Modern College of Arts, Science and Commerce, Pune-05

Department of Statistics

M.Sc. II

Date: Submission date:

Practical No. 5

Practical title: Forecasting using Holt Winter’s method

Q.1 Read the data file dowj.tsm in ITSM. Forecast next 10 observations based on all the data using Holt-Winter’s forecast method and Holt-Winter’s seasonal forecast method.

Q.2 Read the data file Births. Forecast next 24 observations based on all the data using Holt-Winter’s forecast method and Holt-Winter’s seasonal forecast method.

Q1)

> library(tseries)

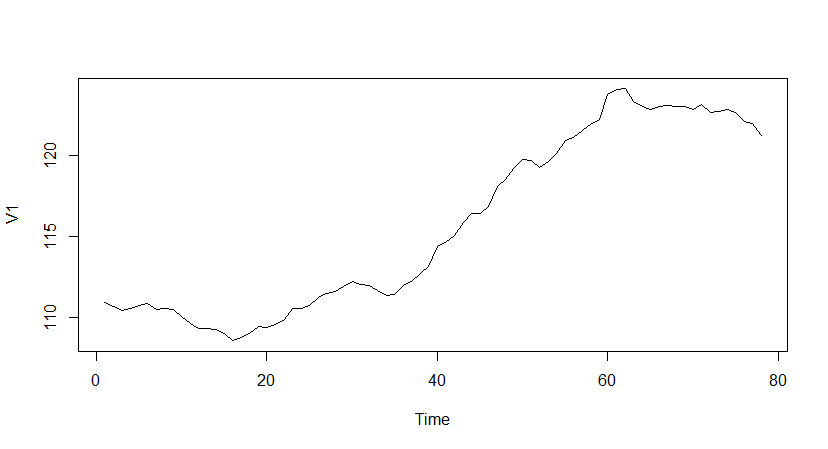
> library(forecast)

> dowj <- read.table("C:/Users/DELL/OneDrive/Desktop/itsm2000/dowj.tsm", quote="\"", comment.char="")

> View(dowj)

> t=ts(dowj)

> plot.ts(t)



From, above graph we can see that the data has increasing trend and there is no seasonality present in the data.

> fit\_data<-HoltWinters(t,gamma=FALSE)

> fit\_data

Holt-Winters exponential smoothing with trend and without seasonal component.

Call:

HoltWinters(x = t, gamma = FALSE)

Smoothing parameters:

alpha: 1

beta : 0.2737748

gamma: FALSE

Coefficients:

[,1]

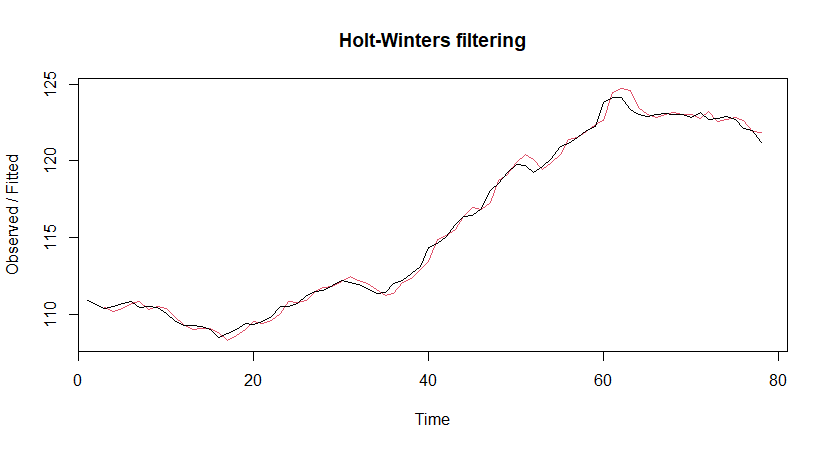
a 121.2300000

b -0.3340033

> fit\_data$SSE

[1] 11.46362

> plot(fit\_data)



> forecasted\_t<-forecast(fit\_data,h=10)

