Package 'nortestARMA'

October 13, 2022

Type Package

Title Neyman Smooth Tests of Normality for the Errors of ARMA Models

Version 1.0.2

Date 2017-04-14

License GPL (> 2)

LazyLoad yes

Author Pierre Duchesne and Pierre Lafaye de Micheaux

Maintainer Pierre Lafaye de Micheaux <lafaye@unsw.edu.au>

Description Tests the goodness-of-fit to the Normal distribution for the errors of an ARMA model.

Depends R (>= 2.10.0), astsa

NeedsCompilation yes

Repository CRAN

Date/Publication 2017-04-14 14:37:35 UTC

R topics documented:

---

nortestARMA   *Neyman Smooth Tests of Normality for the Errors of ARMA Models*

Description

Test statistics are computed to check normality in autoregressive moving-average (ARMA) time series models, adopting the smooth test paradigm. The mean of the ARMA process can be assumed known or unknown. A data-driven order of the family of the BIC type is given.

1

Usage

```
nortestARMA(residuals, sig2hat, d = 2)
```

Arguments

residuals	Vector of residuals of a fitted ARMA(p, q) model.
sig2hat	Estimated variance of the i.i.d. errors of the ARMA(p, q) model.
d	Paramater used in the choice of the order of the Legendre polynomial.

Details

In the 2004 paper, the mean μ of the ARMA process is supposed to be known and thus should not be estimated. In that case, the residuals should be computed without estimation of the mean μ. Moreover, these residuals should not be centered before calling the nortestARMA function because there might be some asymptotic impact of centering the residuals (see Serfling (1980) page 73, case (iv)). In the 2016 paper, we considered that the mean μ is unknown. Thus in that case, the residuals should be computed taking into account the estimation of μ. Moreover, we have showed that there is no asymptotic impact of the centering of the residuals (see Yu (2007)). Nevertheless, in that case, centering the residuals could be a good idea for small sample sizes.

Value

List with the following components:

RKsdest	Values of the test statistics R_k, for $k = 1, ..., 10$ when only the standard deviation is estimated (mean is supposed known).
Koptsdest	Optimal value $\hat{K}$ of K using a BIC type criterion involving the values of RKsdest, $k = d, ..., D = 10$.
pvalsdest	p-value of the test, associated to RKsdest[Koptsdest].
RKsdmuest	Values of the test statistics R_k, for $k = 1, ..., 10$ for the new test (both the standard deviation and the mean are estimated).
Koptsdmuest	Optimal value $\hat{K}$ of K using a BIC type criterion involving the values of RKsdmuest, $k = d, ..., D = 10$.
pvalsdmuest	p-value for the new test, associated to RKsdmuest[Koptsdmuest].

Author(s)

Duchesne P., Lafaye de Micheaux P.

References

Ducharme G., Lafaye de Micheaux P., (2004). Goodness-of-fit tests of normality for the innovations in ARMA models. *Journal of Time Series Analysis* 25 (3), 373–395.

Duchesne P., Lafaye de Micheaux P., Tagne J. (2016). Neyman Smooth Test of Normality for ARMA Time Series Models with Unknown Mean. Submitted.

Piggott J. L., (1980). The use o f Box-Jenkins modelling for the forecasting of daily gas demand. Paper presented to the *Royal Statistical Society*.

Shea B. L., (1987). Estimation of Multivariate Time Series. *Journal of Time Series Analysis* 8 (1), 95–109.

