【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于Python的说法错误的是（ ）。

【A.】Python是一种面向对象的解释型高级编程语言

【B.】可移植性和可扩展性不强

【C.】结构简单，语法和代码定义清晰明确

【D.】易于学习和维护

【答案】B

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于Python中的注释说法正确的是（ ）。

【A.】Python中可以使用“#”作为注释，“#”右边的一切内容均不会被执行

【B.】Python中可以使用“#”进行一行或多行注释

【C.】Python中注释下的一切内容会被执行但不显示

【D.】Python中只有使用“#”这一种注释方法

【答案】A

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

如需在Python中安装第三方工具库或包(packages)，可以在终端使用的命令是（ ）。

【A.】install (package的名称)

【B.】install pip (package的名称)

【C.】pip install (package的名称)

【D.】pip (package的名称)

【答案】C

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于统计和数据挖掘，下列说法错误的是（ ）。

【A.】统计侧重于从样本到总体的推断

【B.】数据挖掘侧重于预测个体记录

【C.】经典统计中强调的推断在数据挖掘中已不存在

【D.】经典统计中强调的推断在数据挖掘中依然存在

【答案】D

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于Python的说法正确的是（ ）。

【A.】Python是一种面向对象的解释型高级编程语言

【B.】可移植性和可扩展性不强

【C.】结构简单，语法和代码定义清晰明确

【D.】易于学习和维护

【答案】ACD

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于统计和数据挖掘，下列说法正确的是（ ）。

【A.】统计侧重于从样本到总体的推断

【B.】数据挖掘侧重于预测个体记录

【C.】经典统计中强调的推断在数据挖掘中已不存在

【D.】经典统计中强调的推断在数据挖掘中依然存在

【答案】ABC

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

统计和数据挖掘的主要区别是（ ）。

【A.】统计侧重于从样本到总体的推断

【B.】统计侧重于预测个体记录

【C.】数据挖掘侧重于从样本到总体的推断

【D.】数据挖掘侧重于预测个体记录

【答案】AD

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于Python的第三方库的是（ ）。

【A.】Statsmodels

【B.】Matplotlib

【C.】Scipy

【D.】Scikit-learn

【答案】ABCD

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于Python中的注释说法错误的是（ ）。

【A.】Python中可以使用“#”作为注释，“#”右边的一切内容均不会被执行

【B.】Python中可以使用“#”进行一行或多行注释

【C.】Python中注释下的一切内容会被执行但不显示

【D.】Python中只有使用“#”这一种注释方法

【答案】BCD

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Python中可以使用“#”作为注释，“#”右边的一切内容均不会被执行。

【A.】√

【B.】×

【答案】A

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Python是一种面向对象的解释型高级编程语言，其结构简单，易于学习和维护，但可移植性和可扩展性较差。

【A.】√

【B.】×

【答案】B

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Python中的注释会被执行但不显示结果。

【A.】√

【B.】×

【答案】B

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

与统计学相比，数据挖掘以开放式方式处理大型数据集，这使得其不可能像推断所需的那样对所分析问题施以严格的限制。

【A.】√

【B.】×

【答案】A

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Python中只能使用“#”作为注释，并且可以对多行注释。

【A.】√

【B.】×

【答案】B

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

经典统计中强调的推断在数据挖掘中依然存在。

【A.】√

【B.】×

【答案】B

【知识点】04399.1.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在Python基本库中，最常见也是最主要的数据结构类型是（ ）。

【A.】表格

【B.】文字

【C.】数组

【D.】序列

【答案】D

【知识点】04399.1.1.2

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Python中常见的序列数据结构类型有（ ）。

【A.】数字

【B.】列表

【C.】元组

【D.】字典

【答案】ABCD

【知识点】04399.1.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在Python基本库中，序列是最常见也是最主要的数据结构类型。

【A.】√

【B.】×

【答案】A

【知识点】04399.1.1.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Python可以使用import语句导入已经安装在系统中的package。

【A.】√

【B.】×

【答案】A

【知识点】04399.1.1.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在调用statsmodels进行数据分析时，往往采用数据分析接口(api)的方式进行。

【A.】√

【B.】×

【答案】A

【知识点】04399.1.1.3

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列操作不属于数据分析和挖掘的一般流程的是（ ）。

【A.】数据预处理

【B.】探索性分析

【C.】从平台或渠道获取数据

【D.】将模型结果进行可视化

【答案】C

【知识点】04399.1.3.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列操作属于数据分析和挖掘的一般流程的是（ ）。

【A.】数据预处理

【B.】探索性分析

【C.】从平台或渠道获取数据

【D.】将模型结果进行可视化

【答案】ABD

【知识点】04399.1.3.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据挖掘的对象是各种各样的数据，针对不同的数据其分析方法可能是不同的。但这并不是说数据挖掘没有章法，其是有一般流程的。

【A.】√

【B.】×

【答案】A

【知识点】04399.1.3.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据挖掘的对象是各种各样的数据，针对不同的数据其分析方法可能是不同的，所以说数据挖掘没有章法。

【A.】√

【B.】×

【答案】B

【知识点】04399.1.3.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据预处理是数据分析工作的（ ）。

【A.】实施阶段

【B.】准备阶段

【C.】总结阶段

【D.】收尾阶段

【答案】B

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据预处理涉及数据整理和整合的各个方面，不包括（ ）。

【A.】数据清洗

【B.】数据筛选

【C.】数据变换

【D.】数据可视化

【答案】D

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

除了填补缺失值，另一个处理缺失值的常用方法就是删除缺失值。这可以通过函数（ ）实现。

【A.】dropna()

【B.】deletena()

【C.】clearna()

【D.】throwna()

【答案】A

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

删除缺失值的函数dropna()，默认是删除行，如果要对列进行删除，则需要添加参数（ ）。

【A.】axis=null

【B.】axis=-1

【C.】axis=0

【D.】axis=1

【答案】D

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

如果想把原始数据中的问号转成Python中缺失值的表示方式NaN，可以输入的语句是（ ）。import numpy as np

【A.】data=data.replace(?,np.NaN)

【B.】data=data.replace('?',np.NaN)

【C.】data=data.replace('?',NaN)

【D.】data=data.replace('?','np.NaN')

【答案】B

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于函数dropna()的说法不正确的是（ ）。

【A.】作用是删除缺失值

【B.】默认是删除行

【C.】默认是删除列

【D.】作用于数据框来实现

【答案】C

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于缺失值的产生原因的是（ ）。

【A.】数据本身为异常值

【B.】某些原因没有收集到信息

【C.】对于这些个体来说这些属性是不可用的

【D.】数据重复出现

【答案】AD

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于函数dropna()的说法正确的是（ ）。

【A.】作用是删除缺失值

【B.】默认是删除行

【C.】默认是删除列

【D.】作用于数据框来实现

【答案】ABD

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

一般数据的预处理，主要包括（ ）。

【A.】异常点检测

【B.】噪声数据检测

【C.】缺失值的处理

【D.】重复数据的处理

【答案】ABCD

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据预处理涉及数据整理和整合的各个方面，主要包括（ ）。

【A.】数据清洗

【B.】数据筛选

【C.】数据变换

【D.】数据可视化

【答案】ABC

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

缺失值的产生，有可能是因为（ ）。

【A.】数据本身为异常值

【B.】某些原因没有收集到信息

【C.】对于这些个体来说这些属性是不可用的

【D.】数据重复出现

【答案】BC

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

缺失值的产生有可能是因为某些原因没有收集到信息，也有可能是对于这些个体来说这些属性是不可用的。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

dropna()函数默认是删除列，要对行进行删除，则需要添加参数axis=1。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据预处理是数据分析工作的实施阶段。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在某些情况下，有些时候并不一定会收集到数据，因而会造成观测值或变量的数据有缺失，这些缺失的数据就称为缺失值。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

除了填补缺失值，另一个处理缺失值的常用方法就是删除缺失值。这可以通过函数dropna()作用于数据框来实现。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据预处理是数据分析工作的准备阶段。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果数据质量不高，则会影响数据挖掘的结果。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

dropna()函数默认是删除行，要对列进行删除，则需要添加参数axis=1。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

dropna()函数默认是删除行，要对列进行删除，则需要添加参数axis=0。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据预处理涉及数据整理和整合的各个方面，包括数据清洗、筛选和数据变换，用于数据挖掘前的模型选择等，有助于提高数据挖掘的效率。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面哪项代码是对数据data进行标准化的过程（ ）。

【A.】Z = (data - data.mean()) / data.std()

【B.】Z = (data.mean() - data) / data.std()

【C.】Z = (data - data.std()) / data.mean()

【D.】Z = (data.std() - data) / data.mean()

【答案】A

【知识点】04399.2.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列代码可以将数据data转换成数值型数据的是（ ）。

【A.】pd.to\_numeric(data)

【B.】pd.to\_num(data)

【C.】pd.to\_number(data)

【D.】pd.to\_string(data)

【答案】B

【知识点】04399.2.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

缺失值会受到标准化过程影响。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

异常点通常是指处于数据边界和数据集中的其他点有着不同模式的点。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

为了去除异常值，我们可以对每个变量计算Z-得分，然后去除那些Z-得分异常高或者异常低的个体。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

缺失值是不受标准化过程影响的。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列代码可以将数据集data中的重复个体检测出的是（ ）。

【A.】data.repeat()

【B.】data.repetition()

【C.】data.duplicated()

【D.】data.duplication()

【答案】C

【知识点】04399.2.1.3

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于函数duplicated()的说法不正确的是（ ）。

【A.】函数结果会返回数据中的所有重复值

【B.】函数结果会返回一个布尔数组

【C.】函数结果表明每一行是否与前面行有重复

【D.】函数结果会返回数据中的所有非重复值

【答案】AD

【知识点】04399.2.1.3

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于函数duplicated()的说法正确的是（ ）。

【A.】函数结果会返回数据中的所有重复值

【B.】函数结果会返回一个布尔数组

【C.】函数结果表明每一行是否与前面行有重复

【D.】函数结果会返回数据中的所有非重复值

【答案】BC

【知识点】04399.2.1.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

duplicated()函数会返回一个布尔数组，表明每一行是否与前面行有重复。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

duplicated()函数可以将数据中的重复值直接筛选出来。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.1.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

一些数据集是由合并数据源形成的，这就使其有可能含有重复数据或者近似重复的个体。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.1.3

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据整合是一个将两个或者多个对象的值合并成一个对象的预处理操作，数据整合不包括（ ）。

【A.】删除数据的异常值

【B.】减少接下来处理数据的大小

【C.】改变粒度分析（从精细到粗放尺度）

【D.】提高数据的解释能力

【答案】A

【知识点】04399.2.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据整合是一个将两个或者多个对象的值合并成一个对象的预处理操作，数据整合包括（ ）。

【A.】删除数据的异常值

【B.】减少接下来处理数据的大小

【C.】改变粒度分析（从精细到粗放尺度）

【D.】提高数据的解释能力

【答案】BCD

【知识点】04399.2.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据整合是一个将两个或者多个对象的值合并成一个对象的预处理操作。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于抽样的说法正确的是（ ）。

【A.】抽样有多种方法，如有放回抽样、无放回抽样等

【B.】有放回抽样则将抽出来的个体从抽样对象中去除

【C.】无放回抽样有可能对同一个个体抽中了多次

【D.】无放回抽样会有重复的个体

【答案】A

【知识点】04399.2.3.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于抽样的说法错误的是（ ）。

【A.】抽样有多种方法，如有放回抽样、无放回抽样等

【B.】无放回抽样则将抽出来的个体从抽样对象中去除

【C.】有放回抽样有可能对同一个个体抽中了多次

【D.】无放回抽样会有重复的个体

【答案】D

【知识点】04399.2.3.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于抽样的说法正确的是（ ）。

【A.】抽样有多种方法，如有放回抽样、无放回抽样等

【B.】无放回抽样则将抽出来的个体从抽样对象中去除

【C.】有放回抽样有可能对同一个个体抽中了多次

【D.】无放回抽样会有重复的个体

【答案】ABC

【知识点】04399.2.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

无放回抽样有可能对同一个个体抽中了多次，有放回抽样则将抽出来的个体从抽样对象中去除，因此不会有重复的个体。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

有放回抽样有可能对同一个个体抽中了多次，无放回抽样则将抽出来的个体从抽样对象中去除，因此不会有重复的个体。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

离散化是一种将连续取值的变量转换成一个分类变量的数据预处理步骤。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

离散化是一种将分类取值的变量转换成一个连续变量的数据预处理步骤。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.4.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于主成分的属性说法错误的是（ ）。

【A.】它们是原始变量的线性组合

【B.】它们相互正交（垂直）

【C.】它们捕获了数据差异最小的方向

【D.】它们捕获了数据差异最大的方向

【答案】C

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于主成分分析的说法错误的是（ ）。

【A.】主成分分析一种经典的减少分析变量的方法

【B.】它将数据从原始的低维空间投影到高维空间

【C.】它将数据从原始的高维空间投影到低维空间

【D.】所有原始变量可通过变换的方式得到综合变量，称为主成分

【答案】B

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于主成分分析的说法错误的是（ ）。

【A.】主成分分析一种经典的减少分析变量的方法

【B.】主成分分析一种经典增加分析变量的方法

【C.】它将数据从原始的高维空间投影到低维空间

【D.】它将数据从原始的低维空间投影到高维空间

【答案】BD

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于主成分分析的说法正确的是（ ）。

【A.】主成分分析一种经典的减少分析变量的方法

【B.】主成分分析一种经典增加分析变量的方法

【C.】它将数据从原始的高维空间投影到低维空间

【D.】它将数据从原始的低维空间投影到高维空间

【答案】AC

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于主成分的属性说法正确的是（ ）。

【A.】它们是原始变量的线性组合

【B.】它们相互正交（垂直）

【C.】它们捕获了数据差异最小的方向

【D.】它们捕获了数据差异最大的方向

【答案】ABD

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于主成分分析的说法正确的是（ ）。

【A.】主成分分析一种经典的减少分析变量的方法

【B.】它将数据从原始的低维空间投影到高维空间

【C.】它将数据从原始的高维空间投影到低维空间

【D.】所有原始变量可通过变换的方式得到综合变量，称为主成分

【答案】ACD

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

主成分分析是一种经典的增加分析变量的方法，它将数据从原始的低维空间投影到高维空间。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

主成分是原始变量的线性组合，它们相互正交（垂直），捕获了数据差异最小的方向。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

主成分分析是一种经典的减少分析变量的方法，它将数据从原始的高维空间投影到低维空间。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

主成分是原始变量的线性组合，它们相互正交（垂直），捕获了数据差异最大的方向。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.5.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于设定虚拟变量时应当遵循的原则说法不正确的是（ ）。

