

全国重点名校系列

新版

全国硕士研究生招生考试 考研专业课精品资料

【电子书】2024年北京大学

838社会学研究方法和分支学科研究方法之社会
统计学考研精品资料

策划：辅导资料编写组

真题汇编 直击考点
考研笔记 突破难点
核心题库 强化训练
模拟试题 查漏补缺

高分学长学姐推荐



【初试】2024 年北京大学 838 社会学研究方法和分支学科研究方法之社会统计学考研精品资料

说明：本套资料由高分研究生潜心整理编写，高清 PDF 电子版支持打印，考研首选资料。

一、重点名校考研真题汇编**1. 附赠重点名校：社会统计学考研真题汇编（暂无答案）**

说明：本科目没有收集到历年考研真题，赠送重点名校考研真题汇编，因不同院校真题相似性极高，甚至部分考题完全相同，建议考生备考过程中认真研究其他院校的考研真题。

二、2024 年北京大学 838 社会学研究方法和分支学科研究方法考研资料**2. 《社会统计学》考研复试相关资料****（1）《社会统计学》[笔记+提纲]****①北京大学 838 社会学研究方法和分支学科研究方法之《社会统计学》考研复习笔记。**

说明：本书重点复习笔记，条理清晰，重难点突出，提高复习效率，基础强化阶段首选资料。

②北京大学 838 社会学研究方法和分支学科研究方法之《社会统计学》复习提纲。

说明：该科目复习重难点提纲，提炼出重难点，有的放矢，提高复习针对性。

（2）《社会统计学》考研题库[仿真+强化+冲刺]**①2024 年北京大学 838 社会学研究方法和分支学科研究方法之社会统计学考研专业课五套仿真模拟题。**

说明：严格按照本科目最新专业课真题题型和难度出题，共五套全仿真模拟试题含答案解析。

②2024 年北京大学 838 社会学研究方法和分支学科研究方法之社会统计学考研强化五套模拟题及详细答案解析。

说明：专业课强化检测使用。共五套强化模拟题，均含有详细答案解析，考研强化复习必备。

③2024 年北京大学 838 社会学研究方法和分支学科研究方法之社会统计学考研冲刺五套模拟题及详细答案解析。

说明：专业课冲刺检测使用。共五套冲刺预测试题，均有详细答案解析，最后冲刺必备资料。

三、电子版资料全国统一零售价**3. 本套考研资料包含以上一、二部分（高清 PDF 电子版，不含教材），全国统一零售价：[¥]****特别说明：**

①本套资料由本机构编写组按照考试大纲、真题、指定参考书等公开信息整理收集编写，仅供考研复习参考，与目标学校及研究生院官方无关，如有侵权、请联系我们将立即处理。

②资料中若有真题及课件为免费赠送，仅供参考，版权归属学校及制作老师，在此对版权所有者表示感谢，如有异议及不妥，请联系我们，我们将无条件立即处理！

四、2024 年研究生入学考试指定/推荐参考书目（资料不包括教材）**4. 北京大学 838 社会学研究方法和分支学科研究方法考研初试参考书**

《社会研究方法教程》，袁方，王汉生，北京大学出版社；

《社会统计学》，卢淑华，北京大学出版社；

《社会调查研究的量化方法》，郭志刚，中国人民大学出版社；

- 《人口社会学》，佟新著，北京大学出版社；
《劳动社会学》，刘艾玉，北京大学出版社；
《群体过程》，RupertBrown，方文校，中国轻工业出版社；
《城市社会学》，郑也夫，上海交通大学出版社；

考研云分享
kaoyany.top

版权声明

编写组依法对本书享有专有著作权，同时我们尊重知识产权，对本电子书部分内容参考和引用的市面上已出版或发行图书及来自互联网等资料的文字、图片、表格数据等资料，均要求注明作者和来源。但由于各种原因，如资料引用时未能联系上作者或者无法确认内容来源等，因而有部分未注明作者或来源，在此对原作者或权利人表示感谢。若使用过程中对本书有任何异议请直接联系我们，我们会在第一时间与您沟通处理。

因编撰此电子书属于首次，加之作者水平和时间所限，书中错漏之处在所难免，恳切希望广大考生读者批评指正。

目录

封面.....	1
目录.....	5
2024 年北京大学 838 社会学研究方法和分支学科研究方法备考信息.....	8
北京大学 838 社会学研究方法和分支学科研究方法考研初试参考书目.....	8
2024 年北京大学 838 社会学研究方法和分支学科研究方法考研核心笔记	9
《社会统计学》考研核心笔记	9
第 1 章 社会学研究与统计分析	9
考研提纲及考试要求	9
考研核心笔记.....	9
第 2 章 单变量统计描述分析	15
考研提纲及考试要求	15
考研核心笔记.....	15
第 3 章 概率	21
考研提纲及考试要求	21
考研核心笔记.....	21
第 4 章 二项分布及其他离散型随机变量的分布	25
考研提纲及考试要求	25
考研核心笔记.....	26
第 5 章 正态分布、常用统计分布和极限定理	31
考研提纲及考试要求	31
考研核心笔记.....	31
第 6 章 参数估计	39
考研提纲及考试要求	39
考研核心笔记.....	39
第 7 章 假设检验的基本概念	44
考研提纲及考试要求	44
考研核心笔记.....	44
第 8 章 单总体假设检验	46
考研提纲及考试要求	46
考研核心笔记.....	46
第 9 章 二总体假设检验(二分变量—二分变量；二分变量—定距变量).....	49
考研提纲及考试要求	49
考研核心笔记.....	49
第 10 章 列联表(定类变量—定类变量).....	54
考研提纲及考试要求	54
考研核心笔记.....	54

