**时间序列分析实验报告书**

**班级：统计2001 姓名：张逸敏 实验日期：2023.4.18**

**实验一 ARMA模型的建模与预测**

**1实验目的**

R语言的基本操作，针对平稳序列建立ARMA模型并做预测。

**2实验条件**

PC机，R语言

**3实验原理**

**3.1 AR模型**

如果误差项是白噪声, 实数使得多项式的零点都在单位圆外:



则称阶差分方程



是一个阶自回归模型, 简称为模型.

满足 AR模型的平稳时间序列称为的平稳解, 也称作AR(p)序列.称是AR(p)模型的自回归系数. 称为模型的特征多项式。模型可用推移算子写成



**3.2 MA模型**

设是WN, 如果实数使得



则称



是阶滑动平均模型, 简称为MA模型。

模型中的显然是平稳列。称平稳序列是滑动平均序列, 简称为MA序列.

**3.3 ARMA模型**

设是WN, 实系数多项式和没有公共根, 满足和



就称差分方程：



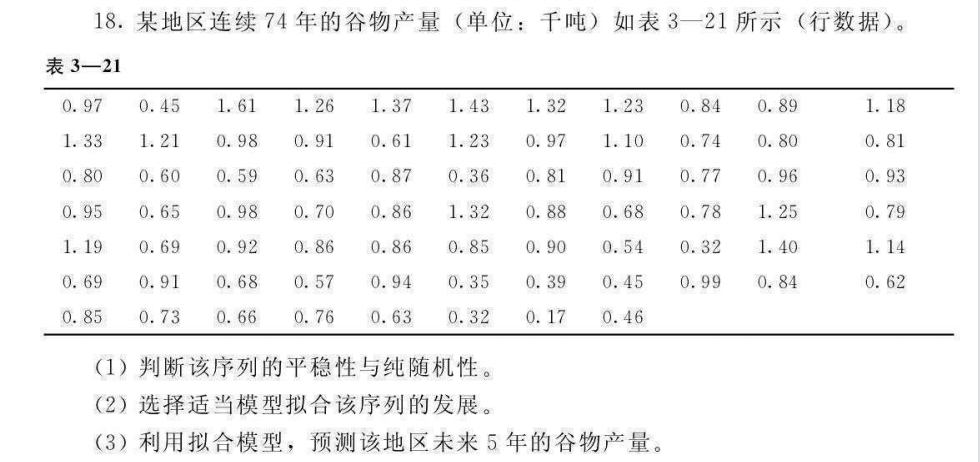
是一个自回归滑动平均模型, 简称为ARMA()模型. 称满足上式的平稳序列为平稳解或ARMA()序列.

