

中国社会科学院大学社会统计学课程大纲

课程基本信息 (Course Information)					
课程编号 (Course ID)	1252041003	*学时 (Credit Hours)	64	*学分 (Credits)	4
*课程名称 (Course Name)	社会统计学				
	Social Statistics				
先修课程 (Prerequisite Courses)	高等数学, 概率论与数理统计				
*课程简介 (Description)	“社会统计学”是社会学专业必修课之一, 与社会研究方法、社会统计软件等课程同为培养学生专业研究基本技能的基础课程。针对所关注的社会现象或社会问题进行实证研究, 在收集资料、整理资料后, 得到的量化资料需要经过计算才能结果和分析结论, 这是社会统计学课程学习的必要性所在。这门课程系统介绍社会统计学的基本概念、原理、思路, 以及多种具体的统计分析方法, 重点在于引导学生根据需求选择恰当的统计方法并运用该方法运算处理数据、得到正确的结果并给出研究结论。主要内容包括(1) 导论, 介绍社会统计学发展历史等, 重点在于变量层次的划分与判断;(2) 单变量统计描述, 介绍针对某一社会现象或问题进行测量后, 所得数据的整理方法和思路, 如何用图表呈现其分布以及分析分布特征;(3) 单变量统计推论, 介绍参数估计和假设检验两种思路及具体方法;(4) 双变量描述与推论, 介绍列联分析、等级相关、回归分析与一元方差分析四个常用的方法;(5) 扩展性介绍多元统计方法, 为进一步学习高级统计技术打下基础。教学以课堂教授为主, 辅之以课堂讨论、提问等环节, 增加与学生互动、交流。				
*课程简介 (Description)	Social Statistics is one of the required courses for students who study sociology . Like Social Research Methods, Social Statistics Software and other courses, its main purpose is to train students' basic skills in professional research. The main content of social statistics is to carry out empirical research on social phenomena or social problems, collect data, sort out data, and especially analyze data. Therefore, it is necessary to use appropriate statistical description and statistical inference methods to get results and draw conclusions. This course systematically expounds and introduces the basic concepts, principles and ideas of social statistics, as well as a variety of specific statistical analysis methods. The focus is to guide students to choose appropriate statistical methods according to their needs and use the methods to calculate and process data, get correct results and give conclusions. The most important of learning social statistics is to provide students with quantitative technical and theoretical support for research practice, to help them to read basic quantitative research literature and to make a preliminary attempt to conduct professional research.				
*教材 (Textbooks)	《社会统计学》, 卢淑华主编, 北京大学出版社, 2009 年 11 月第四版, 书号: 978-7-301-09199-9, 21 世纪社会学系列教材。				