第 11 章 等级相关(定序变量—定序变量).....	68
考研提纲及考试要求	68
考研核心笔记.....	68
第 12 章 回归与相关(定距变量—定距变量).....	71
考研提纲及考试要求	71
考研核心笔记.....	71
第 13 章 方差分析(定类变量—定距变量).....	76
考研提纲及考试要求	76
考研核心笔记.....	76
第 14 章 非参数检验(定类变量—定序变量).....	83
考研提纲及考试要求	83
考研核心笔记.....	83
第 15 章 抽样.....	88
考研提纲及考试要求	88
考研核心笔记.....	88
2024 年北京大学 838 社会学研究方法和分支学科研究方法考研复习提纲	93
《社会统计学》考研复习提纲	93
2024 年北京大学 838 社会学研究方法和分支学科研究方法考研题库[仿真+强化+冲刺]	103
北京大学 838 社会学研究方法和分支学科研究方法之社会统计学考研仿真五套模拟题.....	103
2024 年社会统计学五套仿真模拟题及详细答案解析(一)	103
2024 年社会统计学五套仿真模拟题及详细答案解析(二)	104
2024 年社会统计学五套仿真模拟题及详细答案解析(三)	105
2024 年社会统计学五套仿真模拟题及详细答案解析(四)	107
2024 年社会统计学五套仿真模拟题及详细答案解析(五)	108
北京大学 838 社会学研究方法和分支学科研究方法之社会统计学考研强化五套模拟题.....	110
2024 年社会统计学五套强化模拟题及详细答案解析(一)	110
2024 年社会统计学五套强化模拟题及详细答案解析(二)	111
2024 年社会统计学五套强化模拟题及详细答案解析(三)	113
2024 年社会统计学五套强化模拟题及详细答案解析(四)	115
2024 年社会统计学五套强化模拟题及详细答案解析(五)	116
北京大学 838 社会学研究方法和分支学科研究方法之社会统计学考研冲刺五套模拟题.....	118
2024 年社会统计学五套冲刺模拟题及详细答案解析(一)	118
2024 年社会统计学五套冲刺模拟题及详细答案解析(二)	120
2024 年社会统计学五套冲刺模拟题及详细答案解析(三)	121
2024 年社会统计学五套冲刺模拟题及详细答案解析(四)	123
2024 年社会统计学五套冲刺模拟题及详细答案解析(五)	124
附赠重点名校：社会统计学考研真题汇编(暂无答案)	126
第一篇、2013 年社会统计学考研真题汇编	126

2013 年华中农业大学 612 社会统计学考研专业课真题	126
第二篇、2012 年社会统计学考研真题汇编	130
2012 年华中农业大学 612 社会统计学考研专业课真题	130
第三篇、2011 年社会统计学考研真题汇编	134
2011 年华中农业大学 612 社会统计学考研专业课真题	134
第四篇、2010 年社会统计学考研真题汇编	138
2010 年华中农业大学 612 社会统计学考研专业课真题	138
复旦大学 731 社会统计学考研专业课真题.....	143

考研云分享
kaoyany.top

2024 年北京大学 838 社会学研究方法和分支学科研究方法备考信息

北京大学 838 社会学研究方法和分支学科研究方法考研初试参考书目

- 《社会研究方法教程》，袁方，王汉生，北京大学出版社；
- 《社会统计学》，卢淑华，北京大学出版社；
- 《社会调查研究的量化方法》，郭志刚，中国人民大学出版社；
- 《人口社会学》，佟新著，北京大学出版社；
- 《劳动社会学》，刘艾玉，北京大学出版社；
- 《群体过程》，RupertBrown，方文校，中国轻工业出版社；
- 《城市社会学》，郑也夫，上海交通大学出版社；

考研云分享
kaoyany.top

2024 年北京大学 838 社会学研究方法和分支学科研究方法考研核心笔记

《社会统计学》考研核心笔记

第 1 章 社会学研究与统计分析

考研提纲及考试要求

- 考点：社会统计的方法
- 考点：社会统计学的发展
- 考点：科学的研究逻辑
- 考点：社会科学研究的一般过程
- 考点：单变量和多变量

考研核心笔记

1. 社会统计学的发展

统计学的两大流派：数理统计学派和社会统计学派

(1) 数理统计学派的原创始人是比利时的 A·凯特勒，其最大的贡献就是将法国的古典概率引入统计学，用纯数学的方法对社会现象进行研究；

(2) 社会统计学派的首倡者是德国的 K·克尼斯，他认为统计研究的对象是社会现象，研究方法为大量观察法。

2. 统计学是什么？——概念与内涵

(1) 统计学是描述一系列可用于描述、整理和解释资料或数据的统计工具和技术。

(2) 社会统计学是运用统计的一般原理，对社会各种静态结构与动态趋势进行定量描述或推断的一种专门方法与技术。

(3) 统计学的基本分类：描述统计与推断统计

描述统计：用于整理、描述所收集的数据的特征

推论统计：利用较小群体的数据（样本）来推论可能的较大群体（总体），通过推论统计，我们可以解释数据。

3. 社会统计的方法

(1) 大量观察法

所谓大量观察，即是就总体中足够多的单位进行调查和综合分析，用以反映社会总体的数量特征。

所谓描述统计就是讨论范围仅以搜集资料本身为限，而不予以扩大。

(2) 大数规律

随机现象出现的基本规律，其一般意义是：观察过程中每次取得的结果可能不同(因为具有偶然性)，但大量重复观察结果的平均值却几乎接近某个确定的数值。（青岛掌心博阅电子书）