> forecasted\_t

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

79 120.896 120.3950 121.3970 120.1298 121.6622

80 120.562 119.7507 121.3733 119.3212 121.8028

81 120.228 119.1058 121.3502 118.5117 121.9443

82 119.894 118.4476 121.3404 117.6819 122.1061

83 119.560 117.7729 121.3471 116.8268 122.2932

84 119.226 117.0807 121.3713 115.9450 122.5070

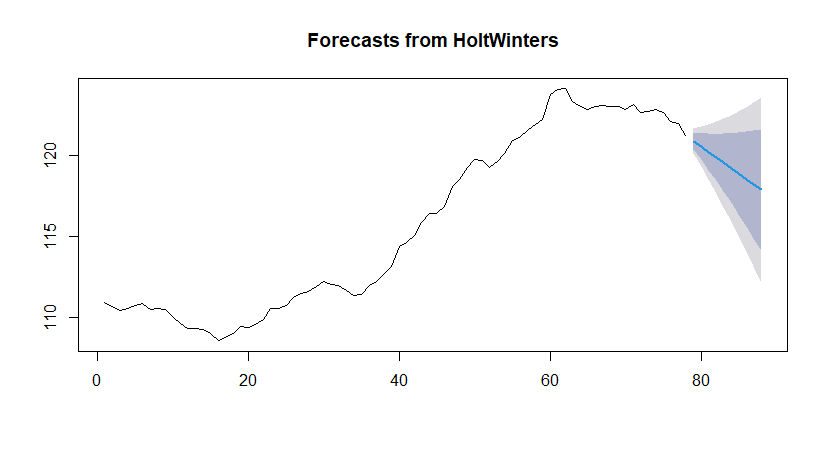
85 118.892 116.3710 121.4129 115.0365 122.7475

86 118.558 115.6442 121.4718 114.1017 123.0143

87 118.224 114.9006 121.5473 113.1413 123.3066

88 117.890 114.1407 121.6392 112.1560 123.6239

> plot(forecasted\_t)



Here the forecasts for 79-88 are plotted as a blue line, the 80% prediction interval as a blue shaded area, and the 95% prediction interval as a grey shaded area.

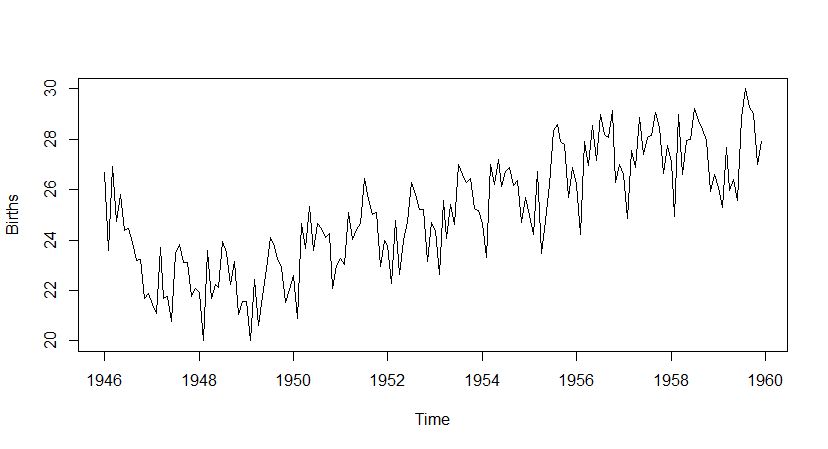
Q2)

> Births <- read\_excel("E:/Downloads/Births.xlsx")

