

考研数学基础30讲

概率论与数理统计分册

- 主 编 张 宇
- 副主编 高昆轮

版权专有 侵权必究

图书在版编目(CIP)数据

考研数学基础30讲. 概率论与数理统计分册 / 张宇

主编. -- 北京 北京理工大学出版社, 2024. 10 (2025.9重印).

ISBN 978-7-5763-4520-9

I. 013

中国国家版本馆CIP数据核字第20246G1K27号

责任编辑: 多海鹏

文案编辑: 多海鹏

责任校对: 周瑞红

责任印制: 李志强

出版发行 / 北京理工大学出版社有限责任公司

社 址 / 北京市丰台区四合庄路6号

邮 编 / 100070

电 话 / (010) 68944451 (大众售后服务热线)

(010) 68912824 (大众售后服务热线)

网 址 / <http://www.bitpress.com.cn>

版 印 次 / 2025年9月第1版第4次印刷

印 刷 / 河北鹏润印刷有限公司

开 本 / 787 mm × 1092 mm 1/16

印 张 / 11.25

字 数 / 281千字

定 价 / 69.90元

图书出现印装质量问题, 请拨打售后服务热线, 负责调换

前言

《考研数学基础 30 讲》(简称《30 讲》)按照《全国硕士研究生招生考试数学考试大纲》,并结合考研数学命题趋势编写,其中高等数学分册分为 18 讲,线性代数分册分为 6 讲,概率论与数理统计分册分为 6 讲,共 30 讲。学完这些内容,即可达到传统意义上的“基础阶段”与“强化阶段”的全部要求。

本书每一讲都由基础知识结构、基础内容精讲、基础习题精练三部分组成,涵盖《全国硕士研究生招生考试数学考试大纲》中所有知识点,循序渐进,由浅入深,最后达到考研数学对试题要求的难度水平。这里需要特别指出的是,数学学习是一环扣一环的,上一个环节的知识和想法,直接影响到下一个环节的知识和想法,任何环节都不能囫圇吞枣或降低要求,必须深刻掌握,否则越往下学习就会越困难。反之,若在初始阶段多下些功夫,搞明白原理,捋清楚关系,看似进度不快,却能为后面的学习加速打下坚实的基础。

《30 讲》是书课包,即本书作者会对书中全部内容进行系统讲解,考生扫描书中二维码即可快速定位对应知识点的视频讲解。2027 版的《30 讲》根据课程讲解,标注了全程的学习笔记,既能防止考生漏记、错记笔记,又节省做笔记的时间,考生只需集中精力认真听即可。

建议考生结合课程反复研读本书,有些知识的掌握是需要反复琢磨的,要养成独立思考的习惯,逐渐达到知识、思路、题型和方法皆会以清晰的结构呈现眼前的效果。本书是作者多年基础阶段教学经验的总结,愿助潜心研读者打好地基、夯实基础,勇攀考研数学高峰。

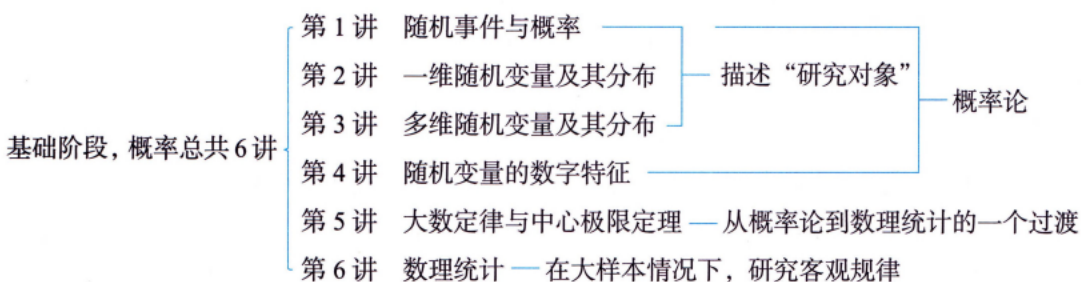
本书自出版以来,承蒙广大师生厚爱,在考研数学基础阶段起到了一定的积极作用。限于作者水平,书中不足或错误之处,望各位不吝赐教,多提意见与建议,特此致谢!

张宇

2025 年 8 月 于北京

概率综述

一 概率引言



第1讲 随机事件与概率

作为整个概率统计的基础内容，它主要研究随机事件 A 的概率问题，即研究 $P(A)$ 。

- 主要涉及
- ① 古典概型；
 - ② 几何概型；
 - ③ 一些重要公式；
 - ④ 事件的独立性

【重要思想】全集分解思想，贯穿整个课程。

★第2、3讲是经典的概率论课程的内容。

第2讲 一维随机变量及其分布

- (1) 引入研究对象“随机变量”。
- (2) 涉及与随机变量相关的分布函数（怎么去描述该随机变量的概率规律）。

(3) 随机变量函数的分布。

-
- ① 离散型随机变量涉及分布律的问题；
 - ② 连续型随机变量涉及概率密度的问题

→ 第3讲 多维随机变量及其分布（在一维的基础上升级到了多维）

第4讲 随机变量的数字特征

- (1) 随机变量的数学期望。
- (2) 随机变量的方差。
- (3) 两个随机变量之间的关系（协方差、相关系数）。

→ 研究原因：① 尽管分布函数非常全面准确地刻画了一个随机变量的概率规律，但本身可能是很难求的（即第2、3讲这两个部分，往往理论意义更强一些，在实践上，分布函数有时不太容易求）——比如，牛顿-莱布尼茨公式很著名，但事实上它不太好用；
② 有时候并不需要知道分布函数，可以通过其他方式或角度来研究随机变量的一些特征，比如期望、方差、协方差、相关系数就能够很好地去体现相应的特征或所研究问题需要知道的信息

第5讲 大数定律与中心极限定理 (过渡: 从概率论到数理统计)

- ① 大数定律, 样本很多(强调大数).
 ② 中心极限定理, 实际上涉及 $n \rightarrow \infty$, 即样本充分多的情况下(强调 $n \rightarrow \infty$).
 综述: 这部分的内容实际上是在研究样本容量, 即样本个数或抽样个数充分大的情况下, 它所体现出来的概率规律

第6讲 数理统计

- (1) 介绍基本统计量及其相应分布 (三大抽样分布). $\rightarrow \begin{cases} \text{① } \chi^2 \text{ 分布;} \\ \text{② } t \text{ 分布;} \\ \text{③ } F \text{ 分布} \end{cases}$
 (2) 参数估计 (考解答题).
 (3) 假设检验.

在统计学中, 分布函数或者相应的概率密度中的量并不全知道, 此时需要根据抽样得到的数据反过来去估计概率规律中的参数. 但实际上是不能求出来真实的客观值的, 所以称之为估计

二 概率考题分布

第1讲 随机事件与概率 5分选择题, 主要集中在怎么求一个复杂事件的概率.

第2、3讲 一/多维随机变量及其分布 5分小题/12分解答题. $\rightarrow$ 常考

第4讲 随机变量的数字特征 5分小题/配合第2、3讲考12分解答题. $\rightarrow$ 考研连续两年考了条件数学期望

第5讲 大数定律与中心极限定理 5分小题.

第6讲 数理统计 5分选择题(抽样分布)/12分解答题(估计问题).

目录

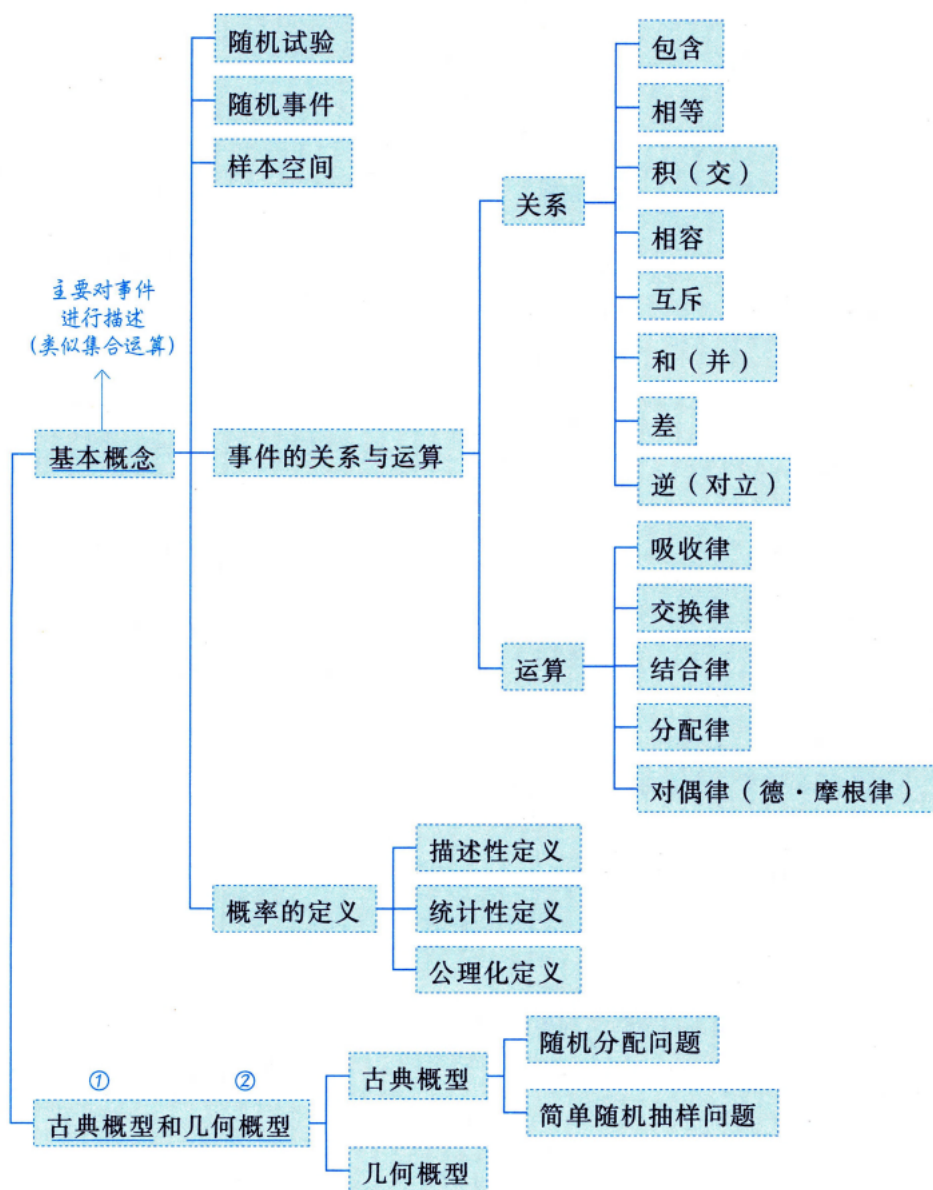
第 1 讲	随机事件与概率	1
第 2 讲	一维随机变量及其分布	33
第 3 讲	多维随机变量及其分布	62
第 4 讲	随机变量的数字特征	103
第 5 讲	大数定律与中心极限定理	125
第 6 讲	数理统计	133

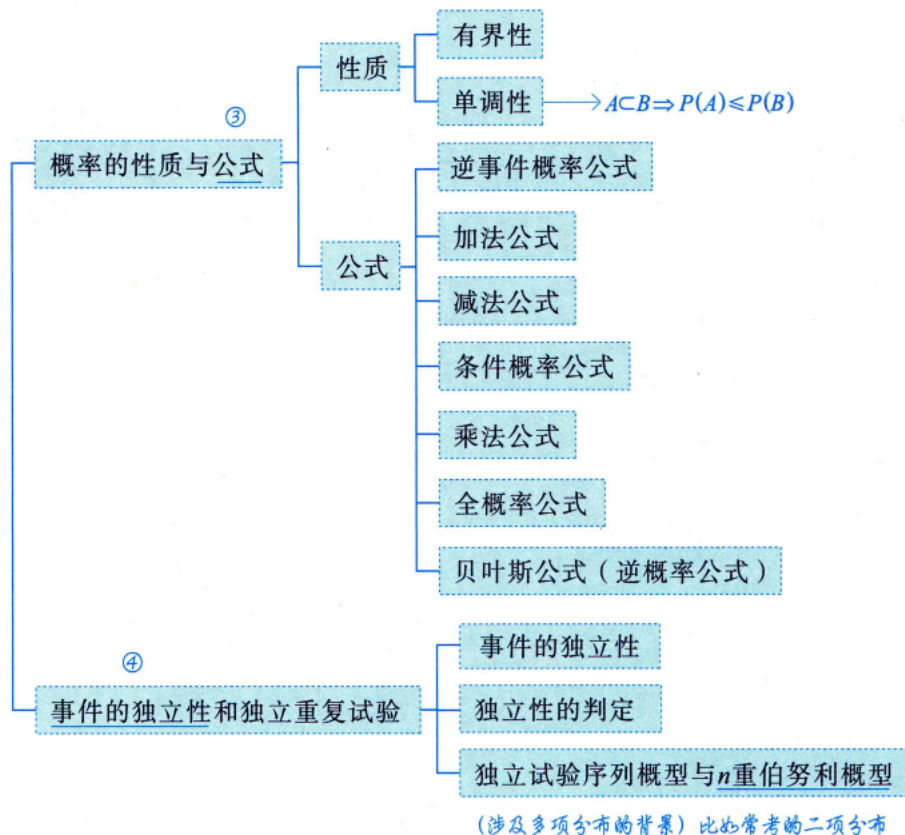
第1讲

随机事件与概率



基础知识结构





基础内容精讲

基本概念

1 随机试验 $\rightarrow E$

如果一个试验满足以下三个条件:

- (1) 试验可以在相同的条件下重复进行;
- (2) 试验所有可能结果明确可知, 且不止一个;
- (3) 每一次试验会出现哪一个结果, 事先并不能确定,

则称此试验为**随机试验**.

我们是通过研究随机试验来研究随机现象的, 随机试验简称**试验**, 用字母 E 或 $E_1, E_2, \dots$ 表示.

示例 1 投篮.

- (1) 相同条件下可重复.

首先这是一个理想化简化, 因为并非真正同条件 (一个篮球运动员投篮或者一个普通人投篮, 假设



已经投 10 次了, 再投第 11 次, 可能就会比较累了. 所以这个同条件在学科中实际上是在近似的相同条件下可重复).

(2) 试验所有可能结果明确可知, 且不止一个, 只有“投中”和“投不中”两种可能.

(3) 每一次投篮, “投中”或“投不中”, 事先未知.

示例 2 掷硬币.

(1) 相同条件下可重复.

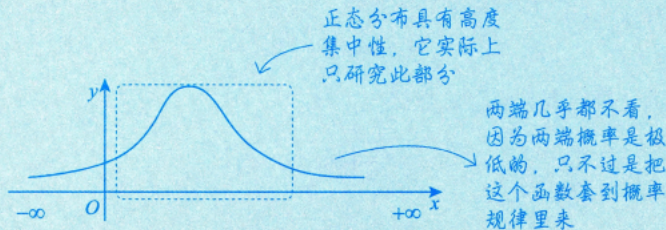
(2) 试验结果明确可知, 且不止一个, 只有“正面”和“反面”两种可能.

(3) 硬币落下之前, 不知道会是“正面”还是“反面”.

注 在不少情况下, 不能确切知道某一随机试验的全部可能结果, 但可以知道它不超出某个范围. 这时, 也可以用这个范围来作为该试验的全部可能结果的集合. 例如, 需要记录某个城市一天的交通事故数量, 则试验结果将是非负整数 x . 无法确定 x 的可能取值的确切范围, 但可以把这个范围取为 $[0, +\infty)$, 将上限取到 $+\infty$, 扩大范围, 它总能包含一切可能的试验结果, 尽管明知某些结果——如 $x > 10\,000$ 是不会出现的. 甚至把这个范围取为 $(-\infty, +\infty)$ 也无妨. 这里体现了一定的数学抽象, 它可以带来很大的方便.

怎样去描述这个取值范围的扩大?

比如正态分布 (两边低中间高).



① 在大样本的情况下, 根据中心极限定理, 所有分布最终都会收敛于正态分布. 这是自然规律中的一个神秘规律.

② 普通现象的概率肯定是高的, 而极端现象的概率是低的.

比如, 人的身高基本上集中在某一个范围内. 那么像姚明的身高, 是少数的, 然后身高往下走, 变成负的, 那是不可能的事情. 但是正态分布里 x 的范围就是从 $-\infty$ 到 $+\infty$, 这相当于做了范围的扩大, 能够使得概率分布完全对应上高斯曲线, 极大方便了数学处理. 实际上, 曲线两端取值的可能性非常小, 几乎取不到.

③ 疑问: 身高怎么可能取负值呢?

答: 并非是身高取负值, 而是对自然现象进行了范围扩大. 比如研究某学校学生身高的分布、体重的分布或者其他健康指标的分布. 用什么函数来研究它呢? 总不能用截断的部分区间的函数去研究, 而将 x 的范围扩大后刚好可以用这样一条从 $-\infty$ 到 $+\infty$ 上的著名曲线来研究概率规律.

2 随机事件

在一次试验中可能出现,也可能不出现的结果称为**随机事件**,简称**事件**,并用大写字母 A, B, C 等表示.将每次试验中一定发生的事件称为**必然事件**,记为 Ω .每次试验中一定不发生的事件称为**不可能事件**,记为 $\emptyset$.

集合中空集的符号

集合中的全集符号

大样本情况下
寻找客观规律

注 随机事件在一次试验中是否发生虽然不能事先确定,但是在大量重复试验的情况下,它的发生呈现出一定的规律性,这门课程正是要研究这种规律性,考生应在学习这门课程后,对此有较为深刻的认识.

3 样本空间

不可再分

随机试验的每一个可能结果称为**样本点**,记为 ω .样本点的全体组成的集合称为**样本空间**(或**基本事件空间**),记为 Ω ,即 $\Omega = \{\omega\}$.由一个样本点构成的事件称为**基本事件**.随机事件 A 总是由若干个基本事件组成,即 A 是 Ω 的子集.

示例 掷骰子,共六个面,即可能会出现1点到6点,记 $\omega_i = \{\text{掷出点数} i\}$,则

(1) 每个 ω_i 都是不可再分的样本,每一个 ω_i 都叫基本事件.