所谓推论统计，主要是依据概率论，研究如何依据有限资料对总体性质作推断，从而使统计的功能大为扩充。

【核心笔记】社会学研究的科学性

1. 科学的研究逻辑

(1) 通过对理论的演绎建立研究假设

- (2) 操作化
- (3) 科学实验
- (4) 得出结论

2. 社会科学研究的一般过程

- (1) 确定课题、了解情况
- (2) 建立研究假设
- (3) 概念的操作化
- (4) 设计问卷、抽样调查
- (5) 统计分析
- (6) 假设检验、得出结论

概念 (Concept)：人们对事物本质的认识，是研究的基本单位，具有抽象性。

定义 (Definition)：将概念呈现、描述出来，从而明确认识对象。

操作化定义：用一系列可以观察、可以测量的指标来模拟或表述定义，使得非量化的概念得以运算并得以操作测量。

概念	抽象定义	操作化定义
都市化	现代都市的生活形态	妇女就业人数、子女数、交通、通讯手段等
个人现代化	一个人由于经济、工业等现代因素所产生的内部变化	对时间、效率、家庭、亲属、消费、自信等具体问题的看法
子代偏重	经济上、感情上和生活上对子代过分的偏重	子代抚育费与家庭平均消费的比较、子代过生日的花费与长辈生日花费的比较、花在子代闲暇时间的多少等

3. 总体与单位

所谓总体，就是作为统计研究对象的、由许多具有共性的单位构成的整体。总体也有人称之为母体。构成总体的每一个个体称为总体单位，简称单位，也称为个体。

总体在抽样调查及推论统计中，还引出了它与样本总体这个概念的联系与区别。推论统计有“部分推断总体”的特征。通过抽样得到的用以推断总体特征的那个“部分”，在统计学上称为样本或样本总体。样本中所含的单位数，在统计学上称为样本大小，也叫做样本容量。从总体抽选出样本的过程叫抽样，也有叫取样的。

(1) 二者是互为存在条件地联结在一起的，二者的地位在某种意义上不是固定不变的。

- ① 相互依存关系。
- ② 当研究目的和研究范围变动，总体和个体可以相互转化。

(2) 总体的特征

- ① 大量性、同质性和差异性。总体的这三个特征是缺一不可的。
- ② 大量性是条件、同质性是基础，差异性的前提。

(3) 样本

- ① 样本是从总体中抽取的一部分个体所组成的集合，也称子样。
- ② 样本容量是指样本所包含的个体数。当样本容量大于 30 时，为大样本。
- ③ 样本个数是指从总体中最多可以抽取的不同样本的套数。样本容量用 n 表示。样本个数用 m 表示。

(4) 总体与样本的关系

- ① 总体是所要研究的对象，而样本则是所要观测的对象。
- ② 样本是用来推断总体的。
- ③ 总体和样本的角色是可以改变的。

④总体与样本都有大量性，同质性和差异性的特征。样本容量用 n 表示。样本个数用 m 表示。

4. 变量

(1) 变量 (Variables) —— 在统计学中表示研究对象的特征，变量下的各种表现形式或不同取值，就是变量值。

变量	变量值
性别	男；女
家庭子女数	1个；2个；3个；…
重要性	非常重要；一般；不重要
文化程度	大学；中学；小学；文盲

变量取值时注意的问题：完备性互斥性

(2) 单变量分析——变量的分布

分布——某个变量的变量值出现的次数或频次，也称频次分布。

例：家庭结构（变量 X ）可能的取值包括

X_1 ——由父母和未婚子女组成的核心家庭

X_2 ——由父母和单一已婚子女组成的三代人家庭，即主干家庭

X_3 ——由父母和多对已婚子女所组成的三代人家庭，即联合家庭

如果我们将调查结果按照被访人的回答，分别归入上述三类，于是有以下三对数据。

(X_1, n_1)

(X_2, n_2)

(X_3, n_3)

…

(X_k, n_k)

其中 X_i 表示变量 X 的第 i 个取值； n_i 表示该取值所对应的次数
所谓分布，实际上就是这些对数据的集合。

当 n 代表不同的含义时，表示不同的分布

如：当 n 表示概率时，称作概率分布

当 n 表示频次时，称作频次分布

当 n 表示百分比时，称作频率或相对频次分布

家庭结构的频次分布

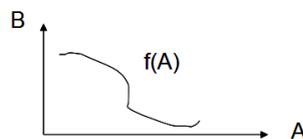
家庭结构	频次
核心家庭	1050
主干家庭	720
联合家庭	110
其 他	250
总 数	2130

家庭结构的百分比分布

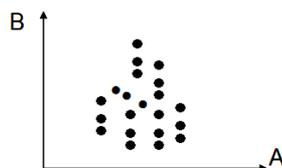
家庭结构	百分比
核心家庭	49.3
主干家庭	33.8
联合家庭	5.2
其 他	11.7
总和百分比	100