模型写成



**4实验过程与结果**

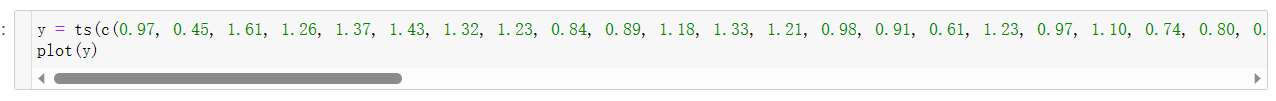
**4.1实验案例表述**



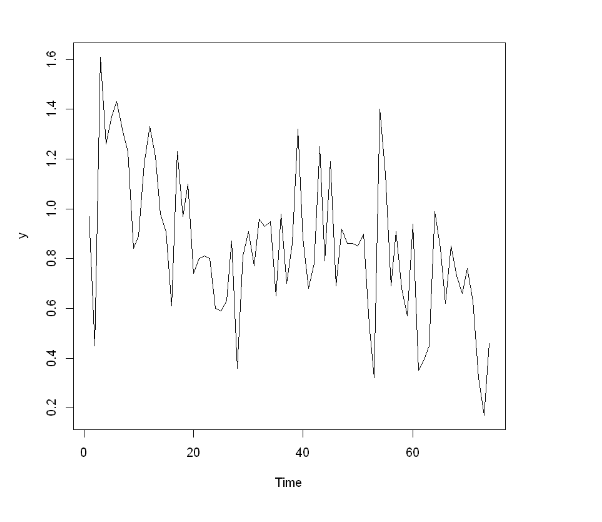
**4.2 实验过程与代码**

**4.2.1 判断的平稳性和纯随机性**

首先读入数据，创建时间序列对象。

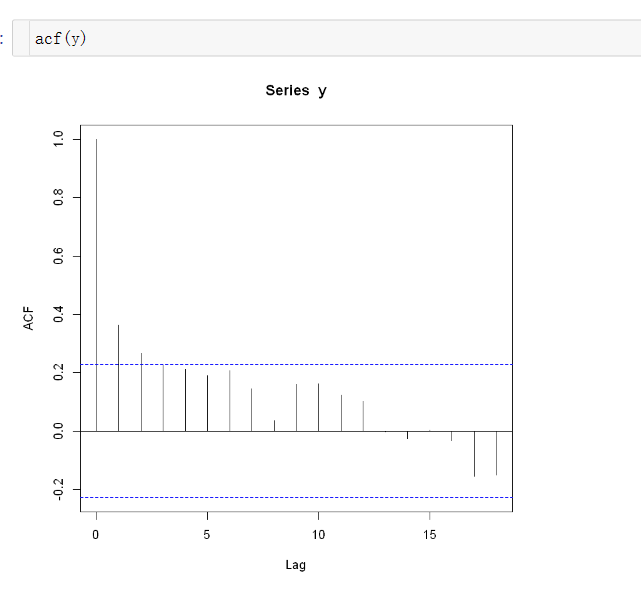


接下来绘制时序图。



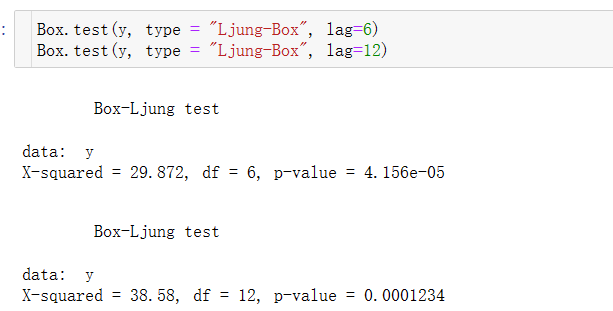
观察时序图，没有发现明显的趋势或者周期，基本上可以视为平稳序列。

为了稳妥起见，还需要进行自相关图检验。



自相关图显示大多数自相关系数在2倍标准差以外，说明序列具有短期相关性，是平稳序列。

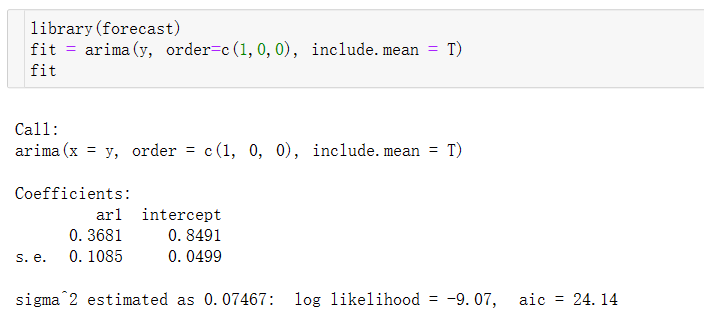
利用LB统计量进行纯随机性检验，结果如下。



延迟6期的LB统计量和延迟12期的LB统计量的p值均，表示延迟期数小于等于6、12的序列值之间存在相关性，时间序列不具有纯随机性，值得继续进行建模。

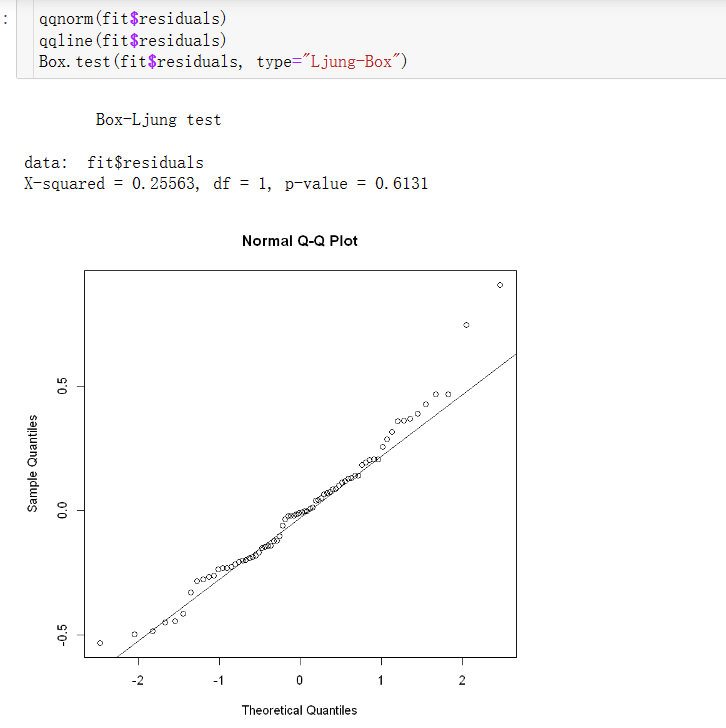
**4.2.2 选择适当模型拟合该序列的发展**

选择模型进行建模。



一般来说，一个模型如果合适，那模型的残差应该满足均值为0的正态分布，并且对于任意的滞后阶数，残差自相关系数都应该为零。换句话说，模型的残差应该满足独立正态分布（即残差间没有关联）。

