**2024年硕士研究生入学考试自命题科目**

**考试大纲**

|  |  |
| --- | --- |
| 考试阶段：复试 | 科目满分值：100 |
| 考试科目：人工智能综合 | 科目代码： |
| 考试方式：闭卷笔试 | 考试时长：180分钟 |

**一、科目的总体要求**

理解概率论、统计学的基本概念，熟练掌握基本理论和基本方法；掌握处理随机现象的基本思想和方法，能运用概率统计方法分析和思考解决实际问题；理解掌握机器学习的基本概念，掌握机器学习的基本算法，具有运用机器学习算法分析数据和解释数据的基本能力。

**二、考核内容与考核要求**

《人工智能综合》共包含2个部分的内容：《概率论与数理统计》、《机器学习算法基础》，所占分值为6：4。

（一）第一部分《概率论与数理统计》

1、随机事件和概率

（1）了解样本空间的概念，理解随机事件的概念，掌握事件的关系及运算.

（2）理解概率、条件概率的概念，掌握概率的基本性质，会计算古典型概率和几何型概率，掌握概率的加法公式、减法公式、乘法公式、全概率公式以及贝叶斯(Bayes)公式.

（3）理解事件独立性的概念，掌握用事件独立性进行概率计算;理解独立重复试验的概念，掌握计算有关事件概率的方法.

2、随机变量及其分布

（1）理解随机变量的概念，理解分布函数的概念及性质，会计算与随机变量相联系的事件的概率.

（2）理解离散型随机变量及其概率分布的概念，掌握0-1分布、二项分布、几何分布、超几何分布、泊松(Poisson)分布及其应用.

（3）理解连续型随机变量及其概率密度的概念，掌握均匀分布、正态分布、指数分布及其应用.

（4）会求随机变量函数的分布.

3、多维随机变量及其分布

（1）理解多维随机变量的概念，理解多维随机变量的分布的概念和性质，理解二维离散型随机变量的概率分布、边缘分布和条件分布，理解二维连续型随机变量的概率密度、边缘密度和条件密度，会求与二维随机变量相关事件的概率.

（2）理解随机变量的独立性及不相关性的概念，掌握随机变量相互独立的条件.

（3）掌握二维均匀分布，了解二维正态分布的概率密度，理解其中参数的概率意义.

（4）会求两个随机变量简单函数的分布，会求多个相互独立随机变量简单函数的分布.

4、随机变量的数字特征

（1）理解随机变量数字特征(数学期望、方差、标准差、矩、协方差、相关系数)的概念，掌握数字特征的计算和性质，并掌握常用分布的数字特征.

（2）会求随机变量函数的数学期望.

5、大数定律和中心极限定理

（1）了解切比雪夫不等式.

（2）了解切比雪夫大数定律、伯努利大数定律和辛钦大数定律(独立同分布随机变量序列的大数定律).

（3）了解棣莫弗-拉普拉斯定理(二项分布以正态分布为极限分布)和列维-林德伯格定理(独立同分布随机变量序列的中心极限定理).

6、数理统计的基本概念

（1）理解总体、简单随机样本、统计量、样本均值、样本方差及样本矩的概念.

（2）了解分布、分布和分布的概念及性质，了解上侧分位数的概念并会查表计算.

（3）了解正态总体的常用抽样分布.

7、参数估计

（1）理解参数的点估计、估计量与估计值的概念.

（2）掌握矩估计法(一阶矩、二阶矩)和最大似然估计法.

（3）了解估计量的无偏性、有效性(最小方差性)和一致性(相合性)的概念，并会验证估计量的无偏性.

（4）理解区间估计的概念，会求单个正态总体的均值和方差的置信区间，会求两个正态总体的均值差和方差比的置信区间.

8、假设检验

（1）理解显著性检验的基本思想，掌握假设检验的基本步骤，了解假设检验可能产生的两类错误.

（2）掌握单个及两个正态总体的均值和方差的假设检验.

（二）第二部分《机器学习算法基础》

1. 统计学习的特点、对象、目的、方法和研究

2. 监督学习：基本概念、问题形式

3. 统计学习三要素：模型、策略、算法

4. 误差与模型评估

5. 过拟合与模型选择

6. 正则化与交叉验证

7. 泛化能力

8. 生成模型与判别模型

9. 分类问题、标注问题与回归问题

10. 感知机：适用条件、输入输出、模型、策略、算法及其对偶形式

11. k近邻法：适用条件、输入输出、模型、策略、算法

12. 朴素贝叶斯法：适用条件、前验概率、后验概率、模型、策略、算法

13. 决策树：适用条件、模型、学习过程（特征选择、决策树生成、决策树修剪）、算法（ID3、C4.5、CART）

14. 逻辑斯谛回归模型

**三、题型结构**

考试包含题型：计算题、简答题、应用题等。

1. **参考书目**

《概率论与数理统计》 第五版 浙江大学编 高等教育出版社

《统计学习方法》第二版 李航著 清华大学出版社