参考资料 (Other References)	- 李沛良:《社会研究的统计应用》, 社会科学文献出版社, 2002 年第二版。 - 莱文、福克斯著, 王卫东译:《社会研究中的基础统计学》(第 1 版), 中国人民大学出版社, 2008 年 3 月。				

		3. 郭志刚主编：《社会统计分析方法与 SPSS 软件应用》，中国人民大学出版社，2017 年第二版。 4. （美）费里斯·里奇：《统计想象》（社会学经典教材影印丛书），北京大学出版社，2006 年 1 月。					
*课程类别 (Course Category)		☐ 公共基础课/全校公共必修课 ☐ 通识教育课 ☒ 专业基础课 ☐ 专业核心课/专业必修课 ☐ 专业拓展课/专业选修课 ☐ 其他_____					
*授课对象 (Target Students)		社会学专业本科 二年级学生		*授课模式 (Mode of Instruction)		☐ 线上，教学平台_____	
*开课院系 (School)		社会与民族学院		*授课语言 (Language of Instruction)		☒ 中文 ☐ 全外语_____语种 ☐ 双语：中文+_____语种（外语讲授不低于 50%）	
*授课教师信息 (Teacher Information)		课程负责人 姓名及简介		刘月，讲师，法学硕士（社会学专业）			
		团队成员 姓名及简介					
学习目标 (Learning Outcomes)		1. 掌握社会统计学的基础知识，了解各种统计方法适用的前提并能正确运算得到结果、给与解释； 2. 激发对于统计技术运用的兴趣，为进一步学习和高级统计技术打下基础； 3. 通过学习统计的思路和方法，提高学生的研究实践能力。					
*考核方式 (Grading)		平时成绩占总成绩 30%，包括两部分：课程作业（20%），共十次课后作业（从教材课后习题与教师补充习题题库中指定），目的是让学生巩固、熟练所学的各种统计方法；课堂测试（10%），通常安排在期中，用于阶段性检验学习成果、提醒学生查漏补缺，并对期末考试有所准备。 期末考试占总成绩 70%，安排两小时闭卷课堂考试，系统考察学生的知识掌握情况。					
*课程教学计划（Teaching Plan）填写规范化要求见附件							
周次	周学时	其中					教学内容摘要 （必含章节名称、讲述的内容提要、实验的名称、教学方法、课堂讨论的题目、阅读文献参考书目及作业等）
		讲授	实验课	习题课	课程讨论	其他环节	
第一周	4						第一讲 导论：社会学研究与统计分析 第一节 社会学研究的基本程序 1. 确定研究课题 它包括三个方面的工作，即选题、查阅文献和进行探索性研究。 2. 提出研究假设 对调查对象的特征以及有关现象之间的相互关系进行推测性判断或设想。 3. 研究课题操作化 通过对抽象概念的定义来选择调查指标，从而将抽象的概念转化为具体的可以观测的变量，将研究假设转化为具体假设。 4. 设计调查研究方案

					通过对研究的程序和实施过程中的各种问题进行详细、全面的考虑，制订出调查研究的总体计划和切实可行的调查大纲。 5. 设计问卷 根据调查指标和项目精心设计问题表格，即问卷，然后把问卷初稿用于试调查，从而发现问卷的缺陷和遗漏，并进行修改。 6. 实施调查 利用各种调查方法收集有关资料，并对资料进行审核与登录。 7. 对资料进行统计分析 对调查资料进行系统的整理、分类、统计和分析，并对研究假设进行检验。 8. 撰写调查研究报告 说明调查结果或研究结论，并对研究过程、研究方法、政策建议以及研究中的一些重要问题或下一步研究的设想等进行系统的叙述和说明。 第二节 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 具有随机性和统计规律性。 2. 统计学的运用价值 社会调查资料的多值性以及由此产生的抽样数据的不确定性，决定了对社会调查资料的分析和推论只能采用统计学的方法。 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作用主要表现在经验层次的大面积数据处理方面。统计分析的内容主要分为统计描述和统计推论两部分。 第三节 统计分析方法的选用 1. 全面调查与抽样调查 全面调查一般使用统计描述。抽样调查既要运用统计描述，又要运用统计推论。 2. 单变量与多变量 单变量研究运用统计描述。多变量研究既要运用统计描述，也要运用统计推论。 3. 变量的层次 定类变量的取值只有类别之分，只具有等于或不等于的性质。定序变量的取值除了有类别之分以外，还有等级、次序的差别，其数学运算特性除具有等于或不等于之外，还有大于或小于之分。定距变量的取值除了有类别、次序之分外，取值之间的距离还可采用标准化的距离去度量，其数学运算特性除了等于、不等于、大于、小于之外，还可以加或减。定比变量的取值除了具有上述三种变量的属性之外，还可以构成一个有意义的比率，即可以乘或除。是否具有实际意义的零点存在，是定比变量和定距变量的唯一区别。 教学方法： 本讲内容和教材第一章对应，阅读文献为陈希孺《数理统计学简史》（电子版） 思考题： - 什么是归纳？什么是演绎？它们在社会研究中的作用是什么？ - 社会学研究的基本程序由哪几个阶段组成？ - 统计分析在社会学研究中的作用是什么？
