《传播统计与数据分析》课程教学大纲

**课程内容：**传播学本质是门社会科学。而统计是社会科学研究中广泛采用的定量分析方法。本课程系统地介绍了传播统计学的基本原理、基本概念和主要内容，按照变量的四个测量层次（定类、定序、定距和定比），课程详细阐述了统计描述和统计推论的操作程序和具体方法，并结合生动的实例说明了统计分析在传播学研究中的作用和地位，并结合上机向学生讲授SPSS软件应用。课程内容适合新闻、广电、广告等专业同学学习。该课程内容和传播学研究方法、市场调查等课程相互印证配合，为同学们的学习、科研实践提供一种有价值的工具。

**指导思想：**

传播现象的独特性和传播研究方法的特点决定了统计在社会研究中的重要地位，统计也因此而成为传播学研究的重要工具和重要手段。近十几年来，统计理论、统计方法和统计手段迅速发展，其应用范围也越来越广泛。本课程的目的就是为深入这一领域建立一个基础和平台，即对统计的基本概念、原理、类型、方法、程序、作用等有基本的和概括了解与把握，并能应用这些知识对研究问题进行简单的统计分析。本课程的教与学强调：

第一，社会研究是一项系统的和严谨的工作，从研究设计 资料收集

 资料整理分析 撰写研究报告，各个步骤之间相互联系、相互影响，密不可分。统计分析作为研究的一个重要环节，只有放在社会研究过程的背景之下，注重其与研究问题及研究方法的联系，才能更准确地掌握每一种统计类型和统计方法的特征，才能针对具体的研究问题选择恰当的统计方法。

第二，作为一门应用性极强的课程，本课程特别强调理论联系实际的原则，在教与学的过程中，一方面教师要通过列举和分析大量研究和应用实例，深化学生对统计原理的和统计思想的理解；另一方面要求学生将学习到的知识不断运用到对实际社会问题的分析中去。为此，要求学生在学习课程讲授的知识的同时，认真完成每一讲后面所指定的“实践性”的练习。

第三，在实际的社会研究中，资料的统计分析都是通过计算机完成的。各种统计描述和统计分析方法被制作成用于计算机的专门的和通用的统计软件。本课程将熟练掌握和灵活运用spss统计软件作为本课程教与学的不可分割的一部分。

第四，课程中介绍的各种具体的统计方法和统计技术，都有其优点和某些局限性，适用于一定的研究目的和分析要求。因此，在课程学习过程中，不仅需要对每一种方法和技术的特点、实施程序和适用范围有清楚的了解，而且也需要认识各种方法与技术之间的异同点，以便能够在面对不同的社会现象和不同的研究目的时，正确、灵活地选择和运用相应的方法与技术。

第五，统计分析是一种定量分析方法，对于统计结果的理解和解释需要联系其它调查资料，如研究对象所处社会的背景状况、所研究问题的特定意义、调查对象的特点等等来进行。对统计结果的解释和使用应当遵循实事求是的原则，杜绝弄虚作假，这是每一个从事社会研究的人员均应该严格遵循的规范。

**主要内容：**

**第一章：导论**

作为社会统计学的入门，本章从社会研究的四种方式入手，通过对统计调查研究的特点、类型、作用和方法论背景的介绍，简要说明传播统计学的含义与特征；进而从社会研究过程的角度，对统计在社会研究中的位置与重要作用进行了讨论。

**第二章：单变量描述统计**

单变量描述统计旨在用简洁和形象的方式勾画出数据的（也就是所研究群体或组织的）基本特征，它是对数据的最基本的和最初步的统计。分布、统计图和统计表是单变量描述的主要工具，而集中趋势和离散趋势则分别以一个数值概括出数据（或群体）的分布特征。本讲依照变量的4个测量层次，分别对上述统计描述内容和具体方法进行了详细的介绍。

**第三章：二变量统计描述 ——列联表**

变量关系是社会研究的重点，其中最简单的是两个变量间的关系，它是分析多变量关系的基础，许多变量之间的关系可以化解为若干个两变量关系进行处理。两变量间是否有关系及关系强弱在统计上用“相关系数”来进行测量，称为统计相关。相关系数取值在0——±1之间，该值刻画出两个变量之间关系的有无、强弱和方向。由于相关系数对变量关系的描述是高度概括的，因此要了解变量关系的具体形态和内容，就需要采用列联表。本讲重点讨论了列联表的概念、制作方式、作用以及变量独立性的条件等。

**第四章：二变量统计描述——相关关系测量**

本讲重点讨论二变量相关关系的测量方法。由于变量的测量层次不同，可以把变量关系以测量层次分为定类——定类、定序——定序、定距——定距三种主要形式，其相关关系分别称为列联相关、等级相关和线性相关。本讲分别讨论了三种相关形式的各种相关系数的计算方法，这些计算方法大多数是利用消减误差比例（PRE）的思想构建出来。如果两个变量的测量层次不同，一般有两种处理方法：发展出专门适于此种情况的相关系数计算方法；或降低（或升高）其中某个变量的测量层次从而能采用上面已知形式的计算方法，不过这会损失某种信息。

