**Meenakshi Nagarajan**

**Assignment 2**

**STT4110/6110**

1. Suppose Find *Var* () . Note any unusual results. In particular,

compare your answer to what would have been obtained if . (Hint: You

may avoid Equation (3.2.3) on page 28 by first doing some algebraic simplification on

*Solution:*

Let

Consider if

Variance obtained from is higher than that obtained from

1. The data file beer sales contains monthly U.S. beer sales (in millions of barrels) for the period January 1975 through December 1990.

**(a)**Display and interpret the plot the time series plot for these data.

**(b)**Now construct a time series plot that uses separate plotting symbols for the   
various months. Does your interpretation change from that in part (a)?

**(c)**Use least squares to fit a seasonal-means trend to this time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

**(d)**Construct and interpret the time series plot of the standardized residuals from   
part (c). Be sure to use proper plotting symbols to check on seasonality in the   
standardized residuals.

**(e)**Use least squares to fit a seasonal-means plus quadratic time trend to the beer   
sales time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

**(f)**Construct and interpret the time series plot of the standardized residuals from   
part

**(g)** Again, use proper plotting symbols to check for any remaining seasonality in the residuals.

*Solution:*

library(TSA)

## Loading required package: leaps

## Loading required package: locfit

## locfit 1.5-9.1 2013-03-22

## Loading required package: mgcv

## Loading required package: nlme

## This is mgcv 1.8-17. For overview type 'help("mgcv-package")'.

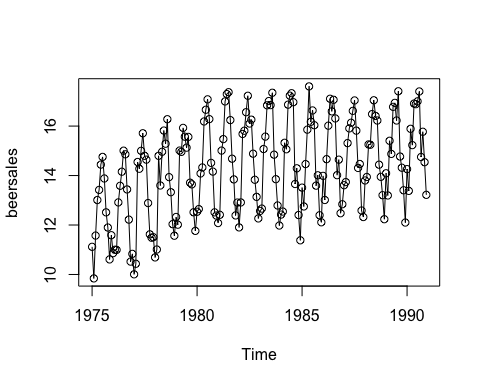
## Loading required package: tseries

##   
## Attaching package: 'TSA'

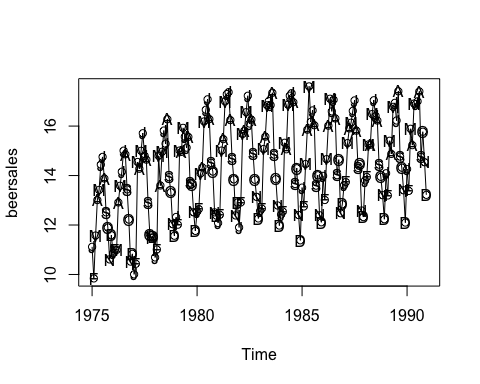
## The following objects are masked from 'package:stats':  
##   
## acf, arima

## The following object is masked from 'package:utils':  
##   
## tar

data("beersales")  
#Time-series plot  
plot.ts(beersales, type='o')



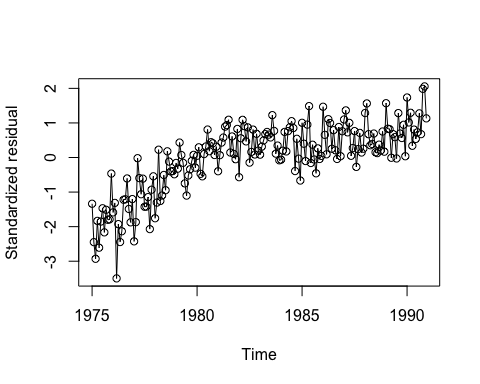
#Separate plotting symbols for various months  
plot.new()  
plot.ts(beersales, type='o')  
points(y = beersales, x= time(beersales), pch = as.vector(season(beersales)))



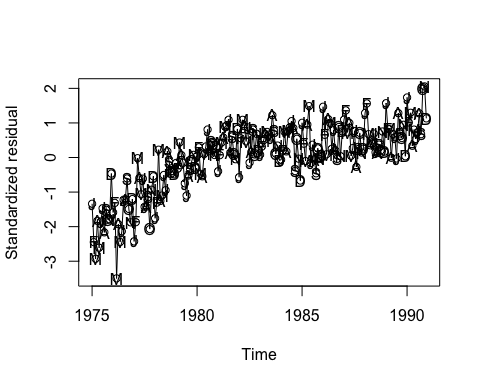
#least square fit  
mon=season(beersales)  
beersalesdata.lm=lm(beersales~mon)  
summary(beersalesdata.lm)