(2) 这些基本事件组成的集合叫作样本空间,即

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}.$$

(3) 随机事件 A 为“掷出偶数点”,显然

$$A = \{\omega_2, \omega_4, \omega_6\}$$

是 Ω 的子集.

4 事件的关系与运算

(1) 关系.

包含: 如果事件 A 发生必导致事件 B 发生,则称事件 B 包含事件 A (或 A 被 B 包含),记为 $A \subset B$.

相等: 如果 $A \subset B$ 且 $B \subset A$,则称事件 A 与 B 相等,记为 $A = B$. A 与 B 相等,事实上也就是说, A 与 B 由一些完全相同的试验结果构成,它不过是同一事件表面上看起来不同的两个说法而已.

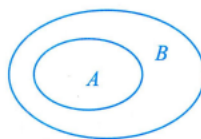
积(交): 称“事件 A 与 B 同时发生”的事件为事件 A 与 B 的积事件(或交事件),记为 $A \cap B$ 或 AB .

如,掷骰子,共六个面, $A = \{\text{掷出的点数是偶数}\}$,

$B = \{\text{掷出的点数不是} 1, 3, 5\}$,

$C = \{\text{掷出的点数是} 2 \text{的倍数}\}$,则 $A = B = C$

同时,还有概率
的单调性,显然
 $P(B) \geq P(A)$



注 称“有限个(或可列个)事件 $A_1, A_2, \dots, A_n(\dots)$ 同时发生”的事件为事件 $A_1, A_2, \dots,$

$A_n(\dots)$ 的积事件(或交事件),记为 $\bigcap_{i=1}^n A_i$ (或 $\bigcap_{i=1}^{\infty} A_i$).

相容: 若 $AB \neq \emptyset$, 则称事件 A 和 B 相容.

互斥: 若 $AB = \emptyset$, 则称事件 A 与 B 互不相容, 也叫互斥. 如果一些事件中任意两个事件都互斥, 则称这些事件是两两互斥的, 或简称互斥的. 加法公式: $P(A \cup B) = P(A) + P(B)$

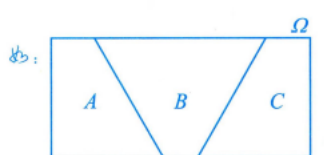
和 (并): 称“事件 A 与 B 至少有一个发生”的事件为事件 A 与 B 的和事件 (或并事件), 记为 $A \cup B$.

注 称“有限个 (或可列个) 事件 $A_1, A_2, \dots, A_n(\dots)$ 至少有一个发生”的事件为事件 $A_1, A_2, \dots, A_n(\dots)$ 的和事件 (或并事件), 记为 $\bigcup_{i=1}^n A_i$ (或 $\bigcup_{i=1}^{\infty} A_i$).

差: 称“事件 A 发生而事件 B 不发生”的事件为事件 A 与 B 的差事件, 记为 $A - B$.

逆 (对立): 称“事件 A 不发生”的事件为事件 A 的逆事件或对立事件, 记为 $\bar{A}$.

由定义易知

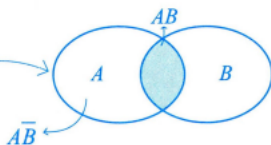


A 与 B , B 与 C , A 与 C 均为互斥事件, 但 $B \cup C$ 才为 A 的对立事件

对立事件一定是互不相容事件, 但互不相容事件不一定是对立事件



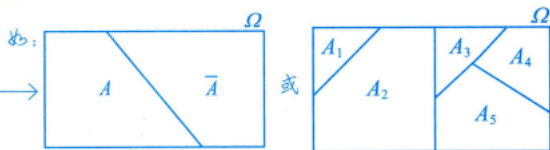
$$A - B = A - AB = \overline{AB},$$



$$B = \bar{A} \Leftrightarrow AB = \emptyset \text{ 且 } A \cup B = \Omega.$$

对立事件

完备事件组: 如果 $\bigcup_{i=1}^n A_i$ (或 $\bigcup_{i=1}^{\infty} A_i$) $= \Omega$, $A_i A_j = \emptyset$ (对一切 $i \neq j$; $i, j = 1, 2, \dots, n(\dots)$), 称有限个 (或可列个) 事件 $A_1, A_2, \dots, A_n(\dots)$ 构成一个完备事件组.



注 完备事件组也称为对样本空间的一种划分, 也叫全集分解.

文氏图：事件的关系与运算可以用文氏图形象地表示出来（见图 1-1），图中的矩形表示必然事件 Ω 。

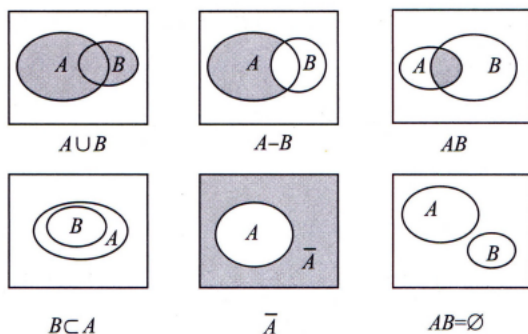


图 1-1

(2) 运算.

①吸收律 若 $A \subset B$, 则 $A \cup B = B$, $A \cap B = A$.

②交换律 $A \cup B = B \cup A$, $A \cap B = B \cap A$.

③结合律 $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$.

④分配律 $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,

$$A \cap (B - C) = (A \cap B) - (A \cap C).$$

★⑤对偶律（德·摩根律） $\overline{A \cup B} = \bar{A} \cap \bar{B}$, $\overline{A \cap B} = \bar{A} \cup \bar{B}$.

→ 口诀：长杠变短杠，开口换方向。

$$\text{如 } \overline{\overline{A \cap B}} = \overline{\bar{A} \cup \bar{B}} = A \cap B.$$

注 (1) 事件运算顺序约定为先进进行逆运算，然后进行交运算，最后进行并或差运算。

(2) 事件的关系、运算与集合的关系、运算相当，且具有相同的运算法则，所以可以对比着理解记忆，并要学会用集合关系去考虑事件关系。

5 概率的定义

(1) 描述性定义.

通常将随机事件 A 发生的可能性大小的度量（非负值）称为事件 A 发生的概率，记为 $P(A)$ 。

如：抛硬币 1 000 次，正面朝上的次数为 527，那么 $\frac{527}{1000}$ 称为这 1 000 次试验中，硬币正面朝上的频率

(2) 统计性定义.

在相同条件下做重复试验，事件 A 出现的次数 k 和总的试验次数 n 之比 $\frac{k}{n}$ 称为事件 A 在这 n 次试验中出现的频率。当试验次数 n 充分大时，频率将“稳定”于某常数 p 。 n 越大，频率偏离这个常数 p 的可能性越小。这个常数 p 就称为事件 A 的概率。

→ 不断趋近 p

比如说投掷硬币正面朝上的频率是 $\frac{527}{1000} = 0.527$ ，但它是硬币正面朝上的概率吗？

对于均匀的硬币，正面朝上的概率应该是 $\frac{1}{2}$ 。这是客观规律。

理想情况下，均匀硬币是真正的均匀，此时就说它的概率是客观上的 $\frac{1}{2}$ 。频率只能是逐渐地靠近 $\frac{1}{2}$ 。

但在硬币不均匀的情况下，不清楚抛出硬币正面朝上的概率，不能直接讲是 $\frac{1}{2}$ 。

注 后面会提到一个概念，就是频率依概率收敛到概率。这实际上涉及大数定律的问题了。

(1) 概率的统计性定义实质上是说，用频率 $\frac{k}{n}$ 作为事件 A 的概率 $P(A)$ 的估计。其直观理解为某事件出现的可能性大小，可由其在多次重复试验中出现的频率去刻画。

(2) 从上述 (1) 可以看出，频率只是概率的估计，而非概率本身。也就是说，概率的统计性定义是无法准确给出某事件的概率的，其重要性主要基于以下两点。

① 它提供了估计概率的方法。比如在一批产品中抽取样品，来估计该批产品的合格率（合格率是客观的数据，抽取样品计算出来的合格率只是一种估计）。

② 它提供了一种检验某结论是否正确的准则。比如，你说某批产品的合格率是 95%，我们做试验，抽取样品进行计算，得出的结果是合格率为 20%，远远低于你所说的 95%，于是毫不犹豫地拒绝你的结论。

统计性定义，它没有理论上的价值，不能精确定义概率。但是从实践上来讲，它有时候是很有用的。

(3) 公理化定义。

设随机试验的样本空间为 Ω ，如果对每一个事件 A 都有一个确定的实数 $P(A)$ ，且事件函数 $P(\cdot)$ 满足：

① 非负性： $P(A) \geq 0$ ；

② 规范性： $P(\Omega) = 1$ ； “归一性”

③ 可列可加性：对任意可列个两两互不相容事件 $A_1, A_2, \dots, A_n, \dots$ （即 $A_i A_j = \emptyset, i \neq j; i, j = 1, 2, \dots$ ），有

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i), \quad \text{事件和的概率等于概率的和}$$

则称 $P(\cdot)$ 为概率， $P(A)$ 为事件 A 的概率。

注 (1) 数学上所说的“公理”，就是一些不加证明而承认的前提，上述公理化定义只是界定了概率这个概念所必须满足的一些一般性质，它不解决具体场合下的概率计算。

(2) 概率 $P(\cdot)$ 是事件的函数。

(3) 虽然公理化定义不解决具体场合下的概率计算，但是我们却常常用它来判断某事件函数 $P(\cdot)$ 是不是概率，这种题型在考研试题中也是经常遇到的。

二 古典概型和几何概型



1 古典概型

(1) 概念.

如果随机试验的样本空间满足:

 ① 只有有限个样本点 (基本事件); $\rightarrow$ 如掷骰子, $\Omega = \{\omega_1, \omega_2, \dots, \omega_6\}$, 一旦出现无穷个, 就不是古典概型

② 每个样本点 (基本事件) 发生的可能性都一样,

 $\rightarrow$ 具有等可能性

则称随机试验的概率模型为古典概型.

 如果古典概型的基本事件总数为 n , 事件 A 包含 k 个基本事件, 也叫作有利于 A 的基本事件有 k 个,

 则 A 的概率为

$$P(A) = \frac{k}{n} = \frac{\text{事件 } A \text{ 所含基本事件的个数}}{\text{基本事件总数}}. \quad \text{如, 掷骰子, 事件 } A = \{\text{掷出偶数点}\}, \text{ 则 } P(A) = \frac{3}{6} = \frac{1}{2}$$

 由上式计算得出的概率称为 A 的古典概率.

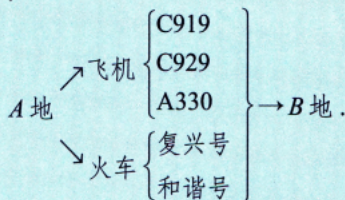
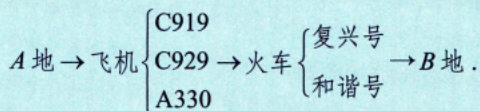
注 计算的关键是基本事件、样本空间的选定以及基本事件数的计算.

计数方法常用的有三种.

(1) 列举法 (直接数数法): 基本事件数不多时常用这种方法.

 (2) 集合对应法: ① 加法原理——完成一件事有 n 类办法, 第一类办法中有 m_1 种方法, 第二类办法中有 m_2 种方法, $\dots$, 第 n 类办法中有 m_n 种方法, 则完成此事共有 $\sum_{i=1}^n m_i$ 种方法.

 ② 乘法原理——完成一件事有 n 个步骤, 第一步有 m_1 种方法, 第二步有 m_2 种方法, $\dots$, 第 n 步有 m_n 种方法, 则完成此事共有 $\prod_{i=1}^n m_i$ 种方法.

 例如, 从 A 地到 B 地的方法.

 看路线, 从加法原理的角度, 可以有 $m_1 + m_2 = 3 + 2 = 5$ (种) 交通方式.

 看路线, 从乘法原理的角度, 需要 2 个步骤. 第 1 步: 坐飞机有 3 种选择; 第 2 步: 坐火车有 2 种选择. 故可以有 $m_1 \cdot m_2 = 3 \times 2 = 6$ (种) 交通方式.

加法原理与乘法原理有一个重要的区分,即在完成这件事情的时候,是否需要分步骤进行.如果需要分步骤进行,即用乘法原理解决.如果不是按照步骤,而是按照完成一件事情有几类方法,把这些方法加起来即可.

③排列——从 n 个不同的元素中取出 $m(\leq n)$ 个元素,并按照一定顺序排成一列,叫作排列.所有排列的个数叫作排列数,记作

有 n 个不同的元素,按顺序取 m 个.取第1个时,有 n 种取法;取第2个时,剩余 $n-1$ 个元素则有 $n-1$ 种可能性,故再乘上 $n-1$,这是第二步;依次类推到取第 m 个元素时,还剩 $n-m+1$ 种取法,所以按乘法原理把它们乘起来即可得到排列数

$$P_n^m = n(n-1)(n-2)\cdots(n-m+1) = \frac{n!}{(n-m)!}.$$

当 $m=n$ 时, $P_n^n = \frac{n!}{0!} = n!$, 叫作全排列.

④组合——从 n 个不同的元素中取出 $m(\leq n)$ 个元素,并成一组,叫作组合.所有组合的个数叫作组合数,记作

$$C_n^m = \frac{P_n^m}{m!}.$$

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

$$C_n^m = C_n^{n-m}.$$

$$0! = 1.$$

$$C_n^0 = 1$$

(3) 逆数法: 先求 $\bar{A}$ 中的基本事件数 $n_{\bar{A}}$, 用基本事件总数 n 减去 $n_{\bar{A}}$ 便得 A 中的基本事件数, 这种方法常用于计算含有“至少”字样的事件的概率. 用对立事件解题: $P(A) = 1 - P(\bar{A})$

(2) 计算方法.

①随机分配问题.

随机分配也叫随机占位, 突出一个“放”字, 即将 n 个可辨质点随机地分配到 N 个盒子中, 区分每盒可以容纳任意多个质点和最多可以容纳一个质点, 不同分法的总数如表所示.

比如在1月1日出生的人, 有很多

→ 一个萝卜一个坑

分配方式	不同分法的总数
每盒容纳任意多个质点	N^n (见注①)
每盒容纳至多一个质点	$P_N^n = N(N-1)\cdots(N-n+1)$ (见注②)

注 ①每个质点均可放到 N 个盒子中的任何一个, 即有 N 种放法, 于是 n 个可辨质点放到 N 个盒子中共有 N^n 种不同放法.

②质点可辨, 且一个盒子至多容纳一个质点, 故 n 个质点放到 $N(N \geq n)$ 个盒子中的所有不同放法即为从 N 个元素中选取 n 个元素的排列数 P_N^n .

→生日问题

例 1.1 将 n 个球随机放入 $N(n \leq N)$ 个盒子中, 每个盒子可以放任意多个球. 求下列事件的概率: $A = \{ \text{某指定 } n \text{ 个盒子各有一球} \}$; $B = \{ \text{恰有 } n \text{ 个盒子各有一球} \}$; $C = \{ \text{指定 } k(k \leq n) \text{ 个盒子各有一球} \}$.

解 这是随机占位的问题, 设 n 个球, N 个盒子是可分辨的 (例如编号), 由于每个盒子可以放任意多个球, 因此每个球都有 N 种不同的放置方法. 将 n 个球随机放入 N 个盒子的一种放法作为基本事件, 则基本事件总数为 N^n . 事件 A 所含的基本事件是 n 个不同球的一种排列, 故

$$n_A = n!, \quad P(A) = \frac{1 \cdot n!}{N^n}.$$

“某指定 n 个”, 只有此一种情况



(n 个盒子的位置是唯一确定的, 现假定就是前 n 个位置)

事件 B 中的基本事件可以设想为先从 N 个盒子中选出 n 个 (共有 C_N^n 种不同选法), 然后把 n 个球随机放入这 n 个盒子中, 每盒一球 (共有 $n!$ 种不同放法), 因此 B 中基本事件数

$$n_B = C_N^n \cdot n!, \quad P(B) = \frac{C_N^n \cdot n!}{N^n}.$$

“恰有 n 个”, 有 C_N^n 种情形



(n 个盒子的位置并非唯一确定的, 是从 N 个里面选出 n 个)

事件 C 的基本事件可以设想为先从 n 个球中选出 k 个球 (有 C_n^k 种不同选法), 再将这 k 个球随机放入指定的 k 个盒子中, 每盒一球 (有 $k!$ 种不同放法), 最后将余下的 $n-k$ 个球随机放入其余的 $N-k$ 个盒子中, 每个球都有 $N-k$ 种放置方法, 因此共有 $(N-k)^{n-k}$ 种不同放法, 故

$$n_C = C_n^k \cdot k! (N-k)^{n-k}, \quad P(C) = \frac{C_n^k k! (N-k)^{n-k}}{N^n} = \frac{n! (N-k)^{n-k}}{(n-k)! N^n}.$$

方法总结 必背基本事件个数, 将 n 个球放入 n 个盒子, 使得每个盒子各有一球, 有 $n!$ 种方法.

注 许多问题的结构形式与分球入盒问题相同, 都属于随机占位问题. 例如生日问题 (n 个人生日, 相当于 n 个球随机放入 365 个盒子中, 每盒可以放多个球); 住房分配问题 (n 个人被分配到 N 个房间中去, 每个房间可住多个人); 乘客下车问题 (n 个乘客在 N 个车站下车的各种可能情况); 等等. 对这些问题的求解都可以用“将 n 个球等可能地投放到 N 个盒子中”的思路来考虑.

比如, 12 个球随机放入 365 个盒子中, 每盒可放多个球.

$$A_1 = \{ \text{指定 12 个盒子, 各放 1 球} \}, \quad \text{则 } P(A_1) = \frac{12!}{365^{12}};$$

$$B_1 = \{ \text{恰有 12 个盒子, 各放 1 球} \}, \quad \text{则 } P(B_1) = \frac{C_{365}^{12} 12!}{365^{12}};$$

$$C_1 = \{ \text{指定 3 个盒子, 各放 1 球} \}, \quad \text{则 } P(C_1) = \frac{C_{12}^3 3! (365-3)^{12-3}}{365^{12}}.$$

又比如 12 个人 $\omega_1, \dots, \omega_{12}$ 回母校参加校庆, 每个人在 365 天中的某一天出生的概率相等. 令

$A_2 = \{ \text{生日全不相同, 且分别为每个月的第一天} \};$

$B_2 = \{ \text{生日全不相同} \}; \bar{B}_2 = \{ \text{至少有两人生日相同} \};$

$C_2 = \{ \text{有且仅有三个人的生日分别在劳动节、儿童节、中秋节} \},$

则 A_2, B_2, C_2 就对应着题中的 A, B, C , 只不过此时 $N=365, n=12, k=3$.

