

806 统计综合考试大纲

一、考试性质

统计综合考试是南京医科大学为招收学术型生物统计专业研究生而设置的具有选拔性质的入学考试科目。

二、考查目标

统计综合的考试目的是科学、公平、有效地测试考生是否具有攻读应用统计学理学硕士学位所需要的医学(生物)统计学基础知识和能力,以利于学校择优选拔,确保硕士研究生的招生质量。要求考生系统掌握概率论与数理统计学、医学(生物)统计学、多元统计学的基本理论、基本知识和基本技能,并运用所学知识分析和解决实际问题。

三、考试形式和试卷结构

1. 试卷满分及考试时间

满分 150 分,时间为 180 分钟。

2. 答题方式

闭卷、笔试。

3. 试卷内容结构

概率论与数理统计学 约 20%

医学统计学 约 40%

多元统计学 约 40%

4. 试卷题型结构

简答题: 10 题,每题 10 分

分析题: 2 题,每题 25 分

四、考查内容

(一) 概率论与数理统计学

1. 概率论

(1) 随机事件和概率

(2) 随机变量及分布: 常见离散型分布、常见连续型分布

- (3) 多维随机变量及其分布
- (4) 随机变量的数字特征：数学期望、方差和协方差
- (5) 大数定律和中心极限定理

2. 数理统计

- (1) 总体、样本、统计量
 - (2) 抽样分布
 - (3) 点估计的概念、矩估计、极大似然估计
- (二) 医学(生物)统计学

1. 绪论及平均数

- (1) 常用统计学的基本概念：同质与异质、变异、总体与样本、参数与统计量、概率、小概率事件和原理、变量的分类
- (2) 频数图和频数表
- (3) 常用描述集中分布的指标及其应用范围

2. 个体变异

- (1) 描述离散趋势的指标及其适用范围
- (2) 定性资料和等级资料的特点
- (3) 相对数及其正确应用
- (4) 率的标准法

3. 正态分布及其应用

- (1) 正态分布的定义及图形
- (2) 正态分布的特征
- (3) 正态分布的曲线下面积
- (4) 标准正态分布
- (5) 利用正态分布估计频数分布范围
- (6) 参考值范围的定义
- (7) 确定参考值范围的原则
- (8) 参考值范围的计算

4. 抽样误差及可信区间

- (1) 抽样误差的概念
- (2) 抽样误差的重要性

- (3) 中心极限定理
- (4) 标准误
- (5) t 分布及其规律性
- (6) 其它的抽样分布： χ^2 分布、 F 分布
- (7) 均数的可信区间
- (8) 均数之差的可信区间
- (9) 可信区间的确切含义
- (10) 可信区间的要素

5. 假设检验

- (1) 假设检验的作用
- (2) 假设检验的基本原理
- (3) 假设检验的基本步骤
- (4) 样本均数与总体均数比较的 t 检验
- (5) 成组 t 检验
- (6) 配对 t 检验
- (7) t 检验的应用条件
- (8) I 类错误、II 类错误和检验效能、两类错误间的关系
- (9) 单双侧检验的区别，如何正确使用单双侧检验
- (10) 假设检验的注意事项
- (11) 可信区间与假设检验的关系

6. 方差分析

- (1) 单因素方差分析的基本原理和基本步骤、方差分析表
- (2) 两因素方差分析的基本原理和基本步骤
- (3) 两两比较
- (4) 方差分析的应用条件

7. 二项分布与 Poisson 分布

- (1) 二项分布的特点和图形
- (2) 率抽样误差的规律性
- (3) Poisson 分布的特点和图形
- (4) 事件数抽样误差的规律性

- (5) 率的假设检验和区间估计
- (6) 事件数的假设检验和区间估计

8. χ^2 检验

- (1) χ^2 检验的基本原理
- (2) 四格表资料 χ^2 检验的基本步骤
- (3) 四格表资料 χ^2 检验的校正
- (4) 两样本率比较的 u 检验与 χ^2 检验的关系
- (5) 四格表的确切概率法
- (6) 行 \times 列表的 χ^2 检验
- (7) 配对四格表的 χ^2 检验

9. 秩和检验

- (1) 非参数统计
- (2) 成组设计两样本比较的秩和检验及其基本原理
- (3) 多组设计资料的秩和检验及两两比较
- (4) 配对秩和检验
- (5) 秩和检验的正确应用

10. 两指标间的直线相关

- (1) 直线相关系数的计算
- (2) 直线相关的图形
- (3) 相关系数的假设检验
- (4) 总体相关系数的区间估计
- (5) 四格表资料的相关系数的计算及其假设检验
- (6) 等级资料的相关

11. 两指标间的线性回归

- (1) 直线回归的定义
- (2) 直线回归方程的形式
- (3) 直线回归方程的求解、最小二乘法
- (4) 直线回归系数的 t 检验
- (5) 直线回归方程的方差分析
- (6) 总体回归系数的可信区间

- (7) 条件总体均数的可信区间
- (8) 给定 X 时个体的容许区间
- (9) 过定点的直线回归
- (10) 两回归线的比较
- (11) 直线相关与回归的关系
- (12) 回归分析的正确应用

12. 实验设计

- (1) 研究设计的作用
- (2) 研究设计的类型
- (3) 研究设计的基本要素：研究因素与混杂因素、受试对象、实验效应
- (4) 对照组及其均衡性
- (5) 随机化及其实现
- (6) 样本含量的影响因素
- (7) 样本含量的计算方法
- (8) 完全随机设计的概念及其特点
- (9) 配对设计和随机区组设计的概念及其特点
- (10) 拉丁方设计的概念及其特点
- (11) 析因设计的概念及其特点

(三) 多元统计学

1. 绪论、多元正态分布

- (1) 多元分析常用统计量
- (2) 多元正态分布
- (3) 多元配对、成组设计 T 检验

2. 多元方差分析

- (1) 多元成组设计方差分析
- (2) 多元区组设计方差分析
- (3) 多元析因设计方差分析
- (4) 多元方差分析的正确应用

3. 多重线性回归

- (1) 多重线性回归模型简介

- (2) 回归系数的估计和方程的假设检验
 - (3) 决定系数与剩余标准差
 - (4) 偏回归系数的假设检验与区间估计
 - (5) 标准偏回归系数与自变量的贡献
 - (6) 逐步回归
 - (7) 回归分析的正确应用
4. 主成分分析和因子分析
- (1) 主成分的定义
 - (2) 主成分的计算
 - (3) 主成分的性质
 - (4) 主成分的应用如人群分层、岭回归、综合评价等
 - (5) 有关的统计推断
 - (6) 主成分分析的正确应用
 - (7) 因子模型
 - (8) 因子模型的估计
5. logistic 回归
- (1) 多重 logistic 回归模型
 - (2) 配比设计的条件 logistic 回归
 - (3) 多类结果变量的 logistic 回归
 - (4) 有序结果的累积比数 logistic 回归
 - (5) 有序结果的相邻比数 logistic 回归模型
 - (6) logistic 族回归模型的正确应用
6. 聚类和判别分析
- (1) 聚类的目的与方法
 - (2) 距离和相似系数
 - (3) 系统聚类法、动态聚类
 - (4) 聚类分析的正确应用
 - (5) 距离判别
 - (6) Bayes 判别
 - (7) 判别分析的正确应用

7. 生存分析

- (1) 生存资料的描述与组间比较
- (2) 生存资料的参数模型
- (3) 生存资料的非参数模型
- (4) 生存资料模型的结果
- (5) 生存资料模型的区别与应用条件