Examples

```
#####################################
# Example in Ducharme et al. (2004) #
#####################################
data(Piggott)
ts.plot(Piggott$temp, main = "Temperature Series")
ts.plot(Piggott$wind.trans, main = "Transformed Wind Speed Series")

# The differenciated series Y_t - Y_{t - 1} of length 366 (taking Y_0 = 0)
diff.temp <- c(Piggott$temp[1], diff(Piggott$temp))

pacf(diff.temp)
# Below, we assumed that the mean E(Y_t) is known and is equal to 0.
fit.MA4 <- arima(diff.temp, order = c(0,0,4), include.mean = FALSE)

resfit.MA4 <- residuals(fit.MA4)
sig2fit.MA4 <- fit.MA4$sigma2
# The residuals should not be centered.
nortestARMA(resfit.MA4, sig2fit.MA4, d = 2)

# Note: In Ducharme et al. (2004), the data were analyzed using the
#  G13DCF routine in NAG.
# This routine gave us slightly different results.
# We obtained the following values:
#  ITERATION NUMBER =      9
#  ESTIMATED CONDITION NUMBER OF HESSIAN MATRIX =     0.181E+01
#  NUMBER OF LIKELIHOOD EVALUATIONS MADE SO FAR =       82
#  VALUE OF LOG LIKELIHOOD FUNCTION = -0.68515E+03
#  NORM OF GRADIENT VECTOR =  0.12052E-03
#  MA PARAMETERS :  0.0715894141 -0.296819534 -0.149662304 -0.19482046
#  VARIANCE :  2.47466874
# For reproducibility purpose, the non centered residuals returned
#  by this routine have been included in the current package in the
# 'resid.temp' object. To obtain exactly the same results as in the
#  original 2004 article, one can thus issue the following commands:
sig2NAG <- 2.47466874
nortestARMA(resid.temp, sig2NAG, d = 2)


#####################################
# Example in Duchesne et al. (2016) #
#####################################
data(potato)
plot(potato, type = "l")

# First difference
potatoe.yield.diff <- diff(potato$potatoeyield)
mean(potatoe.yield.diff)
```

```
potatoe.yield.diff.centered <- scale(potatoe.yield.diff, scale = FALSE)
n <- length(potatoe.yield.diff.centered)
ts.plot(potatoe.yield.diff.centered)

acf2(potatoe.yield.diff.centered)

# -------------------------------------
# ------------- Fitting an AR(1) --------
# -------------------------------------

( fit.AR1 <- arima(potatoe.yield.diff.centered, order = c(1, 0, 0) ,
method= 'ML', include.mean = TRUE) )
( fit.AR1.wo.mean <- arima(potatoe.yield.diff.centered, order = c(1, 0, 0) ,
method= 'ML', include.mean = FALSE) )

resfit.AR1 <- residuals(fit.AR1)
sig2fit.AR1 <- fit.AR1$sigma2
resfit.AR1.centered <- resfit.AR1 - mean(resfit.AR1)
nortestARMA(resfit.AR1.centered, sig2fit.AR1, d = 2)
# Using the (2014) approach where the residuals have been
# computed after estimating the mean (by first removing the average of the
# observations), and doing as if the mean had not
# been estimated and were known (include.mean = FALSE).
# Note that if Y_t = a + bt + X_t for example, differenciating
# will not make the true mean of the differenciated series to be 0.
resfit.AR1.wo.mean <- residuals(fit.AR1.wo.mean)
sig2fit.AR1.wo.mean <- fit.AR1.wo.mean$sigma2
nortestARMA(resfit.AR1.wo.mean, sig2fit.AR1.wo.mean, d = 2)

hist(resfit.AR1/sqrt(sig2fit.AR1))

Box.test(resfit.AR1.centered, lag = 3, type = "Ljung-Box", fitdf = 1)
Box.test(resfit.AR1.centered, lag = 6, type = "Ljung-Box", fitdf = 1)
Box.test(resfit.AR1.centered, lag = 12, type = "Ljung-Box", fitdf = 1)
Box.test(resfit.AR1.centered, lag = 14, type = "Ljung-Box", fitdf = 1)
Box.test(resfit.AR1.centered, lag = 20, type = "Ljung-Box", fitdf = 1)

acf2(resfit.AR1)

# ----------------------------------------------------------
# ------------- Fitting an AR(1) with intervention --------
# ----------------------------------------------------------

matX.AOdsARI11 <- function(n, phi) {
 X <- matrix(0, nrow = n + 2, ncol = n)
 oneminusphi <- c(1, -1 - phi, phi)
 for (i in 1:n) X[i:(i + 2), i] <- oneminusphi
 X <- X[1:n, ]
return(X)
}

findAOdsARI11 <- function(residuals, X, phi) {
 n <- length(residuals)
```

```
 oneminusphi <- c(1, -1 - phi, phi)
 tau2 <- sum(oneminusphi ^ 2)
 sigma2 <- mean(residuals ^ 2)
 res <- matrix(0, nrow = n, ncol = 2)
 for (t in 1:n) {
  colXt <- X[, t]
  fitt <- lm(residuals ~ colXt - 1)
  omegaAt <- coef(fitt)