第二周	4				第二讲 单变量描述统计 第一节 分布、统计表和统计图 1. 分布 一个概念或变量的各个情况出现的次数或频次称为分布。变量值频次对的集合称作频次分布。变量值概率对的集合称作概率分布。变量值百分比对的集合称作百分比分布。 2. 统计表 资料通过统计汇总，得出许多说明社会现象和过程的数字资料，把这些资料按照一定的目的，在表格上表现出来，这种表格就叫做统计表。统计表有简单表、简单分组表和复合分组表之分。 统计表必须具备表号、表头、标识行、主体行和表尾等内容。 3. 统计图 统计图用图形的形式来表示变量的分布。定类变量的分布用圆瓣图或条形图表示，定序变量的分布用条形图表示，定距变量和定比变量的分布用直方图

					或折线图表示。 4. 图和表的累计表示 累计图或累计表表示的是大于某个变量值的频次是多少或小于某个变量值的频次是多少。如果把频次换成频率，还可以作成累计频率分布图或累计频率分布表。 5. 分布图分析 对于曲线可作峰点研究、对称研究、U形曲线和J形曲线研究。 第二节 集中趋势测量法 1. 集中趋势的概念 代表全体变量分布的主要特征或一般水平的某一个典型的变量值或特征值称作集中值或集中趋势。在统计学中集中趋势由均值给出度量。 2. 众值 M_0 众值是指在一个变量数列中出现频数最多的变量值。用单项分组数据资料计算众值一般采用直接观察法。用组距分组资料计算众值采用组中值法和金氏插补法。 3. 中位值 M_d 将一组变量值按大小顺序排列，位于变量数列中间位置的变量值即为中位值。
第三周	3			1	第二讲 单变量描述统计 第二节 集中趋势测量法 4. 算术平均数 $\bar{X}$ 算术平均数是以总体各单位数值之和除以总体单位总数的商。 5. 众值、中位值和算术平均值的比较 众值适用于定类、定序、定距、定比变量，中位值适用于定序、定距、定比变量，算术平均数适用于定距、定比变量。 算术平均数包含的信息最丰富，但要受变量数列中极端值的影响。众值和中位值对资料的使用不完全，但它们的适用范围比算术平均数宽，而且不受极端值的影响。当数列中有极端值时，中位值的代表性好于算术平均数。 第三节 离散趋势测量法 1. 离散趋势的概念 描述变量数列离散程度的特征值叫做离散趋势或离中量数。它反映了一组变量值背离分布中心值的特征。 2. 异众比率 V_R 异众比率是指非众值的频次之和与总体总频次的比率。 3. 极差 R 极差是指在一个变量数列中两个极端数值之差，用公式表示为 $R = \text{最大变量值} - \text{最小变量值}$。 4. 四分位差 Q 把一组数据按大小排列成序列，然后分成四个数据数目相等的段落，各段落分界点上的变量值由小到大分别叫做第一个四分位值、第二个四分位值和第三个四分位值，第三个四分位值和第一个四分位值之差即为四分位差。 5. 方差 σ^2 和标准差 σ 方差是指资料中各个数值与其算术平均数相减之差的平方和的算术平均数。标准差是方差的正平方根。 6. 离散系数 CV 离散系数是标准差与算术平均数的比值，用百分数表示。

					教学方法： 本讲教学内容对应教材第二章 作业： 第二章课后习题 13、14、15、16、17 共 5 题
第四周	4				第三讲 概率论基础与中心极限定理 第一节 基础概率与概率分布 1. 随机现象和概率的概念 所谓随机现象是指在一次观察中可能出现也可能不出现、可能这样出现也可能那样出现、而在大量观察中具有某种规律性的现象。 2. 概率的运算 事件之间的关系分为事件的包含与相等、事件的和、事件的积、事件的互不相容和事件的对立。 概率的加法公式为 $P(A+B) = P(A) + P(B)$ (简化式) 和 $P(A+B) = P(A) + P(B) - P(AB)$ (一般式)。 概率的乘法公式为 $P(AB) = P(A)P(B)$ (简化式) 和 $P(AB) = P(A)P(B/A)$ 或 $P(AB) = P(B)P(A/B)$ (一般式)。 3. 概率分布的概念 概率分布指的是随机现象一共有多少种结果以及每种结果所伴随的概率是多少。概率分布有以下两点性质：①任一取值的概率都是非负的；②所有取值的概率总和为 1。 4. 