北京石油化工学院 硕士研究生课程期末考试试题

2024 — 2025 学年 秋 学期 (闭卷 考试)

课程名称: 应用数理统计 课程编号： Math501

所有试题答案写在答题纸上，答案写在试卷上无效

说明：下列试题均用软件计算，所有问题均要求提供电子答案（包含源代码、输出结果和分析结果），发送至云班课，文件命名：姓名+专业+学号。

1、（本题满分15分）某饮料公司研制出一款新产品，为比较消费者对新旧产品口感的满意程度，随机抽选一组消费者共10人，让每个消费者先品尝一款饮料，再品尝另一款饮料，两款饮料的品尝顺序是随机的，而后每个消费者要对两款饮料分别进行评分（0～10分），评分结果如下表所示。取显著性水平，检验消费者对两款饮料的评分是否有显著差异。

两款饮料评分表

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 旧款饮料 | | | | | 新款饮料 | | | | | |
| 7 | 8 | 6 | 7 | 8 | | 8 | 10 | 9 | 9 | 9 |
| 8 | 7 | 8 | 7 | 9 | | 8 | 9 | 9 | 10 | 7 |

2、（本题满分15分）为分析小麦品种和施肥方式对产量的影响，一家研究机构挑选了3个小麦品种：品种1、品种2、品种3和甲、乙两种施肥方式，然后选择条件和面积相同的30个地块，每个品种在10个地块上试种。除了考虑品种对产量的影响外，还考虑施肥方式对产量的影响，三个品种和两种施肥方式的搭配共有种组合。30地块进行实验，每一种搭配可以做5次实验，也就是每个品种的样本量为5，即相当于每个品种重复做了5次实验。实验取得的数据下表所示。检验品种和施肥方式对产量的影响是否显著()。

|  |  |  |  |
| --- | --- | --- | --- |
| 施肥方式 | 品种1 | 品种2 | 品种3 |
| 甲 | 729 | 639 | 684 |
| 甲 | 738 | 648 | 711 |
| 甲 | 711 | 648 | 693 |
| 甲 | 729 | 594 | 684 |
| 甲 | 702 | 648 | 702 |
| 乙 | 801 | 693 | 801 |
| 乙 | 828 | 729 | 783 |
| 乙 | 783 | 693 | 756 |
| 乙 | 765 | 657 | 783 |
| 乙 | 774 | 711 | 783 |

3、（本题满分15分）餐馆的营业额受多种因素影响，比如客流量、价格、交通便捷程度、服务水平、同业竞争者的数量等。为分析营业额的影响因素，一家市场调查公司在某城市随机抽取25家餐馆，调查得到的有关数据如下表所示。建立多元线性回归模型，并解释各回归系数的含义。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 餐馆编号 |  |  |  |  |  |  |
| 1 | 53.2 | 163 | 168.6 | 6004 | 5 | 6.5 |
| 2 | 18.5 | 14.5 | 22.5 | 209 | 11 | 16 |
| 3 | 11.3 | 88.2 | 109.4 | 1919 | 10 | 18.2 |
| 4 | 84.7 | 151.6 | 277 | 7287 | 7 | 10 |
| 5 | 7.3 | 79.1 | 17.4 | 5311 | 15 | 17.5 |
| 6 | 17.9 | 60.4 | 93 | 6109 | 8 | 3.6 |
| 7 | 2.5 | 53.2 | 21.5 | 4057 | 17 | 18.5 |
| 8 | 27.3 | 108.5 | 114.5 | 4161 | 3 | 4 |
| 9 | 5.9 | 48.7 | 61.3 | 2166 | 10 | 11.6 |
| 10 | 23.9 | 142.8 | 129.8 | 11125 | 9 | 14.2 |
| 11 | 69.4 | 214.7 | 159.4 | 13937 | 2 | 2.5 |
| 12 | 20.6 | 65.6 | 91 | 4000 | 18 | 12 |
| 13 | 1.9 | 13.2 | 6.1 | 2841 | 14 | 12.8 |
| 14 | 3 | 60.9 | 60.3 | 1273 | 26 | 7.8 |
| 15 | 7.3 | 21.2 | 51.1 | 2404 | 34 | 2.7 |
| 16 | 46.2 | 114.3 | 73.6 | 6109 | 12 | 3.2 |
| 17 | 78.8 | 299.5 | 171.7 | 15571 | 4 | 7.6 |
| 18 | 11.1 | 78.9 | 38.8 | 4228 | 11 | 11 |
| 19 | 8.6 | 90 | 105.3 | 3772 | 15 | 28.4 |
| 20 | 48.9 | 160.3 | 161.5 | 6451 | 5 | 6.2 |
| 21 | 22.1 | 84 | 122.6 | 3275 | 9 | 10.8 |
| 22 | 11.1 | 78.9 | 38.8 | 4228 | 10 | 33.7 |
| 23 | 8.6 | 90 | 105.3 | 3772 | 14 | 16.5 |
| 24 | 48.9 | 160.3 | 161.5 | 6451 | 6 | 9.3 |
| 25 | 22.1 | 84 | 122.6 | 3275 | 10 | 11.6 |