【核心笔记】社会调查资料的特点和统计学的运用
1. 社会调查资料的特点
(1) 随机性

确定性现象：“若 A, 则必有 B”，A 与 B 之间，存在确定性的函数关系。



非确定性现象：若 A, 则可能有 B, 或者 C; D; E。


(2) 统计规律性

表面杂乱无章的随机现象，实际上是有内在规律性的。统计规律性表现为大量观察或试验中随机事件发生的频率的稳定性，即一个随机事件发生的频率经常在某个定值附近摆动，而且，试验次数越多，一般摆动越少。

2. 统计学的运用

介绍有关社会调查资料收集、整理、分析和推论的统计方法。

社会统计学的特点：

抽象概念向操作化定义，设计好调查问卷；被测量对象是人，主观意识影响资料收集；低层次变量占较大比重。

3. 统计分析的作用及主要内容

统计分析在社会调查研究中的功用，主要表现在经验层次的大面积数据处理方面。

统计结果不能轻率的作出因果的结论。

【核心笔记】怎样选用统计分析方法
1. 全面调查和非全面调查

(1) 全面调查是对所研究全部对象进行观察与调查，从而掌握整个单位的全部资料。对全面调查，一般可用统计描述。统计描述是将所观察的数据、资料，进行整理、归纳和分析，以期找到某种规律。

(2) 非全面调查是在全部研究对象中只调查其中的一部分。在非全面调查中，以概率抽样法最重要。抽样调查是在一定条件下，科学、客观的抽取总体中一部分单位来加以研究，其目的是获得对于总体性质的正确叙述。