数学期望 离散型变量的数学期望为 $E(\xi) = \sum_{i=1}^n x_i p_i$ (要求级数是绝对收敛的)。 连续型变量的数学期望为 $E(\xi) = \int_{-\infty}^{+\infty} x \varphi(x) dx$ (要求广义积分是收敛的)。 5. 方差与标准差 离散型随机变量的方差为 $D(\xi) = \sum_{i=1}^n [x_i - E(\xi)]^2 p_i$。 连续型随机变量的方差为 $D(\xi) = \int_{-\infty}^{+\infty} [x - E(\xi)]^2 \varphi(x) dx$。 将方差 $D(\xi)$ 开方取正值，即为随机变量 ξ 的标准差 σ。 6. 矩、偏态与峰态 矩是各点对某一固定点离差幂的平均值。偏态就是三阶中心矩，一般用来测量分布偏离对称的程度。峰态就是四阶中心矩，一般用来描述分布尖峰的程度。 第二节 典型随机变量分布 回顾六个典型随机变量分布的性质、条件和计算公式，了解各分布之间的关系。 1. 离散型随机变量分布：二项分布、泊松分布 2. 连续型随机变量分布：正态分布、卡方分布、t 分布、F 分布

					第三节 大数定理与中心极限定理 - 切贝谢夫不等式 - 贝努里大数定理 - 切贝谢夫大数定理 - 中心极限定理 设 $\xi_1, \xi_2, \dots, \xi_n$ 为独立同分布的随机变量, 不管其分布如何, 只要 $E(\xi_i) = \mu, D(\xi_i) = \sigma^2 (i=1, 2, \dots)$ 存在, 则当 n 足够大时, 这些随机变量的和的分布近似地服从标准正态分布。 例题: 一个螺丝钉的重量是一个随机变量, 期望值是 1 两, 标准差是 0.1 两。求一盒 (100 个) 同型号螺丝钉的重量超过 10.2 斤的概率。 5. 棣莫佛—拉普拉斯 (de Moivre - Laplace) 定理 例题: 某地区少数民族占 0.5%。今作 10000 人的随机抽样, 求少数民族人数不多于 70 人的概率。 教学方法: 本讲内容对应教材第三、四、五章, 回顾概率论知识点、衔接后续统计推论 作业: 共 11 题 - 教材习题第 111-112 页, 第三章课后习题第九、十、十二题。 - 教材习题第 142 页, 第四章课后习题第一、三题。 - 教材习题第 183 页, 第三、四、五题。 - 某种呼吸器官传染病的死亡率为 0.002, 求 2000 人中少于 5 人死于此疾病的概率。 - 若在 1000 中平均有 1 人申报所得税时, 填写数字有误。抽查 10000 份申报表, 求其中有 6, 7 或 8 份填错的概率。(用两种方法解题) - 假设 09 年北京市高考成绩服从正态分布 (501, 982), 如果从其中随机抽取 100 人, 问这 100 人的平均成绩超过 510 分的概率是多少?
第五周	3			1	第四讲 参数估计 第一节 名词解释 - 统计推论 所谓统计推论就是根据局部资料 (样本资料) 对总体的特征进行推断。统计推论的内容大体可分两部分: 一是通过样本对总体的未知参数进行估计, 简称参数估计; 二是通过样本对总体的某种假设进行检验, 简称假设检验。 - 总体 所谓总体就是研究对象的全体。 - 样本与简单随机样本 从总体中按一定方式抽出的一部分称为样本。相互独立、遵从同一分布 (即同总体所遵从的分布) 的样本称作简单随机样本。 - 统计量 从总体中抽取容量为 n 的样本 $\xi_1, \xi_2, \dots, \xi_n$, 则函数 $f(\xi_1, \xi_2, \dots, \xi_n)$ 叫作统计量。 第二节 参数的点估计 - 参数的点估计 用样本统计量来估计总体参数的具体值, 叫做参数的点估计。 - 总体参数的点估计公式 用样本平均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ 作为总体均值的点估计值。 用样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ 作为总体方差的点估计值。