【A.】对于有k个表现值的定性变量，只设定（k-1）个虚拟变量

【B.】虚拟变量的值通常用“0”或“1”来表示

【C.】对于每个样本而言，同一个定性变量对应虚拟变量的值之和不超过1

【D.】设定虚拟变量时对其数量一般不做要求

【答案】D

【知识点】04399.2.6.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

设定虚拟变量时应当遵循如下哪些原则（ ）。

【A.】对于有k个表现值的定性变量，只设定（k-1）个虚拟变量

【B.】虚拟变量的值通常用“0”或“1”来表示

【C.】对于每个样本而言，同一个定性变量对应虚拟变量的值之和不超过1

【D.】设定虚拟变量时对其数量一般不做要求

【答案】ABC

【知识点】04399.2.6.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

设定虚拟变量时，虚拟变量的值通常用“0”或“1”来表示。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.6.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

虚拟变量的设定即把对变量的定量描述转化成对定性数据的描述。

【A.】√

【B.】×

【答案】B

【知识点】04399.2.6.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

虚拟变量的设定即把对变量的定性描述转化成对定量数据的描述。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.6.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

设定虚拟变量时，对于有k个表现值的定性变量，只设定（k-1）个虚拟变量。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.6.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

设定虚拟变量时，对于每个样本而言，同一个定性变量对应虚拟变量的值之和不超过1。

【A.】√

【B.】×

【答案】A

【知识点】04399.2.6.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于性能评估的指标说法错误的是（ ）。

【A.】对于预测任务的数据挖掘，性能评估的指标包括平均绝对误差和均方根误差等

【B.】对于分类任务的数据挖掘，评价指标一般可通过一个混淆矩阵计算得到

【C.】数据挖掘方法预测性能的评估是数据挖掘中的一个重要内容

【D.】ROC曲线不是常用的评价方法

【答案】D

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于期望预测误差的说法错误的是（ ）。

【A.】期望预测误差实际上是真实值与预测值在某种损失函数下的差值

【B.】期望预测误差实际上是真实值与预测值在某种损失函数下的平均值

【C.】通常我们使用期望预测误差作为寻优的目标函数

【D.】有监督学习模型的目的是寻找一个可以很好地描述自变量X与因变量Y之间关系的函数f

【答案】A

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在回归学习中常用的损失函数是（ ）。

【A.】0-1损失函数

【B.】平方损失函数

【C.】指数损失函数

【D.】负二项损失函数

【答案】B

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于期望预测误差的是（ ）。

【A.】模型假设失误带来的误差

【B.】采用某种方法估计最优值时，估计值的平均可能偏离了真实值

【C.】估计值自身由于数据的随机性产生一个波动部分

【D.】一些主观人为因素

【答案】D

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于偏差和方差的说法不正确的是（ ）。

【A.】偏差和方差受模型的复杂度影响

【B.】模型越复杂，偏差越小，方差越大

【C.】模型越复杂，偏差越大，方差越小

【D.】模型越简单，偏差越大，而方差越小

【答案】C

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于通过样本估计期望预测误差的方式的是（ ）。

【A.】不重复抽样

【B.】重复抽样

【C.】非交叉验证

【D.】交叉验证

【答案】AC

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于性能评估的指标说法正确的是（ ）。

【A.】对于预测任务的数据挖掘，性能评估的指标包括平均绝对误差和均方根误差等

【B.】对于分类任务的数据挖掘，评价指标一般可通过一个混淆矩阵计算得到

【C.】数据挖掘方法预测性能的评估是数据挖掘中的一个重要内容

【D.】ROC曲线不是常用的评价方法

【答案】ABC

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于通过样本估计期望预测误差的方式的是（ ）。

【A.】不重复抽样

【B.】重复抽样

【C.】非交叉验证

【D.】交叉验证

【答案】BD

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于期望预测误差的说法正确的是（ ）。

【A.】期望预测误差实际上是真实值与预测值在某种损失函数下的差值

【B.】期望预测误差实际上是真实值与预测值在某种损失函数下的平均值

【C.】通常我们使用期望预测误差作为寻优的目标函数

【D.】有监督学习模型的目的是寻找一个可以很好地描述自变量X与因变量Y之间关系的函数f

【答案】BCD

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于偏差和方差的说法正确的是（ ）。

【A.】模型越复杂，偏差越大，方差越小

【B.】模型越复杂，偏差越小，方差越大

【C.】模型越简单，偏差越小，方差越大

【D.】模型越简单，偏差越大，方差越小

【答案】BD

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于偏差和方差的说法不正确的是（ ）。

【A.】模型越复杂，偏差越大，方差越小

【B.】模型越复杂，偏差越小，方差越大

【C.】模型越简单，偏差越小，方差越大

【D.】模型越简单，偏差越大，方差越小

【答案】AC

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于常用的损失函数的是（ ）。

【A.】0-1损失函数

【B.】平方损失函数

【C.】指数损失函数

【D.】负二项损失函数

【答案】ABCD

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

期望预测误差主要包括（ ）。

【A.】模型假设失误带来的误差

【B.】采用某种方法估计最优值时，估计值的平均可能偏离了真实值

【C.】估计值自身由于数据的随机性产生一个波动部分

【D.】一些主观人为因素

【答案】ABC

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于偏差和方差的说法正确的是（ ）。

【A.】偏差和方差受模型的复杂度影响

【B.】模型越复杂，偏差越小，方差越大

【C.】模型越复杂，偏差越大，方差越小

【D.】模型越简单，偏差越大，方差越小

【答案】ABD

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在回归学习中常用的损失函数是指数损失函数。

【A.】√

【B.】×

【答案】B

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

模型越复杂，偏差越小，方差越大；模型越简单，偏差越大，而方差越小。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

通过样本估计期望预测误差一般有重复抽样和交叉验证两种方式。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

有监督学习模型的目的是寻找一个可以很好地描述自变量X与因变量Y之间关系的函数f。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

模型越复杂，偏差越大，方差越小；模型越简单，偏差越小，而方差越大。

【A.】√

【B.】×

【答案】B

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

通过样本估计期望预测误差只有重复抽样这一种方式。

【A.】√

【B.】×

【答案】B

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

对于分类任务的数据挖掘，评价指标一般可通过一个混淆矩阵计算得到，包括准确度、特效度、灵敏度以及考虑错分成本的指标。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

ROC曲线是一种常用的评价方法，可用来展示临界值的选择与算法性能之间的关系，用于评价临界值不同的算法的性能。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对于预测任务的数据挖掘来说，性能评估的指标包括平均绝对误差和均方根误差等。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在回归学习中常用的损失函数是平方损失函数。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.1.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于AUC的说法不正确的是（ ）。

【A.】AUC被定义为ROC曲线下的面积，也就是ROC曲线的积分

【B.】通常取值在0.5~1

【C.】随机挑选一个正样本以及一个负样本，分类器判定正样本的值高于负样本的概率就是AUC值

【D.】AUC值（面积）越小的分类器，性能越好

【答案】D

【知识点】04399.3.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于由混淆矩阵导出的评价准则的是（ ）。

【A.】准确率

【B.】精确率

【C.】召回率

【D.】F1 score

【答案】ABCD

【知识点】04399.3.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于由混淆矩阵导出的评价准则的是（ ）。

【A.】F1 score

【B.】ROC曲线

【C.】AUC

【D.】召回率

【答案】ABCD

【知识点】04399.3.2.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于AUC的说法正确的是（ ）。

【A.】AUC被定义为ROC曲线下的面积，也就是ROC曲线的积分

【B.】通常取值在0.5~1

【C.】随机挑选一个正样本以及一个负样本，分类器判定正样本的值高于负样本的概率就是AUC值

【D.】AUC值（面积）越小的分类器，性能越好

【答案】ABC

【知识点】04399.3.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

AUC值（面积）越小的分类器，性能越好。

【A.】√

【B.】×

【答案】B

【知识点】04399.3.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

随机挑选一个正样本以及一个负样本，分类器判定正样本的值高于负样本的概率就是AUC值。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于回归模型评价指标的是（ ）。

【A.】平均绝对误差

【B.】平均相对误差

【C.】平均平方误差

【D.】决定系数

【答案】B

【知识点】04399.3.2.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于决定系数的说法不正确的是（ ）。

【A.】越接近于0说明模型拟合得越好

【B.】决定系数是指模型解释了的数据波动的比例

【C.】取值区间为[0,1]

【D.】越接近于1说明模型拟合得越好

【答案】A

【知识点】04399.3.2.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于回归模型评价指标的是（ ）。

【A.】平均绝对误差

【B.】平均相对误差

【C.】平均平方误差

【D.】决定系数

【答案】ACD

【知识点】04399.3.2.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于决定系数的说法正确的是（ ）。

【A.】越接近于0说明模型拟合得越好

【B.】决定系数是指模型解释了的数据波动的比例

【C.】取值区间为[0,1]

【D.】越接近于1说明模型拟合得越好

【答案】BCD

【知识点】04399.3.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

决定系数是指模型解释了的数据波动的比例，取值区间为[0,1]，越接近于0说明模型拟合得越好。

【A.】√

【B.】×

【答案】B

【知识点】04399.3.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

决定系数是指模型解释了的数据波动的比例，取值区间为[0,1]，越接近于1说明模型拟合得越好。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.2.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于聚类模型评价的外部指标的是（ ）。

【A.】DB指数

【B.】Jaccard系数

【C.】FM指数

【D.】Rand指数

【答案】A

【知识点】04399.3.3.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于聚类模型评价的外部指标的是（ ）。

【A.】DB指数

【B.】Jaccard系数

【C.】FM指数

【D.】Rand指数

【答案】BCD

【知识点】04399.3.3.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于聚类模型评价的内部指标的是（ ）。

【A.】DB指数

【B.】Jaccard系数

【C.】FM指数

【D.】Rand指数

【答案】A

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于聚类模型评价的外部指标说法正确的是（ ）。

【A.】Dunn指数越小，说明类与类之间分得越开

【B.】Dunn指数越大，说明类与类之间分得越开

【C.】DB指数刻画的是一个平均距离，该指数越小，说明聚类越好

【D.】DB指数刻画的是一个平均距离，该指数越大，说明聚类越好

【答案】BC

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于聚类模型评价的外部指标说法错误的是（ ）。

【A.】Dunn指数越小，说明类与类之间分得越开

【B.】Dunn指数越大，说明类与类之间分得越开

【C.】DB指数刻画的是一个平均距离，该指数越小，说明聚类越好

【D.】DB指数刻画的是一个平均距离，该指数越大，说明聚类越好

【答案】AD

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于聚类模型评价的内部指标的是（ ）。

【A.】DB指数

【B.】Jaccard系数

【C.】Dunn指数

【D.】Rand指数

【答案】BD

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于聚类模型评价的内部指标的是（ ）。

【A.】DB指数

【B.】Jaccard系数

【C.】Dunn指数

【D.】Rand指数

【答案】AC

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DB指数刻画的是一个平均距离，该指数越大，说明聚类越好。

【A.】√

【B.】×

【答案】B

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Dunn指数刻画了所有类两两之间的最短距离与最大类的“直径”之比，该指数越小，说明类与类之间分得越开。

【A.】√

【B.】×

【答案】B

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DB指数刻画的是一个平均距离，该指数越小，说明聚类越好。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Dunn指数刻画了所有类两两之间的最短距离与最大类的“直径”之比，该指数越大，说明类与类之间分得越开。

【A.】√

【B.】×

【答案】A

【知识点】04399.3.3.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关分析根据相关关系表现形式的不同可以分为（ ）。

【A.】简单相关分析和复杂相关分析

【B.】偏相关分析和非偏相关分析

【C.】线性相关分析和非线性相关分析

【D.】参数相关分析和非参数相关分析

【答案】C

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

变量或数据之间的两种主要关系的是（ ）。

【A.】因果关系和相关关系

【B.】因果关系和依存关系

【C.】函数关系和相关关系

【D.】函数关系和依存关系

【答案】C

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于相关分析根据其分析方法和处理对象不同划分的类别是（ ）。

【A.】简单相关分析

【B.】偏相关分析

【C.】非参数相关分析

【D.】线性相关分析

【答案】D

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于变量或数据之间的主要关系的是（ ）。

【A.】因果关系

【B.】函数关系

【C.】依存关系

【D.】以上皆是

【答案】B

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关关系的说法不正确的是（ ）。

【A.】相关关系主要体现为变量之间的相互依存关系

【B.】主要考察变量之间的相互影响，这种影响不存在方向性

【C.】相关关系是确定性的

【D.】相关关系是不确定的

【答案】C

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关分析根据其分析方法和处理对象不同可以分为（ ）。

【A.】简单相关分析、偏相关分析和非参数相关分析

【B.】简单相关分析、复杂相关分析和非参数相关分析

【C.】线性相关分析和非线性相关分析

【D.】简单相关分析和复杂相关分析

【答案】A

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

变量或数据之间的两种主要关系是（ ）。

【A.】依存关系

【B.】因果关系

【C.】函数关系

【D.】相关关系

【答案】CD

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于函数关系和相关关系说法不正确的是（ ）。

【A.】函数关系是确定性的

【B.】函数关系是不确定的

【C.】相关关系是确定性的

【D.】相关关系是不确定的

【答案】BC

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关分析根据相关关系表现形式的不同可以分为（ ）。

【A.】非线性相关分析

【B.】偏相关分析

【C.】非参数相关分析

【D.】线性相关分析

【答案】AD

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于函数关系和相关关系说法正确的是（ ）。

【A.】函数关系是确定性的

【B.】函数关系是不确定的

【C.】相关关系是确定性的

【D.】相关关系是不确定的

【答案】AD

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于变量或数据之间的两种主要关系是（ ）。

【A.】依存关系

【B.】因果关系

【C.】函数关系

【D.】相关关系

【答案】AB

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关分析根据其分析方法和处理对象不同可以分为（ ）。

【A.】简单相关分析

【B.】偏相关分析

【C.】非参数相关分析

【D.】线性相关分析

【答案】ABC

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于相关分析根据相关关系表现形式的不同划分的是（ ）。

【A.】非线性相关分析

【B.】偏相关分析

【C.】非参数相关分析

【D.】线性相关分析

【答案】BC

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关关系的说法正确的是（ ）。

【A.】相关关系主要体现为变量之间的相互依存关系

【B.】主要考察变量之间的相互影响，这种影响不存在方向性

【C.】相关关系是确定性的

【D.】相关关系是不确定的

【答案】ABD

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

函数关系是确定性的，往往将发生变动的变量称为“自变量”，将受自变量变动影响而发生变动的变量称为“因变量”。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关分析根据其分析方法和处理对象不同，可以分为线性相关分析和非线性相关分析。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

函数关系是不确定的，往往将发生变动的变量称为“自变量”，将受自变量变动影响而发生变动的变量称为“因变量”。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

