.02926266
```

Ejemplo 1

```
Forecasts:
     Point Forecast    Lo 95    Hi 95
1931       7.433219 6.337879 8.528560
1932       7.239540 5.587197 8.891883
1933       6.876240 4.925958 8.826521
1934       6.513254 4.507417 8.519091
1935       6.274962 4.264000 8.285924
1936       6.222828 4.125937 8.319719
1937       6.339741 4.113008 8.566475
1938       6.552798 4.243523 8.862072
1939       6.768839 4.442058 9.095619
1940       6.911271 4.583092 9.239451
```

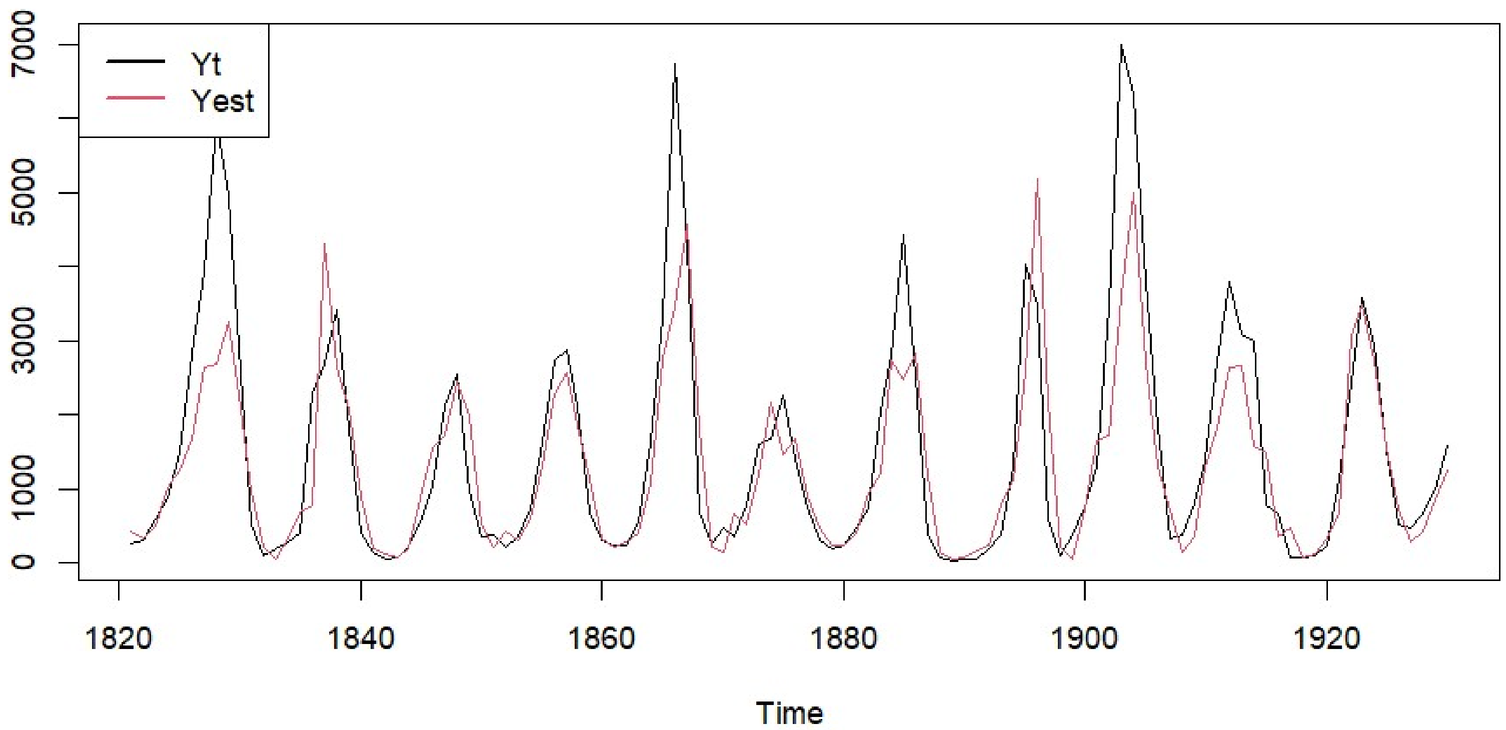
Es importante mencionar que, los valores pronosticados son logaritmos, por tanto, para obtener los verdaderos valores de pronóstico, se tendrán que transformar aplicando la función **exp()**.

Ejemplo 1

SERIE ORIGINAL (Yt) Y PRONOSTICADA (Yest) DEL MODELO ARIMA(3, 0, 0).