②简单随机抽样问题.

设 $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$ 含 N 个元素, 称 Ω 为总体. 如果各元素被抽到的可能性相同, 则总体 Ω 的抽样称作简单随机抽样, 突出一个“取”字.

简单随机抽样分为先后有放回、先后无放回及任取这三种不同的方式. 在每种抽样方式下各种不同抽取方法 (基本事件) 的总数如表所示.

抽样方式	抽取方法总数
先后有放回取 n 次 (样本空间始终无变化)	N^n (见注 (1) ①)
先后无放回取 n 次	$P_N^n = N(N-1) \cdots (N-n+1)$ (见注 (1) ②)
任取 n 个	C_N^n (见注 (1) ③)

注 (1) ①既考虑抽到何元素, 又考虑各元素出现的顺序, 每次从 Ω 中随意抽取一个元素, 并在抽下一元素前将其放回 Ω , 于是每次都有 N 个元素可被抽取, 即有 N 种抽取方法, 抽取 n 次, 即 N^n .

②既考虑抽到何元素, 又考虑各元素出现的顺序, 凡是抽出的元素均不再放回 Ω , 于是每次抽取时都比上一次少了一个元素, 抽 $n(n \leq N)$ 次, 即

$$P_N^n = N(N-1) \cdots (N-n+1).$$

③任取 $n(n \leq N)$ 个是指一次性取 n 个元素, 相当于将 n 个元素无序且无放回地取走, 其抽取方法总数为

$$C_N^n = \frac{P_N^n}{n!}.$$

(2) ①先后有放回与先后无放回的区别. 先后有放回取也不意味是真的放回去了. 当产品的数量足够多的时候, 拿走一个不放回来, 也当作有放回处理 (因为若按照先后无放回去计算, 可能比较复杂, 所以可以当作有放回). 实际上是改变了样本空间, 每取走一个, 样本空间里就少一个元素, 但是如果元素足够多, 取走一个可以近似看作原来的样本空间没有变, 所以如果样本空间不变, 那很多问题研究起来就非常方便了. 故先后有放回取, 往往是用在样本 n 足够大的情况下.

②先后无放回与任取的区别. 先后无放回, 强调“先后”顺序; 任取无顺序, 故计算公式少了排列 $n!$.

例 1.2 袋中有 5 个球, 3 个白球, 2 个黑球.

- (1) 先后有放回取 2 个球;
- (2) 先后无放回取 2 个球;
- (3) 任取 2 个球.

求取的 2 个球中至少 1 个是白球的概率.

分析 考查简单抽样的 3 种抽取方法. 要求必会.

解 用对立事件思想, 计算“两球全黑”的种数, 再用总数减去它.

$$(1) 5^2 - 2^2 = 21(\text{种}).$$

$$(2) 5 \times 4 - 2 \times 1 = 18(\text{种}).$$

$$(3) C_5^2 - C_2^2 = 9(\text{种}).$$

下面计算概率.

$$(1) p_1 = \frac{5^2 - 2^2}{5^2} = 1 - \frac{2^2}{5^2} = \frac{21}{25}.$$

$$(2) p_2 = \frac{5 \times 4 - 2 \times 1}{5 \times 4} = 1 - \frac{2 \times 1}{5 \times 4} = \frac{9}{10}.$$

$$(3) p_3 = \frac{C_5^2 - C_2^2}{C_5^2} = 1 - \frac{C_2^2}{C_5^2} = \frac{9}{10}.$$

注 本题中的 (2) 与 (3) 概率相同, 如何理解?

$\frac{2 \cdot 1}{5 \cdot 4} = \frac{C_2^2 \cdot P_2^2}{C_5^2 \cdot P_2^2} = \frac{C_2^2}{C_5^2}$, 左边上下有序, 右边上下无序, 相当于把顺序“消掉”了, 故“先后无

放回取 k 个球”与“任取 k 个球”的概率相同. 由于 (3) 较方便, 因此计算 (2) 时, 可按 (3) 来计算.

例 1.3 袋中有 100 个球, 40 个白球、60 个黑球.

- (1) 先后有放回取 20 个球, 求取出 15 个白球、5 个黑球的概率;
- (2) 先后无放回取 20 个球, 求取出 15 个白球、5 个黑球的概率;
- (3) 先后有放回取 20 个球, 求第 20 次取到白球的概率;
- (4) 先后无放回取 20 个球, 求第 20 次取到白球的概率.

分析 (1) 转化为投篮模型.

$$\textcircled{1} P\{\text{投中}\} = 40/100, P\{\text{投不中}\} = 60/100.$$

② 这样的投篮试验做 20 次, 15 次投中、5 次投不中 (先后有放回取 20 个球, 每次试验之前, 球放回去了, 它的样本空间没有变化), 这属于二项分布.

(2) 简单随机抽样问题转化为随机占位问题.

解 (1) $p_1 = \frac{40^{15} \cdot 60^5 \cdot C_{20}^{15}}{100^{20}}$.
 20个位置中有15个位置放白球
 每次取都是100种取法, 取20次

(2) 按照“任取20个球”来计算, 即 $p_2 = \frac{C_{40}^{15} C_{60}^5}{C_{100}^{20}}$.

(3) $p_3 = \frac{40}{100} = \frac{2}{5}$ (有放回取球, 每次抽取的样本空间没有变化, 故每次取到白球的概率始终为 $\frac{2}{5}$).

(4) 看作随机占位问题. 设有100个盒子, 每个盒子中放1个球, 则要求第20个盒子中放入白球即

可. 故 $p_4 = \frac{C_{40}^1 \cdot 99!}{100!} = \frac{40}{100} = \frac{2}{5}$.
 剩下99个球放在99个盒子里, 每个盒子各有1球
 40个白球抽取1个放在第20个盒子中

注 (1) 若(4)修改为: 先后无放回取 k 个球, 求第 k 次取到白球的概率. 结果仍为 $\frac{2}{5}$.

第20个盒子放的是白球, 但是前19个盒子放的什么球并不清楚. 若清楚了, 即转化为条件概率问题

(2) $p_4 = p_3$. 例 1.3(4) 是抓阄模型, 即使无放回, 每次取到白球的概率也不会变, 可理解成依概率摸球. 类比的例子: (绝对概率)

① 设有100个灰球, 白的成分: 黑的成分 = 40 : 60, 每次取到球中白的成分是 40%;

② 设有糖水, 糖的成分: 水的成分 = 40 : 60, 每次取一勺糖水, 取到糖的成分是 40%.

2 几何概型

(可测)在测度论的基础上, 概率论会非常精确严谨, 但是在非测度论的基础上, 理解即可

如果随机试验的样本空间满足:

① 样本空间 (基本事件空间) Ω 是一个可度量的有界区域;

【名词解释】可度量: 通俗来讲, 即为可以度量, 就是我们现在理解的长度、面积、体积.

对于数学一: 计算体积, 则要涉及三重积分;

对于数学三: 可度量其实基本上就限定在一维的长度和二维的面积上.

② 每个样本点 (基本事件) 发生的可能性都一样, 即样本点落入 Ω 的某一可度量的子区域 S 的可能性大小与 S 的几何度量成正比, 而与 S 的位置及形状无关, 则称随机试验的概率模型为几何概型.

示例 “天上掉馅饼”.

假如明天早晨 9:00, 在学校操场上会从天上掉一个馅饼下来. 大家拿着食堂的饭盆去接这个馅饼. 现思考两个问题:

9:00



(1) 拿着饭盆到操场上去, 请问大家站在哪里最容易接到这个馅饼?

面积的微分

站在哪儿都行. 在概率论上, 9:00 有一个馅饼要落下来, 这个馅饼落到操场上任意一个 dS , 具有等可能性. 实际上是有无数个 dS 的, 即无数个面积微分. 所以跟你拿着饭盆站哪里没关系.

(2) 跟饭盆的什么有关系?

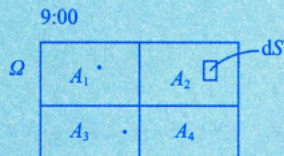
与饭盆的形状没关系, 但是与饭盆的面积有关系. 若饭盆的面积 = 操场面积的 $\frac{1}{2}$, 则 $P\{\text{接到馅饼}\} = \frac{1}{2}$; 若饭盆的面积 = 操场面积, 则 $P\{\text{接到馅饼}\} = 1$.

注 古典概型与几何概型的区别: 基本事件有限、等可能发生的随机试验为**古典概型**; 基本事件无限且具有几何度量、等可能发生的随机试验为**几何概型**.

古典概型与几何概型的联系 (如下例).

操场被平均分成四部分, A_1, A_2, A_3, A_4 . 现假设 4 个人分别拿着与 A_1, A_2, A_3, A_4 同等面积的饭盆, 此时相当于对 Ω 进行均等的划分.

显然, 馅饼落到 4 个饭盆里具有等可能性, 且基本事件数只有四个, 即基本事件数是有限的, 此时就可以把**几何概型转化成古典概型**.



例 1.4 在区间 $(0, 1)$ 中随机地取两个数, 则这两个数之差的绝对值小于 0.5 的概率为_____.

分析 在区间 $(0, 1)$ 上随机取两个数, 这两个数的取值范围就可构成一个可以度量的几何区域, 实际上是二维平面上的一个正方形. 理解为“馅饼落在操场任何一个点具有等可能性”. 记事件

$$A = \{(x, y) | |x - y| < 0.5\} \cap \Omega,$$

$$\text{则 } p = \frac{S_A}{S_\Omega}.$$

解 应填 0.75.

设两个数分别为 x, y , 则依题意 (x, y) 的取值范围为正方形区域 $D = \{(x, y) | 0 < x < 1, 0 < y < 1\}$. 又 $|x - y| < 0.5$, 则所在区域如图 1-2 中阴影部分所示, 于是所求概率为

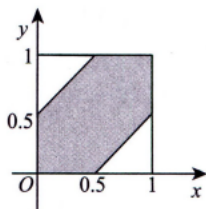


图 1-2

$$p = \frac{1^2 - 0.5^2}{1^2} = 0.75.$$

三 概率的性质与公式



1 性质

不可能事件的概率

全集，也就是必然事件的概率

(1) 有界性：对于任一事件 A ，有 $0 \leq P(A) \leq 1$ ，且 $P(\emptyset) = 0$ ， $P(\Omega) = 1$ 。

8:00~9:00按照时间连续来记，可以认为它是一个可以度量的长度，那么任何一个时刻进入这个教室是具有等可能性的

如，用 Ω 表示一个人在8:00~9:00之间进入教室。 $A = \{8:30 \text{ 到达教室}\}$ ，则 $P(A) = \frac{0}{1} = 0$ 。

8:00~9:00时间长度为1，

8:30是一个点，长度为0。

注 $P(A) = 0$ ，不能断言 $A = \emptyset$ ； $P(A) = 1$ ，不能断言 $A = \Omega$ 。从对立事件理解， $\bar{A} = \{\text{不在8:30进入教室}\}$

对于

$$A = \emptyset \Leftrightarrow P(A) = 0,$$

要知道，在非测度论的基础上，有可能会测不到那个概率。这是我们在现在这个概率学的基础上，数学工具不够完善所导致的，即明明是可能发生的事情，但是测不到它，故它的概率是零了。但是概率是零的事件，却不一定是不可能事件。

两种身份

- ① 它确实确实是不可能事件，所以空集里面什么元素都没有；
- ② 在我们现在所学习的基础上， A 事件的测量是测不到它的，测不到它，那概率依然是零。

示例1 假如，一个男生追求一个女生或者一个女生追求一个男生，如果对方说我们在一起的概率为零。我们知道，这不一定是不可可能事件，还有一丝希望，只要有一丝希望，就要做百倍努力坚持追求下去。**示例2** 有的同学跟你讲，你考上研究生的概率为零。但这也不是不可能实现的。只要有一丝希望，我们都要尽百倍努力，这样才能够在未来越来越好。★ (2) 单调性：设 A, B 是两个事件，若 $A \subset B$ ，则有

$$P(B - A) = P(B) - P(A), \quad P(B) \geq P(A).$$

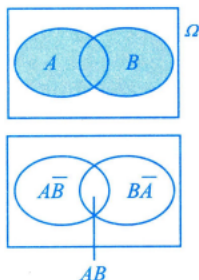
2 公式

看谁好算，灵活运用

(1) 逆事件概率公式：对于任一事件 A ，有 $P(\bar{A}) = 1 - P(A)$ 。(2) 加法公式：对于任意两个事件 A, B ，有

$$P(A \cup B) = P(A) + P(B) - P(AB),$$

$$P(A \cup B) = P(\bar{A}\bar{B} \cup AB \cup B\bar{A}) = P(\bar{A}\bar{B}) + P(AB) + P(B\bar{A}).$$



注 (1) 设 A_1, A_2, A_3 为任意三个事件, 则有

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 A_2) - P(A_1 A_3) - P(A_2 A_3) + P(A_1 A_2 A_3).$$

(2) 若 $A_1, A_2, \dots, A_n$ 是两两互不相容的事件, 则有 两两互斥

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

(3) 减法公式: $P(A - B) = P(A) - P(AB) = P(A\bar{B})$.

排除掉 $P(A)=0$, 是因为它往往出现在分母位置

★★★(4) 条件概率公式: 设 A, B 为任意两个事件, 若 $P(A) > 0$, 我们称在已知事件 A 发生的条件下, 事件 B 发生的概率为条件概率, 记为 $P(B|A)$, 且

$$P(B|A) = \frac{P(AB)}{P(A)}.$$

注 (1) 条件概率 $P(\cdot|A)$ 是概率, 概率的一切性质和结论对条件概率都适用.

例如,

$$P(\bar{B}|A) = 1 - P(B|A),$$

$$P[(B - C)|A] = P(B|A) - P(BC|A),$$

等等.

(2) 条件概率就是在一定的附加条件之下所计算的概率. 当说到“条件概率”时, 总是指另外附加了条件, 其形式可归结为“已知某事件发生了”. (或当……时)

区分: ① 有一个单位长度的质杆, 任意切两刀, 截成三段.

② 一个长度为 1 的质杆, 先切一刀, 把它分成两段之后, 再对剩下的任意一段切一刀.

以上两种说法不一致.

(5) 乘法公式: 如果 $P(A) > 0$, 则 $P(AB) = P(A)P(B|A)$.

一般地, 对于 $n > 2$, 如果 $P(A_1 A_2 \dots A_{n-1}) > 0$, 则

$$P(A_1 A_2 \dots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \dots P(A_n|A_1 A_2 \dots A_{n-1}).$$

如: $P(A_1 A_2) > 0$, 则 $P(A_1 A_2 A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2)$.

全集分解思想示例: 一个村子和三个小偷的故事.

★★★(6) 全概率公式: 如果 $\bigcup_{i=1}^n A_i = \Omega$, $A_i A_j = \emptyset (i \neq j; i, j = 1, 2, \dots, n)$, $P(A_i) > 0$, 则对任一事件

B , 有

存在 $n = \infty$ 的情形

$$B = \bigcup_{i=1}^n A_i B, \quad P(B) = \sum_{i=1}^n P(A_i)P(B|A_i).$$

示例 一个村子和三个小偷的故事.

现实中不会出现成功概率为0的小偷

假设一个村子有3个小偷, 一个叫小张, 一个叫小英, 一个叫小郑, 小张偷东西成功的概率为0, 小英偷东西成功的概率为 $\frac{3}{4}$, 小郑偷东西成功的概率为 $\frac{1}{6}$, 派一个人去偷东西.

现记

$$A_i = \{\text{第}i\text{个人去偷东西}\}, i=1, 2, 3,$$

$$B = \{\text{村子失窃}\}.$$

(1) 求村子失窃的概率;

(2) 若事件 B 已经发生了, 即发生了失窃, 那么现在第 i 个人去偷的概率 $P(A_i|B)$ 是多少呢?

→ 用贝叶斯公式, 也叫逆概率公式

分析 (1) B 事件比较复杂, 它的复杂性在于它不是一个可能原因导致的. 这里有三个原因, 都可以导致结果. 但是每个原因导致结果的能力还不一样. 所以在概率论里, 我们提出了一个重要的思想, 即全集分解思想.

(2) 当 B 事件非常复杂的时候, 它实际上属于整个试验的第二个阶段.

第一个阶段, 选人, 选哪个人去偷; 第二个阶段, 失窃, 偷东西偷成功了.

解 (1)

第一阶段 A_1, A_2, A_3 ,
第二阶段 B .

→ 失窃原因的全集

完备事件组: 保证不重不漏. 所谓不重, 就是任何两个事件的交集是空集或者说没有公共的元素; 不漏也就是所有的事件并起来, 最后是全集, 我们叫这个事件组为一个完备事件组

$$P(B) = P(B\Omega) = P[B(A_1 \cup A_2 \cup A_3)]$$

→ 任何一个集合 B , 交上它的全集 Ω , 当然还是它本身. 集合没有变, 概率不会变

$$= P(BA_1 \cup BA_2 \cup BA_3)$$

→ 根据事件的基本运算律

$$= P(BA_1) + P(BA_2) + P(BA_3)$$

→ 完备事件组是互斥的

$$= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + P(A_3)P(B|A_3)$$

$P(B|A_i)$ 代表第 i 个人偷成功的概率, 且 $\sum P(B|A_i)$ 无要求

$$= \frac{1}{3} \times 0 + \frac{1}{3} \times \frac{3}{4} + \frac{1}{3} \times \frac{1}{6} = \frac{11}{36}.$$

→ 在没有任何附加信息的条件下, 只好认为三个人去偷具有等可能性 (当然有的题目可能不具有等可能性, 但具体的题目会告知), 所以

$$P(A_i) = \frac{1}{3}, i=1, 2, 3$$

$$(2) \quad P(A_i|B) = \frac{P(A_i B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^3 P(A_i)P(B|A_i)}, i=1, 2, 3,$$

故

$$P(A_1|B) = \frac{0}{\frac{11}{36}} = 0, \quad P(A_2|B) = \frac{\frac{1}{4}}{\frac{11}{36}} = \frac{1}{4} \times \frac{36}{11} = \frac{9}{11},$$

$$P(A_3|B) = \frac{\frac{1}{18}}{\frac{11}{36}} = \frac{1}{18} \times \frac{36}{11} = \frac{2}{11}.$$

注 (1) 对比. ① 在村子没有失窃的情况下, 那三个原因具有等可能性, 即在没有任何信息的情况下, 并不知道哪个人去偷, 所以只能认为他们具有等可能性.