变量之间的影响不能够用具体的函数来度量，但变量之间的关系确实存在数量上不是严格对应的相互依存关系，则称这种关系为相关关系。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关关系是确定的，主要考察变量之间的相互影响，这种影响不存在方向性。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关分析根据相关关系表现形式的不同，又可以分为线性相关分析和非线性相关分析。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关关系是不确定的，主要考察变量之间的相互影响，这种影响不存在方向性。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关分析根据相关关系表现形式的不同，可以分为简单相关分析、偏相关分析和非参数相关分析等。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关分析根据其分析方法和处理对象不同，可以分为简单相关分析、偏相关分析和非参数相关分析等。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

变量或数据之前的两种主要关系是函数关系和相关关系。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

变量之间的影响不能够用具体的函数来度量，但变量之间的关系确实存在数量上不是严格对应的相互依存关系，则称这种关系为函数关系。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

变量或数据之前的两种主要关系是因果关系和相关关系。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

当一个或几个变量取一定的值时，另一个变量有确定值与之具体严格相对应，则称这种关系为函数关系。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

当一个或几个变量取一定的值时，另一个变量有确定值与之具体严格相对应，则称这种关系为相关关系。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

相关系数r的取值范围是（ ）。

【A.】0≤r≤1

【B.】0<r<1

【C.】-1<r<1

【D.】-1≤r≤1

【答案】D

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）之间的关系形态不同，可以将相关关系分为（ ）。

【A.】单相关和复相关

【B.】正相关和负相关

【C.】线性相关和非线性相关

【D.】偏相关和参数相关

【答案】C

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关系数r的说法不正确的是（ ）。

【A.】r具有对称性

【B.】r的数值的大小与x和y的计量尺度无关

【C.】r无法反映两个变量之间的因果关系

【D.】r≥0时，表明两个变量之间存在负线性相关关系

【答案】D

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关系数r的说法不正确的是（ ）。

【A.】r具有对称性

【B.】r的数值的大小与x和y的计量尺度有关

【C.】r无法反映两个变量之间的因果关系

【D.】r≥0时，表明两个变量之间存在正线性相关关系

【答案】B

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关系数r的说法不正确的是（ ）。

【A.】当r=0时，表现为完全不相关

【B.】当r=0时，表现为无线性相关

【C.】当r取值为-1时，表示为完全负线性相关

【D.】当r取值为1时，表示为完全正线性相关

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关系数r的说法不正确的是（ ）。

【A.】r具有对称性

【B.】r的数值的大小与x和y的计量尺度无关

【C.】r能够反映两个变量之间的因果关系

【D.】r≥0时，表明两个变量之间存在正线性相关关系

【答案】C

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）的多少，可以将相关关系分为（ ）。

【A.】单相关和复相关

【B.】正相关和负相关

【C.】线性相关和非线性相关

【D.】偏相关和参数相关

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于变量之间的相关关系考察形式的是（ ）。

【A.】单相关和复相关

【B.】正相关和负相关

【C.】线性相关和非线性相关

【D.】偏相关和参数相关

【答案】D

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）之间的相关方向，可以将相关关系分为（ ）。

【A.】单相关和复相关

【B.】正相关和负相关

【C.】线性相关和非线性相关

【D.】偏相关和参数相关

【答案】B

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于Pearson相关系数的显著性检验说法错误的是（ ）。

【A.】如果P值越小则相关系数越显著

【B.】如果P值越大则相关系数越显著

【C.】P值的大小无法看出相关系数是否显著

【D.】P值的大小可以看出相关系数的显著程度

【答案】BC

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）的多少，可以将相关关系分为（ ）。

【A.】单相关

【B.】复相关

【C.】线性相关

【D.】非线性相关

【答案】AB

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于Pearson相关系数的显著性检验说法正确的是（ ）。

【A.】如果P值越小则相关系数越显著

【B.】如果P值越大则相关系数越显著

【C.】P值的大小无法看出相关系数是否显著

【D.】P值的大小可以看出相关系数的显著程度

【答案】AD

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）之间的关系形态不同，可以将相关关系分为（ ）。

【A.】单相关

【B.】复相关

【C.】线性相关

【D.】非线性相关

【答案】CD

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于按相关因素（变量）之间的关系形态不同将相关关系分类的是（ ）。

【A.】单相关

【B.】复相关

【C.】线性相关

【D.】非线性相关

【答案】AB

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）之间的相关方向，可以将相关关系分为（ ）。

【A.】单相关

【B.】复相关

【C.】正相关

【D.】负相关

【答案】CD

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于变量之间的相关关系考察形式的是（ ）。

【A.】单相关和复相关

【B.】正相关和负相关

【C.】线性相关和非线性相关

【D.】偏相关和参数相关

【答案】ABC

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关系数r的说法正确的是（ ）。

【A.】r的取值范围为[-1,1]

【B.】当r=0时，表现为无线性相关

【C.】当r取值为-1时，表示为完全负线性相关

【D.】当r取值为1时，表示为完全正线性相关

【答案】ABCD

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关系数r的说法正确的是（ ）。

【A.】r具有对称性

【B.】r的数值的大小与x和y的计量尺度无关

【C.】r无法反映两个变量之间的因果关系

【D.】r≥0时，表明两个变量之间存在正线性相关关系

【答案】ABCD

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于相关系数r的说法正确的是（ ）。

【A.】当r=0时，表现为完全不相关

【B.】当r=0时，表现为无线性相关

【C.】当r取值为-1时，表示为完全负线性相关

【D.】当r取值为1时，表示为完全正线性相关

【答案】ABCD

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于按相关因素（变量）之间的相关方向将相关关系分类的是（  ）。

【A.】单相关

【B.】复相关

【C.】正相关

【D.】负相关

【答案】AB

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于按相关因素（变量）的多少将相关关系分类的是（ ）。

【A.】单相关

【B.】复相关

【C.】线性相关

【D.】非线性相关

【答案】CD

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

r是两个变量之间线性关系的度量指标，也可以反映两个变量之间的因果关系。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Pearson相关系数是反映两个变量的线性相关程度，但它并不能够度量变量之间的非线性相关程度。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

r的数值的大小与x和y的计量尺度无关，改变x和y的数据的计量尺度，并不改变r的数值。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果P值越大则相关系数越显著，越能说明总体数据之间存在显著关系。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）之间的相关方向，相关关系又可分为正相关和负相关。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

r具有对称性，x与y之间的相关系数与y与x之间的相关系数相等。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

r的数值的大小与x和y的计量尺度有关，改变x和y的数据的计量尺度会改变r的数值。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）之间的关系形态不同，相关关系可分为线性相关和非线性相关。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果P值越小则相关系数越显著，越能说明总体数据之间存在显著关系。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

r是两个变量之间线性关系的度量指标，但无法反映两个变量之间的因果关系。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）的多少，相关关系又可分为正相关和负相关。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）之间的相关方向，可将相关关系分为单相关和复相关。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按相关因素（变量）的多少，可将相关关系分为单相关和复相关。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

r的取值范围为[-1,1]，当r取值为1时，表示为完全正线性相关；当r=0时，表现为无线性相关，当r取值为-1时，表示为完全负线性相关。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在进行相关分析时先剔除其对其他变量的影响，再研究变量之间的相关关系的方法称为偏相关分析。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

简单相关分析有时不能够真实地反映现象之间的关系

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

点二列（点双列）相关分析适用于两列变量中一列是来自正态总体的定距或定比数据，另一列是二分类数据的情况。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.3

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于常见的非参数相关系数计算方法的是（ ）。

【A.】Spearman

【B.】Kendall's tau-b

【C.】Hoeffding's D

【D.】Pearson

【答案】D

【知识点】04399.4.2.4

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

常见的非参数相关系数计算方法有（ ）。

【A.】Spearman

【B.】Kendall's tau-b

【C.】Hoeffding's D

【D.】Pearson

【答案】ABC

【知识点】04399.4.2.4

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

非参数相关系数计算方法较多，常见的主要有Spearman、Kendall's tau-b和Hoeffding's D相关系数等。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.2.4

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于关联规则的说法不正确的是（ ）。

【A.】关联规则的强度可以用支持度和置信度度量

【B.】支持度确定规则可以用于给定数据集的频繁程度

【C.】置信度确定Y在包含X的事务中出现的频繁程度

【D.】支持度确定Y在包含X的事务中出现的频繁程度

【答案】D

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

若关联规则为{牛奶，尿布}→{啤酒}:{牛奶，尿布，啤酒}，项集的支持度计数为2，而事务总数为5，则该项集的支持度为（ ）。

【A.】0.4

【B.】10

【C.】2

【D.】5

【答案】A

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

若关联规则为{牛奶，尿布}→{啤酒}:{牛奶，尿布，啤酒}，项集的支持度计数为2，而事务总数为5，则该项集的置信度为（ ）。

【A.】0.4

【B.】0.67

【C.】2

【D.】5

【答案】B

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列方法可以降低频繁项集的计算复杂度的是（ ）。

【A.】减少候选项集的数目

【B.】减少比较次数

【C.】增加候选项集的数目

【D.】增加比较次数

【答案】AB

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列方法无法降低频繁项集的计算复杂度的是（ ）。

【A.】减少候选项集的数目

【B.】减少比较次数

【C.】增加候选项集的数目

【D.】增加比较次数

【答案】CD

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于关联规则的说法正确的是（ ）。

【A.】关联规则的强度可以用支持度和置信度度量

【B.】支持度确定规则可以用于给定数据集的频繁程度

【C.】置信度确定Y在包含X的事务中出现的频繁程度

【D.】支持度确定Y在包含X的事务中出现的频繁程度

【答案】ABC

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于支持度和置信度的说法正确的是（ ）。

【A.】置信度确定规则可以用于给定数据集的频繁程度

【B.】支持度确定规则可以用于给定数据集的频繁程度

【C.】置信度确定Y在包含X的事务中出现的频繁程度

【D.】支持度确定Y在包含X的事务中出现的频繁程度

【答案】BC

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于支持度和置信度的说法不正确的是（ ）。

【A.】置信度确定规则可以用于给定数据集的频繁程度

【B.】支持度确定规则可以用于给定数据集的频繁程度

【C.】置信度确定Y在包含X的事务中出现的频繁程度

【D.】支持度确定Y在包含X的事务中出现的频繁程度

【答案】AD

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

支持度确定Y在包含X的事务中出现的频繁程度。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

先验原理指的是如果一个项集是频繁的，则它的所有子集也一定是频繁的。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

持度计数是指包含特定项集的事务个数。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

置信度确定规则可以用于给定数据集的频繁程度。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关联分析中，包含0个或多个项的集合被称为项集，若一个项集包含k项，则称其为k-项集。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关联规则是指形如X→Y的蕴含表达式，其中X和Y是不相交的项集。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

置信度确定Y在包含X的事务中出现的频繁程度。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关联规则的强度可以用支持度和置信度度量。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

支持度确定规则可以用于给定数据集的频繁程度。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于FP-growth算法的说法正确的是（ ）。

【A.】FP-growth算法以一种自底向上的方式探索

【B.】FP-growth算法以一种自上而下的方式探索

【C.】FP-growth算法效率较低

【D.】FP-growth算法效率较高

【答案】AD

【知识点】04399.4.4.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于FP-growth算法的说法不正确的是（ ）。

【A.】FP-growth算法以一种自底向上的方式探索

【B.】FP-growth算法以一种自上而下的方式探索

【C.】FP-growth算法效率较低

【D.】FP-growth算法效率较高

【答案】BC

【知识点】04399.4.4.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于Apriori算法和FP-growth算法的说法不正确的是（ ）。

【A.】Apriori算法效率较低

【B.】Apriori算法效率较高

【C.】FP-growth算法效率较低

【D.】FP-growth算法效率较高

【答案】BC

【知识点】04399.4.4.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于Apriori算法和FP-growth算法的说法正确的是（ ）。

【A.】Apriori算法效率较低

【B.】Apriori算法效率较高

【C.】FP-growth算法效率较低

【D.】FP-growth算法效率较高

【答案】AD

【知识点】04399.4.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Apriori算法的效率比FP-growth算法的效率低。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

路径相互重叠越少，使用Fp-tree结构获得的压缩效果就越好。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Apriori算法的效率比FP-growth算法的效率高。

【A.】√

【B.】×

【答案】B

【知识点】04399.4.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

路径相互重叠越多，使用Fp-tree结构获得的压缩效果就越好。

【A.】√

【B.】×

【答案】A

【知识点】04399.4.4.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

P-P图或Q-Q图形在残差符合正态假定条件下，散点图看起来应该像是（ ）。

【A.】一条水平直线

【B.】一条截距为0、斜率为-1的直线

【C.】一条截距为0、斜率为1的直线

【D.】一条垂直直线

【答案】C

【知识点】04399.5.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

只有通过检验的模型才能够充分描述变量之间的关系，建立的模型才有现实意义。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

P-P图或Q-Q图形在残差符合正态假定条件下，散点图看起来应该像是一条截距为0、斜率为1的直线。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

定性因素对因变量的影响在进行回归分析的过程中，需要进行特殊的处理是（ ）。

【A.】把定性变量取对数再做分析

【B.】把定性变量转化为虚拟变量之后再引入回归模型中进行分析

【C.】把定性变量去掉

【D.】无需做任何处理

【答案】B

【知识点】04399.5.1.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

定性因素对因变量的影响在进行回归分析的过程中，需要把定性变量转化为虚拟变量之后再引入回归模型中进行分析。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.1.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

参照方程就是指当所有虚拟变量为0时的方程。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.1.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

定性因素对因变量的影响在进行回归分析的过程中，无需对定性变量做处理。

【A.】√

【B.】×

【答案】B

【知识点】04399.5.1.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对含有虚拟变量的回归方程进行分析，应当先确定分析的参照方程。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.1.3

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

非线性形式的变量关系转化为线性关系可以通过的方式是（ ）。

【A.】变量代换或转换

【B.】变量剔除

【C.】变量筛选

【D.】增加变量

【答案】A

【知识点】04399.5.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果在存在非线性关系的情况下使用线性回归拟合曲线，则会丢失数据之间的大量有用信息，甚至会得出错误的结论。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

非线性形式的变量关系一般可以通过变量代换或转换的方式转化为线性关系。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果在存在非线性关系的情况下使用线性回归拟合曲线也是可以的，不会影响结论。

【A.】√

【B.】×

【答案】B

【知识点】04399.5.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

有些时候变量间的曲线关系比较明显，但是难将其线性化，这个时候可以考虑直接使用非线性最小二乘估计法来估计模型参数。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

多项式回归类似于可线性化的非线性模型，可通过变量代换的方式使用普通最小二乘对参数进行估计。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.2.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

一般的多项式回归模型很少应用到三阶以上。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.2.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

