排版测试

目录

# 数理统计作业-15组-第3周

|  |  |  |
| --- | --- | --- |
| 姓名 | 学号 | 题目 |
| 姜昊成 | 20181977 | 1,2 |
| 杨子涵 | 20181974 | 3,4 |
| 罗少航 | 20171904 | 5,6&排版（组长） |
| 何鑫云 | 20181978 | 7,8 |
| 潘一达 | 20174860 | 9,10,11 |

1. 学习随机变量的数字特征有什么意义？
   * 研究随机变量的数字特征可以总体上掌握随机变量某一侧面的性质,如期望表征随机变量的取值水平即平均数,方差表征随机变量取值的分散或集中程度。
2. 常见的数字特征有哪些？
   * 数学期望
     1. 连续型数据
     2. 离散型数据
   * 方差与矩
     1. 方差
     2. 矩
     3. 协方差
   * 相关系数
3. 均值（数学期望）、方差的定义、计算公式、性质有哪些？
   * 期望的定义：随机变量X的期望(expectation)或均值EX是它全部取值的平均值，也是分布律或密度函数的中心位置，因此，期望有时也称位置参数(location parameter).期望在实际中应用广泛，常被用作随机变量分布的代表性指标。
   * 期望的计算公式：
     + 设二维随机变量的联合分布律为 ，若联合密度函数为，则随机变量 的数学期望的计算公式为
   * 期望的性质
     1. 线性性：若 及 为常数， 为 个随机变量，则
     2. 若 相互独立，则
   * 方差的定义：方差(variance) 刻画了随机变量X取值的“波动”大小或稳定性，度量了随机变量偏离期望的平均幅度。
   * 方差的计算公式：
   * 方差的性质：
     + 随机变量的方差具有如下性质(假设所涉及的方差都存在) ：
     + 若 相互独立， 及 为常数，则
     + 设 是常数，对随机变量和，有
     + **(切比雪夫不等式)** 设的方差存在, 则对任意正数，有
     + 或
     + 若随机变量的方差存在，则的充分必要条件是以概率取某个常数，即。
4. 常见分布的期望、方差是什么？
5. 协方差、相关系数的定义、计算公式是什么？它们有什么作用？
   * 协方差、相关系数的定义
     + 设 为二维随机变量，如果 存在，则称 为与的**协方差**，记为
     + 若 ，则称 为与的**相关系数**，记为 或 ，即
   * 协方差的计算公式
   * 协方差、相关系数的作用
     + 当事件与相互独立时，有
     + 这意味着若 ，则与必然不相互独立，因而与之间存在则某种关联。因此可以用量 ，即**用相关系数来描述与之间的关联程度**
     + 对相关系数，令 ， ，则 。 是一个有量纲的数，而是一个无量纲的数。**相关系数能很好地反映随机变量与之间的关系**
6. 协方差、相关系数的基本性质是什么？
   * 协方差的性质
     1. 设为常数，则
     2. 设 为常数，则
     + 维随机变量推广
       1. 设 为常数
   * 相关系数的性质
     1. 的充要条件为 与 以概率为线性相关，即
7. 怎样认识变量之间的不相关关系与独立关系？
   * 变量间的关系可以用相关系数来刻画。
   * 对二维随机变量 而言，如果 存在，则称 为与的**协方差**，记为
   * 当取值为，则称和为不相关关系。
   * 变量的不相关关系和独立关系之间存在如下推导性质：
   * 但
   * 此外，当随机变量的联合分布是正态分布时，有
8. 随机变量的 k 阶原点矩、中心矩的定义是什么？混合矩的定义是什么？
   * 对随机变量和，若以下的数学期望都存在，
   * 则称 为的阶原点矩；
   * 称 为的阶中心矩；
   * 称为和的阶混合原点矩；
   * 称 为和的阶混合中心矩。
   * 由上述原点矩和中心矩的定义可知，一阶原点矩就是数学期望，二阶中心矩就是方差。同时，因，故随机变量的阶矩存在时，阶矩也存在，从而小于的各阶矩都存在。
9. 极限定理描述什么现象？
   * 极限定理是概率论的基本理论，阐述了大量观测或试验中随机现象的统计规律性，在理论研究和应用中起着重要作用。随机变量序列的极限反映了当实验次数充分大时序列的变化趋势。随机变量序列的极限有两种基本的描述方法，即依概率收敛和依分布收敛。
10. 辛钦大数定律、贝努里大数定律是什么？有什么应用？
    * 辛钦大数定律：
    * 伯努利大数定律：
      + 设是重伯努利试验中事件发生的频率，事件在每次试验中发生的概率为。则对任意正数。有：
      + 记作，即对给定的很小正数，当足够大时，“频率与概率的偏差小于”几乎是一个必然事件，这就是以频率定义概率的理论依据所在。
    * 大数定律不仅可以用于总体的均值估计，还可以用于其他参数的估计问题。例如，用于定积分的估计，当随机变量相互独立，都服从时，则也为独立同分布的随机变量序列，且。那么由大数定律得，当足够大时，可以作为参数的估计值，由此可见，只要能获取来自均匀分布的观测数据，就可以用去估计参数。
11. 独立同分布的中心极限定理是什么？有什么应用？
    * 中心极限定理是研究独立随机变量和的分布在一定条件下以正态分布为极限分布，是数理统计和误差分析的基础。
    * 独立同分布的中心极限定理：
      + 设独立同分布，数学期望存在且，则对于任意实数：
      + 此定理表明：若对某个随机变量进行大量的独立重复观测实验，则实验结果的和近似服从正态分布。
    * 应用：以在误差分析上的应用为例
      + 对一个物理量独立地测量次，每次测量产生的随机误差都服从上的均匀分布，如果取次测量的算术平均值作为测量结果，求它与真值得差小于指定的正数的概率。
      + 解：随机误差的期望与方差为： ，则：