② 当信息增加时, 即当我们获取了一定信息的时候, 比如知道村子丢东西了, 当这个信息进来以后, 第 i 个人偷盗的这三个事件发生的概率就悄然发生了变化.

(2) 总结. 在整个试验过程中, 当我们增加一些信息的时候, 原来所谓的在没有任何信息的条件下所设立的那些可能性, 在信息增加之后, 它们的概率就随之变化. 这是很重要的一个思想, 叫作逆概率公式 (执果索因). 发生变化的原因是权重发生了变化.

(7) 贝叶斯公式 (又称逆概率公式): 如果 $\bigcup_{i=1}^n A_i = \Omega$, $A_i A_j = \emptyset (i \neq j; i, j = 1, 2, \dots, n)$, $P(A_i) > 0$,

则对任一事件 B , 只要 $P(B) > 0$, 就有

$$P(A_j | B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^n P(A_i)P(B|A_i)} \quad (j=1, 2, \dots, n).$$

方法总结 如果一个事件或者说一个结果的发生是复杂的, 它是由多个原因引起的, 那么这个时候一定要考虑到全集分解, 把原因拉进来处理问题.

注 全概率公式用于计算某个“结果” B 发生的可能性大小. 如果一个“结果” B 的发生总是与多个“原因” A_i 相联系, 那么在计算 $P(B)$ 时, 我们一般这样处理:

$$B = B\Omega = B\left(\bigcup_{i=1}^n A_i\right) = \bigcup_{i=1}^n A_i B,$$

然后应用全概率公式计算 $P(B)$, 我们常称这种方法为**全集分解法**. 如果在“结果” B 发生的条件下, 探求导致这一“结果”的各种“原因”, 即 A_i 发生的可能性大小 $P(A_i | B)$, 则要应用贝叶斯公式.

在引入随机变量之后, 还有相应的很多场合要用到全集分解思想.

例 1.5 设 A, B 为两个随机事件, 则 ().

(A) $P(AB) + P(\overline{AB}) \leq 1$

(B) $P(AB) + P(\overline{AB}) \geq 1$

(C) $P(A-B) \leq P(A) - P(B)$

(D) $P(A \cup B) + P(\overline{A \cup B}) \leq 1$

解 应选 (A).

本题涉及事件之间的关系, 由于 $AB \subset A \cup B$, 即有 $P(AB) \leq P(A \cup B)$, 从而有

$$P(AB) \leq P(A \cup B) = 1 - P(\overline{A \cup B}) = 1 - P(\overline{AB}), \quad P(\overline{AB}) = P(\overline{A \cup B})$$

即 $P(AB) + P(\overline{AB}) \leq 1$, 同时有

$$P(A \cup B) - P(AB) = P(A \cup B) + P(\overline{AB}) - 1 = P(A \cup B) + P(\overline{A \cup B}) - 1 \geq 0,$$

即 $P(A \cup B) + P(\overline{A \cup B}) \geq 1, \longrightarrow 1 - P(\overline{A \cup B}) = P(\overline{A \cup B}) = P(AB)$

知选项 (B), (D) 不正确. 又 $AB \subset B$, 有 $P(AB) \leq P(B)$, 从而有

$$P(A - B) = P(A) - P(AB) \geq P(A) - P(B),$$

知选项 (C) 不正确, 故选择 (A).

例 1.6 设随机事件 A, B 满足 $P(A) = P(B) = \frac{1}{2}$ 和 $P(A \cup B) = 1$, 则有 ().

(A) $A \cup B = \Omega$

(B) $AB = \emptyset$

(C) $P(\overline{A \cup B}) = 1$

(D) $P(A - B) = 0$

解 应选 (C).

$\rightarrow$ 不代表它一定是必然事件, 因为在几何模型上有可能测不到

前面已经指出, $P(A \cup B) = 1$ 不能推出 $A \cup B = \Omega$, 且 $P(AB) = P(A) + P(B) - P(A \cup B) = 0$ 不能推出

$AB = \emptyset$. 故选项 (A), (B) 不正确.

$\rightarrow$ 不代表是不可能事件, 比如 8:00~9:00 时段恰好在 8:30 进入教室, 这个事情是可以发生的, 但概率为 0

又 $P(A - B) = P(A) - P(AB) = \frac{1}{2}$, 选项 (D) 不正确. 从而 $P(\overline{A \cup B}) = P(\overline{AB}) = 1 - P(AB) = 1$, 选 (C).

例 1.7 设 A, B, C 是三个随机事件, A 与 C 互斥, $P(AB) = \frac{1}{2}$, $P(C) = \frac{1}{3}$, 则 $P(AB|\overline{C}) =$ _____.

分析 考查全概率公式及单调性.

解 应填 $\frac{3}{4}$.

$$0 \leq P(ABC) \leq P(AC) = 0$$

因为 A 与 C 互斥, 所以 $P(AC) = 0$. 又 $ABC \subset AC$, 所以 $P(ABC) = 0$, 故

$$P(AB|\overline{C}) = \frac{P(\overline{ABC})}{P(\overline{C})} = \frac{P(AB) - P(ABC)}{1 - P(C)} = \frac{\frac{1}{2}}{1 - \frac{1}{3}} = \frac{3}{4}.$$

注 由于 $AC = \emptyset$, 则 $ABC = \emptyset B = \emptyset$, 得到 $P(ABC) = 0$.

例 1.8 已知 $P(A) = \frac{1}{4}$, $P(B|A) = \frac{1}{3}$, $P(A|B) = \frac{1}{2}$, 则 $P(A \cup B) =$ _____.

分析 从求解所需的条件出发, 进行分析.

解 应填 $\frac{1}{3}$.

由乘法公式知

$$P(AB) = P(A)P(B|A) = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}, \quad P(B) = \frac{P(AB)}{P(A|B)} = \frac{1/12}{1/2} = \frac{1}{6},$$

所以

$$P(A \cup B) = P(A) + P(B) - P(AB) = \frac{1}{4} + \frac{1}{6} - \frac{1}{12} = \frac{1}{3}.$$

例 1.9 某人给四位亲友各写一封信, 然后随机地装入 4 个写好地址的信封中, 且每个信封装一封信. 问:

(1) 4 封信都装对了的概率;

(2) 4 封信都装错了的概率.

① 那就是 A_1, A_2, A_3, A_4 四个事件同时发生, 是“且”关系, 即第一封信装对, 且第二封信装对, 且第三封信装对, 且第四封信装对, 用乘法公式.

② 对立事件: 至少有一封信装错了, 而不是全部装错了

解 设事件 $A_i (i=1, 2, 3, 4)$ 为“第 i 封信装对了”, 事件 A 为“4 封信都装对了”, 事件 B 为“4 封信都装错了”. 于是

$$(1) \quad P(A) = P(A_1 A_2 A_3 A_4) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 A_2)P(A_4 | A_1 A_2 A_3)$$

$$= \frac{1}{4} \times \frac{1}{3} \times \frac{1}{2} \times \frac{1}{1} = \frac{1}{24}.$$

古典概型

$$(2) \quad P(B) = P(\overline{A_1} \overline{A_2} \overline{A_3} \overline{A_4})$$

德·摩根公式 $\overline{A \cap B} = \overline{A} \cup \overline{B}$

$$= 1 - P(\overline{\overline{A_1} \overline{A_2} \overline{A_3} \overline{A_4}}) = 1 - P(A_1 \cup A_2 \cup A_3 \cup A_4)$$

$$= 1 - \{P(A_1) + P(A_2) + P(A_3) + P(A_4) -$$

$$= 1 - \left[\sum_{i=1}^4 P(A_i) - \sum_{i \neq j} P(A_i A_j) + \sum_{i \neq j \neq k} P(A_i A_j A_k) - P(A_1 A_2 A_3 A_4) \right]$$

$$[P(A_1 A_2) + P(A_1 A_3) + P(A_1 A_4) + P(A_2 A_3) + P(A_2 A_4) + P(A_3 A_4)] +$$

$$[P(A_1 A_2 A_3) + P(A_1 A_2 A_4) + P(A_1 A_3 A_4) + P(A_2 A_3 A_4)] - P(A_1 A_2 A_3 A_4)\},$$

其中 $P(A_i) = \frac{1}{4}$, $P(A_i A_j) = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$, $P(A_i A_j A_k) = \frac{1}{4} \times \frac{1}{3} \times \frac{1}{2} = \frac{1}{24}$ ($1 \leq i, j, k \leq 4, i \neq j \neq k$), 于是

$$P(B) = 1 - 4 \times \frac{1}{4} + 6 \times \frac{1}{12} - 4 \times \frac{1}{24} + \frac{1}{24} = \frac{3}{8}.$$

注 本题装信过程可看作无放回抽取信件, 依次装入信封. 在计算第 (2) 问时, 直接计算较为困难, 这种情况下, 采用逆向思维方法求解问题更加简便.

例 1.10 从 1, 2, 3, 4 中任取一个数, 记为 X , 再从 1, ..., X 中任取一个数, 记为 Y , 则

$$P\{Y=2\} = \underline{\hspace{2cm}}.$$

解 应填 $\frac{13}{48}$.

$$P\{X=i\} = \frac{1}{4}, \quad i=1, 2, 3, 4,$$

$$P\{Y=2|X=1\} = 0, \quad P\{Y=2|X=2\} = \frac{1}{2}, \quad P\{Y=2|X=3\} = \frac{1}{3}, \quad P\{Y=2|X=4\} = \frac{1}{4},$$

故

$$P\{Y=2\} = \sum_{i=1}^4 P\{X=i\}P\{Y=2|X=i\} = \frac{1}{4} \times 0 + \frac{1}{4} \times \frac{1}{2} + \frac{1}{4} \times \frac{1}{3} + \frac{1}{4} \times \frac{1}{4} = \frac{13}{48}.$$

例 1.11 一批产品有 10 个正品和 2 个次品, 任意抽取两次, 每次抽一个, 抽出后不再放回, 则第二次抽出的是次品的概率为_____.

先后不放回取两次

解 应填 $\frac{1}{6}$.

记 $A_i = \{\text{第 } i \text{ 次取得次品}\}, i=1, 2$, 则由已知得

$$P(A_1) = \frac{2}{12} = \frac{1}{6}, P(\bar{A}_1) = \frac{10}{12} = \frac{5}{6}, P(A_2|A_1) = \frac{1}{11}, P(A_2|\bar{A}_1) = \frac{2}{11}.$$

$$\text{由全概率公式得 } P(A_2) = P(A_1)P(A_2|A_1) + P(\bar{A}_1)P(A_2|\bar{A}_1) = \frac{1}{6} \times \frac{1}{11} + \frac{5}{6} \times \frac{2}{11} = \frac{1}{6}.$$

注 此例为抓阄模型: 有 12 个可辨产品 (10 个正品和 2 个次品) 放入 12 个格子, 每个格子放 1 个产品, 则放法总数为 $12!$, 现只关注第 2 个格子放两个次品中的一个, 则放法总数为 $C_2^1 \cdot 1 \cdot 11!$, 故 $p = \frac{C_2^1 \cdot 1 \cdot 11!}{12!} = \frac{1}{6}$, 此过程与结果表明, 从第 1 次到第 12 次取到次品的概率均为 $\frac{1}{6}$, 即亦可用例 1.3(4) 的结论直接得出结果 $\frac{1}{6}$.

结果已知, 正品可能来自第一批,
也可能来自第二批

例 1.12 设有两批数量相同的零件, 已知有一批产品全部合格, 另一批产品有 25% 不合格. 从两批产品中任取 1 只, 经检验是正品, 放回原处, 并从原所在批次再取 1 只, 求这只产品是次品的概率.

分析 两次抽取情况有不同之处, 第一次是在完全不知情的情况下等可能地从两批产品中抽取, 第二次是在第一次抽取产品并检验合格后, 这时对所抽取是第一批还是第二批产品的概率已有计算, 因此计算要分两步走. 第一步: 在已知抽取产品合格的条件下, 计算抽取的是第一批还是第二批产品的概率, 属于贝叶斯概型; 第二步: 事件“从第一批产品中抽取”和“从第二批产品中抽取”构成一个完备事件组, 计算第二次抽到次品的概率, 属于全概率概型. 两个概型的复合, 常常是该类题型的特点.

解 记 $H_i = \{\text{第一次从第 } i \text{ 批产品中抽取}\}, i=1, 2, P(H_1) = P(H_2) = \frac{1}{2}$. (等可能性)

再记 $A = \{\text{取到正品}\}$, 则 $P(A|H_1) = 1, P(A|H_2) = \frac{3}{4}$, 即

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) = \frac{7}{8}.$$

又记 $C_i = \{\text{第二次从第 } i \text{ 批产品中抽取}\}, i=1, 2$, 则有

→ A 已经发生了

$$\star P(C_1) = P(H_1|A) = \frac{P(H_1A)}{P(A)} = \frac{P(H_1)P(A|H_1)}{P(A)} = \frac{\frac{1}{2}}{\frac{7}{8}} = \frac{4}{7},$$

执果索因 ←

$$\star P(C_2) = P(H_2|A) = \frac{P(H_2A)}{P(A)} = \frac{P(H_2)P(A|H_2)}{P(A)} = \frac{3}{7}.$$

则 $P(\bar{A}) = P(C_1)P(\bar{A}|C_1) + P(C_2)P(\bar{A}|C_2) = \frac{4}{7} \times 0 + \frac{3}{7} \times \frac{1}{4} = \frac{3}{28}.$



四 事件的独立性和独立重复试验

1 事件的独立性

$$\left. \begin{array}{l} P(AB) > P(A)P(B) \\ P(AB) < P(A)P(B) \end{array} \right\} \rightarrow \text{不独立}$$

设 A, B 为两个事件, 如果 $P(AB) = P(A)P(B)$, 则称事件 A 与 B 相互独立, 简称 A 与 B 独立.

(1) 表示两个事件相互独立的其他形式.

① 若 $P(A) > 0$, $P(B|A) = \frac{P(AB)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B).$

A 是否发生, 不影响 B 发生的概率

② $P(A) = 0.$

若 $P(A) = 0$, 且 $AB \subset A$, 则 $0 \leq P(AB) \leq P(A) = 0$, 故 $P(AB) = 0$, 即 $P(AB) = P(A)P(B)$ 依然成立.

(2) 两个事件不相互独立时.

① 若 $P(AB) > P(A)P(B)$, 则 $\frac{P(AB)}{P(A)} > \frac{P(A)P(B)}{P(A)}$

(A 有利于 B 的发生) 相互有利

② 若 $P(AB) < P(A)P(B)$, 则 $P(B|A) < P(B)$. 相互抑止

注 设 $A_1, A_2, \dots, A_n$ 为 $n(n \geq 2)$ 个事件, 如果对其中任意有限个事件 $A_{i_1}, A_{i_2}, \dots, A_{i_k} (2 \leq k \leq n)$, 有

$$P(A_{i_1} A_{i_2} \cdots A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \cdots P(A_{i_k}),$$

则称 n 个事件 $A_1, A_2, \dots, A_n$ 相互独立.

考研中常考的是 $n=3$ 时的情形. 细致说来, 设 A_1, A_2, A_3 为三个事件, 若

$$P(A_1 A_2) = P(A_1) P(A_2), \quad (1-1)$$

$$P(A_1 A_3) = P(A_1) P(A_3), \quad (1-2)$$

$$P(A_2 A_3) = P(A_2)P(A_3), \quad (1-3)$$

$$P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3), \quad (1-4)$$

则称事件 A_1, A_2, A_3 相互独立. 当去掉上述 (1-4) 式后, 称只满足 (1-1), (1-2), (1-3) 式的事件 A_1, A_2, A_3 两两独立. 见例 1.13.

2 独立性的判定

A, B, C 相互独立,
 $\bar{A}, \bar{B}, \bar{C}$ 也相互独立

① A 与 B 相互独立 $\xLeftrightarrow{(*)} A$ 与 $\bar{B}$ 相互独立 $\Leftrightarrow \bar{A}$ 与 B 相互独立 $\Leftrightarrow \bar{A}$ 与 $\bar{B}$ 相互独立. 夫唱妇随

注 (1) 仅证 (*), 由 A, B 独立, 有 $P(AB) = P(A)P(B)$, 于是

$$\begin{aligned} P(A\bar{B}) &= P(A) - P(AB) = P(A) - P(A)P(B) \\ &= P(A)[1 - P(B)] = P(A)P(\bar{B}), \end{aligned}$$

故 $A, \bar{B}$ 独立, 其余证明同理.

(2) 将相互独立的事件组中的任意几个事件换成各自的对立事件, 所得的新事件组仍相互独立.

② 对独立事件组不含相同事件作运算, 得到的新事件组仍独立, 如 A, B, C, D 相互独立, 则 AB 与 CD 相互独立, A 与 $BC-D$ 相互独立.

注 直接使用, 无须证明.

③ 若 $P(A) = 0$ 或 $P(A) = 1$, 则 A 与任意事件 B 相互独立. $\longrightarrow$ 不可能事件或必然事件和任意事件独立

注 证 若 $P(A) = 0$, 由 $P(AB) \leq P(A) \leq P(A+B)$, 有 $0 \leq P(AB) \leq P(A) = 0$, 则 $P(AB) = 0$, 于是 $P(A)P(B) = P(AB)$;

若 $P(A) = 1$, 则 $P(\bar{A}) = 1 - P(A) = 0$, 又由 $P(B\bar{A}) \leq P(\bar{A}) \leq P(\bar{A}+B)$, 有 $0 \leq P(B\bar{A}) \leq P(\bar{A}) = 0$, 知 $P(B\bar{A}) = 0$, 又 $P(B\bar{A}) = P(B) - P(AB)$, 故 $P(B) = P(AB)$, 即 $P(A)P(B) = P(AB)$.

