**《统计计算与软件》课程教学大纲**

**一、课程简介**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **课程中文名** | **统计计算与软件** | | | | | | |
| **课程英文名** | **Statistical Calculation and Software** | | | | **双语授课** | | □是 否 |
| **课程代码** | **06112181** | **课程学分** | **3** | **总学时数** | | 48（含实践学时8） | |
| **课程类别** | □通识教育课程  □公共基础课程  专业教育课程  □综合实践课程  □教师教育课程 | **课程性质** | □必修  选修  □其他 | **课程形态** | | □线上  线下  □线上线下混合式  □社会实践  □虚拟仿真实验教学 | |
| **考核方式** | □闭卷 □开卷 课程论文 □课程作品 □汇报展示 □报告  □课堂表现 □阶段性测试 □平时作业 □其他 （可多选） | | | | | | |
| **开课学院** | 数学与统计学院 | | **开课系**  **(教研室)** | 统计系 | | | |
| **面向专业** | 数据科学与大数据技术 | | **开课学期** | 第 5 学期 | | | |
| **课程负责人** | 李好奇 | | **审核人** | 叶发强 | | | |
| **先修课程** | 概率论、数理统计 | | | | | | |
| **后续课程** | 机器学习、数据挖掘 | | | | | | |
| **选用教材** | 1. 许王莉 朱利平. 数据科学统计计算 [M]. 北京: 中国人民出版社, 2022. | | | | | | |
| **参考书目** | 1. 李东风. 统计计算 [M]. 北京: 高等教育出版社, 2017.  2. 高惠璇. 统计计算 [M]. 北京: 北京大学出版社, 1995. | | | | | | |
| **课程资源** | 无 | | | | | | |
| **课程简介** | 统计计算是应用性很强的一门课程，在科学研究和生产实践的各个领域都有广泛的应用。对于现实生活的大量数据，如何通过已有的统计方法来解决实际问题，如何解决在应用中出现的计算问题，对实际工作者来说是急需解决的问题。本课程力求把统计思想、数值分析步骤及通过软件在计算机上的实现结合起来，使学生掌握用统计计算方法解决实际问题的能力。  本课程是应用统计学专业必修核心课程，内容包含比较基础的随机数抽样内容，也包括前沿统计算法，如Bootstrap、Jackknife、EM算法等。 | | | | | | |

**二、课程目标**

**表 2-1 课程目标**

|  |  |
| --- | --- |
| **序号** | **具体课程目标** |
| **课程目标 1** | 掌握统计计算的基本理论和方法，具有数据处理、统计分析、统计计算等能力。 |
| **课程目标** 2 | 能够应用统计学基础理论和基本方法，设计数据分析方案，并对统计数据进行有效处理，具备常用统计软件的操作能力和定量分析能力。 |
| **课程目标** 3 | 具有从事数据领域科学研究的初步能力和获取新技术知识的基本能力，具有将统计计算的相关理论和方法运用到生活和生产中的创新意识。 |

**表2-2 课程目标与毕业要求对应关系**

| **毕业要求** | **指标点** | **课程目标** |
| --- | --- | --- |
| **毕业要求5：**能够针对大数据应用领域的复杂工程问题，选择和使用恰当的技术、资源和现代工程工具，进行信息表达、建模、设计、模拟、验证，并能够在实践中了解这些工具使用的局限性。【M】 | 5.3 能够应用现代化工具对大数据应用领域相关工程问题进行分析、设计、模拟和预测，并对结果进行合理评价。 | **课程目标 1** |
| **毕业要求6：**基于大数据应用领域相关背景知识，能够合理分析和评价大数据应用实践和复杂工程问题解决方案对社会、健康、安全、法律以及文化的影响，并理解应承担的责任。【L】 | 6.2 能够分析和评价大数据相关的工程实践和复杂工程问题的解决方案对社会、健康、安全、法律以及文化的潜在影响，以及这些因素对项目实施的影响，并理解应承担的责任。 | **课程目标 2** |
| **毕业要求12：**具有自主学习和终身学习的意识，具备在科学研究与技术应用过程中不断学习和适应发展的能力。【L】 | 12.2 能针对个人或职业发展规划，在复杂工程问题的解决方案中体现出一定的自主学习和终身学习的能力，不断适应大数据技术的发展和社会需求。 | **课程目标 3** |