```
Yt_arima1<-exp(mod1$fitted)
grafico_comparativo <- cbind(Yt, Yt_arima1)ts.plot(grafico_comparativo, col = c(1:2), lwd = 1)
legend("topleft", c("Yt", "Yest"), lty = c(1,1), col = c(1:2), lwd = 2)
```

$$\hat{Y}_t = e^{\log(Y_t)} = \exp(\log(Y_t))$$



Ejemplo 1

PRONÓSTICO DE LA SERIE ORIGINAL (Yt)

Deshaciendo la transformación:

```
Pron1$mean <- exp(Pron1$mean)
Pron1$lower <- exp(Pron1$lower)
Pron1$upper <- exp(Pron1$upper)
Pron1$x <- exp(Pron1$x)
Pron1$fitted <- exp(Pron1$fitted)
Pron1$residuals <- exp(Pron1$residuals)
summary(Pron1)
```

```
Forecast method: ARIMA(3,0,0) with non-zero mean

Model Information:
Series: T.Yt
ARIMA(3,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3    mean
      1.1294  -0.3298  -0.2678  6.6754
s.e.  0.0911   0.1376   0.0909  0.1125

sigma^2 = 0.3123:  log likelihood = -91.25
AIC=192.49   AICc=193.07   BIC=205.99
```

Ejemplo 1

```
Error measures:
                   ME     RMSE      MAE       MPE     MAPE      MASE      ACF1
Training set 194.8265 834.1177 510.5462 -16.88696 48.50408 0.6151091 0.2420179

Forecasts:
     Point Forecast      Lo 95     Hi 95
1931      1691.2436 565.59545  5057.157
1932      1393.4526 266.98620  7272.699
1933       968.9759 137.82137  6812.544
1934       674.0160  90.68729  5009.496
1935       531.1063  71.09380  3967.630
1936       504.1269  61.92582  4104.005
1937       566.6496  61.13031  5252.580
1938       701.2033  69.65284  7059.096
1939       870.3008  84.94961  8916.150
1940      1003.5224  97.81640 10295.382
```

Valores en términos de la serie original Yt

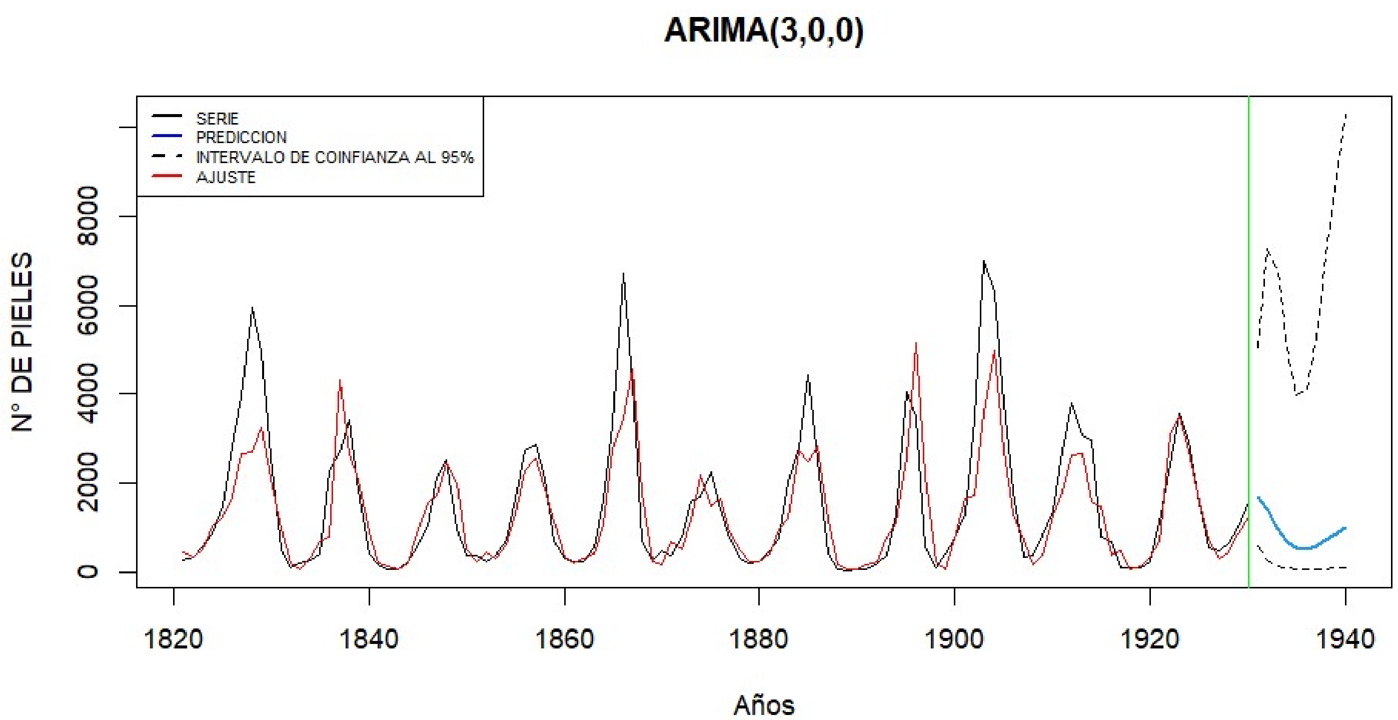
También podemos solicitar independientemente las medidas de error a través de:

```
accuracy(Pron1)
                   ME     RMSE      MAE       MPE     MAPE      MASE      ACF1
Training set 194.8265 834.1177 510.5462 -16.88696 48.50408 0.6151091 0.2420179
```

Ejemplo 1

GRÁFICA DEL AJUSTE Y PRONÓSTICO CON VALORES REALES

```
plot(Pron1, shaded = FALSE, xlab = "Años", ylab = "N° DE PIELES",main = "ARIMA(3,0,0)")
lines(Pron1$fitted, col = "red")
legend("topleft", legend=c("SERIE", "PREDICCION", "INTERVALO DE COINFIANZA AL 95%", "AJUSTE"),
col=c("black", "blue", "black", "red"), lty=c(1,1,2,1), lwd = 2,cex = 0.6)
abline(v=1930, lwd = 1, col="green")
```



Replique el procedimiento de pronóstico los pasos para los modelos 2 y 3.

METRICAS DE EVALUACIÓN PARA PRONÓSTICOS

Son medidas numéricas que cuantifican el rendimiento de un modelo de pronóstico en función de algunos criterios, como la precisión, el error, el sesgo o la incertidumbre.

Se describirán algunas de las métricas de evaluación más importantes para un modelo de pronóstico de series temporales, cómo calcularlas y qué significan.

1. Métricas basadas en el Error

- **Error medio (ME – *Mean Error*)**

El error medio (ME) es el promedio de todos los errores de un conjunto de observaciones.

$$ME = \sum_{t=T+1}^{T+\ell} \frac{\left(\hat{y}_t - y_t\right)}{\ell}$$

- **Error absoluto medio (MAE – Mean Absolute Error)**

La solución más sencilla para la inexactitud del error medio es el uso del error absoluto medio (MAE).

$$MAE = \sum_{t=T+1}^{T+\ell} \frac{\left|\hat{y}_t - y_t\right|}{\ell}$$

1. Métricas basadas en el Error

- **Error cuadrático medio (MSE – Mean Square Error)**

Al igual que el MAE, el error cuadrático medio (MSE) también arregla el problema de la cancelación de errores positivos y negativos.

$$MSE = \sum_{t=T+1}^{T+\ell} \frac{\left(\hat{y}_t - y_t\right)^2}{\ell}$$

- **La raíz del error cuadrático medio (RMSE – Root Mean Square Error)**

La raíz del error cuadrático medio (RMSE) es lo que en estadística se conoce como desviación estándar de los errores.

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

1. Métricas basadas en el Error

- **Error porcentual medio (MPE – Mean Percentage Error)**

El error porcentual medio (MPE) es el promedio de errores porcentuales por los que cada previsión difiere de sus correspondientes valores reales observados.

$$MPE = \sum_{t=T+1}^{T+\ell} \frac{\left(\hat{y}_t - y_t\right)}{\hat{y}_t}.100$$

- **Error porcentual absoluto medio (MAPE – Mean Absolute Percentage Error)**

El error porcentual absoluto medio (MAPE) arregla el problema con la compensación de errores (tal y como lo hacía el MAE) y funciona mejor si no hay extremos en los datos (y no hay ceros).

$$MAPE = \sum_{t=T+1}^{T+\ell} \left|\frac{\hat{y}_t - y_t}{\hat{y}_t}\right|.100$$

1. Métricas basadas en el Error

- **Error medio absoluto escalado (MASE - Mean Absolute Scaled Error)**

Calcula el error medio absoluto escalado (MASE) entre el pronóstico y los eventuales resultados.

$$MASE = \frac{\sum_{t=T+1}^{T+\ell} \frac{\left|\hat{y}_t - y_t\right|}{\ell}}{\sum_{t=2}^{T} \frac{\left|y_t - y_{t-1}\right|}{T-1}} = \frac{MAE}{\sum_{t=2}^{T} \frac{\left|y_t - y_{t-1}\right|}{T-1}}$$

Mide la magnitud media de los errores absolutos entre los valores reales y los predichos, escalada por la magnitud media de los errores absolutos del pronóstico sencillo.

Si el MASE es mayor o igual a 1, la precisión se establece en el 0%, ya que el modelo no mejora sobre el modelo sencillo. Una precisión mayor indica el error de modelo inferior relativo al modelo sencillo.

Ejemplo 1

Continuamos con el Ejemplo 1, cantidad de Pieles de Lince al cual vamos a calcular todas estas métricas en las que tenemos predicciones hechas con 3 modelos distintos.

- **Modelo 1: ARIMA(3, 0, 0):**