3 独立试验序列概型与 n 重伯努利概型

记为 $\text{Ber} - E_n \longleftarrow$ 数字

在同样条件下独立重复地进行一系列完全相同的试验, 即每次试验的可能结果及其发生的概率都不变, 每次试验是相互独立的, 称这种重复试验序列的概率模型为独立试验序列概型. 如果每次试验只有两个结果 A 与 $\bar{A}$, 且在每次试验中 A 发生的概率都相等 (即 $P(A) = p$), 将这种试验独立重复 n 次, 则称这种试验的概率模型为 n 重伯努利概型.

投篮, 投 n 次, 投中 k 次.

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

- ① 每次试验只有 A 与 $\bar{A}$ 两个结果;
- ② 每次试验 A 发生的概率 $p = P(A)$ 不变;
- ③ 试验独立重复进行 n 次.

严格按照定义处理.

例 1.13 将一枚硬币独立地掷两次, 引进事件: $A_1 = \{\text{掷第一次出现正面}\}$, $A_2 = \{\text{掷第二次出现正面}\}$, $A_3 = \{\text{正反面各出现一次}\}$, $A_4 = \{\text{正面出现两次}\}$, 则事件 ().

 (A) A_1, A_2, A_3 相互独立

 (B) A_2, A_3, A_4 相互独立

 $\Downarrow$
 $\Downarrow$

 (C) A_1, A_2, A_3 两两独立

 (D) A_2, A_3, A_4 两两独立

解 应选 (C).

由题可知,

$$\Omega = \{\text{正正, 正反, 反正, 反反}\},$$

则

$$P(A_1) = \frac{1}{2}, P(A_2) = \frac{1}{2}, P(A_3) = \frac{1}{2}, P(A_4) = \frac{1}{4},$$

$$P(A_1 A_2) = \frac{1}{4}, P(A_1 A_3) = \frac{1}{4}, P(A_2 A_3) = \frac{1}{4},$$

但 $P(A_1 A_2 A_3) = 0 \neq P(A_1)P(A_2)P(A_3)$, 故 A_1, A_2, A_3 两两独立, 但三者不相互独立.

又 $A_3 A_4 = \emptyset$, $P(A_4) = \frac{1}{4}$, $P(A_3) = \frac{1}{2}$, $P(A_3 A_4) = 0 \neq P(A_3)P(A_4)$, 知 A_2, A_3, A_4 不两两独立, 也不相互

独立.

例 1.14 对同一目标接连进行 3 次独立重复射击, 假设至少命中目标一次的概率为 $\frac{7}{8}$, 则每次

射击命中目标的概率 $p = \underline{\hspace{2cm}}$.

解 应填 $\frac{1}{2}$.

记 $A_i = \{\text{第 } i \text{ 次命中目标}\}$, $i = 1, 2, 3$, 且 $P(A_1 \cup A_2 \cup A_3) = \frac{7}{8}$, 故

$$1 - P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = 1 - P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3)$$

$$= 1 - (1-p)^3 = \frac{7}{8},$$

则 $p = \frac{1}{2}$.

例 1.15 某人向同一目标独立重复射击, 每次命中目标的概率为 $p (0 < p < 1)$, 则此人第 4 次

射击恰好第 2 次命中目标的概率为 ().

 (A) $3p(1-p)^2$

 (B) $6p(1-p)^2$

 (C) $3p^2(1-p)^2$

 (D) $6p^2(1-p)^2$

分析

√	×	×	√
×	√	×	√
×	×	√	√

解 应选(C).

依题设, 4次射击, 最后一次命中, 前3次有一次命中. 若设前3次射击有 X 次命中, 则 $X \sim B(3, p)$, 且有

$$P\{X=1\} = C_3^1 p(1-p)^2.$$

于是第4次射击恰好第2次命中目标的概率为 $3p^2(1-p)^2$, 故选择(C).



基础习题精练

习题

1.1 设 A, B, C 是任意三个事件, 则下列选项中正确的是().

(A) 若 $A \cup C = B \cup C$, 则 $A = B$

(B) 若 $A - C = B - C$, 则 $A = B$

(C) 若 $AC = BC$, 则 $A = B$

(D) 若 $AB = \emptyset$ 且 $\overline{AB} = \emptyset$, 则 $\overline{A} = B$

1.2 设 A 和 B 是任意两个事件.

①若 $P(A) = P(B)$, 则 $A = B$;

②若 $P(AB) = 0$, 则 $AB = \emptyset$.

则关于两个命题的结论, 正确的是().

(A) ①正确, ②不正确

(B) ①不正确, ②正确

(C) ①, ②均正确

(D) ①, ②均不正确

1.3 考虑一元二次方程 $x^2 + Bx + C = 0$, 其中 B, C 分别是将一枚骰子接连抛两次先后出现的点数, 则该方程有实根的概率为().

(A) $\frac{19}{36}$

(B) $\frac{17}{36}$

(C) $\frac{15}{36}$

(D) $\frac{13}{36}$

1.4 甲口袋有5个白球、3个黑球, 乙口袋有4个白球、6个黑球. 从两个口袋中各任取一个球, 则取到的两个球颜色相同的概率为().

(A) $\frac{19}{40}$

(B) $\frac{21}{40}$

(C) $\frac{1}{4}$

(D) $\frac{9}{40}$

1.5 设 A, B, C 为三个随机事件, 且 A 与 C 相互独立, B 与 C 相互独立, 则 $A \cup B$ 与 C 相互独立的充分必要条件是().

(A) A 与 B 相互独立

(B) A 与 B 互不相容

(C) AB 与 C 相互独立

(D) AB 与 C 互不相容

1.6 随机地向半圆 $0 < y < \sqrt{2ax - x^2} (a > 0)$ 内投掷一点, 点均匀落在半圆内任何一个区域, 则该点和原点连线同 x 轴的夹角 $\theta \leq \frac{\pi}{4}$ 的概率为_____.

1.7 从 $[0, 1]$ 中随机地取两个数, 其积大于 $\frac{1}{4}$, 其和小于 $\frac{5}{4}$ 的概率为_____.

1.8 已知事件 A, B 满足 $P(AB) = P(\bar{A} \cap \bar{B})$, 记 $P(A) = p$, 则 $P(B) =$ _____.

1.9 已知 $P(A) = 0.7, P(A - B) = 0.3$, 则 $P(\bar{A}B) =$ _____.

1.10 已知 $P(\bar{A}) = 0.3, P(B) = 0.4, P(\bar{A}B) = 0.5$, 则 $P(B | A \cup \bar{B}) =$ _____.

1.11 设 A, B 为两个随机事件, 且 $P(A) = 0.4, P(A \cup B) = 0.7$. 若 A, B 互不相容, 则 $P(B) =$ _____; 若 A, B 相互独立, 则 $P(B) =$ _____; 若 A 发生 B 必发生, 则 $P(B) =$ _____.

1.12 三人独立地破译一个密码, 他们能单独译出的概率分别为 $\frac{1}{5}, \frac{1}{3}, \frac{1}{4}$, 则此密码被译出的概率为_____.

1.13 甲、乙两人独立地对同一目标射击一次, 其命中率分别为 0.6 和 0.7. 现已知目标被击中, 则它是甲射中的概率为_____.

1.14 用事件 A, B, C 的运算关系表示事件: ① A, B, C 都不发生; ② A, B, C 不都发生; ③ A, B, C 不多于一个发生.

1.15 设 X, Y 为随机变量, 事件 A 表示 $\{X \geq c\}$, 事件 B 表示 $\{Y \geq c\}$. 用 A, B 表示下列事件:

$D: \{\max\{X, Y\} \geq c\}$. $E: \{\min\{X, Y\} \geq c\}$. $F: \{\max\{X, Y\} < c\}$. $G: \{\min\{X, Y\} < c\}$.

1.16 判断下列命题是否成立, 并说明理由.

(1) $A - (B - C) = (A - B) \cup C$;

(2) 若 $AB = \emptyset$ 且 $C \subset A$, 则 $BC = \emptyset$;

(3) $(A \cup B) - B = A$;

(4) $(A - B) \cup B = A$.

1.17 口袋中有 1 个球, 不知它的颜色是黑的还是白的, 现再往口袋中放入 1 个白球, 然后从口袋中任意取出 1 个, 发现取出的是白球, 求口袋中原来那个球是白球的概率.

1.18 已知装有同种零件的产品两箱, 第一箱内装 50 件产品, 其中一等品 10 件; 第二箱内装 30 件产品, 其中一等品 18 件, 现从两箱中任意挑选一箱. 求: 从中先后取出两件产品 (取后不放回), 先取出的产品是一等品的概率 p ; 已知先取出的产品是一等品, 那么第二次取出的产品仍然是一等品的概率 q .

1.19 每箱产品有 10 件, 其次品数从 0 到 2 是等可能的, 开箱检验时, 从中任取 1 件, 如检验出是次品, 则认为该箱产品不合格而拒收. 假设由于检验有误, 将 1 件正品误认为次品的概率为 2%, 1 件次品被漏查而判为正品的概率为 5%, 求该箱产品通过验收的概率.

解答

1.1 (D) 解 方法一(直接法) 对于选项(D), 由事件运算的对偶律, 有 $\overline{\overline{AB}} = A \cup B = \overline{\emptyset} = \Omega$. 而由 $A \cup B = \Omega$ 且 $AB = \emptyset$, 可见 A 和 B 互为对立事件, 即 $\overline{A} = B$, 因此(D)正确.

方法二(排除法) 前三个选项都不成立, 只需分别举出反例. 例如, 由于 A, B, C 是任意三个事件, 若取 $A \neq B$, 而 $C = \Omega$ 是必然事件, 则 $A \cup C = B \cup C$ 且 $A - C = B - C$, 但 $A \neq B$, 从而(A)和(B)不成立. 取 $A \neq B, C = \emptyset$, 则 $AC = BC$, 但 $A \neq B$, 因此(C)不成立. 从而选(D).

注 本题的结果反映了事件的运算与数的运算的不同之处.

1.2 (D) 解 考虑向区间 $[0, 2]$ 上掷一随机点的试验. 令 $A = \{\text{随机点落入区间 } [0, 1] \text{ 上}\}$, $B = \{\text{随机点落入区间 } [1, 2] \text{ 上}\}$. 显然 $P(A) = P(B)$, $P(AB) = 0$, 然而 $A \neq B$, $AB \neq \emptyset$. 故两个命题均不正确, 选(D).

1.3 (A) 解 本题属于古典概型. 由于很难套用现有概型模式或公式求解, 一个最简单也是最直接的做法是数出总样本数和事件所含样本点数并计算比值, 给出答案. 事件所含样本点数可借助列表等手段列出各种可能的结果, 然后数出事件所含样本点数. 由于接连抛两次, 每次都可能出现 6 个数, 因此其总样本数为 6^2 .

方程有实根, 即事件 $\{B^2 - 4C \geq 0\}$, 列表如下.

$B^2 - 4C$		C					
B		1	2	3	4	5	6
1		-	-	-	-	-	-
2		0	-	-	-	-	-
3		+	+	-	-	-	-
4		+	+	+	0	-	-
5		+	+	+	+	+	+
6		+	+	+	+	+	+

其中事件 $\{B^2 - 4C \geq 0\}$ 所含样本点数为 19, 则 $p = \frac{19}{36}$, 故本题应选择(A).

1.4 (A) 解 从两个口袋中各取一个球, 共有 $C_8^1 C_{10}^1$ 种等可能取法, 而两个球颜色相同有两种情况: 第一种是从甲口袋取出白球, 从乙口袋也取出白球; 第二种是从甲口袋取出黑球, 从乙口袋也

取出黑球, 共有 $C_5^1 C_4^1 + C_3^1 C_6^1$ 种取法. 于是 $P\{\text{取到的两个球颜色相同}\} = \frac{5 \times 4 + 3 \times 6}{8 \times 10} = \frac{19}{40}$.

1.5 (C) 解 $A \cup B$ 与 C 相互独立, 即 $P[(A \cup B) \cap C] = P(A \cup B)P(C)$.

$$\begin{aligned} \text{由} \quad P[(A \cup B) \cap C] &= P(AC \cup BC) \\ &= P(AC) + P(BC) - P(AC \cap BC) \\ &= P(A)P(C) + P(B)P(C) - P(ABC), \end{aligned}$$

$$\begin{aligned} \text{而} \quad P(A \cup B)P(C) &= [P(A) + P(B) - P(AB)]P(C) \\ &= P(A)P(C) + P(B)P(C) - P(AB)P(C), \end{aligned}$$

所以 $A \cup B$ 与 C 相互独立的充分必要条件为 $P(ABC) = P(AB)P(C)$, 即 AB 与 C 相互独立. 应选 (C).

1.6 $\frac{1}{2} + \frac{1}{\pi}$ 解 由题设, 半圆与直线 $y = x$ 的交点坐标为 (a, a) , 如

图 1-3 所示, 阴影区域 D 内任意点与原点连线同 x 轴的夹角 $\theta \leq \frac{\pi}{4}$, 因此所求概率为

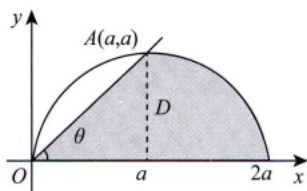


图 1-3

$$P\{(x, y) \in D\} = \left(\frac{\pi a^2}{4} + \frac{a^2}{2} \right) / \frac{\pi a^2}{2} = \frac{1}{2} + \frac{1}{\pi}.$$

1.7 $\frac{15}{32} - \frac{1}{2} \ln 2$ 解 从 $[0, 1]$ 中随机地取两个数 x 与 y , 可以看成从图 1-4

所示正方形区域中任取一点, 满足几何概型的两个特点. 记 $\Omega = \{(x, y) | 0 \leq x \leq 1,$

$0 \leq y \leq 1\}$, $A = \left\{ (x, y) \mid x + y < \frac{5}{4}, xy > \frac{1}{4} \right\}$, 对应图 1-4 中阴影部分.

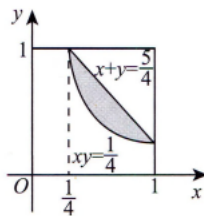


图 1-4

$$\begin{aligned} P(A) &= \frac{S_A}{S_\Omega} = \frac{\int_{1/4}^1 \left(\frac{5}{4} - x - \frac{1}{4x} \right) dx}{1} = \left(\frac{5}{4}x - \frac{1}{2}x^2 - \frac{1}{4} \ln x \right) \Big|_{1/4}^1 \\ &= \frac{15}{32} - \frac{1}{2} \ln 2. \end{aligned}$$

1.8 $1-p$ 解 因为

$$P(AB) = P(\overline{A \cap B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - P(A) - P(B) + P(AB),$$

由此得

$$1 - P(A) - P(B) = 0,$$

所以

$$P(B) = 1 - P(A) = 1 - p.$$

1.9 0.6 解 因为 $0.3 = P(A - B) = P(A) - P(AB) = 0.7 - P(AB)$, 由此得 $P(AB) = 0.7 - 0.3 = 0.4$,

所以

$$P(\overline{AB}) = 1 - P(AB) = 0.6.$$

1.10 0.25 解 由条件概率的定义知 $P(B | A \cup \overline{B}) = \frac{P(AB)}{P(A \cup \overline{B})}$, 其中

$$P(A \cup \bar{B}) = P(A) + P(\bar{B}) - P(\bar{A}\bar{B}) = 0.7 + 0.6 - 0.5 = 0.8.$$

又由 $P(\bar{A}\bar{B}) = P(A) - P(AB)$, 可得

$$P(AB) = P(A) - P(\bar{A}\bar{B}) = 0.7 - 0.5 = 0.2.$$

代回原式, 可得

$$P(B|A \cup \bar{B}) = \frac{P(AB)}{P(A \cup \bar{B})} = \frac{0.2}{0.8} = 0.25.$$

1.11 0.3; 0.5; 0.7 解 由加法公式, 有 $P(A \cup B) = P(A) + P(B) - P(AB)$, 于是:

若 A, B 互不相容, 则 $P(AB) = 0$, $P(B) = P(A \cup B) - P(A) = 0.3$;

若 A, B 相互独立, 则 $P(AB) = P(A)P(B)$, $P(A \cup B) = P(A) + P(B) - P(A)P(B)$, 解得

$$P(B) = \frac{P(A \cup B) - P(A)}{1 - P(A)} = 0.5;$$

若 A 发生 B 必发生, 则 $P(B) = P(A \cup B) = 0.7$.

1.12 $\frac{3}{5}$ 解 记事件 A_i 为“第 i 个人译出密码”, $i=1, 2, 3$, 事件 B 为“密码被译出”. 则

$$\begin{aligned} P(B) &= P(A_1 \cup A_2 \cup A_3) = 1 - P(\bar{A}_1 \bar{A}_2 \bar{A}_3) \\ &= 1 - \frac{4}{5} \times \frac{2}{3} \times \frac{3}{4} = \frac{3}{5}. \end{aligned}$$

注 互不相容可简化并事件的概率计算, 相互独立可简化交事件的概率计算. 这里为了利用相互独立性, 把事件的并在对偶律下转化为事件的交, 这一方法会经常用到.

