北京航空航天大学

2021－2022 学年 第二学期期末

《R语言编程及应用》

考 试 A 卷

任课教师：马 杰

班 级\_ \_ \_学 号 \_\_

姓 名\_\_ \_ \_成 绩 \_\_\_\_\_\_\_\_\_

考试日期:2022年6月15日

班号 学号 姓名 成绩

《R语言编程及应用》期末考试卷

注意事项：

1、因疫情原因，本次考试为在线开卷考试，允许看书、看课件、看参考资料，也允许调用或改编课堂上讲过的代码或函数，答题时请“标清题号”。

2、绝不允许相互抄袭，一旦发现有雷同卷、文字/用词/举例等雷同作弊情形，抄袭与提供抄袭答案的同学，都记为作弊0分。

试题：

一、数据分析与编程应用题………………………………………… (5小题，共100分)

**中国股票市场指数的实证数据分析**

股票市场是中国金融市场的核心组成部分，也是机构投资者与个人投资者展开投资行为的主战场。现搜集了上证综指（股票代码：000001）、深圳成指（股票代码：399001）、创业板指（股票代码：399006）等三只主要股票市场指数的交易数据，分别存于文件“000001.csv、399001.csv、399006.csv”中，请针对这些数据，在线编程展开数据分析、并计算解决以下问题（提示：1）所有分析与计算，均应写出相应的R代码、编程解决，其他操作无效；2）输出的表格或图表，请直接使用R计算结果的截图，必要时按要求辅之以“简洁的文字说明”）：

1．读取三只股票指数的交易数据：

1）以创业交易有效数据的最大时段为准，截取上综指、深成指交易日期，使得三只指数的交易时间可匹配、均有交易数据，然后输出展示上综指的前6行交易数据信息；

1. 在此基础上，新创建一个数据框，其中仅包括交易日期信息、以及三只指数的收盘价与成交量信息，然后输出任意给定的时段内（如本题设定为：2011.4.1~ 2022.3.31）三只指数的后6行交易信息。[提示：编程中的变量名，可自行命名，以“简洁易识别”为原则]

(25分)

1. 在股票市场上，存在着广泛的假说或猜想，认为存在“红周一、黑周五”现象，即周一易上涨、周五易下跌。请以给定时段2011.4.1~ 2022.3.31内的上综指的数据信息为例，计算并输出相应的对数收益率的描述性统计结果（至少应包括均值、方差、偏度、峰度等数值特征），简单验证下“红周一、黑周五”现象是否存在。

(15分)

1. 仍以给定时段2011.4.1~ 2022.3.31内三只股票指数的交易数据为对象，先测算输出三种股指对数收益率的相关性矩阵；然后考虑交互影响与价量结合，以上综指与深成指的收益率、及创业板成交量的对数变化率作为自变量，分析它们对创业板指数收益率的量化影响。

(15分)

4．以给定时段2011.4.1~ 2022.3.31内创业板指数的收盘价数据为基础：

1）将之转化为月度信息，再计算月百分比对数收益、并输出其时序图；

2）对月百分比对数收益序列进行Arch效应检验，采用最基础的Garch(1,1)模型对其建模，并分析建模的效果。

(20分)

5. 股市技术分析中，BOLL线是约翰.布林基于标准差原理设计的常用指标，包含有中轨、上轨、下轨三条线，其计算原理可简单描述为：

其中为时刻之前的个交易日收盘价的标准差。

但通常交易软件中，经验参数只能设为整数（通常为2），这并不完全符合正态分布假定下的分位数值；假定未来股价变化服从正态分布，保险起见，要求BOLL线上下轨以99%的概率动态包括未来股价波动区间，此时参数的理论值应取2.57583较为精准。为此，请自行编写函数Boll(x, date, N, alpha)，以绘制给定时段内的“股价及其布林线走势图”——其中，x为收盘价，date为交易日期，N为移动平均的窗宽，alpha为给定置信水平下的分位数。

新设定时段“2022.1.4~2022.3.31”，利用自编的函数与新给定时段的样本数据，输出创业板指日收盘价的股价及其布林线走势图（要求的置信水平为99%，N=20）。输出的结果中，因交易日期信息字段较长，要求日期刻度标签与横坐标垂直、并合理布局画布结构，以便能清楚标示每个交易日期。

(25分)