2. 单变量和多变量

(1) 单变量主要调查的内容是一个或多个相互独立的概念，需要对每个独立的变量研究其变动情况和频次，及集中和分散的特征。

(2) 二变量研究两个变量之间是否存在关系及关系密切程度，及研究结论是否可以推断整体。

(3) 多变量研究多种关系，一因多果和一果多因。

(4) 两变量关系——命题的研究

两个变量之间是否存在关系？关系的密切程度（强度）如何？

①自变量：也叫独立变量，由研究者主动操纵，或被研究者界定为不受其他变量影响，独立变化的变

2024 年北京大学 838 社会学研究方法和分支学科研究方法考研复习提纲

《社会统计学》考研复习提纲

第一部分 导论

一、社会调查资料的特点（随时掌握）

随机性、统计规律性；

二、统计学的作用：为社会研究提供数据分析和推论的方法

三、统计分析的作用及其前提。

四、统计分析方法的选择

1、全面调查和抽样调查的分析方法

2、单变量和多变量的统计分析方法

五、不同变量层次的比较：定类、定序、定距、定比

定义、数学特征、运算特性、涵盖关系、等

第二部分 单变量统计描述分析

一、统计图表，熟悉不同层次变量对应的分析图表，不能混淆。尤其是直方图的意义。

二、标明组限与真实组限的换算，重要。

三、集中趋势测量法

1、定义、优缺点、注意事项；

2、众值：定义、计算公式、解释、运用，注意事项；

3、中位值：定义、计算公式（频数和比例两种公式）、解释、运用，注意事项；

4、均值：定义、计算公式（分组与加权）、解释、运用，注意事项；

5、众值、中位值和均值的关系及其相互比较，会用众值和中位值估算均值；

四、离散趋势测量法

1、定义、优缺点、注意事项，与集中趋势的关系；

2、异众比例：定义、计算公式、解释、运用，注意事项；

3、质异指数：定义、计算公式、解释、运用，注意事项；

4、四分位差：定义、计算公式（频数和比例两种公式）、解释、运用，注意事项；要会举一反三，如求十分位差、以及根据数据求其在总体中的位置。

4、方差及标准差：定义、计算公式（分组与加权）、解释、运用，注意事项；

第三部分 概率

一、概率：就是指随机现象发生的可能性大小。随机现象具有不确定性和随机性。

二、概率的性质：

1、不可能事件的概率为0；

2、必然事件的概率为1；

3、随机事件的概率在0-1 之间；

三、概率的计算方法：

1、古典法：计算等概率事件， $P = \text{有效样本点数} / \text{样本空间数}$ ；

2、频率法：求随机事件在多次试验后的极限频率。

3、概率是理论值，只有一个，频率是试验值，不同的试验有不同的频率。

四、概率的运算：会画文氏图

1、加法公式：两个或多个随机事件的求和概率 ‘

2、乘法公式：两个或多个随机时间共同发生的概率。分为独立事件的乘法和条件概率的乘法公式。

(1) 独立： $P(AB) = P(A) * P(B)$

(2) 条件： $P(AB) = P(A) * P(A/B) = P(B) * P(B/A)$

3、条件概率：将(2)反过来即可。 $P(B/A)$ 是指在A 发生的条件下B 发生的概率。

4、全概公式：互不相容的完备事件组，求任意一个事件的发生

5、逆概公式：与4相反。

五、离散型随机变量和连续型随机变量的概率分布及密度函数。

六、数学期望：

1、离散型变量数学期望的计算

2、连续型变量数学期望的计算，可以忽略

3、数学期望的性质，6点，重要

七、方差：

1、简化公式，一个变量的方差等于变量平方的期望减去变量期望的平方。

2、方差的性质，4点，重要，经常在参数估计和假设检验中用到。

第四部分 二项分布

一、二点分布，0-1分布，

1、定义，

2、概率分布、期望、方差

二、二项分布，贝努里分布：

1、定义，

2、概率分布公式

3、期望、方差

4、会求不同条件下的概率，如至多、至少出项多少次？

5、二项分布的讨论

三、多项分布，重点是三项分布，了解。

1、三项分布的公式

2、每个变量的期望和方差，注意n项分布，分别有n-1 个期望和方差。

四、超几何分布：

1、定义，跟二项分布的区别

- 2、概率分布、期望、方差。
- 五、泊松分布
- 1、定义、分布形式
 - 2、期望、方差，与二项分布的关系。
 - 3、应用范围及条件。

第五部分 正态分布和极限定理

- 一、正态分布，常态分布：
- 1、定义、密度分布、性质
 - 2、均值、方差，正态曲线下方面积的意义。
 - 3、正态分布标准化及实际意义。
 - 4、正态分布表的查法（注意对称性）。
- 二、常用统计分布
- 1、卡方分布：定义，自由度，均值、方差，性质，换算。
 - 2、t 分布，定义，自由度，均值、方差，性质，换算。
 - 3、F 分布，定义，自由度，均值、方差，性质，换算。
 - 4、三种常用分布适用范围的比较。
- 三、大数定理
- 1、大数定理的含义
 - 2、切贝谢夫不等式：用于保守估计某事件发生的概率
 - 3、贝努里大数定理。
 - 4、切贝谢夫大数定理。
- 四、中心极限定理：重点
- 1、极限定理的含义。
 - 2、中心极限定理的含义，在何种情况下，何种变量趋向于正态分布。
 - 3、中心极限定理的4个推论，灵活运用。
- 五、二项分布、泊松分布、正态分布三者的近似关系

第六部分 参数估计

- 一、统计推论
- 1、统计推论的定义
 - 2、统计推论的特点（优缺点），考题
 - 3、统计推论的理论基础及内容。
- 二、参数的点估计：（定义）
- 1、什么是点估计（样本中称统计值，总体中称为参数）
 - 2、点估计的评价标准：3点或4点。
 - 3、总体均值的点估计
 - 4、总体方差（或标准差）的点估计
 - 5、总体成数的点估计。

三、抽样分布：统计量的抽样分布，如均值、方差的抽样分布

- 1、样本均值的抽样分布：不同的抽样，其均值是不一样的，在具体抽样之前，均值是一个变量，抽样之后，均值就是一个具体的观察指（或统计值）。
- 2、总体分布为正态分布、总体方差已知情况下的均值分布：可以用中心极限定理推演出来。（一般指小样本）
- 3、总体分布为正态分布，总体方差未知、样本方差可知情况下的均值分布：（一般指小样本）。
- 4、大样本、总体未知（或已知都无所谓），总体方差未知（或已知无所谓）情况下的均值分布：凡是提到大样本，均可用正态分布计算，用样本方差替代总体方差
- 5、样本成数的抽样分布：凡是提到样本成数 p ，都是特指大样本，小样本提成数没有意义。在大样本情况下，无论其分布如何，成数的分布都可以确定。
- 6、样本方差的分布，这里特指总体是正态总体的情况。这个运用很多，其分布形式以及卡方换算，重要。

四、区间估计：根据样本大小、总体情况、样本个数情况，待估参数，可以将需要计算的区间估计计划分为9种类型，同假设检验。

- 1、小样本、正态总体、总体方差已知，总体均值的区间估计：Z分布
- 2、小样本、正态总体、总体方差未知，总体均值的区间估计：t分布
- 3、小样本（一般不包括大样本）、正态总体，总体方差的区间估计，卡方分布。
- 4、小样本、正态二总体，总体方差已知，总体均值差的区间估计，Z分布
- 5、小样本，正态二总体，总体方差未知，总体均值差的区间估计，t分布
- 6、大样本，分布未知（或已知无所谓），方差未知（或已知无所谓），总体均值的区间，Z分布。
- 7、大样本，分布未知（或已知无所谓），方差未知（或已知无所谓），二总体均值差的区间，Z分布。
- 8、大样本，分布未知（或已知无所谓），总体成数的区间估计，Z分布
- 9、大样本，分布未知（或已知无所谓），二总体成数差的区间估计，Z分布。

第七部分 假设检验

- 一、假设检验的思想：
- 二、假设检验的原理：小概率原理和大数定理
- 三、基本假定：总体、抽样等假定。
- 四、基本概念：原假设、备择假设、单边检验、双边检验、显著性水平、临界值、接受域、拒绝域、两类错误（是指针对原假设而言的弃真和纳伪错误）。
- 五、假设检验的基本步骤：4步。做题时候，要严格按照步骤及作出解释。
- 六、纳伪错误的计算原理，熟悉，多年没有考，计算相对复杂

2024 年北京大学 838 社会学研究方法和分支学科研究方法考研题库[仿真+强化+冲刺]