  res[t, 1] <- sqrt(tau2) * omegaAt / sqrt(sigma2)
  res[t, 2] <- omegaAt
 }
return(res)
}

findAOandIOdsARI11 <- function(residuals, X, phi) {
 n <- length(residuals)
 oneminusphi <- c(1, -1 - phi, phi)
 tau2 <- sum(oneminusphi ^ 2)
 sigma2 <- mean(residuals ^ 2)
 res <- matrix(0, nrow = n, ncol = 3)
 for (t in 1:n) {
  colXt <- X[, t]
  fitt <- lm(residuals ~ colXt - 1)
  omegaAt <- coef(fitt)
  res[t, 1] <- sqrt(tau2) * omegaAt / sqrt(sigma2)
  res[t, 2] <- omegaAt
  omegaIt <- residuals[t] / sqrt(sigma2)
  res[t, 3] <- omegaIt
 }
return(res)
}

phi <- coef(fit.AR1)[1]
matX <- matX.AOdsARI11(n, phi)
findAO <- findAOdsARI11(resfit.AR1, matX, phi)
findAOandIO <- findAOandIOdsARI11(resfit.AR1, matX, phi)

omega <- rep(0, length(potato$potatoeyield))
findAOandIO[44, 2]
omega[44 + 1] <- -93.8998780  # warning: the residuals in 'resfit.AR1' are in the diff series
potatoe.yieldAO <- potato$potatoeyield - omega

plot(potato, type = "l", xlab = "year", ylab = "Hundredweight - Cwt")
lines(potato$year, potatoe.yieldAO, type = "l", lty = 2, col = "blue")
title('Original time series and with an intervention')

potatoe.yield.diffAO <- diff(potatoe.yieldAO)
mean(potatoe.yield.diffAO)
potatoe.yield.diff.centeredao <- potatoe.yield.diffAO - mean(potatoe.yield.diffAO)
n <- length(potatoe.yield.diff.centeredao)

ts.plot(potatoe.yield.diff.centered)
ts.plot(potatoe.yield.diff.centeredao)
```

```
plot(1:length(potatoe.yield.diff.centered), potatoe.yield.diff, type = "l")
lines(1:length(potatoe.yield.diff.centeredao), potatoe.yield.diffAO,
 type = "l", lty = 2)


acf2(potatoe.yield.diff.centeredao)


( fit.AR1 <- arima(potatoe.yield.diff.centeredao, order = c(1,0,0) ,
method= 'ML', include.mean = TRUE) )
( fit.AR1.wo.mean <- arima(potatoe.yield.diff.centeredao, order = c(1,0,0) ,
method= 'ML', include.mean = FALSE) )

resfit.AR1 <- residuals(fit.AR1)
resfit.AR1.centered <- resfit.AR1 - mean(resfit.AR1)
sig2fit.AR1 <- fit.AR1$sigma2

hist(resfit.AR1/sqrt(sig2fit.AR1))

acf2(resfit.AR1)

phi <- coef(fit.AR1)[1]

nortestARMA(resfit.AR1.centered, sig2fit.AR1, d = 2)
# Using the (2014) approach where the residuals have been
# computed after estimating the mean (by first removing the average of the
# observations), and doing as if the mean had not
# been estimated and were known (include.mean = FALSE).
resfit.AR1.wo.mean <- residuals(fit.AR1.wo.mean)
sig2fit.AR1.wo.mean <- fit.AR1.wo.mean$sigma2
nortestARMA(resfit.AR1.wo.mean, sig2fit.AR1.wo.mean, d = 2)

Box.test(resfit.AR1.centered, lag = 3, type = "Ljung-Box", fitdf = 1)
Box.test(resfit.AR1.centered, lag = 6, type = "Ljung-Box", fitdf = 1)
Box.test(resfit.AR1.centered, lag = 12, type = "Ljung-Box", fitdf = 1)
Box.test(resfit.AR1.centered, lag = 14, type = "Ljung-Box", fitdf = 1)
Box.test(resfit.AR1.centered, lag = 20, type = "Ljung-Box", fitdf = 1)

# ------------------------------------
# ------------ fitting an AR(2) -------
# ------------------------------------

( fit.AR2 <- arima(potatoe.yield.diff.centeredao, order = c(2, 0, 0),
method= 'ML', include.mean = TRUE) )
( fit.AR2.wo.mean <- arima(potatoe.yield.diff.centeredao, order = c(2, 0, 0),
method= 'ML', include.mean = FALSE) )

resfit.AR2 <- residuals(fit.AR2)
sig2fit.AR2 <- fit.AR2$sigma2

hist( resfit.AR2/sqrt(sig2fit.AR2) )
```

```
acf2(resfit.AR2)

resfit.AR2.centered <- resfit.AR2 - mean(resfit.AR2)
nortestARMA(resfit.AR2.centered, sig2fit.AR2, d = 2)
# Using the (2014) approach where the residuals have been
# computed after estimating the mean (by first removing the average of the
# observations), and doing as if the mean had not
# been estimated and were known (include.mean = FALSE).
resfit.AR2.wo.mean <- residuals(fit.AR2.wo.mean)
sig2fit.AR2.wo.mean <- fit.AR2.wo.mean$sigma2
nortestARMA(resfit.AR2.wo.mean, sig2fit.AR2, d = 2)

Box.test(resfit.AR2.centered, lag = 3, type= "Ljung-Box", fitdf = 2)
Box.test(resfit.AR2.centered, lag = 6, type= "Ljung-Box", fitdf = 2)
Box.test(resfit.AR2.centered, lag = 12, type= "Ljung-Box", fitdf = 2)
Box.test(resfit.AR2.centered, lag = 14, type= "Ljung-Box", fitdf = 2)
Box.test(resfit.AR2.centered, lag = 20, type= "Ljung-Box", fitdf = 2)
```