[1] XU M, RHEE S Y. Becoming data-savvy in a big-data world[J]. Trends in Plant Science, 2014, 19(10): 619–622. DOI:[10.1016/j.tplants.2014.08.003](https://doi.org/10.1016/j.tplants.2014.08.003).

[2] JIFA G, LINGLING Z. Data, DIKW, Big Data and Data Science[J]. Procedia Computer Science, 2014, 31(0): 814–821. DOI:[10.1016/j.procs.2014.05.332](https://doi.org/10.1016/j.procs.2014.05.332).

[3] PHILIP CHEN C L, ZHANG C-Y. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data[J]. Information Sciences, 2014, 275(0): 314–347. DOI:[10.1016/j.ins.2014.01.015](https://doi.org/10.1016/j.ins.2014.01.015).

[4] WEICHSELBRAUN A, GINDL S, SCHARL A. Enriching semantic knowledge bases for opinion mining in big data applications[J]. Knowledge-Based Systems, 2014, 69(0): 78–85. DOI:[10.1016/j.knosys.2014.04.039](https://doi.org/10.1016/j.knosys.2014.04.039).

[5] YANG S, GUO K, LI J, 等. Framework Formation of Financial Data Classification Standard in the Era of the Big Data[J]. Procedia Computer Science, 2014, 30(0): 88–96. DOI:[10.1016/j.procs.2014.05.385](https://doi.org/10.1016/j.procs.2014.05.385).

[6] 于秀清. F-数据簇与缺损数据修复-还原[J]. 计算机工程与应用, 2013, 49(6): 139–142. DOI:[10.3778/j.issn.1002-8331.1108-0086](https://doi.org/10.3778/j.issn.1002-8331.1108-0086).

[7] PERNER P. Mining Sparse and Big Data by Case-based Reasoning[J]. Procedia Computer Science, 2014, 35(0): 19–33. DOI:[10.1016/j.procs.2014.08.081](https://doi.org/10.1016/j.procs.2014.08.081).