北京大学 838 社会学研究方法和分支学科研究方法之社会统计学考研仿真五套模拟题

2024 年社会统计学五套仿真模拟题及详细答案解析（一）

一、名词解释

1. 点估计

【答案】所谓点估计，就是根据样本数据算出一个单一的估计值，用它来估计总体的参数值。

2. 调和平均数

【答案】 N 个变量值倒数算术平均数的倒数，也称倒数平均数。

3. 中心极限定理

【答案】中心极限定理具体内容为：不论总体分布是否服从正态分布，从均值为 μ 、方差为 σ^2 的总体中，抽取容量为 n 的随机样本，当 n 充分大时（通常要求 $n \geq 30$ ），样本均值 \bar{X} 的抽样分布近似服从均值为 μ 、方差为 σ^2/n 的正态分布。

4. 离散变量

【答案】如果一个变量的变量值是间断的，可以一一列举的，这种变量称为离散变量。例如，某人兄弟姐妹数、结婚次数、工厂生产产品的数量等，其变量值的取值是 $0, 1, 2, 3, \dots$ 。离散变量的取值是有限个值，而且其取值都是以整数位断开的，是有最小计量单位的。例如，某人的兄弟姐妹数，只能是 1 个、2 个、3 个等，而不能是 1.3 个、2.5 个等。

5. 大量观察法

【答案】大量观察法，就是就总体中足够多的单位进行调查和综合分析，用以反映社会总体的数量特征。大量观察法是统计调查阶段的重要方法

二、简答题

6. 什么是简单一元线性回归分析？其作用是什么？

【答案】一元线性回归就是对具有线性相关关系的两个变量之间（其中一个为自变量）数量变化的一般关系进行预测，确定相应的数学关系式，以便进行估计或预测。通常将描述因变量 y 如何依赖自变量 x 和误差项 e 的方程称为回归模型。一元线性回归模型表达式为： $y = \beta_0 + \beta_1 x + \varepsilon$

7. 一元线性回归模型中有哪些基本假定

【答案】一元回归模型的三个基本假定

(1) 误差项 ε 是一个期望值为 0 的随机变量，即 $E(\varepsilon)=0$ 。在一个回归模型中， β_0 和 β_1 都是常数，

因而，对于一个特定的 x 值， y 的期望值为 $E(y)=\beta_0 + \beta_1 x$ 。

(2) 对于所有的 x 值， ε 的方差 σ^2 都是相同的。

(3) 误差项 ε 是一个服从正态分布的随机变量，而且互相独立，即 $\varepsilon \sim N(0, 1)$ 。

8. 方差分析中有哪些基本假定

【答案】每个总体都应服从正态分布；每个总体的方差必须相同；观测值是独立的

2024 年社会统计学五套仿真模拟题及详细答案解析（二）

一、名词解释

1. 总变差

【答案】记作 SS_T ，它表示 Y_{ij} 对于总均值 \bar{Y} 的偏差之平方和，即

$$SS_T = \sum_{i=1}^c \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2$$

2. 众数

【答案】在一组资料中，出现次数（或频数）呈现“峰”值的那些变量值。

3. 置信水平

【答案】置信水平就是将构造置信区间的步骤重复很多次，置信区间包含总体参数真值的次数所占的比例。

4. 总体与样本

【答案】总体是构成它的所有个体的集合，个体则是构成总体的最基本的单位。样本就是从总体中按照一定方式抽取的一部分个体的集合。例如，要从某省所有育龄妇女中抽取 1000 人进行调查，那么，该省所有育龄妇女就是研究总体，其中每一位育龄妇女就是个体，而抽取出的 1000 名育龄妇女就构成了该总体的一个样本。

5. 指标体系

【答案】指标体系就是一系列有内在联系得统计指标集合体。

二、简答题

6. 简要举例说明在分析双变量的关系时，t 检验和卡方检验的主要区别。

【答案】T 检验（TTest）是最常见的一种假设检验类型，主要验证总体均值间是否存在显著性差异。T 检验属于参数假设检验，所以它适用的范围是数值型的数据，在网站分析中可以是访问数、独立访客数、停留时间等，电子商务的订单数、销售额等。T 检验还需要符合一个条件——总体符合正态分布。

卡方检验（chi-square test），也就是 χ^2 检验，用来验证两个总体间某个比率之间是否存在显著性差异。卡方检验属于非参数假设检验，适用于布尔型或二项分布数据，基于两个概率间的比较，早期用于生产企业的产品合格率等，在网站分析中可以用于转化率、BounceRate 等所有比率度量的比较分析，

7. 概率抽样方法主要包括哪几种？

【答案】概率抽样主要的方式有：简单随机抽样、系统抽样、分层抽样、整群抽样、多阶段抽样等。

8. 理解原假设与备择假设的含义，并归纳常见的几种建立原假设与备择假设的原则。

【答案】原假设通常是研究者想收集证据予以反对的假设；而备择假设通常是研究者想收集证据予以支持的假设。

建立两个假设的原则有：

- （1）原假设和备择假设是一个完备事件组，而且相互对立。
- （2）一般先确定备择假设。再确定原假设。
- （3）等号“=”总是放在原假设上。
- （4）假设的确定带有一定的主观色彩。
- （5）假设检验的目的主要是收集证据来拒绝原假设。

2024 年社会统计学五套仿真模拟题及详细答案解析（三）

一、名词解释

1. 虚拟变量

【答案】当品质标志的变异性用离散变量来表达时，这个变量可称虚拟变量。

2. 零假设

【答案】概率分布的具体形式是由假设决定的，假设肯定不止一个。在统计检验中，通常把被检验的那个假设称为零假设(或称原假设，用符号 H_0 表示)，并用它和其他备择假设(用符号 H_1 表示)相对比。