多项式回归模型常常会应用到三阶以上。

【A.】√

【B.】×

【答案】B

【知识点】04399.5.2.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

多项式回归当阶数过高时，待估参数过多，但这并不影响对于结果的判断。

【A.】√

【B.】×

【答案】B

【知识点】04399.5.2.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

多项式回归当阶数过高时，待估参数过多，在样本量不大的情况下会比较困难，这是多项式回归的一大缺陷。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.2.3

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

当数据具有尖峰厚尾的分布特征或有离群点（即异常值）时，模型的稳健性（ ）。

【A.】没有影响

【B.】较好

【C.】较差

【D.】无法判断

【答案】C

【知识点】04399.5.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

当数据具有尖峰厚尾的分布特征或有离群点（即异常值）时，模型稳健性较差。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

分位数回归是对以古典条件均值模型为基础的延伸，它用几个分位函数来估计整体模型。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

当数据具有尖峰厚尾的分布特征或有离群点（即异常值）时，模型稳健性较好。

【A.】√

【B.】×

【答案】B

【知识点】04399.5.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

与一般的回归分析过程一样，分位数回归模型进行参数估计之后，也需要对模型进行评价以及进行显著性检验。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.3.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

分析自变量对计数因变量的影响时所建立的模型为（ ）。

【A.】线性回归模型

【B.】Logistic回归模型

【C.】分位数回归

【D.】计数模型

【答案】D

【知识点】04399.5.4.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

讨论是否购买手机与哪些变量有关，采用离散选择模型来建模，那么响应概率就是指（ ）

【A.】做出“购买”决策的概率

【B.】做出“不购买”决策的期望

【C.】做出“购买”决策的对数

【D.】做出“不购买”决策的倒数

【答案】A

【知识点】04399.5.4.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

要考虑与运动会上获得的奖牌数相关的因素，比如参赛队总人数、运动员平均年龄、参与项目数等，此时最适合的模型是（ ）。

【A.】线性回归模型

【B.】Logistic回归模型

【C.】分位数回归

【D.】计数模型

【答案】D

【知识点】04399.5.4.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

属于离散因变量模型的是（ ）。

【A.】线性概率模型

【B.】计数模型

【C.】二元选择模型

【D.】多重选择模型

【答案】ABCD

【知识点】04399.5.4.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

假定隐变量的随机误差项服从Logit分布的模型是（ ）。

【A.】线性概率模型

【B.】二元Probit模型

【C.】二元Logit模型

【D.】计数模型

【答案】C

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

假定隐变量的随机误差项服从标准正态分布的模型是（ ）。

【A.】线性概率模型

【B.】二元Probit模型

【C.】二元Logit模型

【D.】计数模型

【答案】B

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

离散因变量具有两个选项或两种属性，离散因变量的属性往往是对立或互斥的模型是（ ）

【A.】线性概率模型

【B.】计数模型

【C.】二元选择模型

【D.】多重选择模型

【答案】C

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于线性概率模型，说法错误的是（ ）

【A.】解释变量的合理变化会导致预测概率溢出[0,1]区间；

【B.】使用加权最小二乘法修正异方差性可以保证概率预测值在[0,1]区间。

【C.】当样本量充分大时，普通最小二乘法参数估计量的结果近似服从正态分布，可以克服随机误差项分布未知的缺陷。

【D.】模型误差具有异方差性，但异方差性不影响参数估计。

【答案】BD

【知识点】04399.5.4.2

【解析】B：使用加权最小二乘法修正异方差性也无法保证概率预测值在[0,1]区间；D; 模型误差具有异方差性，异方差性使参数估计不具有时效性。

【题目】

【题型：多选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于二元选择模型的模型检验方法的是（ ）.

【A.】Z统计量对回归系数检验

【B.】LR似然比对模型的拟合优度进行检验

【C.】AIC、BIC信息指数

【D.】Wald统计量对多个系数进行检验

【答案】ABCD

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

线性概率模型的缺点是（ ）

【A.】解释变量的合理变化会导致预测概率溢出[0,1]区间；

【B.】使用加权最小二乘法修正异方差性也无法保证概率预测值在[0,1]区间；

【C.】随机误差项的分布未知；

【D.】模型误差具有异方差性，异方差性使参数估计不具有时效性。

【答案】ABCD

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于线性概率模型，说法正确的是（ ）

【A.】解释变量的合理变化会导致预测概率溢出[0,1]区间；

【B.】使用加权最小二乘法修正异方差性可以保证概率预测值在[0,1]区间。

【C.】当样本量充分大时，普通最小二乘法参数估计量的结果近似服从正态分布，可以克服随机误差项分布未知的缺陷。

【D.】模型误差具有异方差性，但异方差性不影响参数估计。

【答案】AC

【知识点】04399.5.4.2

【解析】B：使用加权最小二乘法修正异方差性也无法保证概率预测值在[0,1]区间；D; 模型误差具有异方差性，异方差性使参数估计不具有时效性。

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

二元选择模型的因变量中的“0”和“1”可以直接进行数学运算。

【A.】√

【B.】×

【答案】B

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

隐变量是不能直接进行观测，但可以通过其他直接观测得到的变量（显变量）进行描述和反映的变量。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

二元选择模型的的参数估计结果不能理解为自变量变动对因变量的边际影响，应当理解为自变量的变动，对因变量取“1”的概率的影响有多大。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

二元选择模型的因变量中的“0”和“1”只是对应属性的标注或符号，不具备任何数值上的意义，不直接进行数学运算。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

二元选择模型的的参数估计结果可以理解为自变量变动对因变量的边际影响，应当理解为自变量的变动。

【A.】√

【B.】×

【答案】B

【知识点】04399.5.4.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在多重选择模型中，对于无序选择模型，一般假定随机误差项是独立同分布的随机变量，且假定服从（ ）。

【A.】卡方分布

【B.】均匀分布

【C.】标准正态分布

【D.】韦布尔分布

【答案】D

【知识点】04399.5.4.3

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对于分析多个属性的离散因变量与自变量的关系的模型是（ ）。

【A.】线性概率模型

【B.】计数模型

【C.】二元选择模型

【D.】多重选择模型

【答案】D

【知识点】04399.5.4.3

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

依据离散因变量选项的含义和次序的不同，可以将多重选择模型分为（ ）.

【A.】ordinal probit模型

【B.】有序模型

【C.】无序模型

【D.】ordinal logit模型

【答案】BC

【知识点】04399.5.4.3

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对于某个时间、空间等范围内事情发生次数的计数数据，一般都认为其近似服从（ ）。

【A.】卡方分布

【B.】韦布尔分布

【C.】标准正态分布

【D.】泊松分布

【答案】D

【知识点】04399.5.4.4

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面关于Poisson回归模型说法不正确的是（ ）

【A.】又称为计数模型

【B.】假设观测因变量数据服从Poisson分布

【C.】使用极大似然法进行参数估计

【D.】模型不需要进行检验

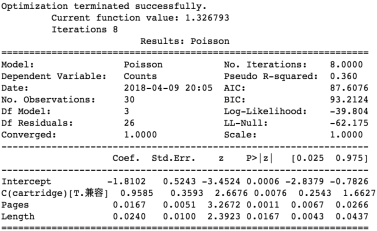
【答案】ABC

【知识点】04399.5.4.4

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

为监测某厂家生产的某款激光打印机的质量问题，考察该款打印机发生故障的次数。其发生故障的次数可能会受到打印纸张数量（千页）、打印机使用时长（千小时）、硒鼓（原装/兼容）等因素的影响。收集数据后的分析结果如下：        【图片】请问关于上面的结果说法正确的是：（ ）

【A.】此次分析构建了一个计数模型

【B.】收集了30个观测数据

【C.】对数似然值为-39.804

【D.】自变量都不显著

【答案】ABC

【知识点】04399.5.4.4

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：理解】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

对于Poisson回归模型，可以使用极大似然估计进行参数估计。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.4.4

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

计数模型的离散因变量的数字是有数值含义的，即次数之间可以进行数学运算。

【A.】√

【B.】×

【答案】A

【知识点】04399.5.4.4

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

计数模型的离散因变量的数字是没有数值含义的。

【A.】√

【B.】×

【答案】B

【知识点】04399.5.4.4

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

根据从不同总体中随机抽取出来的不同样本，在分析样本特征的基础上建立一定的判别法则，根据新的样本特征和判别法则判别新样本应该来自哪一个总体的是：（ ）

【A.】回归分析

【B.】聚类分析

【C.】主成分分析

【D.】判别分析

【答案】D

【知识点】04399.6.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于分类算法的是（ ）。

【A.】决策树

【B.】Kmeans

【C.】最近邻分类

【D.】支持向量机

【答案】B

【知识点】04399.6.1.1

【解析】Kmeans属于聚类算法。

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于分类算法的是（ ）。

【A.】决策树

【B.】Kmeans

【C.】最近邻分类

【D.】支持向量机

【答案】ACD

【知识点】04399.6.1.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

针对于多分类问题，我们可以采取哪些方法。（ ）

【A.】“一对一”，对于两两的类别组合，建立【图片】Generated个二分类模型；

【B.】选择可以直接进行多分类的模型；

【C.】直接使用针对二分类问题的模型；

【D.】“一对其他”，给每个类别与其他非此类别的样本建立二分类模型。

【答案】ABD

【知识点】04399.6.1.1

【解析】C：专门针对于二分类问题的模型，不能直接应用于多分类问题。

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

根据已知或假定总体是否服从一定的分布（如多元正态分布），可以将判别分析的判别规则分为（ ）。

【A.】两组判别分析

【B.】参数判别规则

【C.】非参数判别规则

【D.】多组判别分析

【答案】BC

【知识点】04399.6.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

分类任务就是通过学习得到一个目标函数f，一个把自变量空间映射到因变量空间的函数，我们称该目标函数为决策函数或判别函数。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

分类和回归的根本区别在于，分类的解释变量Y是一个一维的定性变量或离散变量。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

判别分析的核心是建立判别法则。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果已知或假定总体服从一定的分布，则可以使用参数判别规则，反之则可以使用非参数判别规则。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

所有针对于二分类问题的模型都可以直接应用于多分类问题。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.1.1

【解析】专门针对于二分类问题的模型，不能直接应用于多分类问题。

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

判别规则所依据的最简单的原则是，新样本离哪一个类别中心的距离最近，那么它就属于哪一类。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于混淆矩阵，如果一个点属于正类并且被预测到负类中，称为：（ ）

【A.】真正类；

【B.】假正类；

【C.】真负类；

【D.】假负类。

【答案】D

【知识点】04399.6.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于混淆矩阵，如果一个点属于负类并且被预测到正类中，称为：（ ）

【A.】真正类；

【B.】假正类；

【C.】真负类；

【D.】假负类。

【答案】B

【知识点】04399.6.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于混淆矩阵，如果一个点属于正类并且被预测到正类中，称为：（ ）

【A.】真正类；

【B.】假正类；

【C.】真负类；

【D.】假负类。

【答案】A

【知识点】04399.6.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于混淆矩阵，如果一个点属于负类并且被预测到负类中，称为：（ ）

【A.】真正类；

【B.】假正类；

【C.】真负类；

【D.】假负类。

【答案】C

【知识点】04399.6.1.2

【解析】无

【题目】

【题型：配伍题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

请为名词选择①-④表示其含义的描述，将配好的A-D填写到括号中。

【A.】①真正类 -> -> 如果一个点属于负类并且被预测到正类中

【B.】②假正类 -> -> 如果一个点属于正类并且被预测到正类中

【C.】③真负类 -> -> 如果一个点属于正类并且被预测到负类中

【D.】④假负类 -> -> 如果一个点属于负类并且被预测到负类中

【答案】B,A,D,C

【知识点】04399.6.1.2

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

样品和哪个总体距离最近，就判它属于哪个总体的分类方法是：（ ）

【A.】支持向量机

【B.】决策树

【C.】最近邻分类

【D.】线性判别

【答案】D

【知识点】04399.6.2.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

马氏距离实质上是经过标准化的变量的欧氏距离。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.2.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

通常情况下，线性判别（距离判别）一般采用马氏距离。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.2.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

马氏距离不受总体空间大小的影响，也不受计量单位的影响，反映了按平均水平计算被判定样本到中心的相对距离。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.2.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

马氏距离既受总体空间大小的影响，也受计量单位的影响，反映了按平均水平计算被判定样本到中心的相对距离。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.2.1

【解析】马氏距离不受总体空间大小的影响，也不受计量单位的影响。

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对每个样本，先计算判别函数得分，然后根据先验概率和判别函数得分的条件概率，计算出该样本被判为每一类的后验概率，被判入哪一类的后验概率最大，则把样本判为哪一类。该分类方法是（ ）

【A.】聚类判别

【B.】决策树

【C.】贝叶斯判别

【D.】线性判别

【答案】C

【知识点】04399.6.2.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于贝叶斯判别和线性判别，说法正确的是（ ）

【A.】贝叶斯判别考虑每个分类的观察值不同时，每类出现的机会是不同的。

【B.】贝叶斯判别与距离判别的结果总是一样的。

【C.】线性判别简单、直观且实用。

【D.】线性判别考虑了误判所造成的损失差异。

【答案】AC

【知识点】04399.6.2.2

【解析】B：贝叶斯判别与距离判别的结果不一定相同；D：贝叶斯判别考虑了误判所造成的损失差异。

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不是判别分析的方法（ ）。

【A.】距离判别

【B.】协方差阵判别

【C.】聚类判别

【D.】贝叶斯判别

【答案】BC

【知识点】04399.6.2.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于贝叶斯判别和线性判别，说法错误的是（ ）。

【A.】贝叶斯判别考虑每个分类的观察值不同时，每类出现的机会是不同的。

【B.】贝叶斯判别与距离判别的结果总是一样的。

【C.】线性判别简单、直观且实用。

【D.】线性判别考虑了误判所造成的损失差异。

【答案】BD

【知识点】04399.6.2.2

【解析】B：贝叶斯判别与距离判别的结果不一定相同；D：贝叶斯判别考虑了误判所造成的损失差异。

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于下面的贝叶斯判别结果说法正确的是（ ）。Setosaversicolorvirginica01.000000e+005.048809e-171.869461e-2516.250263e-1712.635294e-029.736471e-0121.937522e-3152.179357e-121.000000e+0033.729519e-1169.868228e-011.317724e-0241.000000e+005.634438e-184.545093e-26

【A.】编号0的观测将被判为Setosa一类

【B.】编号4的观测将被判为Setosa一类

【C.】编号3的观测将被判为versicolor一类

【D.】编号1的观测将被判为versicolor一类

【答案】ABC

【知识点】04399.6.2.2

【解析】编号1的观测将被判为virginica一类的最大概率为9.736471e-01，故将其归为“virginica”类。

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

贝叶斯判别既没有考虑每个分类的观察值不同时，每类出现的机会是不同的，也没有考虑误判所造成的损失差异。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.2.2

