

清华大学 432 《统计学》考试大纲

2026.9

考查目标

432《统计学》是清华大学非全日制应用统计硕士专业学位研究生入学统一考试的科目之一，同时也部分适用于其他选择该科目作为考试科目的专业。本考试旨在科学、公平、有效地考查考生是否具备攻读非全日制应用统计专业硕士及相关专业硕士研究生所必需的基本素养、学术潜力和综合能力。本考试结合清华大学统计与数据科学系在数理统计、生物统计、金融统计、机器学习与大数据分析和人工智能等领域的学科优势，特别注重考查考生对概率论与统计学基本概念的正确理解，以及对统计学核心理论的理解深度、逻辑推理的严密性，以及运用统计思维分析和解决实际问题的综合能力。考试要求考生掌握概率论与数理统计的基础理论与核心方法，具备良好的数学推导能力和统计建模素养，能够将统计方法应用于数据驱动的决策分析之中。

考试形式

一、**试卷满分及考试时间**：试卷满分为 150 分，考试时间为 180 分钟。

二、**答题方式**：答题方式为闭卷、笔试。禁止使用计算器。

三、题型结构

1. **概念题**：着重考查考生对概率论与统计学基本概念、定义和原理的准确理解与正确辨析能力；
2. **计算题**：着重考查考生运用统计方法进行数据分析和解决实际问题的综合能力；
3. **证明题**：着重考查考生对统计理论方法的逻辑推导与论证能力；
4. **综合应用题**：着重考查考生综合运用概率论与统计学方法，对实际问题进行统计建模、数据分析和决策推断的能力。

试卷内容结构

概率论 约 50 分

统计学 约 100 分

考试内容

一、概率论

1. 概率论基础

样本空间与事件的定义；事件的关系、运算及性质；概率的公理化定义及其基本推论；条件概率、乘法公式、全概率公式与贝叶斯公式及其应用；事件的独立性及其性质。

2. 随机变量与分布

随机变量的定义及分类；离散型随机变量的概率分布列与累积分布函数；连续型随机变量的概率密度函数与累积分布函数；常见离散分布：二项分布、泊松分布、几何分布、负二项分布、超几何分布的定义、性质、均值、方差及矩母函数；常见连续分布：均匀分布、正态分布、指数分布、伽马分布、贝塔分布、柯西分布、对数正态分布、双指数分布的定义、性质、均值、方差及矩母函数；指数族分布的形式与基本性质。

3. 随机向量

随机向量的联合分布函数与联合概率质量/密度函数；边际分布与条件分布的计算；随机变量的独立性及其判定；随机向量函数的分布（包括和、差、乘积、商的分布）；次序统计量的分布；协方差与相关系数的定义、性质与计算；多项分布的定义、性质与协方差结构。

4. 数字特征与变换

期望、方差、协方差、相关系数的定义、性质与计算；条件期望与条件方差及其性质；随机变量的变换及其分布。

5. 极限定理

依概率收敛与依分布收敛的定义及其关系；大数定律（弱大数定律）及其应用；中心极限定理及其在近似计算中的应用；Delta Method 的基本思想与应用。

二、统计学

1. 随机样本与抽样分布

总体与样本的概念；随机样本（iid 样本）的定义；统计量的定义；样本均值与样本方差的定义及性质（期望与方差）；抽样分布的基本概念；三大抽样分布： χ^2 分布、 t 分布与 F 分布的定义、导出、性质及其相互之间的关系；正态总体下常用统计量的抽样分布：样本均值的精确分布、样本方差的分布、样本均值与样本方差的独立性及其证明、正态总体下样本均值和样本方差的联合分布；两个正态总体下样本均值差与样本方差比的抽样分布；次序统计量的联合分布与边际分布。

2. 数据缩减原理

充分统计量的定义与直观含义；因子分解定理及其在判定充分统计量中的应用；完备统计量的定义；指数族分布及其充分完备统计量的识别。

3. 点估计

估计方法：矩估计法的原理与计算；极大似然估计法的原理、计算与不变性；贝叶斯估计法的基本思想与后验均值的计算。

评价准则：无偏性的定义；有效性的概念；均方误差（MSE）的分解（方差+偏差²）及其在估计量比较中的应用；一致最小方差无偏估计（UMVUE）的概念；Cramér-Rao 不等式（信息不等式）的陈述、适用条件及其在确定方差下界中的应用；Rao-Blackwell 定理及其在改进无偏估计中的应用；完备充分统计量与 UMVUE 的关系（利用完备充分统计量构造 UMVUE）。

4. 假设检验

基本概念：原假设与备择假设；检验统计量与拒绝域；两类错误（Type I Error 与 Type II Error）及其概率；显著性水平与检验的势函数（功效函数）的定义与计算； p 值的定义与计算。

检验方法：似然比检验（LRT）的构造、统计量及拒绝域确定；与正态总体有关的各种假设检验。

5. 区间估计

基本概念：区间估计的定义；置信系数与覆盖率；枢轴量法构造置信区间的基本原理。

构造方法：枢轴量法。

最优性：最短置信区间的基本思想（在给定置信水平下最小化区间长度）。

6. 渐近评价

相合性：点估计量相合性的定义；均方误差收敛与相合性的关系（充分条件）；MLE 的相合性（结论）。

渐近正态性与有效性：渐近正态分布的定义；Delta Method 及其在导出函数渐近分布中的应用；MLE 的渐近有效性与渐近方差（基于 Fisher 信息）；渐近相对效率（ARE）的概念及其在估计量比较中的应用。

参考书

1. Casella, G. & Berger, R. L. (2024). *Statistical Inference* (2nd ed.). CRC Press. 教材核心内容覆盖考试范围，不包括第 11 章和第 12 章。
2. 茆诗松，程依明，濮晓龙，倪蓓. 概率论与数理统计教程（第 4 版）. 高等教育出版社。

本大纲规定范围为考试主要内容，少量内容略超出此范围属正常。参考教材供备考参考，非指定教材。考试范围、内容和参考书如有调整，以清华大学统计与数据科学系后续发布的官方公告为准。