Piggott   *Daily gas demand study by Piggott*

Description

Two series consisting of daily temperature values and the (square root transformed) wind speed series obtained from Piggott (1980). This data set was first studied by Shea (1987) and then by Ducharme et al. (2004).

Usage

```
data(Piggott)
```

Format

A data frame with 366 observations on the following 2 variables.

temp numeric vector of daily temperatures.

wind.trans numeric vector of the (square root transformed) wind speed series.

Author(s)

Duchesne P., Lafaye de Micheaux P.

References

Ducharme G., Lafaye de Micheaux P., (2004). Goodness-of-fit tests of normality for the innovations in ARMA models. *Journal of Time Series Analysis* 25 (3), 373–395.

Duchesne P., Lafaye de Micheaux P., Tagne J. (2016). Neyman Smooth Test of Normality for ARMA Time Series Models with Unknown Mean. Submitted.

Piggott J. L., (1980). The use o f Box-Jenkins modelling for the forecasting of daily gas demand. Paper presented to the *Royal Statistical Society*.

Shea B. L., (1987). Estimation of Multivariate Time Series. *Journal of Time Series Analysis* 8 (1), 95–109.

Examples

```
data(Piggott)
```

---

potato	*Prince Edward Island; Average Yield, Potatoes (Hundredweight per harvested acres (1957-2014))*

Description

We retrieved the data from the Canadian Socio-Economic Information Management System (CAN-SIM), which is a computerized database of Statistics Canada. The time series consists of annual data on the productivity of potatoes (per acre) in Prince Edward Island for the time period 1957-2014 (CANSIM Series V47152).

Usage

```
data(potato)
```

Format

A data frame with 58 observations on the following 2 variables.

`year` numeric vector of the years of harvest

`potatoeyield` numeric vector of the productivity of potatoes (Hundredweight per harvested acres)

Author(s)

Duchesne P., Lafaye de Micheaux P.

References

Cheng H., (2005). Competitive relationship among potato production areas in northeastern america. *Journal of Food Distribution Research Proceedings* 36 (1), 27–32.

Ducharme G., Lafaye de Micheaux P., (2004). Goodness-of-fit tests of normality for the innovations in ARMA models. *Journal of Time Series Analysis* 25 (3), 373–395.

Duchesne P., Lafaye de Micheaux P., Tagne J. (2016). Neyman Smooth Test of Normality for ARMA Time Series Models with Unknown Mean. Submitted.

Examples

```
data(potato)
```

Index