【解析】距离判别既没有考虑每个分类的观察值不同时，每类出现的机会是不同的，也没有考虑误判所造成的损失差异。

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

先验概率与后验概率一定不相同。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.2.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

距离判别既没有考虑每个分类的观察值不同时，每类出现的机会是不同的，也没有考虑误判所造成的损失差异。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.2.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

先验概率是根据以往经验和分析得到的概率。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.2.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

贝叶斯判别与距离判别的结果总是一样的。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.2.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

每个样本都可以用它最接近的k个邻居来代表的算法是：（ ）

【A.】支持向量机

【B.】决策树

【C.】最近邻分类

【D.】线性判别

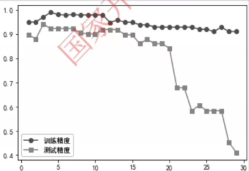
【答案】C

【知识点】04399.6.3.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如图是某数据集训练集和测试集的KNN算法的训练精度和测试精度，从图中可以判断k值取（ ）最为合适。        【图片】

【A.】1

【B.】5

【C.】12

【D.】29

【答案】B

【知识点】04399.6.3.2

【解析】当k取5附近时，KNN算法的测试精度达到了最高点，因此取5是较好的选择。

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于k近邻算法的k值，说法正确的是（ ）

【A.】随着k值的增加，模型的复杂程度减小，模型越稳健。

【B.】随着k值的增加，模型的复杂程度增加，模型越稳健。

【C.】随着k值的减少，模型的复杂程度增加，模型越稳健。

【D.】随着k值的减少，模型的复杂程度减少，模型越稳健。

【答案】A

【知识点】04399.6.3.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列哪一个方法适用于类域的交叉或重叠较多的待分样本集（ ）

【A.】决策树

【B.】随机森林

【C.】KNN算法

【D.】支持向量机

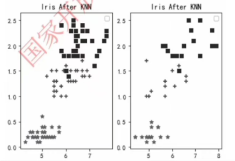
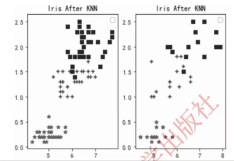
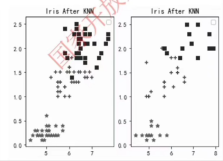
【答案】C

【知识点】04399.6.3.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如图分别为k=1，k=10，k=30时，训练集数据与测试集数据的分布，说法正确的是（ ）        【图片】k=1        【图片】k=10        【图片】k=30

【A.】近邻选得越多，模型越容易受敏感点的影响。

【B.】近邻选得越少，模型越容易受敏感点的影响。

【C.】近邻选得越少，模型越稳健，但是不容易发现一些敏感部分。

【D.】近邻选得越少，模型越稳健，越容易发现一些敏感部分。

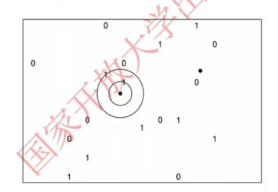
【答案】B

【知识点】04399.6.3.2

【解析】近邻选得越少，模型越容易受敏感点的影响。近邻选得多，模型则比较稳健，但是不容易发现一些敏感部分。

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

试想有如图的两类数据，“1”点表示一类数据，“0”点表示另一类数据，当k=1时，请问图中实心的圆点（从左到右）各自属于哪一类。（ ）【图片】

【A.】“1”；“1”

【B.】“1”；“0”

【C.】“0”；“0”

【D.】“0”；“1”

【答案】B

【知识点】04399.6.3.2

【解析】k=1，距离左边实心的圆点最近的一个是“1”，则被分到“1”类；距离右边实心的圆点最近的一个是“0”，则被分到“0”类。

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于KNN算法说法错误的是（ ）。

【A.】随着k值的增加，模型的复杂度增大。

【B.】在实际应用中，一般采用交叉验证的方法来估计预测误差，找出使得估计的预测误差最小的k。

【C.】最近邻分类每个样本都可以用它最接近的k个邻居来代表。

【D.】近邻选得少，模型则比较稳健。

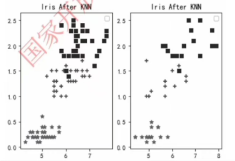
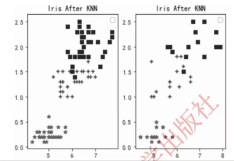
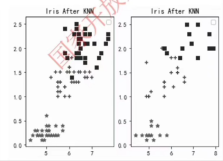
【答案】AD

【知识点】04399.6.3.2

【解析】A： 随着k值的增加，模型的复杂度减小；D：近邻选得多，模型则比较稳健。

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如图分别为k=1，k=10，k=30时，训练集数据与测试集数据的分布，说法正确的是（ ）。        【图片】k=1        【图片】k=10        【图片】k=30

【A.】近邻选得越多，模型越容易受敏感点的影响。

【B.】近邻选得越少，模型越容易受敏感点的影响。

【C.】近邻选得越多，模型越稳健，但是不容易发现一些敏感部分。

【D.】近邻选得越少，模型越稳健，越容易发现一些敏感部分。

【答案】BC

【知识点】04399.6.3.2

【解析】近邻选得越少，模型越容易受敏感点的影响。近邻选得多，模型则比较稳健，但是不容易发现一些敏感部分。

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于KNN算法说法正确的是（ ）。

【A.】随着k值的增加，模型的复杂度增大

【B.】在实际应用中，一般采用交叉验证的方法来估计预测误差，找出使得估计的预测误差最小的k。

【C.】最近邻分类每个样本都可以用它最接近的k个邻居来代表

【D.】近邻选得少，模型则比较稳健。

【答案】BC

【知识点】04399.6.3.2

【解析】A： 随着k值的增加，模型的复杂度减小；D：近邻选得多，模型则比较稳健。

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

最近邻算法中，随着近邻个数的增加，模型的复杂度增加。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.3.2

【解析】最近邻算法中，随着近邻个数的增加，模型的复杂度减小。

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

KNN算法在类别决策时，与极少量的相邻样本有关。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.3.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

近邻选得多，模型则比较稳健。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.3.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

KNN算法主要靠判别类域的方法来确定所属类别的，而不是靠周围有限的邻近的样本。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.3.2

【解析】KNN算法主要靠周围有限的邻近的样本，而不是靠判别类域的方法来确定所属类别的。

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

KNN算法在类别决策时，与极多数量的相邻样本有关。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.3.2

【解析】KNN算法在类别决策时，与极少量的相邻样本有关。

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

KNN算法主要靠周围有限的邻近的样本，而不是靠判别类域的方法来确定所属类别的。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.3.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在决策树中，节点熵越大，那么（ ）

【A.】节点的不纯度越低

【B.】节点的不纯度越高

【C.】节点越靠近根节点

【D.】节点越靠近叶节点

【答案】B

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

（假定在某个节点）如果数据已经只有一类了，则该节点为叶节点，否则进行下一步。该算法是：（ ）

【A.】Hunt

【B.】ID3

【C.】CART

【D.】C4.5

【答案】A

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在决策树中，节点误分率越大，那么（ ）

【A.】节点的不纯度越低

【B.】节点的不纯度越高

【C.】节点越靠近根节点

【D.】节点越靠近叶节点

【答案】B

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

采用将样本递归划分建立对象属性与对象值之间映射的算法是：（ ）

【A.】支持向量机

【B.】决策树

【C.】最近邻分类

【D.】线性判别

【答案】B

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树算法依据数据类型的不同和树状结构的不同有不同版本的决策树，下列属于决策算法的是：（ ）

【A.】Hunt

【B.】ID3

【C.】CART

【D.】C4.5

【答案】ABCD

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树是一种由节点和有向边组成的层次结构。下列属于决策树的节点的是：（ ）

【A.】根节点

【B.】内部节点

【C.】叶节点

【D.】终节点

【答案】ABCD

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中结点的最优划分是依据（ ）来确定的。

【A.】局部最优性

【B.】信息增益大的划分

【C.】信息增益小的划分

【D.】GINI增益大的划分

【答案】ABD

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，我们需要比较父节点与子节点的不纯度，他们的差越小，测试条件的效果就越好。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.4.1

【解析】决策树中，我们需要比较父节点与子节点的不纯度，他们的差越大，测试条件的效果就越好。

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树只有单一输出。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，我们需要比较父节点与子节点的不纯度，他们的差越大，测试条件的效果就越好。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，ID3算法的一个变量使用过后还能再次使用。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.4.2

【解析】决策树中，ID3算法只能处理定性变量，且一个变量使用过后就不能再次使用了。

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，ID3算法只能处理定性变量，且一个变量使用过后就不能再次使用了。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.4.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，ID3算法既能处理定性变量，也能处理连续变量、缺失值等。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.4.2

【解析】决策树中，ID3算法只能处理定性变量，且一个变量使用过后就不能再次使用了。

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，由于信息增益最大的划分原则容易产生小而纯的子集，所以C4.5算法提出了增益比。其中k为划分的总数，关于k的说法正确的是（ ）

【A.】代表如果某个属性产生了大量的划分，它的划分信息将会很小，从而降低增益比。

【B.】代表如果某个属性产生了大量的划分，它的划分信息将会很小，从而提高增益比。

【C.】代表如果某个属性产生了大量的划分，它的划分信息将会很大，从而降低增益比。

【D.】代表如果某个属性产生了大量的划分，它的划分信息将会很大，从而提高增益比。

【答案】C

【知识点】04399.6.4.3

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树算法的C4.5算法可以处理（ ）。

【A.】定性变量

【B.】缺失值

【C.】连续变量

【D.】剪枝

【答案】ABCD

【知识点】04399.6.4.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，由于信息增益最大的划分原则容易产生小而纯的子集，所以C4.5算法提出了增益比。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.4.3

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，由于信息增益最大的划分原则容易产生小而纯的子集，所以ID3算法提出了增益比。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.4.3

【解析】决策树中，由于信息增益最大的划分原则容易产生小而纯的子集，所以C4.5算法提出了增益比。

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，C4.5算法是ID3算法的延伸，可以处理缺失值、连续变量及剪枝等。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.4.3

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面关于决策树的剪枝操作理解错误的是（ ）。

【A.】对决策树进行剪枝是为了降低模型复杂度

【B.】决策树的剪枝操作可以在生成决策树后再剪枝

【C.】不可以在决策树生成过程中剪枝

【D.】一般用复杂度参数设定门限来进行剪枝

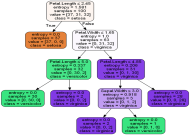
【答案】C

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于下面的决策树说法不正确的是（ ）。【图片】

【A.】此决策树根节点的熵为1.581

【B.】根节点的子节点的熵为0与1

【C.】此根节点的子节点都需要继续划分

【D.】落入Petal.Length≤2.45的数据将为判为setosa一类

【答案】C

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于CART算法的优点的是（ ）。

【A.】CART对于自变量和因变量不做任何形式的分布假定；

【B.】CART的结果受自变量的单调变换的影响;

【C.】CART分析所产生的树的结构很容易被任何领域的人理解;

【D.】CART能探索和揭示数据中的交互作用。

【答案】B

【知识点】04399.6.4.4

【解析】B：CART的结果不受自变量的单调变换的影响。

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

减枝准则是为了在精确性和复杂性中找到平衡，一般采用（ ）来度量两者之间的关系。

【A.】代回误分损失

【B.】惩罚常数

【C.】叶节点数目

【D.】复杂性损失

【答案】D

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树算法中，将当前样本划分为两个子样本，使得生成的每个非叶节点都有两个分支的算法是（ ）

【A.】Hunt

【B.】ID3

【C.】CART

【D.】C4.5

【答案】C

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树的cp值越小，那么（ ）

【A.】模型越复杂

【B.】训练误差越小

【C.】预测误差越小

【D.】模型更简单

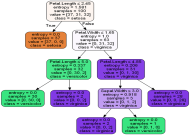
【答案】C

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

关于下面的决策树说法正确的是（ ）。        【图片】

【A.】此决策树根节点的熵为1.581

【B.】根节点的子节点的熵为0与1

【C.】此根节点的子节点都需要继续划分

【D.】落入Petal.Length≤2.45的数据将为判为setosa一类

【答案】ABD

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列属于CART算法的优点的是（ ）。

【A.】CART对于自变量和因变量不做任何形式的分布假定；

【B.】CART的结果受自变量的单调变换的影响;

【C.】CART分析所产生的树的结构很容易被任何领域的人理解;

【D.】CART能探索和揭示数据中的交互作用。

【答案】ACD

【知识点】04399.6.4.4

【解析】B：CART的结果不受自变量的单调变换的影响。

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面关于决策树的剪枝操作理解正确的是（ ）。

【A.】对决策树进行剪枝是为了降低模型复杂度

【B.】决策树的剪枝操作可以在生成决策树后再剪枝

【C.】不可以在决策树生成过程中剪枝

【D.】一般用复杂度参数设定门限来进行剪枝

【答案】ABD

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面关于决策树的剪枝操作理解不正确的是（ ）。

【A.】对决策树进行剪枝的目的是提高模型的预测误差

【B.】对决策树进行剪枝是为了提高模型的解释能力

【C.】只能在生成决策树后再剪枝

【D.】剪枝就是从下往上的剪掉一些节点

【答案】ABD

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，CART算法生成的决策树是结构简洁的二叉树，在每一步的决策树只能选择“是”或“否”，即使一个feature有多个取值，也只能把数据分为两部分。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

CART的结果受自变量的单调变换的影响。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

CART的结果不受自变量的单调变换的影响。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.4.4

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：否】【自动判分：是】【应用类别：终考】【特殊标签：】

决策树中，CART算法生成的决策树是结构简洁的二叉树，在每一步的决策树只能选择“是”或“否”但当一个feature有多个取值，可以把数据分为多个部分。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.4.4

【解析】决策树中，CART算法生成的决策树是结构简洁的二叉树，在每一步的决策树只能选择“是”或“否”，即使一个feature有多个取值，也只能把数据分为两部分。

【题目】

【题型：单选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于随机森林算法，说法错误的是（ ）

【A.】建立决策树的过程中，需要注意采样和完全分裂两点；

【B.】建立决策树过程中需要两个随机采样，即行采样和列采样；

【C.】对于行采样，采样无放回的方式；

【D.】采用完全分裂的方式构建决策树，这样决策树的某个叶节点要么无法继续分裂，要么里面的所有样本都指向同一个分类。

【答案】C

【知识点】04399.6.5.1

【解析】对于行采样，采样有放回的方式，也就是在采样得到的样本集合中，可能有重复的样本。

【题目】

【题型：单选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

用随机的方式建立一个森林。森林由很多决策树组成的算法是：（ ）

【A.】支持向量机

【B.】最近邻分类

【C.】随机森林

【D.】线性判别

【答案】C

【知识点】04399.6.5.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考,形考,模拟】【特殊标签：】

