**云南大学数学与统计学院**

**实验报告**

**实验课名称： 应用时间序列分析实验**

**指导教师： 周建军**

**专业（年级）： 2021级统计学**

**学生姓名： 枫叶 学号:**

**实验名称： 线性平稳时间序列建模**

**实验时间： 2024.5.4**

**实验成绩：**

1. **实验目的和要求：**

线性平稳时间序列建模

1. **实验内容和原理**

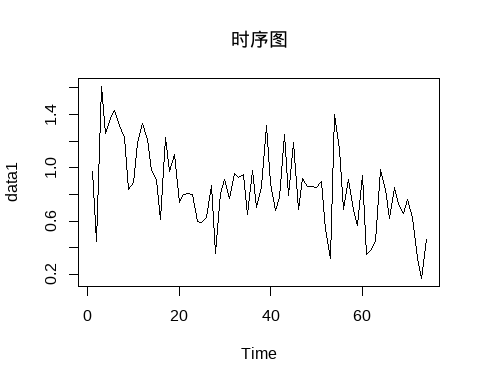
习题3.18数据，利用某地区连续74年谷物产量数据，完成：

1. 建立一个合理的时间序列模型；
2. 预测该地区未来5年的产量。
3. 习题3.19数据，连续201个生产记录数据，对该数据完成：
4. 建立一个合理的时间序列模型；
5. 预测未来2期的95%的置信区间。
6. **实验步骤及方法（包含具体的程序）**

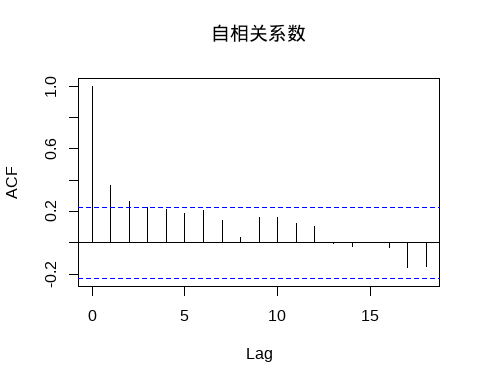
library(dplyr)  
library(forecast)  
library(ggplot2)  
library(modelsummary)

## 第一题

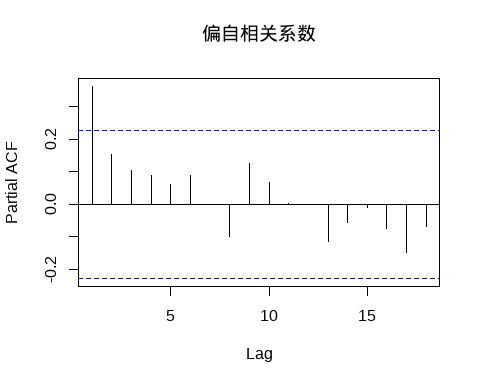
data1 <- read.table("D:/预删除文件夹/大三下/时间序列/习题3.18数据.txt",sep = "\t") %>%  
 as.matrix() %>%  
 t() %>%  
 as.vector() %>%  
 ts() %>%  
 na.omit()  
plot(data1,main="时序图")



acf(data1,main="自相关系数")



acf(data1,type = "partial",main="偏自相关系数")



Box.test(data1,lag = 5)

##   
## Box-Pierce test  
##   
## data: data1  
## X-squared = 24.8, df = 5, p-value = 0.0001523

从时序图来看，没有明显的趋势或周期性，从ACF图和PACF图来看，样本自相关系数呈现出一定的拖尾特征，偏自相关系数一阶截尾，为了稳健起见，先对数据作纯随机性检验，基于LB统计量的检验结果显示在0.05的显著性水平下可以认为不是纯随机序列（后文中统一以0.05为显著性水平，并不再给出说明），故尝试AR(1)模型进行拟合

model1 <- Arima(data1,order = c(1,0,0))  
modelsummary(list("AR1模型"=model1),gof\_map = c("rmse"))

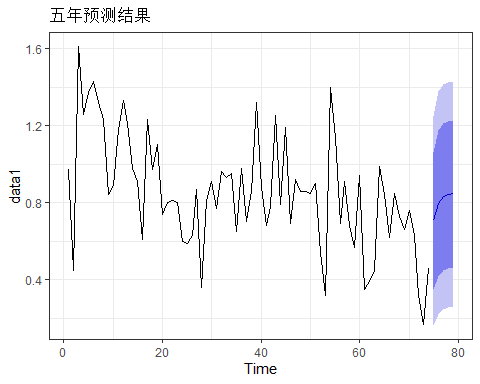
|  | AR1模型 |
| --- | --- |
| ar1 | 0.368 |
|  | (0.109) |
| intercept | 0.849 |
|  | (0.050) |
| RMSE | 0.27 |

Box.test(model1$residuals,type = "Ljung-Box")