1.13 0.682 解 记事件 A 为“目标被击中”, 事件 B_1 为“甲射中目标”, 事件 B_2 为“乙射中目标”, 因为 $A = B_1 \cup B_2$, 所以

$$\begin{aligned} P(A) &= P(B_1 \cup B_2) = P(B_1) + P(B_2) - P(B_1 B_2) \\ &= 0.6 + 0.7 - 0.6 \times 0.7 = 0.88. \end{aligned}$$

考虑到 $B_1 \subset A$, 故有

$$P(B_1|A) = \frac{P(AB_1)}{P(A)} = \frac{P(B_1)}{P(A)} = \frac{0.6}{0.88} \approx 0.682.$$

1.14 解 ① $\{A, B, C \text{ 都不发生}\} = \{A \text{ 不发生, 且 } B \text{ 不发生, 且 } C \text{ 不发生}\}$

$$= \bar{A}\bar{B}\bar{C} = \overline{A \cup B \cup C}.$$

② $\{A, B, C \text{ 不都发生}\} = \{A, B, C \text{ 至少有一个不发生}\}$

$= \{A, B, C \text{ 都发生}\}$ 的逆事件

$$= \bar{A} \cup \bar{B} \cup \bar{C} = \overline{ABC}.$$

③ $\{A, B, C \text{ 不多于一个发生}\} = \{A, B, C \text{ 至多有一个发生}\}$

$$= \{A, B, C \text{ 至少有两个不发生}\}$$

$$= \bar{A}\bar{B} \cup \bar{A}\bar{C} \cup \bar{B}\bar{C} = \bar{A}\bar{B}\bar{C} \cup \bar{A}B\bar{C} \cup \bar{A}\bar{B}C \cup \bar{A}\bar{B}\bar{C}.$$

1.15 解 $D = A \cup B$ (即 X, Y 中至少有一个 $\geq c$).

$$E = A \cap B \text{ (即 } X, Y \text{ 都要 } \geq c \text{)}.$$

$$F = \bar{D} = \bar{A} \cap \bar{B} \text{ (即 } X, Y \text{ 都要 } < c \text{)}.$$

$$G = \bar{E} = \bar{A} \cup \bar{B} \text{ (即 } X, Y \text{ 中至少有一个 } < c \text{)}.$$

1.16 解 (1) 不成立. $A - (B - C) = A - B\bar{C} = \overline{AB\bar{C}} = \overline{A(B \cup C)} = \bar{A}\bar{B} \cup \bar{A}\bar{C} = (A - B) \cup \bar{A}\bar{C} \neq (A - B) \cup C$;

也可由图 1-5 快速判断.

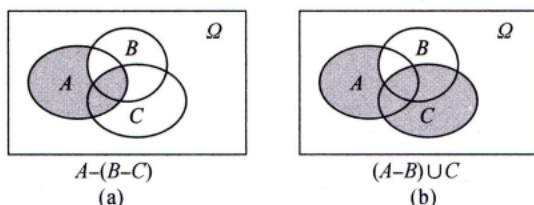


图 1-5

(2) 成立. 因 $C \subset A$, 有 $BC \subset AB = \emptyset$, 故 $BC = \emptyset$.

(3) 不成立. 因 $(A \cup B) - B = (A \cup B)\bar{B} = \bar{A}\bar{B} \cup B\bar{B} = \bar{A}\bar{B} = A - B \neq A$.

(4) 不成立. 因 $(A - B) \cup B = \bar{A}\bar{B} \cup B = (A \cup B)(\bar{B} \cup B) = A \cup B \neq A$.

注 考生可再做练习, 证明下列事件的运算公式:

$$(1) A = AB \cup A\bar{B};$$

$$(2) A \cup B = A \cup \bar{A}B.$$

证 (1) $AB \cup A\bar{B} = A(B \cup \bar{B}) = A\Omega = A$.

$$(2) A \cup \bar{A}B = (A \cup \bar{A})(A \cup B) = \Omega(A \cup B) = A \cup B.$$

1.17 解 记事件 A 为“取出的是白球”, 事件 B 为“原来那个球是白球”. 容易看出

$$P(A|B) = 1, P(A|\bar{B}) = \frac{1}{2}.$$

另外, 由于不知道袋中原来那个球的颜色, 故 $P(B) = P(\bar{B}) = \frac{1}{2}$, 由贝叶斯公式得

$$\begin{aligned}
 P(B|A) &= \frac{P(B)P(A|B)}{P(B)P(A|B) + P(\bar{B})P(A|\bar{B})} \\
 &= \frac{\frac{1}{2} \times 1}{\frac{1}{2} \times 1 + \frac{1}{2} \times \frac{1}{2}} = \frac{2}{3}.
 \end{aligned}$$

1.18 解 记 $A = \{\text{第一次取出的产品是一等品}\}$, $B = \{\text{第二次取出的产品是一等品}\}$, 则

$$p = P(A), \quad q = P(B|A) = \frac{P(AB)}{P(A)}.$$

要计算 p, q , 仅需求出 $P(A), P(AB)$. 而事件 A, AB 都是试验的结果, 它与其前提条件产品取自哪一箱有关. 因此我们自然想到: 写出前提条件, 应用全概率公式计算 $P(A), P(AB)$.

记 $C_i = \{\text{产品取自第 } i \text{ 箱}\} (i=1, 2)$, 则

$$C_1 C_2 = \emptyset, \quad C_1 \cup C_2 = \Omega, \quad A = AC_1 + AC_2, \quad AB = ABC_1 + ABC_2,$$

故

$$\begin{aligned}
 P(A) &= P(AC_1) + P(AC_2) \\
 &= P(C_1)P(A|C_1) + P(C_2)P(A|C_2) \\
 &= \frac{1}{2} \times \frac{10}{50} + \frac{1}{2} \times \frac{18}{30} = \frac{2}{5}, \\
 P(AB) &= P(ABC_1) + P(ABC_2) \\
 &= P(C_1)P(AB|C_1) + P(C_2)P(AB|C_2) \\
 &= \frac{1}{2} \times \frac{10}{50} \times \frac{9}{49} + \frac{1}{2} \times \frac{18}{30} \times \frac{17}{29} = \frac{276}{1421}, \\
 p = P(A) &= \frac{2}{5}, \quad q = \frac{P(AB)}{P(A)} = \frac{276}{1421} \times \frac{5}{2} = \frac{690}{1421}.
 \end{aligned}$$

1.19 解 题中有两个完备事件组: 一是箱中次品的件数为 0, 1, 2; 二是通过验收是在抽取正品和抽取次品这两种情况下发生. 设 $A_i (i=0, 1, 2)$ 为“箱中有 i 个次品”, B 为“通过验收”, B_1 为“抽取正品”, 于是有

$$\begin{aligned}
 P(A_i) &= \frac{1}{3} (i=0, 1, 2), \\
 P(B_1|A_i) &= \frac{10-i}{10},
 \end{aligned}$$

从而由全概率公式得

$$\begin{aligned}
 P(B_1) &= \sum_{i=0}^2 P(A_i)P(B_1|A_i) \\
 &= \frac{1}{3} \sum_{i=0}^2 \frac{10-i}{10} = 0.9,
 \end{aligned}$$

$$P(\bar{B}_1) = 0.1,$$

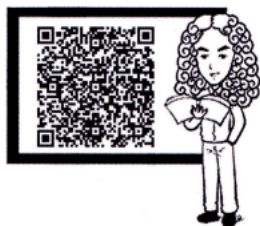
因此由全概率公式得

$$\begin{aligned} P(B) &= P(B_1)P(B|B_1) + P(\bar{B}_1)P(B|\bar{B}_1) \\ &= 0.9 \times 0.98 + 0.1 \times 0.05 = 0.887. \end{aligned}$$

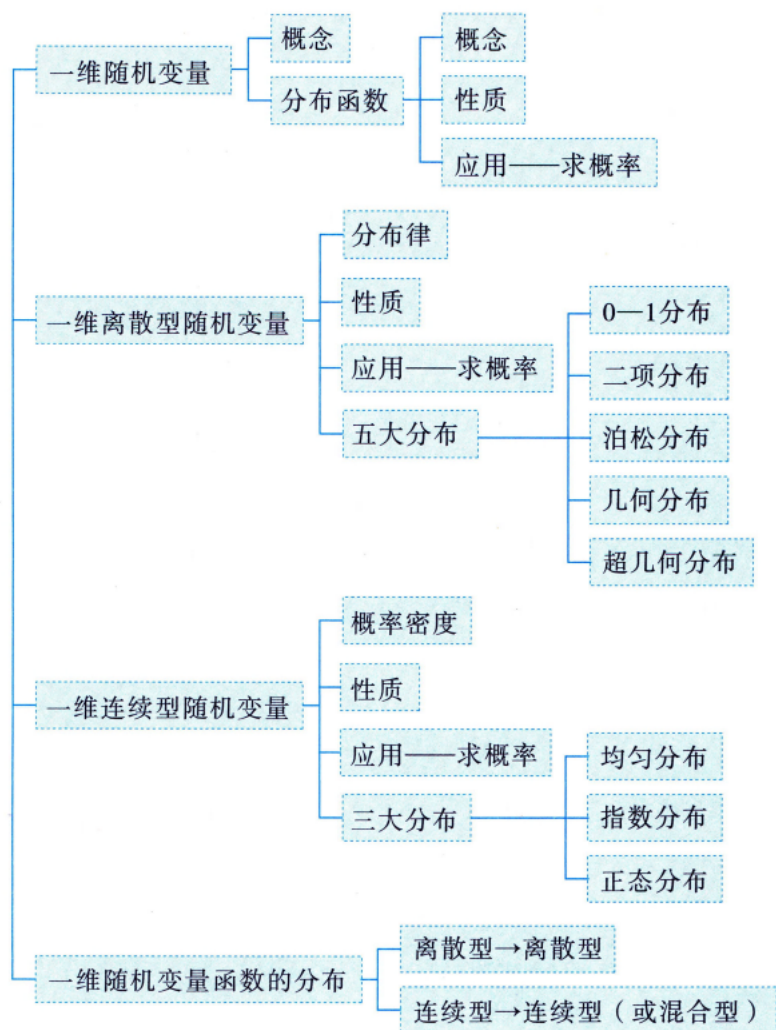
注 本题中，通过验收是在完备事件组 $B_1, \bar{B}_1$ 背景下发生的，抽取正品则是在完备事件组 A_0, A_1, A_2 背景下实现的，因此，具备全概率模型的特征。

第2讲

一维随机变量及其分布



基础知识结构



基础内容精讲

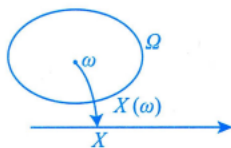
一 随机变量及其分布函数的概念、性质及应用



1 随机变量的概念

对 $y = y(x)$ 而言, x 是确定性变量, y 是确定性函数值.

对 $X = X(\omega)$ 而言, X 是不确定性变量. 具有可能性, 例如投篮是否投中



随机变量就是“其值会随机而定”的变量. 设随机试验 E 的样本空间为 $\Omega = \{\omega\}$, 如果对每一个 $\omega \in \Omega$, 都有唯一的实数 $X(\omega)$ 与之对应, 并且对任意实数 x , $\{\omega | X(\omega) \leq x, \omega \in \Omega\}$ 是随机事件, 则称定义在 Ω 上的实值单值函数 $X(\omega)$ 为随机变量, 简记为随机变量 X . 一般用大写字母 $X, Y, Z, \dots$ 或希腊字母 $\xi, \eta, \zeta, \dots$ 来表示随机变量.

注 (1) 随机事件是从静态的观点来研究随机现象, 而随机变量则是一种动态的观点, 如高等数学中常量与变量的区别.

(2) 随机变量的实质是“实值单值函数”, 这个定义不同于高等数学中函数的定义(其“定义域”一般是实数集), 它的“定义域”不一定是实数集, 这点需要考生注意.

2 分布函数的概念及性质

(1) 概念.

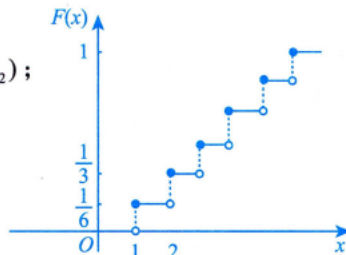
设 X 是随机变量, x 是任意实数, 称函数 $F(x) = P\{X \leq x\} (x \in \mathbf{R})$ 为随机变量 X 的分布函数, 或称 X 服从 $F(x)$ 分布, 记为 $X \sim F(x)$.

当 x 从 $-\infty$ 到 $+\infty$ 取遍所有实数时, $\{X \leq x\}$ 就相应从不可能事件 $\emptyset$ 取到必然事件 Ω , 即分布函数完整地描述了随机变量的概率规律

如掷骰子, $\Omega: \{\omega_1, \omega_2, \dots, \omega_6\}$, $F(x) = P\{X \leq x\}$ 分布函数 $F(x)$ 图像如图所示

(2) 性质. 充要条件

① $F(x)$ 是 x 的单调不减函数, 即对任意实数 $x_1 < x_2$, 有 $F(x_1) \leq F(x_2)$;



② $F(x)$ 是 x 的右连续函数, 即对任意 $x_0 \in \mathbf{R}$, 有 $\lim_{x \rightarrow x_0^+} F(x) = F(x_0 + 0) = F(x_0)$;

③ $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$, $F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$. $\longrightarrow$ 分布函数是事件的概率, 由此知 $0 \leq F(x) \leq 1$, 即 $F(x)$ 是有界函数

注 满足以上三条性质的函数 $F(x)$ 必是某个随机变量 X 的分布函数, 所以, 这三条性质也是判断函数 $F(x)$ 是否为某一随机变量 X 的分布函数的充要条件.

3 分布函数的应用——求概率

$$P\{X \leq a\} = F(a);$$

$$P\{X < a\} = F(a-0);$$

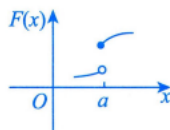
$$P\{X = a\} = F(a) - F(a-0);$$

$$P\{a < X < b\} = F(b-0) - F(a);$$

$$P\{a \leq X < b\} = F(b-0) - F(a-0);$$

$$P\{a < X \leq b\} = F(b) - F(a);$$

$$P\{a \leq X \leq b\} = F(b) - F(a-0).$$



$$\star\star\star F(x) = P\{X \leq x\}, x \in (-\infty, +\infty),$$

针对离散型: $F(a) = P\{X \leq a\} = P\{X < a\} + P\{X = a\}$
 $\downarrow \qquad \qquad \downarrow$
 $F(a-0) \quad F(a) - F(a-0)$

例 2.1 设随机变量 X 的分布函数为

$$X \sim F(x) = \begin{cases} a + be^{-x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

求常数 a, b 及概率 $P\{|X| < 2\}$.

分析 $F(x)$ 中含有 2 个未知参数, 依据分布函数的性质建立两个独立方程解之.

解 由分布函数的性质, 有

$$F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = \lim_{x \rightarrow +\infty} (a + be^{-x}) = 1,$$

得 $a = 1$. 又 $F(x)$ 在 $x = 0$ 处右连续, 有

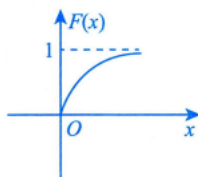
$$\lim_{x \rightarrow 0^+} F(x) = \lim_{x \rightarrow 0^+} (a + be^{-x}) = a + b = 0,$$

得 $b = -1$. 所以

$$F(x) = \begin{cases} 1 - e^{-x}, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

从而有

$$P\{|X| < 2\} = P\{-2 < X < 2\} = F(2-0) - F(-2) = 1 - e^{-2} - 0 = 1 - e^{-2}.$$



满足: ①单调不减;
②右连续;
③极限性质

二 常见的两类随机变量——离散型随机变量和连续型随机变量



1 离散型随机变量及其概率分布

如果随机变量 X 只可能取有限个或可列无限个值 $x_1, x_2, \dots$, 则称 X 为离散型随机变量, 称

$$P\{X=x_i\}=p_i, i=1, 2, \dots$$

为 X 的分布列、分布律或概率分布, 记为 $X \sim p_i$, 概率分布常常用表格形式或矩阵形式表示, 即

X	x_1	x_2	$\dots$
P	p_1	p_2	$\dots$

或 $X \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$.

如掷骰子, X 的概率分布为 $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$.

数列 $\{p_i\}(i=1, 2, \dots)$ 是离散型随机变量的概率分布的充要条件: $p_i \geq 0 (i=1, 2, \dots)$, 且 $\sum_i p_i = 1$.

设离散型随机变量 X 的概率分布为 $P\{X=x_i\}=p_i, i=1, 2, \dots$, 则 X 的分布函数

归一性

某点概率 $F(x) = P\{X \leq x\} = \sum_{x_i \leq x} P\{X=x_i\},$

$$P\{X=x_i\} = P\{X \leq x_i\} - P\{X < x_i\} = F(x_i) - F(x_i - 0),$$

并且对实数轴上的任一集合 B , 有

区间概率 $P\{X \in B\} = \sum_{x_i \in B} P\{X=x_i\},$

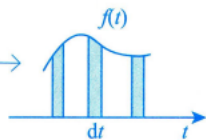
特别地,

$$P\{a < X \leq b\} = P\{X \leq b\} - P\{X \leq a\} = F(b) - F(a).$$

2 连续型随机变量及其概率密度

如果随机变量 X 的分布函数可以表示为

$$F(x) = \int_{-\infty}^x f(t) dt \quad (x \in \mathbf{R}),$$



其中 $f(x)$ 是非负可积函数, 则称 X 为连续型随机变量, 称 $f(x)$ 为 X 的概率密度函数, 简称概率密度, 记为 $X \sim f(x)$.

离散型随机变量 $X \sim p_i, F(x) = P\{X \leq x\} = \sum_{x_i \leq x} p_i,$

对比

连续型随机变量 $X \sim f(x), F(x) = P\{X \leq x\} = \int_{-\infty}^x f(t) dt.$

归一性

$f(x)$ 为某一随机变量 X 的概率密度的充分必要条件: $f(x) \geq 0$, 且 $\int_{-\infty}^{+\infty} f(x) dx = 1$ (由此可知, 在保

证非负的条件下, 改变 $f(x)$ 有限个点的值, $f(x)$ 仍然是概率密度).

一个点的值改变, 不会改变面积

设 X 为连续型随机变量, $X \sim f(x)$, 则对任意实数 c , 有 $P\{X=c\}=0$; 对实数轴上任一集合 B ,

有

$$P\{X \in B\} = \int_B f(x) dx,$$

特别地,

$$P\{a < X < b\} = P\{a \leq X < b\} = P\{a < X \leq b\} = P\{a \leq X \leq b\} = \int_a^b f(x) dx = F(b) - F(a).$$

注 (1) “概率密度”这个名词的由来可解释如下: 取定一个点 x , 则按分布函数的定义, 事件 $\{x < X \leq x+h\}$ ($h > 0$, 为常数) 的概率应为 $F(x+h) - F(x)$. 所以, 比值 $[F(x+h) - F(x)]/h$ 可以解释为在点 x 附近长为 h 的区间 $(x, x+h]$ 内, 单位长所占的概率. 令 $h \rightarrow 0$, 则这个比值的极限, 即 $F'(x) = f(x)$, 也就是在点 x 处 (无穷小区间内) 单位长的概率. 或者说, 它反映了概率在点 x 处的“密集程度”. 设想一条极细的无穷长的金属杆, 总质量为1, 概率密度相当于杆上各点的质量密度.