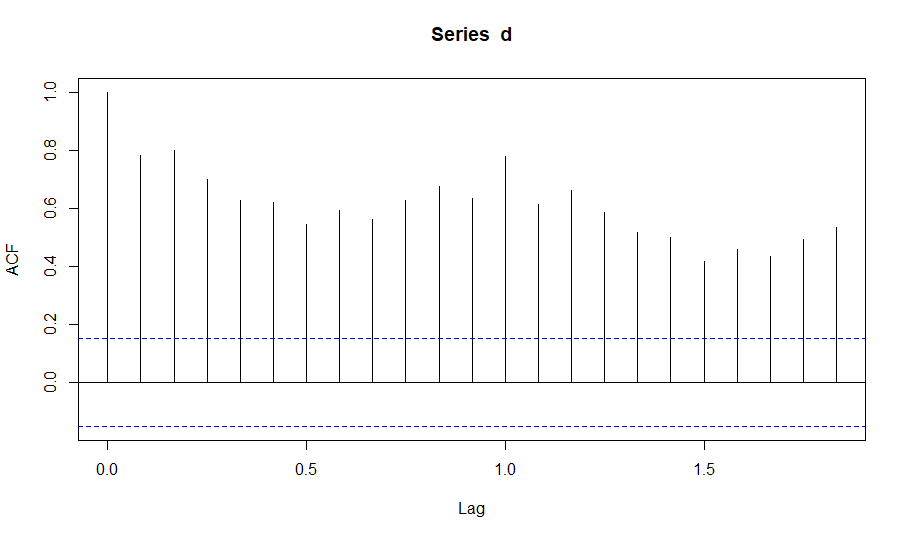
> View(Births)

> d<-ts(Births,start=c(1946,1),frequency=12)

> plot.ts(d)

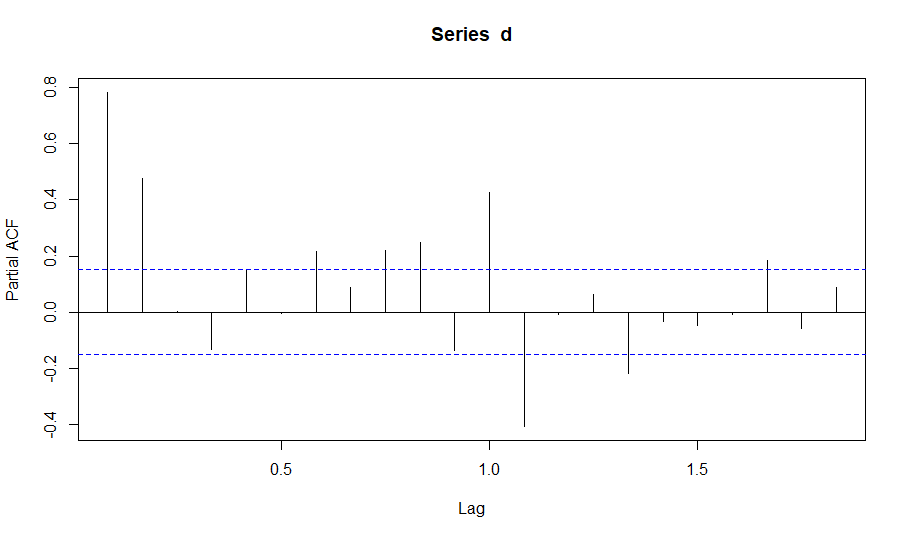


> acf(d ,type = "correlation")



From the above ACF plot we can conclude that as most of the observation lie outside the limit the process is not stationary.

> acf(d,type = "partial")