**三、课程学习内容与方法**

**（一）理论学习内容及要求**

**表3-1 课程目标、学习内容和教学方法对应关系**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **序号** | **课程模块** | **学习内容** | **学习任务** | **课程目标** | **学习重点难点** | **教学方法** | **学时** |
| 1 | 随机数的产生及抽样 | 1.常用连续非均匀随机数的产生 | 1.个人作业：课后习题2.拓展阅读：相关文献 | 课程目标1 | 重点：  1. 随机数的产生  2. 随机数抽样法  难点：  3. 随即向量随机数的抽样法 | **讲授法：**能够引导学生建立知识间的联结，促进学生高效理解知识内涵。 | 8 |
| 2.离散分布随机数抽样法 | 课程目标3 |
| 3.连续随机向量随机数抽样法 | 课程目标2 |
| 4.离散随机向量随机数抽样法 | 课程目标1 |
| 2 | 参数估计和假设检验的数值计算 | 1.点估计数值计算 | 1.个人作业：课后习题2.拓展阅读：相关文献 | 课程目标1 | 重点：  1. 点估计区间估计数值计算  2. 参数检验数值计算  难点：  3. 参数检验数值计算 | **小组讨论：**能够引导学生自学能力的发展 | 12 |
| 2.置信区间估计数值计算 | 课程目标2 |
| 3.参数检验数值计算 | 课程目标3 |
| 3 | 重抽样方法 | Bootstrap估计的思想 | 1.个人作业：课后习题  2.拓展阅读：相关文献 | 课程目标3 | 重点：  1. Bootstrap估计  2.Jackknife法的估计  难点：  3. Jackknife估计 | **实验指导：**促进学生实践实验能力的发展 | 10 |
| 基于Jackknife法的估计 | 课程目标1 |
| 基于Bootstrap法的置信区间估计 | 课程目标2 |
| 4 | EM算法 | EM算法的步骤和原理 | 1.个人作业：课后习题  2.拓展阅读：相关文献 | 课程目标3 | 重点：  1. EM算法  难点：  2. EM算法估计应用 | **查阅文献：**促进学生文献搜索能力的发展 | 10 |
| 几个特殊分布参数的EM算法 | 课程目标2 |

1. **实验学习内容及要求**

**表3-2 课程目标、学习内容和教学方法对应关系**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **序号** | **项目名称** | **项目来源** | **教学目标（观测点、重难点）** | **学时数** | **项目类型** | **要求** | **每组人数** | **教学方法** | **课程目标** |
| 1 | 实验1：随机数的产生及抽样 | 教师  开发 | 学生能够利用计算机，根据需要产生随机数，掌握随机数抽样的编程能力  （观测点：随机数编程实现代码  重难点：随机数产生机制） | 2 | 综合性 | 必做 | 3-6人 | 实验指导小组讨论 | 课程目标1 |
| 2 | 实验2：参数估计和假设检验的数值计算 | 教师  开发 | 学生能够利用计算机，进行数值计算，掌握参数估计和假设检验的计算编程能力  （观测点：参数估计和假设检验计算实现  重难点：数值计算内在机制） | 2 | 综合性 | 必做 | 3-6人 | 实验指导小组讨论 | 课程目标2 |
| **3** | 实验3：重抽样方法 | 教师  开发 | 学生能够利用计算机，进行重抽样，掌握Bootstrap和Jackknife的编程能力  （观测点：Bootstrap和Jackknife代码实现  重难点：Bootstrap和Jackknife内在原理） | **2** | 设计研究 | 必做 | 3-6人 | 实验指导小组讨论 | 课程目标3 |
| **4** | 实验4：EM算法 | 教师  开发 | 学生能够利用计算机，进行EM计算，掌握EM算法解决实际问题的编程能力  （观测点：EM算法代码实现  重难点：EM算法内在原理） | **2** | 设计研究 | 必做 | 3-6人 | 实验指导小组讨论 | 课程目标2 |

## 四、课程考核

**（一）考核内容与考核方式**

**表4-1 课程目标、考核内容与考核方式对应关系**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **课程目标** | **考核内容** | **所属**  **学习模块/项目** | **考核占比** | **考核方式** |
| 课程  目标 1 | 1.常用连续非均匀随机数的产生 | 1.随机数的产生及抽样 | 30% | 课堂表现  课程论文  平时作业 |
| 2.离散随机向量随机数抽样法 | 1.随机数的产生及抽样 |
| 3.点估计数值计算 | 2.参数估计和假设检验的数值计算 |
| 课程  目标 2 | 1. 连续随机向量随机数抽样法 | 1.随机数的产生及抽样 | 30% | 课堂表现  平时作业  课程论文 |
| 2. 置信区间估计数值计算 | 2.参数估计和假设检验的数值计算 |
| 3. 基于Jackknife法的估计 | 3.重抽样方法 |
| 4. 基于Bootstrap法的置信区间估计 | 3.重抽样方法 |
| 5. 几个特殊分布参数的EM算法 | 4. EM算法 |
| 课程  目标 3 | 1.离散分布随机数抽样法 | 1.随机数的产生及抽样 | 40% | 课堂表现  课程论文  平时作业 |
| 2.参数检验数值计算 | 2.参数估计和假设检验的数值计算 |
| 3.Bootstrap估计的思想 | 3.重抽样方法 |
| 4.EM算法的步骤和原理 | 4. EM算法 |