##   
## Call:  
## lm(formula = beersales ~ mon)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.5745 -0.4772 0.1759 0.7312 2.1023   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 12.48568 0.26392 47.309 < 2e-16 \*\*\*  
## monFebruary -0.14259 0.37324 -0.382 0.702879   
## monMarch 2.08219 0.37324 5.579 8.77e-08 \*\*\*  
## monApril 2.39760 0.37324 6.424 1.15e-09 \*\*\*  
## monMay 3.59896 0.37324 9.643 < 2e-16 \*\*\*  
## monJune 3.84976 0.37324 10.314 < 2e-16 \*\*\*  
## monJuly 3.76866 0.37324 10.097 < 2e-16 \*\*\*  
## monAugust 3.60877 0.37324 9.669 < 2e-16 \*\*\*  
## monSeptember 1.57282 0.37324 4.214 3.96e-05 \*\*\*  
## monOctober 1.25444 0.37324 3.361 0.000948 \*\*\*  
## monNovember -0.04797 0.37324 -0.129 0.897881   
## monDecember -0.42309 0.37324 -1.134 0.258487   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.056 on 180 degrees of freedom  
## Multiple R-squared: 0.7103, Adjusted R-squared: 0.6926   
## F-statistic: 40.12 on 11 and 180 DF, p-value: < 2.2e-16

#time series plot of residuals  
plot(y=rstandard(beersalesdata.lm), x=as.vector(time(beersales)), type = 'o',xlab="Time",ylab="Standardized residual")



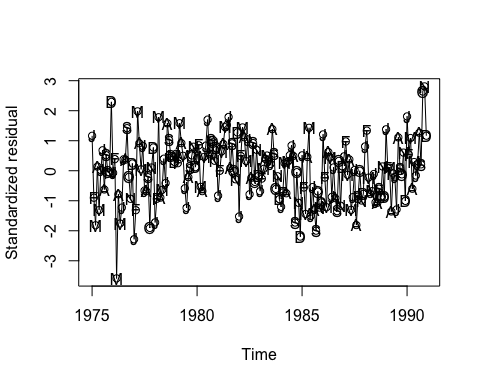
#separating seasonal trend with symbols  
plot.new()  
plot(y=rstandard(beersalesdata.lm), x=as.vector(time(beersales)), type = 'o',xlab="Time",ylab="Standardized residual")  
points(y = rstandard(beersalesdata.lm), x= as.vector(time(beersales)), pch = as.vector(season(beersales)),xlab="Time",ylab="Standardized residual")



#Quadratic model  
beersalesdata.qm = lm(beersales ~ mon + I(time(beersales)^2)+time(beersales))  
summary(beersalesdata.qm)

##   
## Call:  
## lm(formula = beersales ~ mon + I(time(beersales)^2) + time(beersales))  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.03203 -0.43118 0.04977 0.34509 1.57572   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -7.150e+04 8.791e+03 -8.133 6.93e-14 \*\*\*  
## monFebruary -1.579e-01 2.090e-01 -0.755 0.45099   
## monMarch 2.052e+00 2.090e-01 9.818 < 2e-16 \*\*\*  
## monApril 2.353e+00 2.090e-01 11.256 < 2e-16 \*\*\*  
## monMay 3.539e+00 2.090e-01 16.934 < 2e-16 \*\*\*  
## monJune 3.776e+00 2.090e-01 18.065 < 2e-16 \*\*\*  
## monJuly 3.681e+00 2.090e-01 17.608 < 2e-16 \*\*\*  
## monAugust 3.507e+00 2.091e-01 16.776 < 2e-16 \*\*\*  
## monSeptember 1.458e+00 2.091e-01 6.972 5.89e-11 \*\*\*  
## monOctober 1.126e+00 2.091e-01 5.385 2.27e-07 \*\*\*  
## monNovember -1.894e-01 2.091e-01 -0.906 0.36622   
## monDecember -5.773e-01 2.092e-01 -2.760 0.00638 \*\*   
## I(time(beersales)^2) -1.810e-02 2.236e-03 -8.096 8.63e-14 \*\*\*  
## time(beersales) 7.196e+01 8.867e+00 8.115 7.70e-14 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5911 on 178 degrees of freedom  
## Multiple R-squared: 0.9102, Adjusted R-squared: 0.9036   
## F-statistic: 138.8 on 13 and 178 DF, p-value: < 2.2e-16

#standardized residuals time series plot  
plot(y=rstandard(beersalesdata.qm), x= as.vector(time(beersales)), type='o',xlab="Time",ylab="Standardized residual")  
points(y = rstandard(beersalesdata.qm), x= as.vector(time(beersales)), pch = as.vector(season(beersales)),xlab="Time",ylab="Standardized residual")



1. (Continuation of Exercise 3.6) Consider the time series in the data file beer-sales.

(a)  Obtain the residuals from the least squares fit of the seasonal-means plus quadratic time trend model.   
(b)  Perform a runs test on the standardized residuals and interpret the results.   
(c)  Calculate and interpret the sample autocorrelations for the standardized residuals.   
(d)  Investigate the normality of the standardized residuals (error terms). Consider   
histograms and normal probability plots. Interpret the plots.