下面进行对模型残差进行检验。

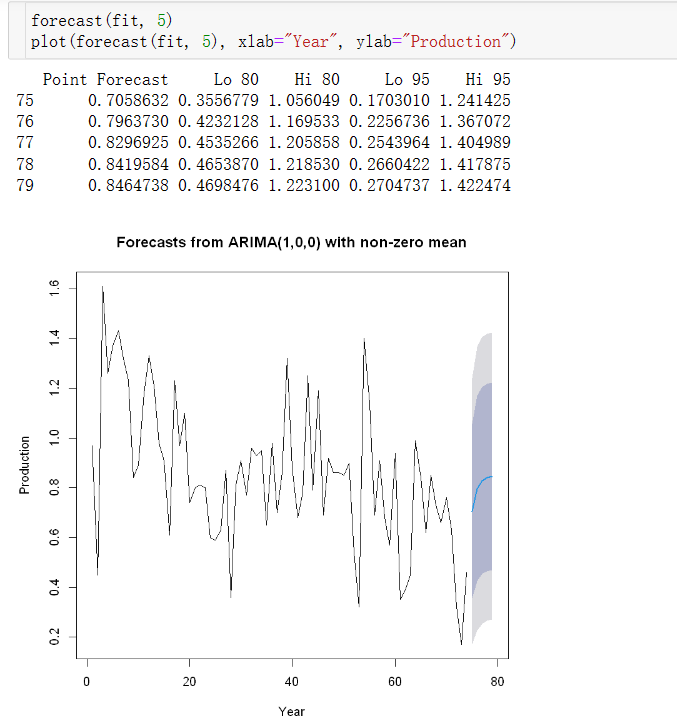


如果数据呈现正态分布，那么数据点会落在图中直线上。显然，本模型的残差大致符合正态分布，模型效果不错。

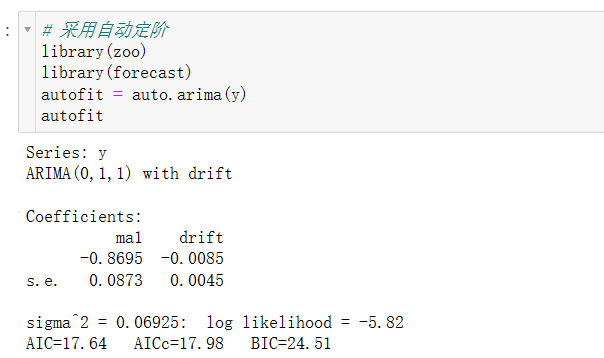
对残差的自相关检验，，接受原假设，即认为残差之间的自相关系数为0，符合残差独立的假设。