3. 众数

【答案】众数是一组数据中出现频数最多的数值，用 M_o 表示。例如，一个城市有多种产业，但如果以旅游业为最多，那么旅游业就是众数，这个城市也被称为旅游城市。

4. 组内变差

【答案】记作 SS_w ，它是各观测值 Y_{ij} 对其所属类别均值 \bar{Y}_i 的偏差的平方和，即 $\sum_{i=1}^c \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2$

5. 组间平方和

【答案】记作 SS_B ，是自变量因素所没有解释的 Y_{ij} 的变异，即 $\sum_{i=1}^c n_i (\bar{Y}_i - \bar{Y})^2$

二、简答题

6. 怎样理解均值在统计中的地位？

【答案】均值是对所有数据平均后计算的一般水平的代表值，数据信息提取得最充分，具有良好的数学性质，是数据误差相互抵消后的客观事物必然性数量特征的一种反映，在统计推断中显示出优良特性，由此均值在统计中起到非常重要的基础地位。受极端数值的影响是其使用时存在的问题。

7. 等距分组和不等距分组有什么区别？请举例说明。

【答案】在分组时，如果分组组距相等，就是等距分组，如：学生平均每天上网时间：组距设 3，分成 3-6，6-9，9-12，12-15 共 4 个小组。如果分组组距不相等，就是不等距分组。如：人口分析，分婴幼儿组（0-6），少儿组（7-17），中青年组（18-59），老人组（60-130），组距不相等。

8. 简要说明卡方的拟合优度检验和独立性检验的含义。

【答案】（1）拟合优度检验

从总体的不同类别中抽取元素构成样本，样本包含总体中各个类别的元素，对不同类别的目标量之间是否存在显著性差异进行的检验称为拟合优度检验。拟合优度检验是 χ^2 检验中重要的一部分，可以同时多个总体进行比较。在 χ^2 检验中，如果 χ^2 统计量小于给定的 χ^2_{α} ，那么就不能拒绝原假设，如果大于则拒绝。即 χ^2 检验的拒绝域是 $\chi^2 > \chi^2_{\alpha}(k-1)$ ，其中 $k-1$ 是自由度。自由度是可以自由变动的变量个数，在列联表中，自由度 = $(R-1) * (C-1)$ 。

（2）独立性检验： χ^2 检验还可用于判断两个分类变量之间是否存在联系，例如，对宿舍管理改革的

态度是否与专业有关、对性行为的态度是否与学历有关等。如果两个分类变量之间没有关系，则称为独立，我们用 χ^2 判断它们之间是否关联，这时称为独立性检验。

考研云分享
kaoyany.top

附赠重点名校：社会统计学考研真题汇编（暂无答案）

第一篇、2013 年社会统计学考研真题汇编

2013 年华中农业大学 612 社会统计学考研专业课真题

华中农业大学二〇一三年硕士研究生入学考试
试 题 纸

课程名称：612 社会统计学

第 1 页 共 4 页

注意：所有答案必须写在答题本上，不得写在试题纸上，否则无效。

一、不定项选择题（共 20 分，每小题 2 分）

1. 工资收入若按高薪、低薪、一般水平来测量，它属于（ ）。
A. 定类变量 B. 定序变量 C. 定距变量 D. 定比变量
2. 适合于定距变量的统计图有（ ）。
A. 条形图 B. 圆瓣图 C. 直方图 D. 折线图
3. 以下相关系数，取值在 $[-1, +1]$ 之间的有（ ）。
A. R B. λ C. G D. Tau-y
4. $PRE=0$ ，说明（ ）。
A. X 与 Y 之间完全相关 B. X 与 Y 之间完全无关
C. X 能百分之百解释 Y 的变化 D. X 对 Y 无解释力
5. 两变量的线性相关系数为 0，表明两变量之间（ ）。
A. 完全相关 B. 无关系 C. 不完全相关 D. 不存在线性相关
6. 样本容量既定，则置信度越高，置信区间会越（ ）。
A. 宽 B. 窄 C. 不变 D. 不能确定
7. 抽样误差是指（ ）。
A. 样本均值标准差 B. 总体均值标准差
C. 样本均值方差 D. 总体均值方差
8. 某一研究的虚无假设是： $H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$ ，那么，最恰当的检验是（ ）。
A. T 检验 B. 卡方检验 C. 方差检验 D. Z 检验
9. 回归分析和相关分析的关系是（ ）。
A. 回归分析可用于估计和预测
B. 相关分析是研究变量之间的相互依存关系的密切程度
C. 回归分析中自变量和因变量可以互相推导并进行预测
D. 相关分析需区分自变量和因变量
E. 相关分析是回归分析的基础
10. 反映某一线性回归方程 $y=a+bx$ 好坏的指标有（ ）。
A. 相关系数 B. 判定系数 C. b 的大小 D. 估计标准误 E. a 的大小