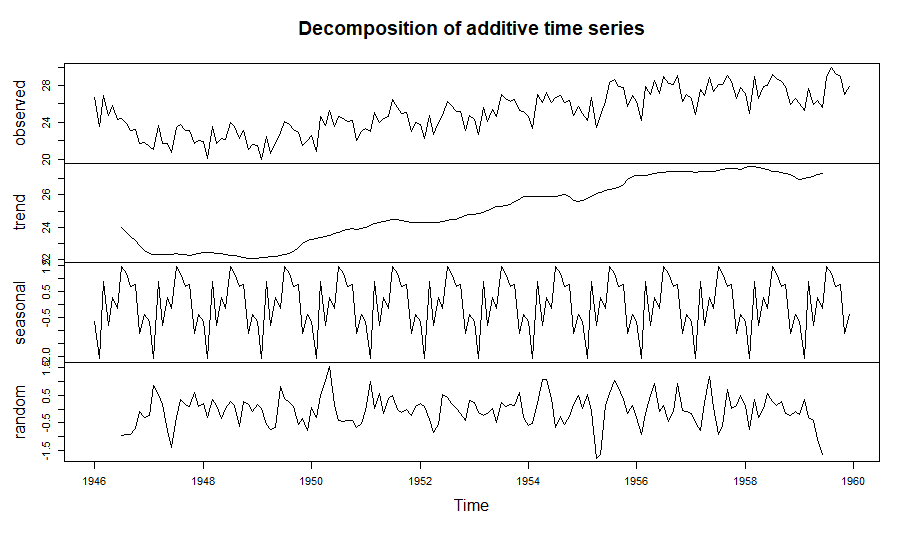


From the above ACF and PACF plots it can be observed the data contains both trend and seasonality. It can also be observed that the pattern of the data is of decreasing trend which concludes that ARIMA model is appropriate for the given data.

> dcomp=decompose(d)

> dcomp$seasonal

> plot(dcomp)



> Af=forecast.auto.arima(d)

> Af

Series: d

ARIMA(2,1,2)(1,1,1)[12]

Coefficients:

ar1 ar2 ma1 ma2 sar1 sma1

0.6539 -0.4540 -0.7255 0.2532 -0.2427 -0.8451

s.e. 0.3004 0.2429 0.3228 0.2879 0.0985 0.0995

sigma^2 estimated as 0.4076: log likelihood=-157.45

AIC=328.91 AICc=329.67 BIC=350.21

> Ad=tseries.adf.test(d,alternative="stationary")

> Ad

Augmented Dickey-Fuller Test

data: d

Dickey-Fuller = -5.9547, Lag order = 5, p-value = 0.01

alternative hypothesis: stationary

H0: Process is not stationary

V/s H1: Process is stationary

Here p-value > 0.01 Hence accept H0 and conclude that process is not stationary.

After removing the seasonality component from the data :

> dadj=d-dcomp$seasonal

> af=forecast::auto.arima(dadj)

> af

Series: dadj

ARIMA(2,1,1)(1,0,2)[12] with drift

Coefficients:

ar1 ar2 ma1 sar1 sma1 sma2 drift

0.3894 -0.2033 -0.4471 -0.1050 -0.2195 -0.1200 0.0192

s.e. 0.2227 0.0843 0.2248 0.6952 0.6989 0.2087 0.0195

sigma^2 estimated as 0.336: log likelihood=-143.28

AIC=302.56 AICc=303.47 BIC=327.5

> fit=HoltWinters(d)

> fit

Holt-Winters exponential smoothing with trend and additive seasonal component.

Call:

HoltWinters(x = d)

Smoothing parameters:

alpha: 0.4823655

beta : 0.02988495

gamma: 0.563186

Coefficients:

[,1]

a 28.04366357

b 0.04199921

s1 -0.78546221

s2 -2.19944507

s3 0.87813012

s4 -0.65164728

s5 0.63427267

s6 0.21182821

s7 2.23177191

s8 2.17167733

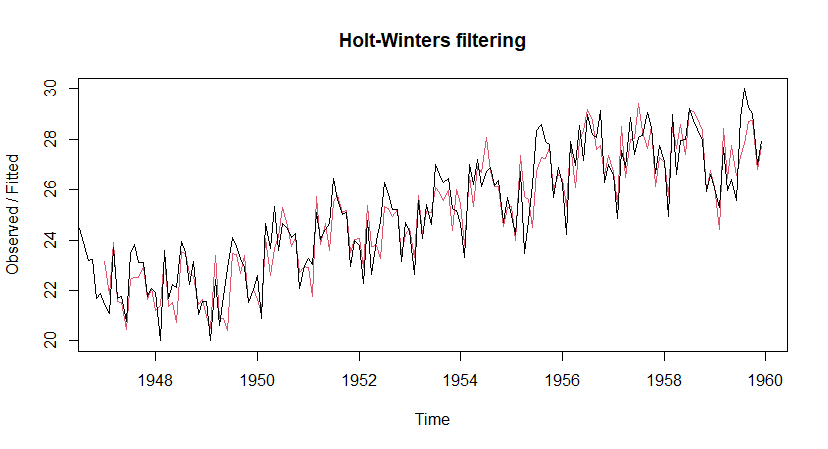
s9 1.52077678

s10 1.16900861

s11 -0.97500043

s12 -0.18636055

> plot(s\_fit)