						用样本成数 $p = \frac{m}{n}$ 作为总体成数的点估计值。 3. 评价估计值的标准 无偏性、有效性和充分性。 第三节 抽样分布 1. 抽样分布 统计量的分布即为抽样分布。 2. 样本均值 $\bar{X}$ 的分布 若 $\xi_1, \xi_2, \dots, \xi_n$ 是从总体分布为 $N(\mu, \sigma^2)$ 中抽出的一个样本，则样本均值 $\bar{X}$ 仍然服从正态分布 $N(\mu, \frac{\sigma^2}{n})$。 总体分布为正态分布 $N(\mu, \sigma^2)$，但方差 σ^2 为未知，则统计量 $\frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$ 服从自由度 K 为 $n-1$ 的 t 分布。 在大样本情况下，不管总体是什么分布，$\bar{X}$ 的分布将接近正态分布。 3. 样本方差 S^2 的分布 样本方差 S^2 乘以 $\frac{n-1}{\sigma^2}$ 满足自由度 $K=n-1$ 的 χ^2 分布。 第四节 正态总体的区间估计 1. 区间估计 区间估计是要估计出两点来定出一个区间，并要指出所给区间包含未知参数的概率是多少。 2. 正态总体均值的区间估计 3. 正态总体方差的区间估计 教学方法： 本讲内容对应教材第六章第 1-5 节 讨论题目： 1. 置信区间和置信度之间的关系如何？ 2. 区间估计精度和置信度之间的关系如何？ 3. 设 X_1, X_2, X_3 为简单随机抽样的 3 个观测值，采用不等权的平均值 $\bar{X}' = \frac{2}{5}X_1 + \frac{2}{5}X_2 + \frac{1}{5}X_3$ 作为总体均值的点估计值是否比采用等权的平均值 $\bar{X} = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$ 作为总体均值的点估计值要差？为什么？
第六周		4				第四讲 参数估计 第五节 大样本区间估计 1. 大样本总体均值 μ 的区间估计 2. 大样本总体成数 p 的区间估计 3. 大样本二总体均值差的区间估计 4. 大样本二总体成数差的区间估计 第五讲 假设检验

					第一节 假设检验的基本概念、原理和步骤 1. 原假设和备择假设、统计量、显著性水平、临界值、接受域和拒绝域、双边检验和单边检验 2. 小概率原理 所谓小概率原理即小概率事件在一次观测中几乎是不可能发生的原理。通常把概率不超过 0.01 或 0.05 的事件称为小概率事件。 3. 假设检验的步骤和两类错误 根据实际问题作出原假设和备择假设；根据样本找出合适的、检验的统计量及其分布；确定显著性水平α，并依据 H_1 的性质选用单边或双边检验；计算检验统计量的值，并进行判断，若统计量的值落入拒绝域，便拒绝 H_0，接受 H_1，否则，接受 H_0。 第一类错误是指否定 H_0，但实际上 H_0 为真的错误，即弃真的错误。 第二类错误是指接受 H_0，但实际上 H_0 并不真实的错误，即纳伪的错误。 教学方法： 本讲内容对应教材第六章第六节，第七章 作业：共 7 题 1. 教材习题第 226 页，第六章课后习题第二、三、五题。 2. 从某医院育婴室（总体）随机的抽取 4 个婴儿，他们的出生体重（公斤）分别为 3.1、2.8、3.6、3.7。已知总体标准差为 0.4 公斤，求：总体平均出生体重的 99% 的置信区间。（假定体重服从正态分布） 3. 在一次广告宣传图片比赛中，给每副参赛图片都评了分，综合得分的满分为 100 分。假定一个统计学家并不知道总体中每副图片的得分，但他从该批参赛图片中随机的抽取了 30 幅，求出这 30 幅的平均综合得分值为 83 分，标准差为 20 分。问：参赛的所有图片综合得分的平均值的 95% 的置信区间是多少？ 4. 从两个出租汽车公司随机的抽取了两个独立样本得到如下数据： <table><tr><td></td><td>甲公司</td><td>乙公司</td></tr><tr><td>样本数</td><td>100 人</td><td>100 人</td></tr><tr><td>平均年收入</td><td>6000 元</td><td>6400 元</td></tr><tr><td>标准差</td><td>2800 元</td><td>3100 元</td></tr></table> 求：两个公司司机平均年收入的差异的 95% 的置信区间。 5. 1954 年在美国为了检验一种新的小儿麻痹疫苗的效果进行了一次大规模的实验。在全美二年级的学生中选出了 740,000 名儿童，其中有 400,000 是自愿接种的。从这志愿者中随机的选取一半来注射疫苗；剩下的一半注射生理盐水作为对照组，结果如下： <table><tr><td>分组情况</td><td>儿童数</td><td>小儿麻痹病例数</td></tr><tr><td>接种疫苗</td><td>200000</td><td>57</td></tr><tr><td>对照组</td><td>200000</td><td>142</td></tr><tr><td>拒绝接种</td><td>340000</td><td>157</td></tr></table> 估计接种疫苗后所造成的发病率（每 100,000 人中的病例数）的减少量的 95% 的置信区间。		甲公司	乙公司	样本数	100 人	100 人	平均年收入	6000 元	6400 元	标准差	2800 元	3100 元	分组情况	儿童数	小儿麻痹病例数	接种疫苗	200000	57	对照组	200000	142	拒绝接种	340000	157