华中农业大学二〇一三年硕士研究生入学考试
试 题 纸

课程名称：612 社会统计学

第 2 页 共 4 页

注意：所有答案必须写在答题本上，不得写在试题纸上，否则无效。

二、填空题（每空 2 分，共 20 分）

1. 各组频数与组距之比称为_____。
2. 从样本中计算出来数值，通常称为_____，在总体中的数值，则称为_____。
3. 在一个右偏分布中，三个位置平均数的关系是_____。
4. 分布函数 $F(x)$ 和 $P(x)$ 的关系，就像向上累计频数和频数的关系一样。所不同的是， $F(x)$ 累计的是_____。
5. 推论统计有两个基本内容：参数估计和_____。
6. 正态总体均值的假设检验， $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$ ，显著性水平 α ，这种检验称作_____侧检验，若总体方差 σ^2 已知，小样本，则检验统计量_____，拒绝域_____。
7. 如果从 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 两个总体中分别抽取容量为 n_1 和 n_2 的独立随机样本，那么两个样本的均值差 $(\bar{X}_1 - \bar{X}_2)$ 的抽样分布就是 N _____。

三、判断题(对的打“√”，错的打“×”，每小题 2 分，共 10 分)

1. 在已经通过调查获得的资料中，定序变量经过适当处理，可以变为定距测量层次，定距变量不可能转化为定序测量层次。（ ）
2. 95%的置信度是说每一次抽样都有 95%估计正确的可能性，而不能指出哪一次估计是错误的。（ ）
3. 在其他条件不变的情况下，减少弃真的错误，会增大纳伪的错误。（ ）
4. 假设检验的拒绝域正是区间估计的置信区间。（ ）
5. 把 S_1^2 和 S_2^2 中的较大者放在分子上，那么无论是单侧检验还是双侧检验，F 的临界值都只在右侧，这样就可以统一使用右侧检验的方法得出检验的结论。（ ）

华中农业大学二〇一三年硕士研究生入学考试 试 题 纸

课程名称：612 社会统计学

第 3 页 共 4 页

注意：所有答案必须写在答题本上，不得写在试题纸上，否则无效。

四、计算题（每小题 20 分，共 100 分）

1. 根据下面资料计算：家庭成员数的众数、中位数、平均数及其相应的离散趋势。

家庭成员数	2	3	4	5	6	8
家庭比重数 (%)	10	37	28	15	9	1

2. 青年歌手大奖赛评委会对 10 名决赛选手的演唱水平 (X) 和综合素质 (Y) 进行打分，评价结果如下表（表中已先将选手按演唱水平作了次序排列）所示，试计算选手的演唱水平和综合素质间的肯德尔等级相关系数 Tau-a 及斯皮尔曼等级相关系数。

选手名	A	B	C	D	E	F	G	H	I	J
演唱水平 (X)	1	2	3	4	5	6	7	8	9	10
综合素质 (Y)	3	1	5	2	7	4	10	8	6	9

3. 甲、乙两地各作 1000 户抽样调查。其中甲地拥有电视机 825 户，乙地拥有电视机为 760 户，求置信度为 0.95 两地电视机拥有成数差的置信区间。

4. 对某工段 8 名工人进行的技能培训前后的产量数据如下表所示

工人	甲	乙	丙	丁	戊	己	庚	辛
培训后 (A)	86	87	56	93	84	93	75	79
培训前 (B)	80	87	58	91	77	82	74	66

试问此项培训是否有效？（ $\alpha = 0.05$ ）

5. 某连锁商业企业在同城三个不同地点开设了三个分店，从这三家分店随机抽出 5 天的营业额的数据如下表所示：

华中农业大学二〇一三年硕士研究生入学考试
 试 题 纸

课程名称：612 社会统计学

第 4 页 共 4 页

注意：所有答案必须写在答题本上，不得写在试题纸上，否则无效。

单位：万元

	第一家分店	第二家分店	第三家分店
第一天	12	24	18
第二天	14	15	18
第三天	15	25	27
第四天	18	16	19
第五天	16	15	18

要求：

- (1) 计算分店地点与营业额的相关系数；
- (2) 检验三个分店的地点不同对每天的营业额是否有显著的影响。
($\alpha=0.05$)

参考临界值：

$Z_{0.05}=1.65$	$Z_{0.025}=1.96$	$Z_{0.01}=2.33$	$Z_{0.005}=2.58$
$T_{0.025}(6)=2.447$	$T_{0.025}(7)=2.365$	$T_{0.025}(8)=2.306$	
$T_{0.025}(9)=2.2262$			
$T_{0.05}(6)=1.9432$	$T_{0.05}(7)=1.8946$		$T_{0.05}(8)=1.8595$
$T_{0.05}(9)=1.8331$			
$F_{0.025}(3, 15)=4.15$	$F_{0.025}(3, 12)=4.47$	$F_{0.025}(2, 12)=5.1$	
$F_{0.025}(2, 15)=4.77$			
$F_{0.05}(3, 15)=3.29$	$F_{0.05}(3, 12)=3.49$		$F_{0.05}(2, 12)=3.89$
$F_{0.05}(2, 15)=3.68$			

以上为本书摘选部分页面仅供预览，如需购买全文请联系卖家。

全国统一零售价： **¥ 198.00元**

卖家联系方式： 客服电话： 17165966596（同微信）

微信扫码加卖家好友：

考研云分享-精品资料库

真题汇编 | 考研笔记 | 模拟题库



长按二维码加Q仔6号微信
有疑问直接私聊我

考研云分享-官方网站

免费真题 | 免费笔记 | 全科资源



长按二维码跳转至官网
还有更多内容和服务访问查看