下列关于随机森林算法，说法正确的是（ ）。

【A.】建立决策树的过程中，需要注意采样和完全分裂两点；

【B.】建立决策树过程中需要两个随机采样，即行采样和列采样；

【C.】对于行采样，采样无放回的方式；

【D.】采用完全分裂的方式构建决策树，这样决策树的某个叶节点要么无法继续分裂，要么里面的所有样本都指向同一个分类。

【答案】ABD

【知识点】04399.6.5.1

【解析】对于行采样，采样有放回的方式，也就是在采样得到的样本集合中，可能有重复的样本。

【题目】

【题型：判断题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

随机森林算法建立决策树过程中，对于行采样，采样有放回的方式，也就是在采样得到的样本集合中，可能有重复的样本。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.5.1

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

随机森林算法建立决策树过程中，对于行采样，采样无放回的方式。

【A.】√

【B.】×

【答案】B

【知识点】04399.6.5.1

【解析】随机森林算法建立决策树过程中，对于行采样，采样有放回的方式。

【题目】

【题型：多选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考,形考,模拟】【特殊标签：】

随机森林的优势是（ ）。

【A.】能够处理高维度数据，并且不用做特征选择；

【B.】在训练过程中，能够检测到特征间的交互作用；

【C.】训练速度快；

【D.】容易做成并行化方法。

【答案】ABCD

【知识点】04399.6.5.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对于解决小样本、非线性及高维模式识别问题，（ ）表现出许多特有的优势。

【A.】支持向量机

【B.】最近邻分类

【C.】随机森林

【D.】线性判别

【答案】A

【知识点】04399.6.6.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

支持向量机算法中，超平面到所有两类数据点的距离之和称为分类器的边缘。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.6.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

以寻找最大边缘超平面为目的，以及构建相应的分类决策函数来分类的算法是（ ）

【A.】决策树

【B.】线性判别分析

【C.】线性可分支持向量机

【D.】随机森林

【答案】C

【知识点】04399.6.6.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

构成随机森林的决策树越多，那么（ ）

【A.】随机森林的预测精度越高

【B.】随机森林的预测精度越低

【C.】随机森林的训练精度越低

【D.】随机森林的训练精度越高

【答案】D

【知识点】04399.6.6.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考,形考,模拟】【特殊标签：】

常用的核函数有（ ）。

【A.】线性核

【B.】高斯核

【C.】神经网络核

【D.】多项式核

【答案】ABCD

【知识点】04399.6.6.3

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

非线性支持向量机通过引进核函数解决线性不可分问题。

【A.】√

【B.】×

【答案】A

【知识点】04399.6.6.3

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按照经济发展、教育水平、面积大小、人口等诸多方面对我国地市级以上城市进行分类。通常采用（ ）。

【A.】回归分析

【B.】分类

【C.】关联分析

【D.】聚类分析

【答案】D

【知识点】04399.7.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

当不知道哪些花是同一品种，哪些花不同品种时，考虑按照花瓣长宽与花萼长宽将不同品种的花分类。通常采用（ ）。

【A.】聚类分析

【B.】分类

【C.】回归分析

【D.】关联分析

【答案】A

【知识点】04399.7.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

不知道哪些客户有同样的购物习惯与特征等。考虑按照客户的个人资料、购物特征、购物习惯等将客户进行分类。通常采用（ ）。

【A.】回归分析

【B.】聚类分析

【C.】分类

【D.】关联分析

【答案】B

【知识点】04399.7.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

多元统计分析中的聚类分析方法可以对反映事物特征的指标或变量进行分类，记为（ ）。

【A.】S型分类

【B.】H型分类

【C.】Q型分类

【D.】R型分类

【答案】D

【知识点】04399.7.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列哪一项不属于聚类分析（ ）。

【A.】系统聚类

【B.】决策树

【C.】Kmeans

【D.】DBSCAN聚类

【答案】B

【知识点】04399.7.1.1

【解析】决策树属于分类算法。

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

不知道哪些学生有相同的某些特征。考虑按照学生的考试成绩、社会实践、思想品德等方面划分奖学金的等级。通常采用（ ）。

【A.】回归分析

【B.】分类

【C.】聚类分析

【D.】关联分析

【答案】C

【知识点】04399.7.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

多元统计分析中的聚类分析方法可以对样本进行分类，记为（ ）。

【A.】S型分类

【B.】H型分类

【C.】Q型分类

【D.】R型分类

【答案】C

【知识点】04399.7.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

多元统计分析中的聚类分析方法既可以对样本进行分类，记为R型分类，也可以对反映事物特征的指标或变量进行分类，记为Q型分类。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

多元统计分析中的聚类分析方法既可以对样本进行分类，记为Q型分类，也可以对反映事物特征的指标或变量进行分类，记为R型分类。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Mahalanobis距离指的是：（ ）

【A.】明氏距离

【B.】车比雪夫距离

【C.】欧氏距离

【D.】马氏距离

【答案】D

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类分析中，如果新类与其他类别之间存在多个点与点之间的距离，通过计算两类重心之间的欧氏距离平方作为两类之间的距离的方法为（ ）。

【A.】中间距离法

【B.】离差平方和法

【C.】重心法

【D.】类平均法

【答案】C

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类分析中，如果新类与其他类别之间存在多个点与点之间的距离，取这些距离最小者作为两类的距离，该方法为（ ）。

【A.】重心法

【B.】离差平方和法

【C.】最短距离法

【D.】类平均法

【答案】C

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类过程中的类需要遵从的原则中，说明了每一个个体在同一次分类过程中只能分在一个类别当中的原则是（ ）。

【A.】同质性原则

【B.】互斥性原则

【C.】完备性原则

【D.】充分性原则

【答案】C

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类分析中，如果新类与其他类别之间存在多个点与点之间的距离，取这些距离最大者作为两类的距离，该方法为（ ）。

【A.】重心法

【B.】离差平方和法

【C.】最长距离法

【D.】类平均法

【答案】C

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

首先考虑在没有进行聚类之前，所有参加聚类过程的个体没有归入任何类别，即对于每个个体而言，自成一类的方法是（ ）。

【A.】关联分析

【B.】系统聚类

【C.】快速聚类

【D.】DBSCAN聚类

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类分析中，如果新类与其他类别之间存在多个点与点之间的距离，将公式作为两类的距离，按照最小距离原则吧类别之间距离最小的两类合并为一类，直至把所有样本归为一类，该方法为（ ）。

【A.】中间距离法

【B.】离差平方和法

【C.】重心法

【D.】类平均法

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Euclidean距离指的是：（ ）

【A.】明氏距离

【B.】车比雪夫距离

【C.】欧氏距离

【D.】马氏距离

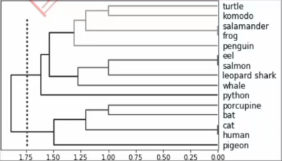
【答案】C

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下述谱系图使用的是类平均法，横轴表示类别之间的平均距离，纵轴为具体的每个样本。如在图中距离为1.75 处画一条竖线，该直线与谱系聚类图有2个交点，即可把所有样本分为两类。与第1个交点相连的样本是 （ ），它们形成第一类；而与第2个交点相连的样本是（ ），它们形成另一类。        【图片】

【A.】turtle、komodo、salamander、frog、penguin、eel；salmon、leopard shark、whale、python、porcupine、bat、cat、human、pigeon

【B.】turtle、komodo、salamander、frog、penguin、eel、salmon、leopard shark；whale、python、porcupine、bat、cat、human、pigeon

【C.】turtle、komodo、salamander、frog、penguin、eel、salmon、leopard shark、whale、python；porcupine、bat、cat、human、pigeon

【D.】turtle、komodo、salamander、frog、penguin、eel、salmon、leopard shark、whale；python、porcupine、bat、cat、human、pigeon

【答案】C

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类分析中，如果新类与其他类别之间存在多个点与点之间的距离，通过用不同类的样本点两两之间的平均距离作为类间距离的方法为（ ）。

【A.】中间距离法

【B.】离差平方和法

【C.】重心法

【D.】类平均法

【答案】D

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于类间距确定方法说法错误的是（ ）。

【A.】离差平方和法要求样本间距离可以采用欧式距离、马氏距离等；

【B.】重心法考虑了每一类中所包含的样本点数目；

【C.】类平均法充分利用各个样本的信息；

【D.】使用离差平方和作为类间距离时，如果聚类聚得恰当，类内样本点之间的离差平方和应该较小，类间离差平方和应该较大。

【答案】A

【知识点】04399.7.2.1

【解析】A：离差平方和法要求样本间距离必须采用欧氏距离。

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类过程中的类需要遵从的原则中，保证了类别之内个体特征的共性的原则是（ ）。

【A.】同质性原则

【B.】互斥性原则

【C.】完备性原则

【D.】充分性原则

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Chebychev距离指的是：（ ）

【A.】明氏距离

【B.】车比雪夫距离

【C.】欧氏距离

【D.】马氏距离

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类过程中的类需要遵从的原则中，保证了类别之间的差异性的原则是（ ）。

【A.】同质性原则

【B.】互斥性原则

【C.】完备性原则

【D.】充分性原则

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在聚类分析中，如果新类与其他类别之间存在多个点与点之间的距离，先让每个样本自身各成一类，然后并类，每并一类离差平方和就要增大，选择使其增量最小的两类合并，直到所有的样本聚为一类，该方法为（ ）。

【A.】中间距离法

【B.】离差平方和法

【C.】重心法

【D.】类平均法

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

使用离差平方和作为类间距离时，如果聚类聚得恰当，类内样本点之间的离差平方和应该（ ），类间离差平方和应该（ ）。

【A.】较小；较小

【B.】较小；较大

【C.】较大；较小

【D.】较大；较大

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面哪个原则不是聚类过程中的类需要遵从的原则（ ）。

【A.】同质性原则

【B.】互斥性原则

【C.】完备性原则

【D.】充分性原则

【答案】D

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于类间距确定方法说法正确的是（ ）。

【A.】离差平方和法要求样本间距离必须采用欧氏距离；

【B.】重心法没有考虑每一类中所包含的样本点数目；

【C.】类平均法没有充分利用各个样本的信息；

【D.】使用离差平方和作为类间距离时，如果聚类聚得恰当，类内样本点之间的离差平方和应该较大，类间离差平方和应该较小。

【答案】A

【知识点】04399.7.2.1

【解析】B：重心法考虑了每一类中所包含的样本点数目；C：类平均法充分利用各个样本的信息；D：使用离差平方和作为类间距离时，如果聚类聚得恰当，类内样本点之间的离差平方和应该较小，类间离差平方和应该较大。

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Minkowski距离指的是：（ ）

【A.】闵可夫斯基距离

【B.】车比雪夫距离

【C.】欧氏距离

【D.】马氏距离

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

block距离指的是：（ ）

【A.】明氏距离

【B.】车比雪夫距离

【C.】绝对距离

【D.】马氏距离

【答案】C

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

马氏距离是由（ ）计算出来的相对距离，（ ）指标之间的相关性。

【A.】协方差矩阵；考虑

【B.】协方差矩阵；不考虑

【C.】相关系数矩阵；考虑

【D.】相关系数矩阵；不考虑

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于类间距确定方法说法正确的是（ ）。

【A.】离差平方和法要求样本间距离必须采用欧氏距离；

【B.】重心法没有考虑每一类中所包含的样本点数目；

【C.】类平均法没有充分利用各个样本的信息；

【D.】使用离差平方和作为类间距离时，如果聚类聚得恰当，类内样本点之间的离差平方和应该较小，类间离差平方和应该较大。

【答案】BC

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面哪个原则是聚类过程中的类需要遵从的原则（ ）。

【A.】同质性原则

【B.】互斥性原则

【C.】完备性原则

【D.】充分性原则

【答案】ABC

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

常用的类间距确定的方法有（ ）。

【A.】最短距离法

【B.】离差平方和法

【C.】重心法

【D.】类平均法

【答案】ABCD

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面哪些距离是明氏距离：（ ）

【A.】绝对距离

【B.】欧氏距离

【C.】车比雪夫距离

【D.】马氏距离

【答案】ABC

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于类间距确定方法说法正确的是（ ）。

【A.】离差平方和法要求样本间距离必须采用欧氏距离；

【B.】重心法没有考虑每一类中所包含的样本点数目；

【C.】类平均法没有充分利用各个样本的信息；

【D.】使用离差平方和作为类间距离时，如果聚类聚得恰当，类内样本点之间的离差平方和应该较小，类间离差平方和应该较大。

【答案】AD

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

系统聚类中的类间距离是：（ ）

【A.】合并后新出现的类与其他类之间的距离

【B.】不能通过前一步的类间距离递推得到

【C.】可以通过前一步的类间距离递推得到

【D.】有多重可以选择的方式

【答案】ACD

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Demirmen 提出了一些在决定聚类方法取舍时应遵循的原则:（ ）

【A.】任何类必须在邻近的各类中是突出的，即各类重心(常用平均数衡量)之间应该有最大的距离；

【B.】确定的类中，各类所包含的元素都不宜过分多；

【C.】聚类数目应符合实际；

【D.】当用许多方法进行分类时，应选出现次数最多的那种分类结果。

【答案】ABCD

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

设dij表示第i个样本与第j个样本之间的距离。那么距离必须满足的原则有：（ ）。

【A.】

【B.】

【C.】

【D.】

【答案】ABCD

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

系统聚类中使用的类间距离不同，聚类结果可能是不同的。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，重心法考虑了每一类中所包含的样本点数目。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，个体与个体之间的距离越远越有可能是一类。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

闵可夫斯基距离的大小与个体指标的观测单位有关，考虑了指标之间的相关性。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

系统聚类中的类间距离是合并后新出现的类与其他类之间的距离。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

系统聚类可以给出聚类过程，但是要事先确定聚类个数。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

马氏距离是由协方差矩阵计算出来的相对距离，考虑了指标的相关性。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

马氏距离是由协方差矩阵计算出来的相对距离，没有考虑指标之间的相关性。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

闵可夫斯基距离的大小与个体指标的观测单位有关，没有考虑指标之间的相关性。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，类平均法充分利用各个样本的信息。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，离差平方和法要求样本间距离必须采用欧氏距离。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，个体与个体之间的距离越近越有可能是一类。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，类平均法没有充分利用各个样本的信息。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，重心法没有考虑每一类中所包含的样本点数目。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

系统聚类可以给出聚类过程，并且不用事先确定聚类个数。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：配伍题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

请为名词选择①-④表示其含义的描述，将配好的A-D填写到括号中。设*dij*表示第*i*个样本与第*j*个样本之间的距离。

【A.】①绝对距离 -> ->

【B.】②切比雪夫距离 -> ->

【C.】③马氏距离 -> ->

【D.】④欧氏距离 -> ->

【答案】B,A,C,D

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：配伍题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