	甲公司	乙公司																											
样本数	100 人	100 人																											
平均年收入	6000 元	6400 元																											
标准差	2800 元	3100 元																											
分组情况	儿童数	小儿麻痹病例数																											
接种疫苗	200000	57																											
对照组	200000	142																											
拒绝接种	340000	157																											
第七周	3		1	第五讲 假设检验 第二节 单总体假设检验 1. 大样本均值检验 包括大样本总体均值检验和成数检验。 2. 小样本正态总体假设检验																									

					(1) 单总体均值检验（总体方差已知和未知两种情况） (2) 单总体方差检验 教学方法： 本讲内容对应教材第八章 讨论题目： 1. 假设检验的目的是什么？如何来做假设检验（以单总体假设检验为例）？ 2. 在单总体假设检验中，当样本足够大时，如果由样本计算得到的统计量 $t = 2.00$，而在 0.05 显著性水平下的 t 值为 1.96，是否意味着该样本肯定不是来自于这个给定的总体？ 作业：共 5 题 1. 教材习题第 254-255 页，第八章课后习题第一、二、三、四题。 2. 1968 年在波士顿，一名反对越战的著名医生兼活动家受到了控告和审判。负责审判该医生的法官有一段有趣的记录：在最后几次审判中，他选出的 700 名陪审员中只有 15%是妇女。而在作为总体的波士顿市，有被选资格的陪审员中，29%是妇女。评价法官在选择女陪审员时是否公平（即是否按比例来选取的）？									
第八周		4			第六讲 二总体假设检验 第一节 大样本二总体假设检验 1. 大样本二总体均值差检验 2. 大样本二总体成数差检验 第二节 小样本二总体假设检验 1. 小样本总体均值差检验 分两种情况：二总体方差已知、二总体方差未知但相等 2. 小样本二总体方差比检验 第三节 配对样本的比较 教学方法： 本讲内容对应教材第九章，介绍四种情况下均值比较的公式和应用 讨论题目： 1. 小样本总体均值检验有哪些假定？ 2.为什么对于来自配对样本的数据必须使用配对数据的检验公式，而不能使用二总体均值差的检验公式？ 作业：共 5 题 1. 教材习题第 272-273 页，第九章课后习题第一、二、四、五题。 2. 从死于汽车碰撞事故的司机中抽取 2000 名司机的随机样本，根据他们的血液中是否含有酒精以及他们是否对事故负有责任将数据整理如下表所示。问：在整个总体中，血液中含有酒精和没有酒精的司机之间在对事故负有责任方面有差异么？ <table><tr><td></td><td>有责任</td><td>无责任</td></tr><tr><td>有酒精</td><td>650</td><td>150</td></tr><tr><td>无酒精</td><td>700</td><td>500</td></tr></table>		有责任	无责任	有酒精	650	150	无酒精	700	500
	有责任	无责任												
有酒精	650	150												
无酒精	700	500												
第九周		3		1	第七讲 双变量分析思路与详析模式 第一节 变量关系概述 在统计学上，变量有关指的是一个变量的值和另一变量的值有连带性。双变量关系描述以相关系数来呈现结果；社会学研究更关心的是总体中两变量是否有关，根据变量层次的组合形成五种最常用的方法（图示）。 第二节 测量方法的选择原则和分析思路 1. 首先确定测量层次，选取相应的测量方法 2. 其次注意变量间关系是否对称（共变？纯粹相关？因果？） 3. 在样本中选择相应的系数表达关系的强度、方向（对称？是否符合 PRE 法则？.....）									