**第五章：二项分布及其他离散性随机变量的分布**

本章分别对几种重要的离散型随机变量的概率分布，即二项分布、多项分布、泊松分布和超几何分布的来源、分布形态与分布特征等进行详细的讨论。这些讨论有助于加深对“概率分布”概念的理解，并熟悉处理“随机现象”的方式和思路。

**第六章：正态分布、常用统计分布和极限定理**

正态分布是连续型随机变量概率分布中的一种最重要和应用最广泛的概率分布。本章将对正态分布的概念、特征、刻画方式、转化为标准正态分布的方法，以及标准正态分布表的构造和使用方法等做了系统的介绍，同时还将简要介绍统计中经常使用的三种连续型随机变量的概率分布：t分布、χ²（卡方）分布、F分布。

上面所介绍的各种概率分布都是一种理论分布，而实际调查所形成的变量分布是一种经验分布，大数定理建立了概率分布与统计分布，即理论分布与经验分布的关系，从而为统计推论奠定了基础。大数定理从另一个角度为统计推论提供了理论依据，它指出，无论总体的分布如何，在大样本的情况下，其和的分布趋向于正态分布。除此之外，本章还要学习中心极限定理。

**第七章：推论性统计——参数估计**

推论性统计分为参数估计与假设检验两大类，其中参数估计又分为点估计和区间估计。本章从抽样和抽样分布入手，重点讨论参数估计的概念、相关术语、估计值的评价标准、以及各种特征值的点估计和区间估计的方法，并以实例说明正态总体和大样本的区间估计的方法与步骤。

**第八章：推论性统计——假设检验**

本章重点讨论统计推论的另一种类型——假设检验。假设检验采用的推论方法是所谓的小概率原理，其基本逻辑是：先做出关于总体的一个假设（原假设），然后随机从总体中抽取一个样本，如果样本调查的结果在原假设成立的情况几乎是不可能发生的，就拒绝原假设，而接受原假设的对立面（备择假设）。小概率原理这种推理方法决定了无论是接受或是拒绝原假设都会犯错误，但两者的机会不同。本章对假设检验的概念、原理、相关术语、方法、步骤、与区间估计的关系等进行详细阐述，并分别对单总体、二总体及二变量关系的假设检验方法进行介绍。

**第九章：线性回归分析**

变量间的因果关系是较相关关系更强的一种关系，也是社会研究分析和讨论的重点。变量因果关系通常采用建构因果模型的方式进行描述和分析，回归模型是其中的一种，该模型在社会研究中被广泛采用，回归模型中最简单的是线性回归模型。本章通过两变量线性回归模型的讨论，说明回归分析的基本思想、前提假设、回归方程的建构与检验，回归系数与相关系数的关系，以及如何利用回归方程进行预测等。

**第十章：方差分析**

方差分析在统计中的重要作用在于：一方面它提供了一种分析与检验多个变量间复杂关系的重要方法，二是这种分析方法适用面广泛，可用于各种测量层次的自变量。根据自变量的个数，方差分析可分为一元方差分析、二元方差分析、三元方差分析……等等。本章重点讨论一元方差分析，并对二元方差分析做一简单介绍。具体内容包括方差分析的基本逻辑、前提假设、方差分析的检验等。

**第十一章：非参数检验**

前面介绍的在假设检验中使用的检验方法，如Z检验、t检验、F检验等，均是针对总体分布的某个参数（如：均值、方差、成数、相关系数等）进行的检验，所以又可以称作参数检验，参数检验往往要求总体的分布满足一定的条件。非参数检验，顾名思义，不是对总体分布的某个参数进行的检验，而是对总体有关性质的检验。由于非参数检验对总体的分布无须加以限制且计算简单，所以获得广泛的应用。本章从原理和方法角度介绍了一些社会研究中常用的非参数检验方法：符号检验、符号秩检验、秩和检验、游程检验、单向和双向方差秩检验等。

**教材与主要参考书**

**教材：**卢淑华：《社会统计学》 北京大学出版社，1997。

**主要参考书：**

1、R. Mark Sirkin 《Statistics for the Social Sciences》 SAGA Publications, London Delhi 1999年2版。

2、里奇 .《统计想象》 .北京大学出版社，2006年2月。

3、布莱洛克 .《社会统计学》 .中国社会科学出版社，1988年。

4、David Freedmam、 Robert Pisani 等著 《统计学》 中国统计出版社，1997年。

5、达瑞尔. 哈夫 《统计陷阱》 上海财经大学出版社，2002年。

6、谢宇. 《回归分析》，社会科学出版社，2013年。

7、戴维·K·希尔德布兰德等著《社会统计方法与技术》，社会科学文献出版社， 2005年6月。