请为名词选择①-⑥表示其含义的描述，将配好的A-F填写到括号中。

【A.】①最小距离法 ->  ->

【B.】②最长距离法 -> ->

【C.】③中间距离法 -> ->

【D.】④重心法 -> ->

【E.】⑤类平均法 -> ->

【F.】⑥离差平方和法 -> ->

【答案】F,D,B,C,A,E

【知识点】04399.7.2.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

有些情况下，研究者对于研究的对象事先知道分为几类，即已知类别的个数，只是不知道这些类别当中的具体样本，该方法是（ ）。

【A.】关联分析

【B.】系统聚类

【C.】快速聚类

【D.】DBSCAN聚类

【答案】C

【知识点】04399.7.3.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

快速聚类中，要确定数据到底聚成几个类合适，我们可以设定K从1到K的变化，计算相应的（ ），选择其最（ ）的类别个数。

【A.】SSE；大

【B.】SSE；小

【C.】SSR；大

【D.】SSR；小

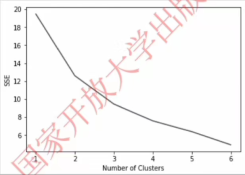
【答案】B

【知识点】04399.7.3.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

快速聚类中，要确定数据到底聚成几个类合适，我们可以设定K从1到K的变化，计算相应的SSE，下图设定了K的取值为1，2，3，4，5，6。K选（ ）最为合适。        【图片】

【A.】1

【B.】2

【C.】3

【D.】6

【答案】D

【知识点】04399.7.3.1

【解析】要确定数据到底聚成几个类合适，我们可以设定K从1到K的变化，计算相应的SSE，选择SSE最小的类别个数。

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，快速聚类一般用于小样本情况下的样本聚类。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

快速聚类要事先确定聚类个数，但是运算速度很快，适合少量数据的情况。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.3.1

【解析】快速聚类要事先确定聚类个数，但是运算速度很快，适合大量数据的情况。

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，快速聚类一般用于大样本情况下的样本聚类。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.3.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

快速聚类要事先确定聚类个数，但是运算速度很快，适合大量数据的情况。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.3.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法中的直接密度可达、密度可达与密度相连理解错误的是（ ）

【A.】只有核心对象之间相互密度可达

【B.】密度相连是对称关系

【C.】DBSCAN目的是找到密度相连对象的最大集合

【D.】密度可达也是对称关系

【答案】D

【知识点】04399.7.4.1

【解析】密度可达是非对称关系

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于聚类方法说法正确的是（ ）

【A.】DBSCAN聚类可以用于数据分布不均匀，呈块状分布，并且聚集形态是任意形状的情况。

【B.】快速聚类要事先确定聚类个数，但是运算速度很快，适合少量数据的情况。

【C.】系统聚类可以给出聚类过程，但是要事先确定聚类个数。

【D.】聚类分析中，快速聚类一般用于小样本情况下的样本聚类。

【答案】A

【知识点】04399.7.4.1

【解析】B：快速聚类要事先确定聚类个数，但是运算速度很快，适合大量数据的情况。C：系统聚类可以给出聚类过程，并且不用事先确定聚类个数。D：聚类分析中，快速聚类一般用于大样本情况下的样本聚类。

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

利用基于密度的聚类的概念，即要求聚类空间中的一定区域所包含对象的数目不小于某一给定的阈值，该方法是（ ）。

【A.】关联分析

【B.】系统聚类

【C.】快速聚类

【D.】DBSCAN聚类

【答案】D

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN 算法需要用户输入两个参数:一个参数是（ ），表示以给定点P为中心的圆形邻域;另一个参数是以点P为中心的邻域内最（ ）点的数量(MinPts)。

【A.】半径；小

【B.】半径；大

【C.】直径；小

【D.】直径；大

【答案】A

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法中，密度可达是直接密度可达的传递闭包，并且这种关系是（ ）的，该算法的目的是找到密度相连对象的（ ）集合。

【A.】对称；最小

【B.】对称；最大

【C.】非对称；最小

【D.】非对称；最大

【答案】D

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法的优点是（ ）。

【A.】当数据量增大时，要求较大的内存支持，I/Q消耗也很大。

【B.】能够有效处理噪声点和发现任意形状的空间聚类。

【C.】DBSCAN算法聚类速度快。

【D.】当空间聚类的密度不均匀、聚类间距相差很大时 聚类质量较差。

【答案】BC

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法中的直接密度可达、密度可达与密度相连理解正确的是（ ）

【A.】只有核心对象之间相互密度可达

【B.】密度相连是对称关系

【C.】DBSCAN目的是找到密度相连对象的最大集合

【D.】密度可达也是对称关系

【答案】ABC

【知识点】04399.7.4.1

【解析】D：密度可达是非对称关系

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于聚类方法说法错误的是（ ）

【A.】DBSCAN聚类可以用于数据分布不均匀，呈块状分布，并且聚集形态是任意形状的情况。

【B.】快速聚类要事先确定聚类个数，但是运算速度很快，适合大量数据的情况。

【C.】系统聚类可以给出聚类过程，但是要事先确定聚类个数。

【D.】聚类分析中，快速聚类一般用于小样本情况下的样本聚类。

【答案】CD

【知识点】04399.7.4.1

【解析】C：系统聚类可以给出聚类过程，并且不用事先确定聚类个数。D：聚类分析中，快速聚类一般用于大样本情况下的样本聚类。

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于聚类方法说法正确的是（ ）

【A.】DBSCAN聚类可以用于数据分布不均匀，呈块状分布，并且聚集形态是任意形状的情况。

【B.】快速聚类要事先确定聚类个数，但是运算速度很快，适合大量数据的情况。

【C.】系统聚类可以给出聚类过程，但是要事先确定聚类个数。

【D.】聚类分析中，快速聚类一般用于小样本情况下的样本聚类。

【答案】AB

【知识点】04399.7.4.1

【解析】C：系统聚类可以给出聚类过程，并且不用事先确定聚类个数。D：聚类分析中，快速聚类一般用于大样本情况下的样本聚类。

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

聚类分析中，关于内部评价指标说法正确的是（ ）。

【A.】DBI指数越大，说明聚类结果越好

【B.】DBI指数越小，说明聚类结果越好

【C.】DI指数越大，说明聚类结果越好

【D.】DI指数越小，说明聚类结果越好

【答案】BC

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法的缺点是（ ）。

【A.】当数据量增大时，要求较大的内存支持，I/Q消耗也很大。

【B.】能够有效处理噪声点和发现任意形状的空间聚类。

【C.】DBSCAN算法聚类速度快。

【D.】当空间聚类的密度不均匀、聚类间距相差很大时 聚类质量较差。

【答案】AD

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN聚类可以用于数据分布均匀，呈块状分布，并且聚集形态是固定形状的情况。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.4.1

【解析】DBSCAN聚类可以用于数据分布不均匀，呈块状分布，并且聚集形态是任意形状的情况。

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

密度相连是对称关系，密度可达是非对称关系。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法的目的是找到密度相连对象的最小集合。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法的目的是找到密度相连对象的最大集合。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法基于密度的聚类的概念，即要求聚类空间中的一定区域所包含对象的数目不小于某一给定的阈值。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

密度相连是对称关系，密度可达也是对称关系。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.4.1

【解析】密度相连是对称关系，密度可达是非对称关系。

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN聚类可以用于数据分布不均匀，呈块状分布，并且聚集形态是任意形状的情况。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法中，密度可达是直接密度可达的传递闭包，并且这种关系是对称的。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法中，密度可达是直接密度可达的传递闭包，并且这种关系是非对称的。

【A.】√

【B.】×

【答案】A

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

DBSCAN算法基于密度的聚类的概念，即要求聚类空间中的一定区域所包含对象的数目不大于某一给定的阈值。

【A.】√

【B.】×

【答案】B

【知识点】04399.7.4.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

事物或现象在较长时间内持续发展变化的一种趋向或状态，称为( )

【A.】长期趋势

【B.】季节变动

【C.】循环波动

【D.】不规则变动

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

要通过某超市历史1年的客流量数据预测未来的客流量，通常采用( )

【A.】分类

【B.】时间序列模型

【C.】关联分析

【D.】聚类分析

【答案】B

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

事物或现象在一年内随着季节更换而形成的有规律变动，称为( B )

【A.】长期趋势

【B.】季节变动

【C.】循环波动

【D.】不规则变动

【答案】B

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

要通过某气象站历史10年的逐日降水数据预测未来的降水量，通常采用( )

【A.】时间序列模型

【B.】分类

【C.】关联分析

【D.】聚类分析

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据反映若干事物或现象在同一时刻或时间上所处的状态或特征，或者反映其与时间无关的特征，这些数据反映了事物或现象之间存在内在数值联系，称为（ ）。

【A.】横截面数据

【B.】时间序列数据

【C.】虚拟变量数据

【D.】随机数据

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据是随时间变化而变化，反映了事物、现象在时间上的发展变动情况，描述数据是相同事物或现象在不同时刻或时期所形成的数据，称为（ ）。

【A.】横截面数据

【B.】时间序列数据

【C.】虚拟变量数据

【D.】随机数据

【答案】B

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

要通过某城市历史10年的用电量数据预测未来的用电量，通常采用( )

【A.】分类

【B.】聚类分析

【C.】关联分析

【D.】时间序列模型

【答案】D

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

指事物或现象呈周而复始的变动，且是无固定规律的交替波动，称为( )

【A.】长期趋势

【B.】季节变动

【C.】循环波动

【D.】不规则变动

【答案】C

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

无法用组成部分解释或不可控的随机变动，称为( )。

【A.】长期趋势

【B.】季节变动

【C.】循环波动

【D.】不规则变动

【答案】D

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

时间序列的组成部分包括（ ）。

【A.】长期趋势

【B.】季节变动

【C.】循环波动

【D.】不规则变动

【答案】ABCD

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

空调销售量随着季节不同而发生较大变动，夏季的销售量一般高于冬季的销售量。空调销售量数据的这种变动称为季节波动。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据是随时间变化而变化，反映了事物、现象在时间上的发展变动情况，称为横截面数据。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据反映了事物或现象之间存在内在数值联系，称为横截面数据。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

由于受到各种偶然因素的影响，时间序列往往表现出某种随机性，彼此之间存在着统计上的依赖关系。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

1880年1月到2010年8月每月的全球温度异常值有一种无固定规律的交替波动，这种变动被称为循环波动。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据反映了事物或现象之间存在内在数值联系，称为时间序列数据。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

数据是随时间变化而变化，反映了事物、现象在时间上的发展变动情况，称为时间序列数据。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

时间序列数据的分析主要研究目的是总结过去预测未来。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

全球温度异常数据中有呈直线上升的趋势，这种变动是长期趋势波动。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对非平稳的时间序列进行转化，使之成为平稳时间序列的数据转换过程，称为（ ）。

【A.】零均值化

【B.】标准化

【C.】平稳化

【D.】归一化

【答案】C

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对均值不为零的时间序列进行转化，使其均值为零的数据转换过程，称为（ ）。

【A.】零均值化

【B.】标准化

【C.】平稳化

【D.】归一化

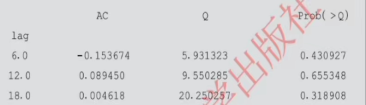
【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面白噪声检验结果表示，在α=0.1的条件下，白噪声检验的p值均大于α，表明白噪声检验结果（ ），所以该数据是（ ）。        【图片】

【A.】显著；平稳

【B.】显著；非平稳

【C.】不显著；平稳

【D.】不显著；非平稳

【答案】C

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果时间序列的一阶矩、二阶矩存在，而且对任意时刻满足均值为常数，协方差为时间间隔的函数，则称该序列为（ ）。

【A.】绝对数时间序列

【B.】宽平稳时间序列

【C.】非平稳时间序列

【D.】严平稳时间序列

【答案】B

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

一个时间序列的自相关系数：（ ）

【A.】取值在-1到1之间

【B.】取值在0到1之间

【C.】取值在-1到0之间

【D.】取值在-3到3之间

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面不能进行时间序列平稳性检验的是（ ）。

【A.】自相关系数

【B.】自相关图检验

【C.】单位根检验

【D.】DW检验

【答案】D

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果一个时间序列的概率分布与时间t无关，则称该序列为（ ）。

【A.】绝对数时间序列

【B.】平稳时间序列

【C.】非平稳时间序列

【D.】相对数时间序列

【答案】B

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

序列均值或协方差与时间有关的序列为（ ）。

【A.】绝对数时间序列

【B.】平稳时间序列

【C.】非平稳时间序列

【D.】相对数时间序列

【答案】C

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面白噪声序列说法错误的是（ ）。

【A.】具有零均值

【B.】具有同方差性

【C.】协方差为零

【D.】不是平稳时间序列

【答案】D

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列不属于要研究平稳时间序列的原因为（ ）

【A.】在平稳的保证情况下，对历史时序数据进行分析的参数估计结果也比较稳定；

【B.】平稳时间序列可以直接用于对未来时序数据的预测；

【C.】平稳时间序列存在“伪回归”的情况；

【D.】非平稳时间序列在分析时，本来没有什么关系的变量之间可能出现“伪回归”的情况。

【答案】C

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

为了能进行ARIMA时间序列分析建模，通常将非平稳时间序列进行（ ）和（ ），将其转化为零均值平稳时间序列。

【A.】零均值化；平稳化

【B.】零均值化；标准化

【C.】标准化；平稳化

【D.】标准化；归一化

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

利用白噪声检验时间序列平稳性，如果白噪声结果显著，则表明时间序列总体自相关是（ ）的，即表现为（ ）。

【A.】显著；平稳

【B.】显著；非平稳

【C.】不显著；平稳

【D.】不显著；非平稳

【答案】B

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果销售额数据与时间有密切相关的联系，即销售额数值随时间的推进而不断上升，则称该序列为（ ）。