(2) $P\{a < X < b\} = \int_a^b f(x) dx$ 意味着 X 落入某一区间的概率等于该区间上概率密度曲线之下的曲边梯形的面积. 应用概率的这种几何意义, 常常有助于问题的分析与求解.

(3) 设 $X \sim f(x)$, 则 X 的分布函数 $F(x)$ 是 x 的连续函数, 在 $f(x)$ 的连续点 x_0 处有 $F'(x_0) = f(x_0)$. 如果 $F(x)$ 是连续函数, 除有限个点外, $F'(x)$ 存在且连续, 则 X 为连续型随机变量, 且 $f(x) = F'(x)$ (在 $F'(x)$ 不存在的地方可以令 $f(x) = 0$ 或取其他值).

例 2.2 设 $F_1(x), F_2(x)$ 是随机变量的分布函数, $f_1(x), f_2(x)$ 是相应的概率密度, 则 ().

(A) $F_1(x) + F_2(x)$ 是分布函数

(B) $F_1(x) \cdot F_2(x)$ 是分布函数

(C) $f_1(x) + f_2(x)$ 是概率密度

(D) $f_1(x) \cdot f_2(x)$ 是概率密度

分析

$$F(x) \text{ 是分布函数} \Leftrightarrow \begin{cases} \text{① } F(x) \text{ 单调不减,} \\ \text{② } F(x_0+0) = F(x_0), \\ \text{③ } \begin{cases} F(+\infty) = 1, \\ F(-\infty) = 0. \end{cases} \end{cases}$$

$$f(x) \text{ 是概率密度} \Leftrightarrow \begin{cases} \text{① } f(x) \geq 0, \\ \text{② } \int_{-\infty}^{+\infty} f(x) dx = 1. \end{cases}$$

解 应选 (B).

根据连续型随机变量分布函数和概率密度的性质, 有

$$F_1(+\infty) + F_2(+\infty) = 2, \int_{-\infty}^{+\infty} [f_1(x) + f_2(x)] dx = 2,$$

$$\int_{-\infty}^{+\infty} f_1(x)f_2(x)dx \neq \int_{-\infty}^{+\infty} f_1(x)dx \int_{-\infty}^{+\infty} f_2(x)dx = 1,$$

应排除选项 (A), (C), (D), 故选择 (B). 事实上, 由于 $F_1(x) \cdot F_2(x)$ 是单调不减的右连续函数, 且

$$F_1(-\infty) \cdot F_2(-\infty) = 0, F_1(+\infty) \cdot F_2(+\infty) = 1,$$

可直接判断选项 (B) 正确.

注 判断 $\sqrt{f_1(x)f_2(x)}$ 是不是概率密度. 易知 $\frac{f_1(x)+f_2(x)}{2} \geq \sqrt{f_1(x)f_2(x)}$, $f_1(x) \neq f_2(x)$, 故

$$1 > \int_{-\infty}^{+\infty} \sqrt{f_1(x)f_2(x)} dx,$$

不满足归一性, 所以 $\sqrt{f_1(x)f_2(x)}$ 不是概率密度.

例 2.3 已知随机变量 X 的概率分布为

X	1	2	3
$P\{X=k\}$	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

且 $P\{X \geq 2\} = \frac{3}{4}$, 求未知参数 θ 及 X 的分布函数 $F(x)$.

分析 此题若利用归一性, $\theta^2 + 2\theta(1-\theta) + (1-\theta)^2 = 1$, 发现解不出 θ .

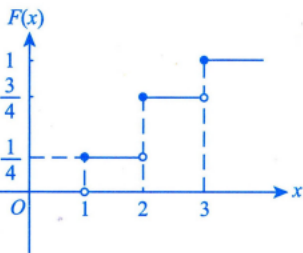
解 由 $P\{X \geq 2\} = P\{X=2\} + P\{X=3\} = 2\theta(1-\theta) + (1-\theta)^2 = 1 - \theta^2 = \frac{3}{4}$, 解得 $\theta = \pm \frac{1}{2}$.

又 $P\{X=2\} = 2\theta(1-\theta) \geq 0$, 故取 $\theta = \frac{1}{2}$, 从而得 X 的概率分布为 利用概率不能为负值

X	1	2	3
$P\{X=k\}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

X 的分布函数

$$F(x) = P\{X \leq x\} = \sum_{i \leq x} P\{X=i\} = \begin{cases} 0, & x < 1, \\ \frac{1}{4}, & 1 \leq x < 2, \\ \frac{3}{4}, & 2 \leq x < 3, \\ 1, & x \geq 3. \end{cases}$$



例 2.4 已知 X 的分布函数为 $F(x)$, 概率密度为 $f(x)$. 当 $x \leq 0$ 时, $f(x)$ 连续, 且 $f(x) = F(x)$,

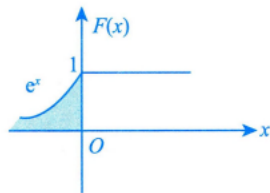
若 $F(0) = 1$, 则 $f(x) = \underline{\hspace{2cm}}$.

解 应填 $\begin{cases} e^x, & x \leq 0, \\ 0, & x > 0. \end{cases}$

由于 $F(x)$ 是单调不减的函数, 且 $F(0) = 1$, 故对任意的 $x > 0$, 均有 $F(x) = 1$. 又当 $x \leq 0$ 时, $f(x)$ 连续, 故 $f(x) = F'(x)$. 已知 $f(x) = F(x)$, 由微分方程 $F'(x) = F(x)$ 及 $F(0) = 1$, 解得 $F(x) = e^x (x \leq 0)$, 从而得

$$F(x) = \begin{cases} e^x, & x < 0, \\ 1, & x \geq 0, \end{cases}$$

$$f(x) = \begin{cases} e^x, & x \leq 0, \\ 0, & x > 0. \end{cases}$$



注 当分布函数是不连续的分段函数时, 为保证其右连续性, 自变量 x 的分段小区间除第一个区间之外, 其余都应是左闭右开的, 即等号跟着大于号. 连续情况下等号跟谁都可以.

例 2.5 已知随机变量 X 的概率密度为

$$f(x) = \begin{cases} 2\left(1 - \frac{1}{x^2}\right), & 1 \leq x \leq 2, \\ 0, & \text{其他.} \end{cases}$$

→ 正概率密度区间

求 X 的分布函数 $F(x)$.

解 当 $x < 1$ 时,

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x 0 dt = 0;$$

当 $1 \leq x < 2$ 时,

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt = \int_1^x 2\left(1 - \frac{1}{t^2}\right) dt = 2\left(t + \frac{1}{t}\right) \Big|_1^x \\ &= 2\left(x + \frac{1}{x}\right) - 4 = 2\left(x + \frac{1}{x} - 2\right); \end{aligned}$$

当 $x \geq 2$ 时,

$$F(x) = \int_{-\infty}^x f(t) dt = \int_1^2 2\left(1 - \frac{1}{t^2}\right) dt = 2\left(t + \frac{1}{t}\right) \Big|_1^2 = 2 \times \frac{1}{2} = 1.$$

故

$$F(x) = \begin{cases} 0, & x < 1, \\ 2\left(x + \frac{1}{x} - 2\right), & 1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

→ 归一性



三 常见的随机变量分布类型

1 离散型

(1) 0—1 分布 $B(1, p)$ (Ber- E_1) . → 只有0与1两个点, 0—1分布是二项分布的一种特殊情况

→ 伯努利计数变量——发生计1, 不发生计0

如果 X 的概率分布为 $X \sim \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$, 即 $P\{X=1\}=p$, $P\{X=0\}=1-p$, 则称 X 服从参数为 p 的

0—1 分布, 记为 $X \sim B(1, p)$ ($0 < p < 1$).

(2) 二项分布 $B(n, p)$ (Ber- E_n) . → n 重伯努利试验中事件 A 发生的次数, 其中 $p=P(A)$

如果 X 的概率分布为 $P\{X=k\}=C_n^k p^k (1-p)^{n-k}$ ($k=0, 1, \dots, n$; $0 < p < 1$), 则称 X 服从参数为 (n, p)

的二项分布, 记为 $X \sim B(n, p)$.

举例: 投篮投中打 $\checkmark$, 投不中打 $\times$, $X=k$ 即表示有 k 个 $\checkmark$, $n-k$ 个 $\times$.

则 $X \sim B(n, p)$, $P=\{X=k\}=C_n^k p^k (1-p)^{n-k}$, $k=0, 1, \dots, n$.

若 $n=1$, 则 $X_i \sim B(1, p)$, $i=1, 2, \dots, n$, $X=X_1+X_2+\dots+X_n$ → 随机变量的分解

稀有事件发生概率

(3) 泊松分布 $P(\lambda)$. →

如果 X 的概率分布为

$$P\{X=k\} = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k=0, 1, \dots; \lambda > 0),$$

λ 表示来流强度, $EX=\lambda$, 期望代表随机变量取值的中心水平, 可以理解为强度

则称 X 服从参数为 λ 的泊松分布, 记为 $X \sim P(\lambda)$. → 定义 k 为非负整数

注 泊松定理 若 $X \sim B(n, p)$, 当 n 很大, p 很小, $\lambda=np$ 适中时, 二项分布可用泊松分布近似表示, 即

$$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}.$$

例如, $C_{100}^1 \left(\frac{1}{500}\right)^1 \left(1-\frac{1}{500}\right)^{99} \approx \frac{\left(\frac{1}{5}\right)^1}{1!} e^{-\frac{1}{5}} = \frac{1}{5} e^{-\frac{1}{5}}$

一般地, 当 $n \geq 20$, $p \leq 0.05$ 时, 用泊松近似公式逼近二项分布效果比较好, 特别当 $n \geq 100$, $np \leq 10$ 时, 逼近效果更佳.

此处《全国硕士研究生招生考试数学考试大纲》的要求是“会用泊松分布近似表示二项分布”, 考生应予以重视.

参数只有 p , 首中即停止(4) 几何分布 $G(p)$ (Ber- E_∞).如果 X 的概率分布为

$$P\{X=k\} = (1-p)^{k-1} p \quad (k=1, 2, \dots; 0 < p < 1),$$

$\rightarrow$ n 重伯努利试验中, 事件 A 首次出现即停止时的试验次数, 其中 $p = P(A)$

$\rightarrow$ 例: 记投篮投中为 $\checkmark$, 投不中为 $\times$, 即 $\times \cdots \times \checkmark$ $\xrightarrow{k-1}$ 第 k 次投中

则称 X 服从参数为 p 的几何分布, 记为 $X \sim G(p)$.

$\rightarrow$ 离散型等待分布: 守株待兔 

(5) 超几何分布 $H(n, N, M)$. $\rightarrow$ 见注

如果 X 的概率分布为 $P\{X=k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} \quad (\max\{0, n-N+M\} \leq k \leq \min\{M, n\}; M, N, n \text{ 为正整数且}$

$M \leq N, n \leq N, k \text{ 为整数})$, 则称 X 服从参数为 (n, N, M) 的超几何分布, 记为 $X \sim H(n, N, M)$.

 $\rightarrow$ 古典概型

注 设由 N 个产品组成的总体, 其中含有 M 个不合格品. 若从中随机不放回地抽取 n 个, 则其中含有不合格品的个数 X 是一个离散型随机变量. 假如 $n \leq M$, 则 X 可能取 $0, 1, \dots, n$; 若 $n > M$, 则 X 可能取 $0, 1, \dots, M$. 由古典概型概率公式容易算得

$$P\{X=k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n},$$

其中 $\max\{0, n-N+M\} \leq k \leq \min\{M, n\}; M, N, n$ 为正整数且 $M \leq N, n \leq N, k$ 为整数.

这个分布就是超几何分布, 它含有三个参数 N, M 和 n , 记为 $H(n, N, M)$.

当 $n \ll N$ (即抽取个数 n 远远小于产品总数 N) 时, 每次抽取后, 总体中的不合格品率 $p = \frac{M}{N}$ 改变甚微, 这时不放回抽样可近似看作放回抽样, 超几何分布可用二项分布近似.

2 连续型

(1) 均匀分布 $U(a, b)$.如果随机变量 X 的概率密度和分布函数 (见图 2-1, 图 2-2) 分别为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & \text{其他}, \end{cases} \quad F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x < b, \\ 1, & x \geq b, \end{cases}$$

$\rightarrow F(x) = \int_{-\infty}^x f(t) dt$ 面积的思想

则称 X 在区间 (a, b) 上服从均匀分布, 记为 $X \sim U(a, b)$.

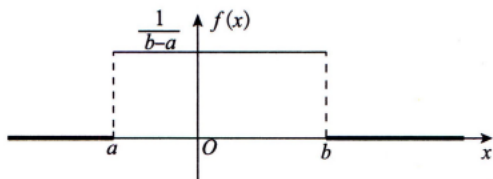


图 2-1

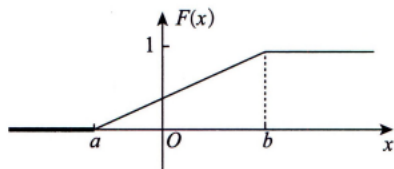


图 2-2

注 区间 (a, b) 可以是闭区间 $[a, b]$, 记为 $X \sim U[a, b]$; 几何概型是均匀分布的实际背景, 几何概型可以用均匀分布函数计算概率. 对于连续型随机变量, 取一点的概率是测不到的

(2) 指数分布 $E(\lambda)$. → 连续型等待分布

如果 X 的概率密度和分布函数 (见图 2-3, 图 2-4) 分别为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & \text{其他}, \end{cases} \quad F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0, \end{cases} \quad (\lambda > 0),$$

λ 叫失故频率, 后面会知道 $EX = \frac{1}{\lambda}$

则称 X 服从参数为 λ 的指数分布, 记为 $X \sim E(\lambda)$.

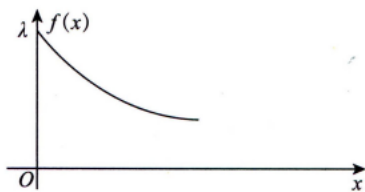


图 2-3

记 $\{N_t, t \geq 0\}$, N_t —— 质点来流的个数, 泊松分布中 λ 表示单位时间质点来流的强度, 所以 t 时间内质点来流的强度为 λt .

令 T 表示等待时间 (第一个质点出现), 则

$$T \sim F(t) = P\{T \leq t\} = 1 - P\{T > t\},$$

由 $\{N_t, t \geq 0\}$ 知 $[0, t]$ 时,

$$P\{T > t\} = P\{N_t = 0\} = e^{-\lambda t},$$

$$\text{故 } T \sim F(t) = P\{T < t\} = 1 - P\{T > t\}$$

$$= \begin{cases} 1 - e^{-\lambda t}, & t > 0, \\ 0, & t < 0 \end{cases}$$

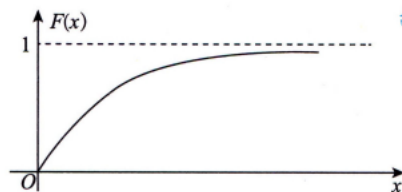


图 2-4

注 当 $t > 0, s > 0$ 时, $P\{X \geq t+s | X \geq t\} = P\{X \geq s\}$ 称为指数分布的无记忆性 (Memoryless Property).

(3) 正态分布 $N(\mu, \sigma^2)$.

如果 X 的概率密度为

这是由于 λ 是常数, 失故频率不变, 也即研究对象 (如元件、产品等) 被视为理想状态, 无任何损耗.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (-\infty < x < +\infty),$$

其中 $-\infty < \mu < +\infty, \sigma > 0$, 则称 X 服从参数为 (μ, σ^2) 的正态分布或称 X 为正态变量, 记为 $X \sim N(\mu, \sigma^2)$. 此时 $f(x)$ 的图形关于直线 $x = \mu$ 对称, 即 $f(\mu-x) = f(\mu+x)$, 并在 $x = \mu$ 处有唯一最大

值 $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$. 如图 2-5 所示.

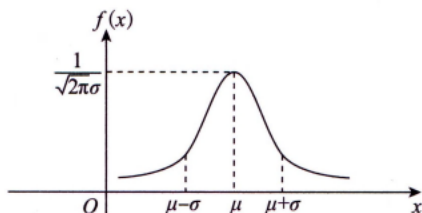


图 2-5

称 $\mu=0, \sigma=1$ 时的正态分布 $N(0, 1)$ 为标准正态分布, 通常记标准正态分布的概率密度为 $\varphi(x) =$

$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, 分布函数为 $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$. 显然 $\varphi(x)$ 为偶函数, 则

$$\Phi(0) = \frac{1}{2}, \quad \Phi(-x) = 1 - \Phi(x).$$

如图 2-6 所示.

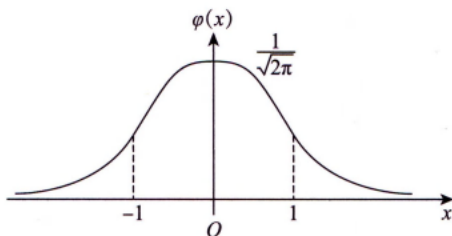


图 2-6

若 $X \sim N(0, 1)$, $P\{X > \mu_\alpha\} = \alpha$, 则称 μ_α 为标准正态分布的上侧 α 分位数 (上 α 分位点).

如果 $X \sim N(\mu, \sigma^2)$, 则

为了查表方便, 所以将正态分布标准化

$$F(x) = P\{X \leq x\} = \Phi\left(\frac{x-\mu}{\sigma}\right),$$

$$F(\mu-x) + F(\mu+x) = 1,$$

$$P\{a < X < b\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right),$$

$$aX + b \sim N(a\mu + b, a^2\sigma^2) (a \neq 0).$$

例 2.6 通过某交叉路口的汽车流可以看作服从泊松分布. 已知在 1 分钟内有汽车通过的概率为 0.7, 则 1 分钟内最多有 1 辆汽车通过的概率为 ().