					4. 适用正确方法中的恰当的统计量、给出检验结果。注意基本的套路：双变量分析的原假设都是不存在显著性相关、备择假设是存在显著性相关，然后根据计算出的统计量结果与临界值（也可以 sig.与α）的比较结果确定接受还是拒绝原假设。 第三节 详析模式 两个变量的关系常由于样本量的影响很容易通过检验，但通过检验的两个变量间的关系只是统计相关，实际上只是虚伪关系，这时需要做详析分析。引入第三变量，观察该变量引入后自变量与因变量原有关系的变化情况，以深化对原关系的认识，该过程为详析分析。 1. 原理：阐明分析、因果分析、条件分析 2. 示例： 例：阐明分析（中介变量）女性教育水平越高、子女数越少（$G=-0.7$）；控制结婚年龄后，晚婚-0.71，早婚-0.68，晚婚解释不成立；控制“重男轻女”，分别为-0.45，-0.50，部分解释；控制“社会意识（关心社会的政策和前途的程度）”，高中低分别为-0.18，-0.20，-0.23，解释力更强。 教学方法： 阅读文献：李沛良著《社会研究的统计应用》，社会科学文献出版社，2002年第二版，第211-228页。 讨论题目： 1. 绘制自己的双变量分析方法框图，并尝试结合参考书绘制多变量分析方法框图。 2. 二分变量是特殊的定类变量，其分析方法的特殊之处在哪里？既有的大样本二总体均值和成数分析，和列联分析、方差分析的关系如何？
第十周	2			2	期中随堂测验（两小节） 第八讲 列联分析 第一节 列联表的基本知识 1. 列联表的概念 2. 列联表中变量分布的不同形式 3. 列联表中变量的互相独立性 教学方法： 本讲内容对应教材第十章第一节 讨论题目： 1. 对于2×2列联表，如果格数过少，为什么要作连续性修正？ 2. 为什么要求列联表中每一格的期望值$E_{ij} \geq 5$？ 3. 当样本容量增加K倍时，如果相对频数不变，χ^2值将增加多少倍？为什么？
第十一周	4				第八讲 列联分析 第二节 列联表的检验 1. 列联表检验的原假设 2. 列联表检验的统计量和分析过程 第三节 列联强度 1. ϕ系数和Q系数 2. λ系数和τ系数 教学方法： 本讲内容对应教材第十章第二、三节

						作业：共 5 题 1. 教材第 312-313 页，第十章课后习题第二、三、四、五题。 2. 对于 10 年前出生的一个 10000 名婴儿的样本，以及 10 年后的另一个 10000 名婴儿的样本，得到如下列联表。通过卡方检验，能否得出“10 年间男孩出生的概率保持不变”这一结论（$\alpha=0.05$）？（用两种方法解答） <table><tr><td></td><td>10 年前</td><td>10 年后</td><td>合计</td></tr><tr><td>男</td><td>5060</td><td>5180</td><td>10240</td></tr><tr><td>女</td><td>4940</td><td>4820</td><td>9760</td></tr><tr><td>合计</td><td>10000</td><td>10000</td><td>20000</td></tr></table>		10 年前	10 年后	合计	男	5060	5180	10240	女	4940	4820	9760	合计	10000	10000	20000													
	10 年前	10 年后	合计																																
男	5060	5180	10240																																
女	4940	4820	9760																																
合计	10000	10000	20000																																
第十二周	4				第九讲 等级相关 第一节 斯皮尔曼等级相关 1. 斯皮尔曼等级相关系数 r_s 计算公式 2. r_s 统计检验应用 第二节 Gamma 等级相关 1. 同序对、异序对、同分对 2. GAMMA 系数公式及检验应用 第三节 其他等级相关系数 1. 肯氏 τ 系数系列 2. d 系数 3. 各系数比较分析 教学方法： 本讲内容对应教材第十一章 讨论题目： 1. 斯皮尔曼等级相关适用于分析什么特点的数据？GAMMA 等级相关呢？ 2. 适合处理多列等级变量资料的方法是本章介绍的哪种方法？ 作业：共 5 题 1. 教材第 337-338 页，第十一章课后习题第一、二、三、四题。 2. 一位研究者有兴趣研究工作满意度与上班时间的长短是否有关。他访问了某家公司的 185 名雇员，调查了他们上班的时间（X）与工作满意度（Y）。对于以下数据，计算上班的时间与工作满意度的 GAMMA 相关系数，并指出 X 与 Y 之间的关系是否显著。