4、（本题满分15分）下表给出了 2014 - 2020 年金砖国家和七国集团总发电量数据（单位：太瓦时(10亿度)）。根据这些数据，试在最短距离法、最长距离法、离差平方和法、重心法、中间距离法5种方法中选择至少3种方法把这些地区聚成4类，并分别指明4类中各元素。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 国家\年份 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
| 中国 | 5794.5 | 5814.6 | 6133.2 | 6604.4 | 7166.1 | 7503.4 | 7779.1 |
| 俄罗斯 | 1064.2 | 1067.5 | 1091.0 | 1091.2 | 1109.2 | 1118.1 | 1085.4 |
| 印度 | 1262.2 | 1317.3 | 1401.7 | 1471.3 | 1579.2 | 1603.7 | 1560.9 |
| 巴西 | 590.5 | 581.2 | 578.9 | 589.3 | 601.4 | 626.3 | 620.1 |
| 南非 | 254.8 | 250.1 | 252.7 | 255.1 | 256.3 | 252.6 | 239.5 |
| 美国 | 4363.3 | 4348.7 | 4347.9 | 4302.5 | 4461.6 | 4411.2 | 4286.6 |
| 加拿大 | 647.6 | 659.3 | 663.7 | 660.1 | 655.8 | 648.7 | 643.9 |
| 德国 | 627.8 | 647.6 | 649.7 | 652.9 | 642.9 | 609.4 | 571.9 |
| 法国 | 564.9 | 571.8 | 556.2 | 554.0 | 574.0 | 562.8 | 524.9 |
| 英国 | 338.1 | 338.9 | 339.2 | 338.2 | 332.8 | 324.8 | 312.8 |
| 意大利 | 279.8 | 283.0 | 289.8 | 295.8 | 289.7 | 293.9 | 282.7 |
| 日本 | 1062.7 | 1030.1 | 1035.1 | 1042.1 | 1053.2 | 1030.3 | 1004.8 |

1. （本题满分15分）在某沉积盆地一坳陷区的17个取样点经勘探测定了六个地质变量： *x*1为有机碳（%）； *x*2为生油层埋深（米）；*x*3油层孔隙度（%）；*x*4为储层厚度（米）；*x* 5为地下水含碘量（p.p.m）；*x*6为地下水矿化度(克/升)。根据这些数据进行主成分分析。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 点号 | *x*1 | *x*2 | *x*3 | *x*4 | *x*5 | *x*6 |
| 1 | -0.9142 | -0.7119 | -0.9293 | -0.4385 | -0.5710 | 0.7361 |
| 2 | -0.3095 | -0.5206 | -1.3309 | -0.2764 | -0.5710 | 0.5714 |
| 3 | -1.0654 | -0.7119 | 0.2756 | -0.7626 | -1.0957 | 0.9007 |
| 4 | -1.3073 | -0.9511 | 1.2574 | 0.3718 | -1.0957 | 1.3946 |
| 5 | 0.1743 | -0.4727 | 0.3203 | -0.9895 | -0.0463 | -0.2518 |
| 6 | -0.8235 | -0.5923 | 0.4095 | 1.3441 | -0.8333 | 0.4068 |
| 7 | 0.9000 | 2.1583 | -0.1260 | -0.8598 | 1.7901 | -1.8983 |
| 8 | -0.0071 | -0.3532 | -1.4201 | -1.0219 | -0.0463 | -0.5811 |
| 9 | 1.2023 | 1.6799 | -0.7508 | -0.6005 | 2.3148 | -1.2397 |
| 10 | 0.1743 | -0.3532 | -0.9739 | 1.3441 | -0.0463 | 0.2421 |
| 11 | 2.2606 | 1.4407 | 0.7219 | 2.6405 | 0.7407 | -1.0750 |
| 12 | -1.4282 | -0.9511 | 0.0079 | -0.7950 | -1.0957 | 1.0653 |
| 13 | -0.3397 | -0.5206 | 2.1499 | -0.1144 | -0.5710 | 0.4068 |
| 14 | 0.7790 | -0.2336 | 1.1970 | 0.6959 | 0.2161 | 0.9104 |
| 15 | 0.4162 | 0.7232 | 1.0789 | -0.3088 | 0.4784 | 0.7457 |
| 16 | -0.6118 | -0.7119 | 0.3649 | 0.0477 | -0.5710 | 1.5593 |
| 17 | 0.9000 | 1.0820 | 0.1418 | -0.2764 | 1.0031 | -0.5811 |

6、（本题满分15分）考虑某四因子二水平试验，除考虑因子外，还需考察交互作用.选用正交表，表头设计及试验数据如表所示，所考虑指标越小越好。试用极差分析方法指出因子的主次顺序和较优工艺条件。



|  |  |  |
| --- | --- | --- |
| 列号  试验号 | A B C D  1 2 3 4 5 6 7 | 实验数据 |
| 1  2  3  4  5  6  7  8 | 1 1 1 1 1 1 1  1 1 1 2 2 2 2  1 2 2 1 1 2 2  1 2 2 2 2 1 1  2 1 2 1 2 1 2  2 1 2 2 1 2 1  2 2 1 1 2 2 1  2 2 1 2 1 1 2 | 1.20  0.87  1.64  0.99  1.21  1.41  0.95  0.99 |

7（本题满分10分）文件7中是学生的期末成绩，试利用合适的图表进行描述性分析。