**表4-2 课程目标与考核方式矩阵关系**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 课程  目标 | 考核方式 | | | | 考核占比 |
| 期末考试成绩比例60% | 课堂表现成绩比例12% | 作业完成情况成绩比例12% | 课程论文成绩比例16% |
| 课程目标1 | 30% | 30% | 40% | 30% | 31.2%=60%\*30%+12%\*30%+12%\*40%+16%\*30% |
| 课程目标2 | 30% | 30% | 30% | 30% | 30.0%=60%\*30%+12%\*30%+12%\*30%+16%\*30% |
| 课程目标3 | 40% | 40% | 30% | 40% | 38.8%=60%\*40%+12%\*40%+12%\*30%+16%\*40% |

## （二）成绩评定

**1.平时成绩评定**

**（1）课堂表现（30%）**：通过学生在课堂上的出勤、表现、发言与提问情况，来评价学生相关的能力。

**（2）作业情况（30%）**：围绕课程的学习目标进行作业的设计。如让学生简述对知识的认识，考核学生对于概念的理解情况，帮助学生将定义转化为自己的理解。

**（3）课程论文（40%）**：学生收集资料能力，研究设计能力，解决实际问题能力和合作研究能力。

平时总成绩（100%）=课堂表现（30%）+作业情况（30%）+课程论文（40%）

**2.期末成绩评定**

期末考核主要考察学生对基本概念、操作程序和具体方法的理解与运用等。方式为开卷或闭卷考试等。要求学生掌握基本概念、操作程序，运用具体方法解决相关问题。

**3.总成绩评定**

总成绩应由平时考核成绩和期末考核成绩构成，其构成比例应科学合理。

总成绩（100%）=平时成绩（40%）+期末成绩（60%）

## （三）评分标准

**表4-3 评分标准（非试卷考核项目）**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **考核项目** | **评分标准** | | | | |
| **优秀**  **(100>x≥90)** | **良好**  **(90> x≥80)** | **中等**  **(80> x≥70)** | **及格**  **(70> x≥60)** | **不及格**  **(x <60)** |
| 课程论文 | （1）论文选题符合课程性质，选题范围适中，具有较高的研究价值和意义，表现出很强的问题意识。（2）论证过程严谨，所使用的证据或材料充分，结论清晰，具有相当的说服力和解释力。（3）文章结构合理，组织严密，连贯一致。（4）语言表达准确，叙述清楚，所使用的教育专业术语规范。（5）论文符合学术规范。 | （1）论文选题恰当合理，具有较高的研究价值和意义，表现出较强的问题意识。（2）论证过程较为严谨，所使用的证据或材料较为充分，结论清晰，具有较强的说服力和解释力。（3）文章结构合理，组织较为严密，连贯一致。（4）语言表达较为准确，叙述清楚，所使用的教育专业术语较为规范。（5）论文基本符合学术规范，无明显错误。 | （1）论文选题较为合理，具有一定的研究价值和意义，表现出一定的问题意识。（2）论证过程具有一定的严谨性，所使用的证据或材料较为充分，结论清晰，具有一定的说服力和解释力。（3）文章结构较为合理，组织较为严密。（4）语言表达较为准确，叙述较为清楚，所使用的教育专业术语较为规范。  （5）论文基本符合学术规范，有部分错误。 | （1）论文主题具有一定的研究价值和意义，但选题凝练不够，问题意识欠佳。（2）论证过程较为合理但不太严谨，具有一定的证据或材料但不够充分，结论基本清晰。（3）文章结构较为合理，组织具有一定的严密性，但存在部分不连贯现象。（4）语言表达基本清楚，所使用的教育专业术语基本规范。（5）论文基本符合学术规范，有部分错误。 | （1）论文选题不符合课程性质，或主题不明确（2）论证过程随意，所使用的证据或材料极其不充分，结论不清晰。（3）文章结构混乱，存在前后不连贯现象。（4）语言不通顺，所使用的教育专业术语不规范。（5）论文明显不符合学术规范，或存在抄袭现象。 |

## 五、其它说明

本课程大纲依据2023版应用统计学专业人才培养方案，由数学与统计学院统计教学系讨论制定，数学与统计学院教学工作委员会审定，教务处审核批准，自2023级开始执行。