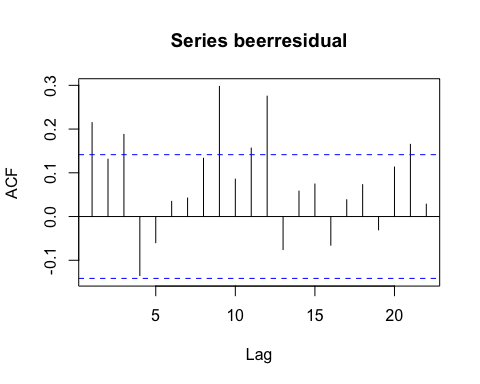
*Solution:*

beerresidual <- rstudent(beersalesdata.qm)  
runs(beerresidual)

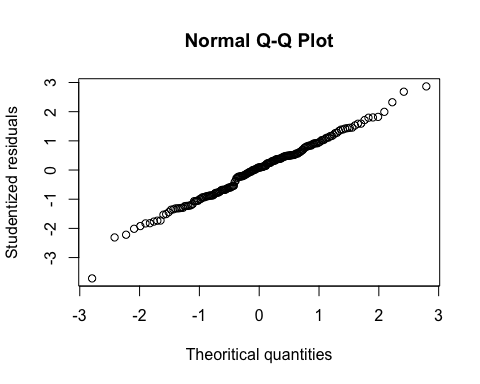
## $pvalue  
## [1] 0.0127  
##   
## $observed.runs  
## [1] 79  
##   
## $expected.runs  
## [1] 96.625  
##   
## $n1  
## [1] 90  
##   
## $n2  
## [1] 102  
##   
## $k  
## [1] 0

The test looks significant since p value is less than 0.05

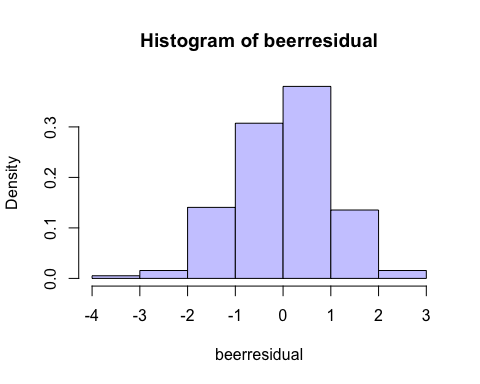
#autocorrelation function  
acf(beerresidual)

 Correlation looks significant for several lags

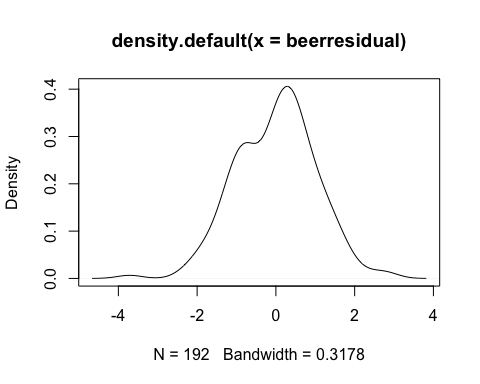
#normal qualtile plot  
qqnorm(beerresidual,xlab="Theoritical quantities",ylab="Studentized residuals")



#Residual vs. fitted  
hist(beerresidual, breaks=8, col="#CCCCFF", freq=FALSE)



plot(density(beerresidual))



1. Suppose that a stationary time series,, has an autocorrelation function of the form , where is a constant in the range
2. Show that

(Hint: Use Equation (3.2.3) on page 28)

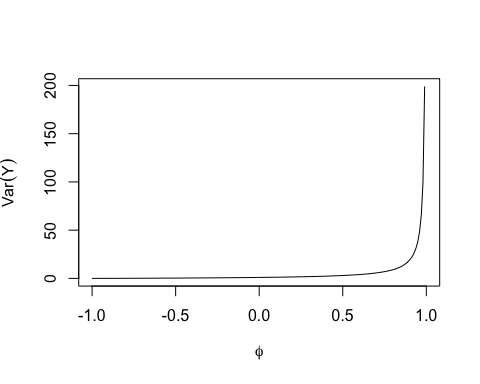
]

1. If n is large, argue that

1. Plot for φ over the range −1 to +1. Interpret the plot in terms of the precision in estimating the process mean.

#Rcode

phivalue<-seq(-1,1,0.01)  
varybar<-(1+phivalue)/(1-phivalue)  
plot(varybar ~ phivalue,ylab=expression(Var(bar(Y))),xlab=expression(phi),type="l")



1. Verify Equation (3.2.6) on page 29. (Hint: You will need the fact that

*Solution:*

]