```
accuracy(Pron1)
                   ME     RMSE      MAE       MPE     MAPE      MASE      ACF1
Training set 194.8265 834.1177 510.5462 -16.88696 48.50408 0.6151091 0.2420179
```

- **Modelo 2: ARIMA(0, 0, 2):**

```
accuracy(Pron2)
                  ME     RMSE      MAE       MPE     MAPE      MASE      ACF1
Training set 398.588 1097.889 680.8917 -30.15231 69.92053 0.8203424 0.3650378
```

- **Modelo 3: ARIMA(2, 0, 1):**

```
accuracy(Pron3)
                   ME    RMSE      MAE       MPE     MAPE      MASE      ACF1
Training set 189.5478 843.877 518.1915 -17.15388 49.11714 0.6243202 0.2406269
```

Ejemplo 1

Conclusión:

- En primer lugar, el error medio (ME) del modelo 2 está mucho más lejos de cero, lo que sugiere un mayor sesgo en la predicción y el modelo 3, presenta el menor sesgo.

- Luego, con el propósito de elegir uno de los tres modelos, tendremos que ver con más detalle las medidas del RMSE que en el primer caso es de 834.1177, el segundo 1097.889 y el tercero 843.877, siendo el primero menor a los demás; así mismo, también es un buen indicar de error de pronóstico el MAE siendo estos 510.5462, 680.8917 y 518.1915 para el primero, segundo y tercer modelo respectivamente de error en la predicción; por tanto, tomando en cuenta el MAE también se puede elegir el **modelo 1**.

.

2. Evaluación utilizando AIC y BIC

Si los modelos contienen diferentes números de parámetros, el principio de parsimonia conduce a la selección del modelo más sencillo. Sin embargo, el modelo con más parámetros puede tener un error cuadrático medio considerablemente menor.

- **El criterio de información de Akaike (AIC - *Aikaike Information Criterion*)**

El AIC fue propuesto por Akaike (1974) como un estimador insesgado asintótico de la información de Kullback-Leibler esperada, entre un modelo candidato ajustado y el verdadero modelo.

$$AIC = \ln \hat{\sigma}^2 + \frac{2}{n} r$$

- donde:

ln	: logaritmo natural
$\hat{\sigma}^2$	: suma residual de cuadrados dividida entre el número de observaciones
n	: número de observaciones (residuos)
r	: número total de parámetros (incluyendo el término constante) en el modelo ARIMA

2. Evaluación utilizando AIC y BIC

- **El criterio de información bayesiano (BIC -** ***Bayesian Information Criterion*****)**

  El BIC fue derivado por Schwarz en 1978 como una aproximación a una transformación de la probabilidad posterior de un modelo candidato.

$$BIC = \ln \hat{\sigma}^2 + \frac{\ln n}{n} r$$

- El segundo término tanto en el criterio AIC como en el criterio BIC es un factor que “penaliza” la inclusión de parámetros adicionales en el modelo.

- Como el criterio BIC impone una penalización mayor para el número de parámetros que el criterio AIC, el uso del BIC mínimo para la selección del modelo tendrá como resultado un modelo cuyo número de parámetros no será mayor que el seleccionado por el AIC.

- A menudo los dos criterios producen el mismo resultado.

A2

Ejemplo 1

En el ejemplo 1 los tres modelos ARIMA parecían dar una descripción adecuada de las cantidades de Pieles de Lince. Se tiene:

- modelo 1: ARIMA(3,0,0) con r = 4 parámetros estimados (incluyendo un término constante),
- modelo 2: ARIMA(0,0,2) con r = 3 parámetros y
- Modelo 3: ARIMA(2,0,1) con r = 4 parámetros.

De los cálculos, los criterios AIC y BIC es como sigue:

Criterios AIC y BIC para los modelos de pronóstico.

Modelo	AIC	BIC
1) ARIMA(3,0,0)	192.49	205.99
2) ARIMA(0,0,2)	241.16	251.96
3) ARIMA(2,0,1)	192.51	206.01

El AIC es menor para el modelo ARIMA(3,0,0), así como el BIC. Recordemos que el BIC no seleccionaría un modelo con más parámetros que el AIC a causa de la mayor penalización por incluir parámetros adicionales. También a través de los criterios se selecciona el modelo 1: ARIMA(3,0,0), con base en el principio de parsimonia, estaba apoyada por el criterio BIC.

Diapositiva 34

A2 ALCIO, 24/06/2024

GRACIAS

https://aulavirtual2.unap.edu.pe/