（$\alpha=0.05$） <table><tr><td></td><td colspan="4">上班时间（小时/周）</td></tr><tr><td>工作满意度</td><td>60 以上</td><td>30-59</td><td>15-29</td><td>15 以下</td></tr><tr><td>非常满意</td><td>8</td><td>12</td><td>25</td><td>22</td></tr><tr><td>一般</td><td>9</td><td>20</td><td>23</td><td>11</td></tr><tr><td>不满意</td><td>13</td><td>18</td><td>17</td><td>7</td></tr><tr><td>总数</td><td colspan="4">185</td></tr></table>		上班时间（小时/周）				工作满意度	60 以上	30-59	15-29	15 以下	非常满意	8	12	25	22	一般	9	20	23	11	不满意	13	18	17	7	总数	185			
	上班时间（小时/周）																																		
工作满意度	60 以上	30-59	15-29	15 以下																															
非常满意	8	12	25	22																															
一般	9	20	23	11																															
不满意	13	18	17	7																															
总数	185																																		
第十三周	4				第十讲 一元回归 第一节 一元回归线性方程 1. 回归研究的对象 2. 散点图 3. 回归方程的描述																														

					第二节 一元回归方程的建立 1. 最小二乘法 2. 方程求解过程演示 第三节 一元回归方程的假定与检验 1. 一元回归方程的基本假定 2. 回归方程的 F 检验 第四节 简单相关 1. 皮尔逊相关系数公式及检验 2. 判定系数 3. 相关与回归的比较 第五节 利用回归方程进行预测 教学方法： 本讲内容对应教材第十二章 讨论题目： 1. 什么是相关分析？什么是回归分析？二者之间的关系是什么？ 2. 构成事物间因果关系的条件是什么？ 3. 怎样评价一个回归方程对于散点的代表性？ 作业：共 3 题 教材第 373-374 页，第十二章课后习题第二、三、四题。
第十四周		4			第十一讲 一元方差分析 第一节 方差分析的基本知识 1. 方差分析的概念 2. 一元方差分析与多元方差分析的关系 第二节 一元方差分析 1. 一元方差分析的假定 2. 一元方差分析的检验 3. 相关比率 η^2 教学方法： 本讲内容对应教材第十三章第一、二节 讨论题目： 1. 对两个以上平均数两两之间的差异检验为什么不能两两之间进行 t 检验？ 作业：共 2 题 教材第 420-421 页，第十三章课后习题第一、二题。
第十五周		4			第十二讲 非参数检验 第一节 非参数检验方法 1. 非参数检验的概念 2. 非参数检验的优缺点 第二节 符号检验 1. 符号检验的原理 2. 符号检验的步骤 第三节 符号秩检验 1. 符号秩检验的原理 2. 符号秩检验表的使用 第四节 秩和检验

						1. 秩和检验法的原理 2. 秩和检验表的使用 第五节 游程检验 1. 游程检验的原理 2. 游程数检验表的使用 第六节 累计频次检验等其他方法简介 1. 累计频次检验 2. 单向方差秩检验 3. 双向方差秩检验 教学方法： 本讲内容对应教材第十四章 作业：共 6 题 教材第 446-448 页，第十四章课后习题第一、二、三、四、五、六题。
第十六周		3			1	第十三讲 多元统计方法简介 第一节 多元回归的原理简介和 SPSS 统计指令 1. 多元回归方程与标准化回归方程 2. 偏回归系数与偏相关系数 3. 回归方程的整体检验与每个自变量的单独检验 4. 操作指令与演示 第二节 多元方差分析 1. 二元方差分析的数学模型 2. 无重复情况下的二元方差分析 3. 重复情况下的二元方差分析 4. 多元方差分析操作指令与演示 教学方法： 本讲内容对应教材第十三章第三、四节，最后进行课程内容的回顾总结 阅读文献： 郭志刚主编：《社会统计分析方法 SPSS 软件应用》，中国人民大学出版社，2017 年第二版，第 35-90 页。 讨论题目： 1. 多元回归中自变量之间的多重共线性问题如何解决？ 2. 多元回归分析有哪里前提假定？与一元回归分析的异同？ 3. 什么是非参数检验？与参数方法比较，它有哪些特点？ 4. 符号检验法的基本思想是什么？ 5. 秩和检验的基本思想是什么？
总计						
备注（Notes）						