##   
## Box-Ljung test  
##   
## data: model1$residuals  
## X-squared = 0.25563, df = 1, p-value = 0.6131

从系数估计值的置信区间来看，和都显著不为零，进一步对模型残差作纯随机检验，检验结果显示可以认为残差序列是纯随机的，这说明AR(1)模型较好地提取了序列信息

forecast(model1,5) %>%  
 autoplot() +  
 labs(title = "五年预测结果") +  
 theme\_bw()



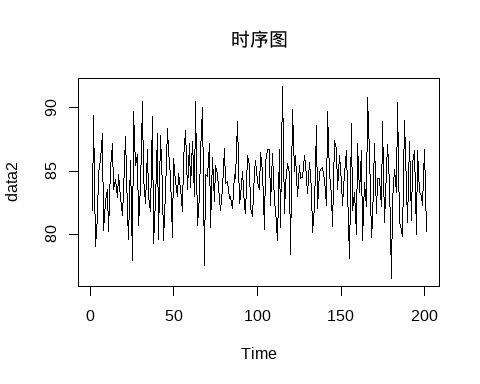
mean(data1)

## [1] 0.8512162

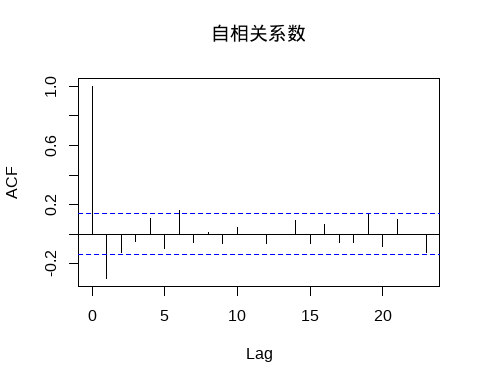
五年预测结果如上图所示，预测时间越长，可以看到AR(1)模型的预测结果越接近历年数据的均值

## 第二题

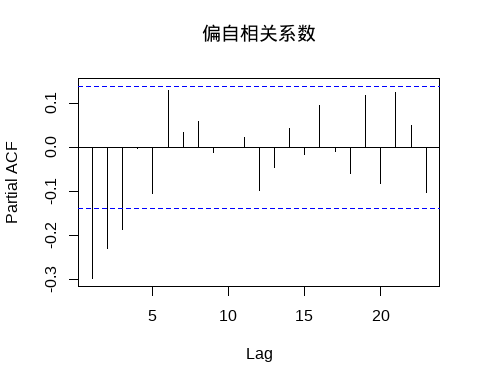
data2 <- read.table("D:/预删除文件夹/大三下/时间序列/习题3.19数据.txt",sep = "\t") %>%  
 as.matrix() %>%  
 t() %>%  
 as.vector() %>%  
 ts() %>%  
 na.omit()  
plot(data2,main="时序图")



acf(data2,main="自相关系数")



acf(data2,type = "partial",main="偏自相关系数")



Box.test(data2,lag = 5)

##   
## Box-Pierce test  
##   
## data: data2  
## X-squared = 25.392, df = 5, p-value = 0.000117

由于数据过多，时序图不太直观，从ACF图和PACF图中可以发现样本自相关系数一阶截尾，偏自相关系数呈现一定的拖尾特征，稳健起见同样地作纯随机性检验，结果显示显著不随机，下面尝试用MA(1)模型拟合

model2 <- Arima(data2,order = c(0,0,1))  
modelsummary(list("MA1模型"=model2),gof\_map = c("rmse"))

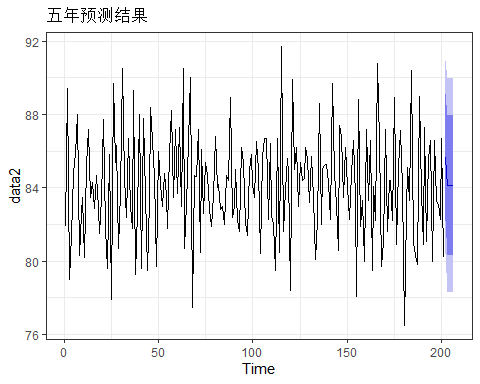
|  | MA1模型 |
| --- | --- |
| ma1 | -0.478 |
|  | (0.067) |
| intercept | 84.130 |
|  | (0.099) |
| RMSE | 2.68 |

Box.test(model2$residuals,lag = 5)

##   
## Box-Pierce test  
##   
## data: model2$residuals  
## X-squared = 5.3213, df = 5, p-value = 0.3779

从置信区间来看，和都显著不为零，对残差的纯随机检验显示可以认为残差是纯随机序列，说明MA(1)模型较好地提取了序列信息

forecast(model2,5) %>%  
 autoplot() +  
 labs(title = "五年预测结果") +  
 theme\_bw()



mean(data2)

## [1] 84.1194

五年预测结果如上图所示，在预测年份大于1之后，MA(1)模型的预测值直接为序列均值，本数据的样本均值约为84

1. **实验结果分析说明**

已附于图下