> aa=forecast::forecast(fit,h=24)

> aa

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

Jan 1960 27.30020 26.32517 28.27523 25.80902 28.79139

Feb 1960 25.92822 24.83950 27.01694 24.26316 27.59327

Mar 1960 29.04779 27.85040 30.24518 27.21654 30.87905

Apr 1960 27.56001 26.25756 28.86247 25.56808 29.55195

May 1960 28.88793 27.48307 30.29280 26.73938 31.03649

Jun 1960 28.50749 27.00220 30.01278 26.20535 30.80963

Jul 1960 30.56943 28.96521 32.17365 28.11598 33.02288

Aug 1960 30.55133 28.84929 32.25338 27.94828 33.15439

Sep 1960 29.94243 28.14338 31.74148 27.19102 32.69384

Oct 1960 29.63266 27.73720 31.52813 26.73380 32.53153

Nov 1960 27.53065 25.53918 29.52213 24.48496 30.57635

Dec 1960 28.36129 26.27407 30.44852 25.16916 31.55342

Jan 1961 27.80419 25.52190 30.08648 24.31373 31.29466

Feb 1961 26.43221 24.05832 28.80609 22.80166 30.06275

Mar 1961 29.55178 27.08594 32.01762 25.78061 33.32295

Apr 1961 28.06400 25.50582 30.62219 24.15159 31.97641

May 1961 29.39192 26.74095 32.04289 25.33761 33.44623

Jun 1961 29.01148 26.26726 31.75569 24.81457 33.20839

Jul 1961 31.07342 28.23548 33.91136 26.73317 35.41367

Aug 1961 31.05533 28.12315 33.98750 26.57096 35.53969

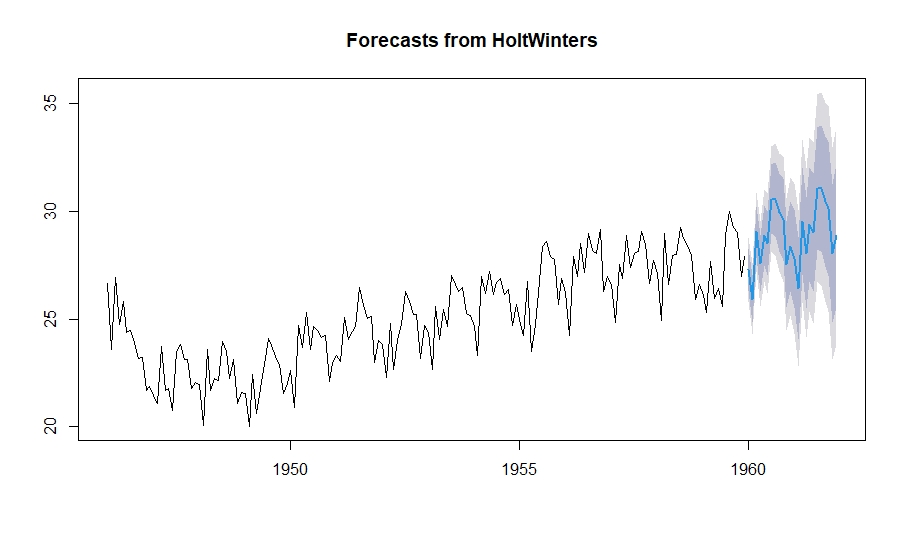
Sep 1961 30.44642 27.41950 33.47335 25.81714 35.07571

Oct 1961 30.13665 27.01444 33.25887 25.36163 34.91168

Nov 1961 28.03464 24.81658 31.25271 23.11304 32.95625

Dec 1961 28.86528 25.55081 32.17976 23.79623 33.93434

> plot(aa)



Here the forecasts for the year 1961 are plotted as a blue line, the 80% prediction interval as a blue shaded area, and the 95% prediction interval as a grey shaded area.