【A.】绝对数时间序列

【B.】宽平稳时间序列

【C.】非平稳时间序列

【D.】严平稳时间序列

【答案】C

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

一个时间序列的自相关系数：（ ）

【A.】其值越小，说明时间序列的自相关程度越高

【B.】其值越大，说明时间序列的自相关程度越高

【C.】其绝对值越小，说明时间序列的自相关程度越高

【D.】其绝对值越大，说明时间序列的自相关程度越高

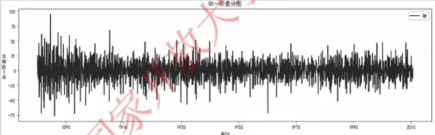
【答案】D

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

        从下面的时序图不能得到的结论是（ ）【图片】

【A.】该序列具有零均值

【B.】该序列均值不随时间变化而变化

【C.】这是一个平稳时间序列

【D.】这是一个非平稳时间序列

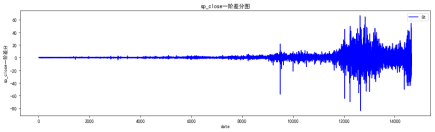
【答案】D

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

从下面的时序图一定可以得到的结论是（ ）        【图片】

【A.】该序列具有零均值

【B.】该序列具有同方差性

【C.】这是一个白噪声序列

【D.】这是一个平稳时间序列

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

我们可以通过（ ）方式将非平稳时间序列进行零均值化和平稳化。

【A.】标准化

【B.】函数

【C.】差分

【D.】归一化

【答案】BC

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于时间序列说法正确的是（ ）。

【A.】宽平稳时间序列其特征即均值和协方差不随时间变化而变化。

【B.】白噪声序列是平稳时间序列。

【C.】平稳时间序列的特征表现为在整体上或局部上有明显的上升或下降的趋势。

【D.】如果时间序列的一阶矩、二阶矩存在，而且对任意时刻满足均值为常数，协方差为时间间隔的函数，则称该序列为严平稳时间序列。

【答案】AB

【知识点】04399.8.1.2

【解析】C：非平稳时间序列的特征表现为在整体上或局部上有明显的上升或下降的趋势。D：如果时间序列的一阶矩、二阶矩存在，而且对任意时刻满足均值为常数，协方差为时间间隔的函数，则称该序列为宽平稳时间序列。

【题目】

【题型：多选题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

按照不同的性质和特征，可以对时间序列进行分类，从统计特性上来看，时间序列可以分为（ ）。

【A.】绝对数时间序列

【B.】平稳时间序列

【C.】非平稳时间序列

【D.】相对数时间序列

【答案】BC

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面能够进行时间序列平稳性检验的是（ ）。

【A.】自相关系数

【B.】自相关图检验

【C.】单位根检验

【D.】DW检验

【答案】ABC

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下面白噪声序列说法正确的是（ ）。

【A.】具有零均值

【B.】具有同方差性

【C.】协方差为零

【D.】不是平稳时间序列

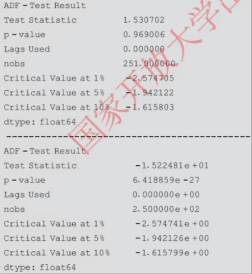
【答案】ABC

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下图分别为原始数据和一阶差分数据的单位根检验的结果，根据其p值，可以看出原始数据的单位根检验（ ），即p值非常大，没有充分的理由拒绝原假设，即原始序列是（ ）的序列;而一阶差分后的序列的单位根检验的p值（ ），故可以拒绝原假设，认为一阶差分序列是（ ）。        【图片】

【A.】不显著；平稳

【B.】不显著；非平稳

【C.】显著；平稳

【D.】显著；非平稳

【答案】BC

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：多选题】【难度：难】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列关于时间序列说法错误的是（ ）。

【A.】宽平稳时间序列其特征即均值和协方差不随时间变化而变化。

【B.】白噪声序列是平稳时间序列。

【C.】平稳时间序列的特征表现为在整体上或局部上有明显的上升或下降的趋势。

【D.】如果时间序列的一阶矩、二阶矩存在，而且对任意时刻满足均值为常数，协方差为时间间隔的函数，则称该序列为严平稳时间序列。

【答案】CD

【知识点】04399.8.1.2

【解析】C：非平稳时间序列的特征表现为在整体上或局部上有明显的上升或下降的趋势。D：如果时间序列的一阶矩、二阶矩存在，而且对任意时刻满足均值为常数，协方差为时间间隔的函数，则称该序列为宽平稳时间序列。

【题目】

【题型：多选题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

要研究平稳时间序列的原因为（ ）

【A.】在平稳的保证情况下，对历史时序数据进行分析的参数估计结果也比较稳定；

【B.】平稳时间序列可以直接用于对未来时序数据的预测；

【C.】平稳时间序列存在“伪回归”的情况；

【D.】非平稳时间序列在分析时，本来没有什么关系的变量之间可能出现“伪回归”的情况。

【答案】ABD

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

当一个时间序列具有单位根时是平稳的。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

对原始数据进行一次差分的过程称为一阶差分。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

在一般情况下，非平稳的时间序列在经过一阶差分或二阶差分之后都可以平稳化。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

序列均值或协方差与时间有关的序列为非平稳时间序列。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

平稳时间序列的特征表现为在整体上或局部上有明显的上升或下降的趋势。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

白噪声序列是平稳时间序列。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

一个时间序列的自相关系数，其绝对值越接近于1，说明时间序列的自相关程度越高。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

非平稳时间序列的特征表现为在整体上或局部上有明显的上升或下降的趋势。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

当一个时间序列具有单位根时是非平稳的。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

单位根检验用于检验一个时间序列的平稳性。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果时间序列的一阶矩、二阶矩存在，而且对任意时刻满足均值为常数，协方差为时间间隔的函数，则称该序列为严平稳时间序列。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.1.2

【解析】应为宽平稳

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

严平稳时间序列其特征即均值和协方差不随时间变化而变化。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

一个时间序列的自相关系数，其绝对值越接近于0，说明时间序列的自相关程度越高。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

白噪声序列不是平稳时间序列。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

一个时间序列如果能通过差分的方式平稳化，则可称其具有单位根。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

宽平稳时间序列其特征即均值和协方差不随时间变化而变化。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果一个时间序列的概率分布与时间t无关，则称该序列为平稳时间序列。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

差分用于将非平稳的时间序列平稳化。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：理解】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如果时间序列的一阶矩、二阶矩存在，而且对任意时刻满足均值为常数，协方差为时间间隔的函数，则称该序列为宽平稳时间序列。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

白噪声序列可以对时序模型拟合进行检验。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.1.2

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

具体表现为某个观测值不仅与其以前p个时刻的自身观测值有关，还与其以前时刻进入系统的q个随机误差存在一定的依存关系的模型是（ ）。

【A.】AR

【B.】ARIMA

【C.】ARMA

【D.】MA

【答案】C

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

请问下面的哪个模型是移动平均模型（ ）

【A.】

【B.】

【C.】

【D.】

【答案】C

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

具体表现为某个观测值与其滞后p期的观测值的线性组合再加上随机误差项的模型是（ ）。

【A.】AR

【B.】ARIMA

【C.】ARMA

【D.】MA

【答案】A

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：单选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

ARIMA模型也被叫做( )

【A.】自回归移动平均模型

【B.】自回归模型

【C.】移动平均模型

【D.】整合自回归移动平均模型

【答案】D

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

请问下面的哪个模型是自回归模型（ ）

【A.】

【B.】

【C.】

【D.】

【答案】B

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

具体表现为某个观测值与其先前的t-1，t-2，t-q个时刻进入系统的q个随机误差项的线性组合的模型是（ ）。

【A.】AR

【B.】ARIMA

【C.】ARMA

【D.】MA

【答案】D

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：多选题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

ARIMA模型也叫整合自回归平均模型，可以分为（ ）模型、（ ）模型；（ ）模型。

【A.】AR

【B.】Box-Jenkins

【C.】ARMA

【D.】MA

【答案】ACD

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

MA（q）模型的基本假设有（ ）。

【A.】假设仅与有关；

【B.】与无关；

【C.】随机误差项是一个白噪声；

【D.】与有关；

【答案】ABC

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

AR（p）模型的基本假设有（ ）。

【A.】假设仅与有线性关系；

【B.】在已知的条件下，与线性相关

【C.】随机误差项是一个白噪声；

【D.】在已知的条件下，与无关；

【答案】ACD

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

进行ARMA模型建模之前，分析的时间序列必须满足平稳性条件。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

Box-Jenkins法的基本思想是用时间序列过去值和现在的值的线性组合来预测其未来的值。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.2.1

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

平稳时间序列的自相关系数图拖尾, 偏自相关系数图p阶截尾,可以识别为（ ）模型

【A.】MA(p)

【B.】ARIMA(p,q)，q>0

【C.】AR(p)

【D.】ARMA(p,q)，q>0

【答案】C

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

采用AIC准则找最优模型得到如下的结果：The AIC of ARMA(0,0) is 13679.401951273543The AIC of ARMA(0,1) is 13332.135879407648The AIC of ARMA(0,2) is 13289.758498467238The AIC of ARMA(1,0) is 13458.95993448334The AIC of ARMA(1,1) is 13266.651183460379The AIC of ARMA(1,2) is 13241.110101949496The AIC of ARMA(2,0) is 13396.594961413079The AIC of ARMA(2,1) is 13245.885887026114The AIC of ARMA(2,2) is 13242.201821971108请问哪个模型最优（ ）

【A.】ARMA(0,2)

【B.】ARMA(2,2)

【C.】ARMA(1,0)

【D.】ARMA(1,2)

【答案】D

【知识点】04399.8.2.2

【解析】13241.110101949496最小判断题

【题目】

【题型：单选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

偏自相关系数的实质是使得残差的（ ）达到（ ）的k阶AR模型的第k项系数。

【A.】均值；最小

【B.】均值；最大

【C.】方差；最小

【D.】方差；最大

【答案】C

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

平稳时间序列的自相关系数图p阶截尾, 偏自相关系数图拖尾,可以识别为（ ）

【A.】MA(p)

【B.】ARIMA(p,q)，q>0

【C.】AR(p)

【D.】ARMA(p,q)，q>0

【答案】A

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

时间序列经过1阶差分后平稳，其自相关系数图p阶截尾, 偏自相关系数图q阶截尾,可以识别为（ ）。

【A.】MA(p)模型

【B.】ARIMA(p,1,q)模型

【C.】AR(p)模型

【D.】ARMA(p,q)模型

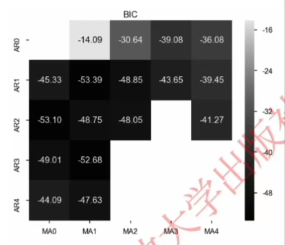
【答案】B

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：单选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

如图是某模型最小信息数值，根据结果可以选择最优模型是（ ）。        【图片】

【A.】MA(1)

【B.】ARMA(1,1)

【C.】AR(1)

【D.】ARMA(0,1)

【答案】B

【知识点】04399.8.2.2

【解析】p=1与q=1的时候值-53.39最小。

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列说法错误的是（ ）。

【A.】偏自相关系数的实质是使得残差的方差达到最大的k阶AR模型的第k项系数。

【B.】相关系数图与偏相关系数图得到的最优模型和 AIC、BIC选择的最优模型不一定一致，需要多次进行建模比较寻优。

【C.】进行ARMA模型建模之前，分析的时间序列必须满足平稳性条件。

【D.】拖尾是指在自相关系数图或偏自相关系数图中，自相关系数或偏自相关系数的前几期处于置信区间之外，而之后是系数基本上都落入置信区间内，且逐渐趋于零。

【答案】AD

【知识点】04399.8.2.2

【解析】A：偏自相关系数的实质是使得残差的方差达到最小的k阶AR模型的第k项系数。D：截尾是指在自相关系数图或偏自相关系数图中，自相关系数或偏自相关系数的前几期处于置信区间之外，而之后是系数基本上都落入置信区间内，且逐渐趋于零。综合题

【题目】

【题型：多选题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

下列说法正确的是（ ）。

【A.】偏自相关系数的实质是使得残差的方差达到最大的k阶AR模型的第k项系数。

【B.】相关系数图与偏相关系数图得到的最优模型和 AIC、BIC选择的最优模型不一定一致，需要多次进行建模比较寻优。

【C.】进行ARMA模型建模之前，分析的时间序列必须满足平稳性条件。

【D.】拖尾是指在自相关系数图或偏自相关系数图中，自相关系数或偏自相关系数的前几期处于置信区间之外，而之后是系数基本上都落入置信区间内，且逐渐趋于零。

【答案】BC

【知识点】04399.8.2.2

【解析】A：偏自相关系数的实质是使得残差的方差达到最小的k阶AR模型的第k项系数。D：截尾是指在自相关系数图或偏自相关系数图中，自相关系数或偏自相关系数的前几期处于置信区间之外，而之后是系数基本上都落入置信区间内，且逐渐趋于零。

【题目】

【题型：多选题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

确定ARMA模型p、q的过程即为模型的识别过程，也称ARMA模型的定阶。下列属于模型识别的方法是（ ）。

【A.】自相关系数

【B.】协方差矩阵

【C.】偏自相关系数图

【D.】最小信息准则

【答案】ACD

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

拖尾是指在自相关系数图或偏自相关系数图中，自相关系数或偏自相关系数的前几期处于置信区间之外，而之后是系数基本上都落入置信区间内，且逐渐趋于零。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

自相关系数的是为给定中间观测值的条件下，观测值与前面某个间隔的观测值之间的相关系数。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

偏自相关系数的实质是使得残差的方差达到最小的k阶AR模型的第k项系数。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

偏自相关系数的实质是使得残差的方差达到最大的k阶AR模型的第k项系数。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关系数图与偏相关系数图得到的最优模型和 AIC、BIC选择的最优模型一定一致。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.2.2

【解析】无综合应用题（客观化处理）

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

拖尾是指在自相关系数图或偏自相关系数图中的系数有指数型、正弦型或震荡型衰减的波动，并不会都落入置信区间内。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

截尾是指在自相关系数图或偏自相关系数图中，自相关系数或偏自相关系数的前几期处于置信区间之外，而之后是系数基本上都落入置信区间内，且逐渐趋于零。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

截尾是指在自相关系数图或偏自相关系数图中的系数有指数型、正弦型或震荡型衰减的波动，并不会都落入置信区间内。

【A.】√

【B.】×

【答案】B

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

相关系数图与偏相关系数图得到的最优模型和 AIC、BIC选择的最优模型不一定一致，需要多次进行建模比较寻优。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：易】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

确定ARMA模型p、q的过程即为模型的识别过程，也称ARMA模型的定阶。

【A.】√

【B.】×

【答案】A

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：判断题】【难度：中】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

偏自相关系数的是为给定中间观测值的条件下，观测值与前面某个间隔的观测值之间的相关系数。

【A.】√

【B.】×

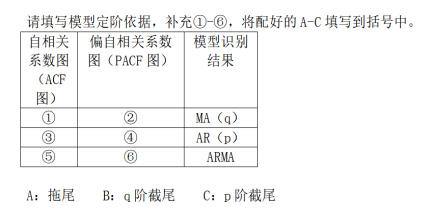
【答案】A

【知识点】04399.8.2.2

【解析】无

【题目】

【题型：配伍题】【难度：难】【考核能力：应用】【漏选得分：是】【自动判分：是】【应用类别：终考】【特殊标签：】

【图片】

【A.】① -> 拖尾

【B.】② -> q阶截尾

【C.】③ -> p阶截尾

【D.】④ ->

【E.】⑤ ->

【F.】⑥ ->

【答案】B,A,A,C,A,A

【知识点】04399.8.2.2

【解析】无