**4.2.3利用拟合模型预测该地区未来5年的谷物产量**

利用forecast()函数，并画出80%和95%的置信区间。

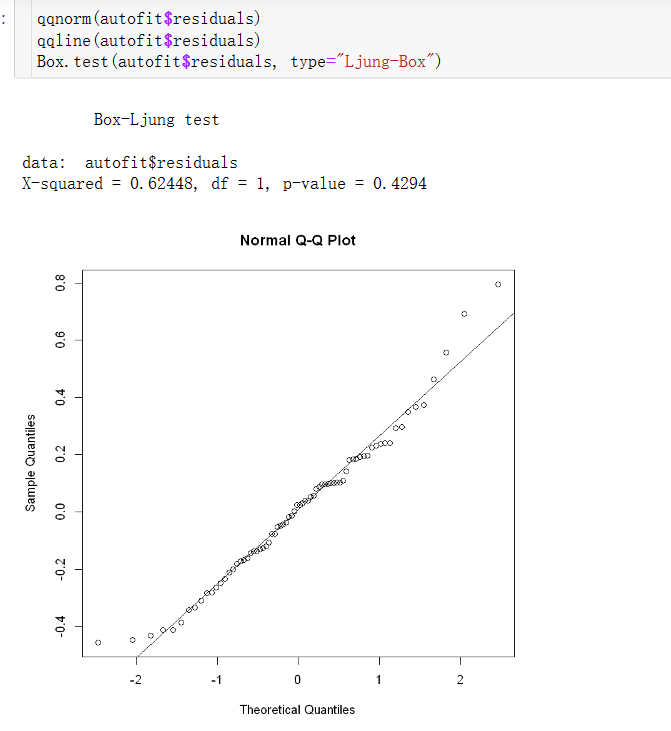


**4.2.4 采用自动定阶函数建立模型**



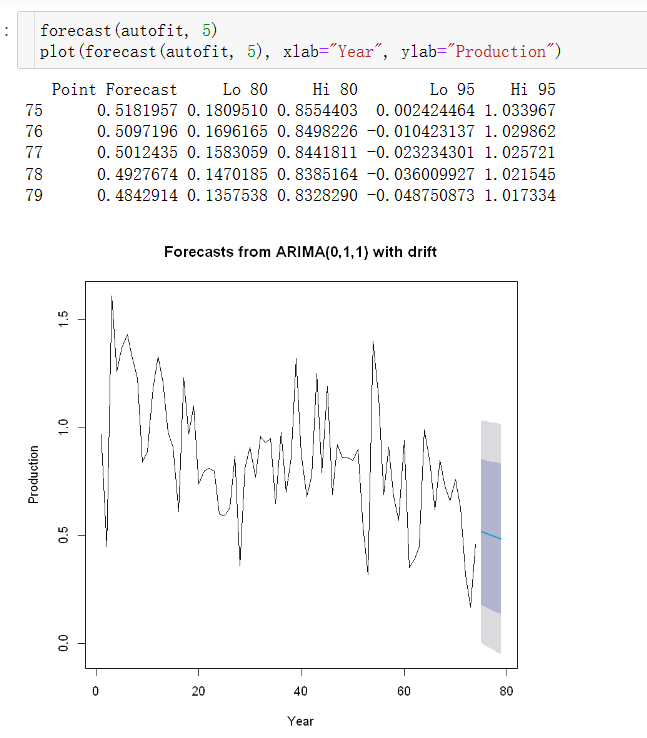
利用自动定阶函数auto.arima()建立模型，模型，小于模型的24.14，模型更优。

对ARIMA模型进行残差检验：



符合残差独立同分布于正态分布的假设。

模型预测结果如下：



**5 实验分析与总结**

**5.1实验分析**

（1）时序图显示该地区连续74年的谷物产量的波动是比较平稳的。

（2）自相关图检验，考察该样本序列的自相关图，可以发现延迟2阶以后，自相关系数都在两倍标准差范围左右，没有明显超出两倍标准差。可以看出，这是一个典型的短期相关的样本自相关图。由时序图和样本自相关图的性质，可以认为该序列平稳。

（3）纯随机性检验。利用LB统计量进行纯随机性检验，计算在6延迟阶数和12延迟阶数下的LB统计量的p值。在6和12延迟阶数下的LB统计量的p值都为0()，所以有大于的置信水平认为该地区连续74年的谷物产量属于非白噪声序列。

总结（1）、（2）、（3）的结果可以发现，这个序列是一列平稳的、蕴含相关信息的序列，值得去分析和研究。

**5.2 实验总结**

（1）建立ARIMA模型的步骤

1.1确保时序是平稳的

1.2 找到一个（或几个）合理的模型（即选定可能的*p*值和*q*值）

1.3拟合模型

1.4从统计假设和预测准确性等角度评估模型

1.5预测

（2）判断不同模型效果的准则

本文采用AIC准则来评价不同模型的效果好坏。AIC是衡量统计模型拟合优良性的一种标准，由于它为[日本](https://baike.baidu.com/item/%E6%97%A5%E6%9C%AC/111617?fromModule=lemma_inlink)[统计学](https://baike.baidu.com/item/%E7%BB%9F%E8%AE%A1%E5%AD%A6/1175?fromModule=lemma_inlink)家赤池弘次创立和发展的，因此又称[赤池信息量准则](https://baike.baidu.com/item/%E8%B5%A4%E6%B1%A0%E4%BF%A1%E6%81%AF%E9%87%8F%E5%87%86%E5%88%99/130462?fromModule=lemma_inlink)。它建立在熵的概念基础上，可以权衡所估计模型的复杂度和此模型拟合数据的优良性。

AIC的假设条件是模型的误差服从独立正态分布，其公式为



其中k是参数的数量n为样本数，SSR为残差平方和。