(A) $0.7(1 - \ln 0.7)$

(B) $0.3(1 - \ln 0.7)$

(C) $0.3(1 - \ln 0.3)$

(D) $0.7(1 - \ln 0.3)$

解 应选 (C).

1分钟内通过路口的汽车数量 X 服从参数为 λ 的泊松分布.由已知,

$$P\{X \geq 1\} = 1 - \frac{\lambda^0}{0!} e^{-\lambda} = 0.7, \quad \text{首先求出参数 } \lambda, \text{ 再根据具体问题分析}$$

得 $\lambda = -\ln 0.3$, 于是 $P\{X \leq 1\} = \frac{\lambda^0}{0!} e^{-\lambda} + \frac{\lambda^1}{1!} e^{-\lambda} = e^{\ln 0.3} (1 - \ln 0.3) = 0.3(1 - \ln 0.3)$, 故选择(C).

前中停止

例 2.7 设随机变量 X 的概率密度为 $f(x) = \begin{cases} 2^{-x} \ln 2, & x > 0, \\ 0, & \text{其他.} \end{cases}$ 对 X 进行独立重复的观测, 直到

第2个大于3的观测值出现时停止, 记 Y 为观测次数. 求 Y 的概率分布.

分析 $2^{-x} \ln 2 = e^{-x \ln 2} \cdot \ln 2$. 先求大于3的概率.

解 记 p 为“观测值大于3”的概率, 则 $p = P\{X > 3\} = \int_3^{+\infty} 2^{-x} \ln 2 dx = \frac{1}{8}$.

依题意, Y 为离散型随机变量, 而且取值为2, 3, ..., 则 Y 的概率分布为

$$P\{Y = k\} = C_{k-1}^1 p(1-p)^{k-2} \cdot p = C_{k-1}^1 p^2 (1-p)^{k-2} = (k-1) \left(\frac{1}{8}\right)^2 \left(\frac{7}{8}\right)^{k-2}, \quad k = 2, 3, \dots$$

注 k 只能从2开始. $\underbrace{\times \cdots \times \sqrt{\cdots} \times \sqrt{\cdots}}_{k-1}$, 前 $k-1$ 次包含一个 $\sqrt{\cdots}$, 第 k 次出现第2个 $\sqrt{\cdots}$.

例 2.8 已知 $X \sim U(a, b) (a > 0)$, 且 $P\{0 < X < 3\} = \frac{1}{4}$, $P\{X > 4\} = \frac{1}{2}$, 求 X 的概率密度及

$P\{1 < X < 5\}$.

解 如图2-7所示, $X \sim U(a, b)$.

由 $P\{X > 4\} = \frac{1}{2}$, 且 $P\left\{X > \frac{a+b}{2}\right\} = \frac{1}{2}$, 得 $\frac{a+b}{2} = 4$, 即 $a+b=8$.

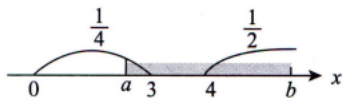


图 2-7

由于 $P\{3 < X < 4\} = 1 - P\{X \geq 4\} - P\{X \leq 3\} = 1 - P\{X > 4\} - P\{0 < X < 3\} = 1 - \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$, 又 $P\{3 < X <$

$4\} = \frac{4-3}{b-a}$, 故 $\frac{4-3}{b-a} = \frac{1}{4}$, 得 $b-a=4$.

联立得 $\begin{cases} a+b=8, \\ b-a=4, \end{cases}$ 解得 $\begin{cases} a=2, \\ b=6. \end{cases}$

a, b 值的另一种解法. 因为只在 (a, b) 内取值且 $P\{0 < X < 3\}$ 不为0, 所以3只能大于 a . 同理, 4只能小于 b , 则有

$$\begin{cases} P\{0 < X < 3\} = P\{a < X < 3\} = \frac{1}{4}, \\ P\{X > 4\} = P\{4 < X < b\} = \frac{1}{2}, \end{cases} \quad \text{即} \quad \begin{cases} \frac{3-a}{b-a} = \frac{1}{4}, \\ \frac{b-4}{b-a} = \frac{1}{2}, \end{cases} \quad \begin{matrix} \text{①} \\ \text{②} \end{matrix}$$

①与②联立, 解得 $a=2, b=6$

所以 X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{4}, & 2 < x < 6, \\ 0, & \text{其他.} \end{cases}$$

于是

$$P\{1 < X < 5\} = P\{2 < X < 5\} = \frac{5-2}{6-2} = \frac{3}{4}.$$

例 2.9 某元件的工作寿命 X (小时) 服从参数为 $\lambda (\lambda > 0)$ 的指数分布.

(1) 求该元件正常工作 t 小时的概率;

正常工作 t 小时, 即工作寿命肯定大于 t 小时.

(2) 已知该元件已正常工作 10 小时, 求在此基础上再工作 10 小时的概率 ($\lambda = 0.01$).

解 由题意 $X \sim F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0, & \text{其他.} \end{cases}$

$$(1) P\{X > t\} = 1 - P\{X \leq t\} = 1 - F(t) = 1 - (1 - e^{-\lambda t}) = e^{-\lambda t}.$$

$$(2) \quad P\{X > 20 | X > 10\} = \frac{P\{X > 20, X > 10\}}{P\{X > 10\}} = \frac{P\{X > 20\}}{P\{X > 10\}} = \frac{e^{-0.01 \times 20}}{e^{-0.01 \times 10}} = e^{-0.1} = P\{X > 10\}$$

注 本题第 (2) 问的结果, $P\{X > 20 | X > 10\} = P\{X > 10\}$, 即元件正常工作 10 小时的概率与在这之前的工作状态无关, 说明指数分布具有无记忆性.

例 2.10 设 $X \sim N(2, \sigma^2)$, 且 $P\{2 < X < 4\} = 0.3$, 则 $P\{X < 0\} = \underline{\hspace{2cm}}$.

解 应填 0.2.

已知 $X \sim N(2, \sigma^2)$, 故

把正态分布进行标准化处理:

$$P\{2 < X < 4\} = P\left\{\frac{2-2}{\sigma} < \frac{X-2}{\sigma} < \frac{4-2}{\sigma}\right\}$$

$$P\{2 < X < 4\} = \Phi\left(\frac{4-2}{\sigma}\right) - \Phi\left(\frac{2-2}{\sigma}\right) = \Phi\left(\frac{2}{\sigma}\right) - 0.5 = 0.3, \quad \Phi\left(\frac{2}{\sigma}\right) = 0.8,$$

$$\text{所以 } P\{X < 0\} = \Phi\left(\frac{0-2}{\sigma}\right) = 1 - \Phi\left(\frac{2}{\sigma}\right) = 0.2.$$

注 结合图 2-8, 本题应用面积求解更为直观、简便.

已知 $P\{2 < X < 4\} = 0.3$, 故

也可利用对称性解题 $\rightarrow$

$$P\{X < 0\} = 0.5 - 0.3 = 0.2.$$

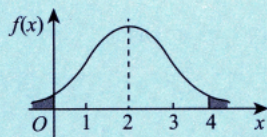


图 2-8

例 2.11 设随机变量 X 服从正态分布 $N(\mu, \sigma^2)$, 则随着 σ 的增大, $P\{|X - \mu| < 1\}$ ().

(A) 单调增加

(B) 单调减少

(C) 保持不变

(D) 先增后减

解 应选 (B). 只要不是标准正态分布, 就往标

准正态上想: $\frac{X - \mu}{\sigma} \sim N(0, 1)$

讨论 $P\{|X - \mu| < 1\}$ 的单调性应在将其标准化之后进行. 由 $P\{|X - \mu| < 1\} = P\left\{\left|\frac{X - \mu}{\sigma}\right| < \frac{1}{\sigma}\right\} =$

$2\Phi\left(\frac{1}{\sigma}\right)-1$, 又 $\Phi(u)$ 为单调增加函数, $u=\frac{1}{\sigma}$ 为单调减少函数, 则随着 σ 的增大, $\Phi\left(\frac{1}{\sigma}\right)$ 单调减少,

从而知, 随着 σ 的增大, $P\{|X-\mu|<1\}$ 单调减少, 故选择 (B).

例 2.12 设随机变量 X 服从正态分布 $N(0, 1)$, 对给定的 $\alpha(0<\alpha<1)$, 数 u_α 满足 $P\{X>u_\alpha\}=\alpha$,

若 $P\{|X|<x\}=\alpha$, 则 x 等于 ().

(A) $\frac{u_\alpha}{2}$

(B) $u_{1-\frac{\alpha}{2}}$

(C) $\frac{u_{1-\alpha}}{2}$

(D) $u_{1-\alpha}$

解 应选 (C).

如图 2-9 所示, 从几何直观看, 条件中给出的 u_α 即为上侧 α 分位数. 类比可得 $P\{|X|<u_{1-\frac{\alpha}{2}}\}=\alpha$,

故选择 (C).

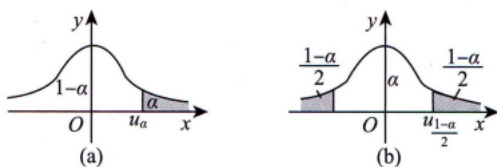


图 2-9

例 2.13 设随机变量 X 满足不等式 $1\leq X\leq 4$, 且 $P\{X=1\}=\frac{1}{4}$, $P\{X=4\}=\frac{1}{3}$, 当 X 在区间

$(1, 4)$ 时, X 服从均匀分布. 求 X 的分布函数.

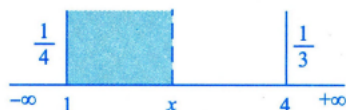
分析 X 的分布是混合型的, 则用定义法求分布函数.

条件分布

解 当 $x<1$ 时, $F(x)=P\{X\leq x\}=0$;

当 $x=1$ 时, $F(x)=P\{X\leq x\}=\frac{1}{4}$;

当 $1<x<4$ 时,



分布函数 $F(x)=P\{X\leq x\}$, 且 x 需取遍 $-\infty$ 到 $+\infty$

$$F(x)=P\{X\leq 1\}+P\{1<X\leq x\},$$

其中

$$P\{1<X\leq x\}=P\{1<X\leq x, \Omega\}=P\{1<X\leq x, 1<X<4\}+P\{1<X\leq x, \overline{1<X<4}\}$$

$$=P\{1<X\leq x|1<X<4\}P\{1<X<4\}+P\{1<X\leq x|\overline{1<X<4}\}P\{\overline{1<X<4}\}$$

$$=\frac{x-1}{4-1}\cdot\frac{5}{12}+0=\frac{5(x-1)}{36},$$

注意到 X 只在 $[1, 4]$ 取值, 所以 X 在 $(1, 4)$ 内的概率即为 X 在 $[1, 4]$ 的概率减去两个端点.

$$P\{1<X<4\}=1-P\{X=1\}-P\{X=4\}=1-\frac{1}{4}-\frac{1}{3}=\frac{5}{12}$$

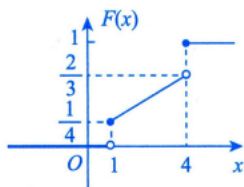
故

$$F(x)=\frac{1}{4}+\frac{5(x-1)}{36}=\frac{1}{9}+\frac{5}{36}x;$$

当 $x\geq 4$ 时, $F(x)=1$.

从而得 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < 1, \\ \frac{1}{9} + \frac{5}{36}x, & 1 \leq x < 4, \\ 1, & x \geq 4. \end{cases}$$



注 (1) 将 $\{1 < X \leq x\}$ 写成 $\{1 < X \leq x, \Omega\}$, 然后作全集分解, 是解决本题的关键.

(2) 由图可知, X 既不是离散型随机变量, 也不是连续型随机变量, 称为混合型随机变量.

(3) 还可知, 当 $1 < x < 4$ 时, 记 $A = \{1 < X < 4\}$, 则

$$f(x) = f(x|A)P(A) + f(x|\bar{A})P(\bar{A}) = \frac{1}{3} \times \frac{5}{12} + 0 \times \frac{7}{12} = \frac{5}{36}.$$

或在 $1 < x < 4$ 时, $f(x) = F'(x) = \frac{5}{36}$.
 对立事件, 所以概率为0

四 一维随机变量函数的分布



1 概念

设 X 为随机变量, 函数 $y = g(x)$, 则以随机变量 X 作为自变量的函数 $Y = g(X)$ 也是随机变量, 称为

随机变量 X 的函数. 例如: $Y = aX^2 + bX + c$, $Y = |X - a|$, $Y = \begin{cases} X, & X \leq 1, \\ 1, & X > 1, \end{cases}$ 等等.

示例 设随机变量 $X \sim \begin{pmatrix} 1 & 2 & 5 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$, 令 $Y = 10\,000X + 8$, 则 $Y \sim \begin{pmatrix} 10\,008 & 20\,008 & 50\,008 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$.

2 随机变量函数的分布

(1) 离散型 $\rightarrow$ 离散型.

设 X 为离散型随机变量, 其概率分布为 $P\{X = x_i\} = p_i (i = 1, 2, \dots)$, 则 X 的函数 $Y = g(X)$ 也是离散型随机变量, 其概率分布为 $P\{Y = g(x_i)\} = p_i (i = 1, 2, \dots)$, 即

$$Y \sim \begin{pmatrix} g(x_1) & g(x_2) & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}.$$

如果有若干个 $g(x_i)$ 值相同, 则合并诸项为一项 $g(x_k)$, 并将相应概率相加作为 Y 取 $g(x_k)$ 值的概率.

示例 设随机变量 $X \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$, 令 $Y = X^2$, 则 $Y \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$.

(2) 连续型→连续型(或混合型). 注意下标的书写, 说明是谁的分布函数

设 X 为连续型随机变量, 其分布函数、概率密度分别为 $F_X(x)$ 与 $f_X(x)$, 随机变量 $Y = g(X)$ 是 X 的函数, 则 Y 的分布函数或概率密度可用下面两种方法求得.

① 分布函数法(定义法). 最可靠

直接由定义求 Y 的分布函数

此不等式的几何意义: 曲线 $Y = g(X)$ 在直线 $Y = y$ 下方, 由此可通过作图得出 X 的取值范围, 在 $Y = g(X)$ 是非单调函数时, 一般比解析法方便

$$F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = \int_{g(x) \leq y} f_X(x) dx,$$

反解出 X 即可

如果 $F_Y(y)$ 连续, 且除有限个点外, $F'_Y(y)$ 存在且连续, 则 Y 的概率密度 $f_Y(y) = F'_Y(y)$.

② 公式法.

根据上面的分布函数法, 若 $y = g(x)$ 在 (a, b) 上是关于 x 的严格单调可导函数, 则存在 $x = h(y)$ 是 $y = g(x)$ 在 (a, b) 上的可导反函数.

若 $y = g(x)$ 严格单调增加, 则 $x = h(y)$ 也严格单调增加, 即 $h'(y) > 0$, 且

$$F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = P\{X \leq h(y)\} = \int_{-\infty}^{h(y)} f_X(x) dx,$$

故 $f_Y(y) = F'_Y(y) = f_X[h(y)] \cdot h'(y)$.

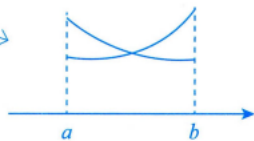
若 $y = g(x)$ 严格单调减少, 则 $x = h(y)$ 也严格单调减少, 即 $h'(y) < 0$, 且

$$F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\} = P\{X \geq h(y)\} = \int_{h(y)}^{+\infty} f_X(x) dx,$$

故 $f_Y(y) = F'_Y(y) = -f_X[h(y)] \cdot h'(y) = f_X[h(y)] \cdot [-h'(y)]$.

综上,

$$f_Y(y) = \begin{cases} f_X[h(y)] \cdot |h'(y)|, & \alpha < y < \beta, \\ 0, & \text{其他,} \end{cases}$$



其中 $\alpha = \min\left\{\lim_{x \rightarrow a^+} g(x), \lim_{x \rightarrow b^-} g(x)\right\}$, $\beta = \max\left\{\lim_{x \rightarrow a^+} g(x), \lim_{x \rightarrow b^-} g(x)\right\}$.

例 2.14 设随机变量 X 的分布律为

$$P\{X = -1\} = \frac{1}{4}, \quad P\{X = 0\} = \frac{1}{2}, \quad P\{X = 1\} = \frac{1}{4},$$

求 $Y = X^2$ 的分布.

分析 离散型随机变量的函数仍属于离散型随机变量范畴, 求解其分布问题仍然按照计算离散型随机变量分布的三个步骤进行.

解 注意到 Y 的可能取值为 0, 1, 于是有

$$P\{Y = 0\} = P\{X^2 = 0\} = P\{X = 0\} = \frac{1}{2},$$

$$P\{Y = 1\} = P\{X^2 = 1\} = P\{X = \pm 1\}$$

$$= P\{X=1\} + P\{X=-1\} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$

或

$$P\{Y=1\} = 1 - P\{Y=0\} = \frac{1}{2},$$

即有

$$Y \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

★★★ **例 2.15** 设随机变量 X 的分布函数 $F_X(x)$ 是严格单调增加函数, 其反函数 $F_X^{-1}(y)$ 存在,

$Y = F_X(X)$. 证明: Y 服从区间 $(0, 1)$ 上的均匀分布.

证 $Y = F_X(X)$ 是在区间 $(0, 1)$ 上取值的随机变量, 故

当 $y < 0$ 时, $F_Y(y) = 0$;

当 $y \geq 1$ 时, $F_Y(y) = 1$;

当 $0 \leq y < 1$ 时,

$$F_Y(y) = P\{Y \leq y\} = P\{F_X(X) \leq y\} = P\{X \leq F_X^{-1}(y)\} = F_X[F_X^{-1}(y)] = y.$$

综上所述, $Y = F_X(X)$ 的分布函数为

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ y, & 0 \leq y < 1, \\ 1, & y \geq 1, \end{cases}$$

注意到分布函数处处右连续

这是在区间 $(0, 1)$ 上的均匀分布函数, 所以 $Y \sim U(0, 1)$.

注 (1) 题设条件中的“ $F_X(x)$ 严格单调增加”是充分条件, 事实上只需要 $F_X(x)$ 在 X 的正概率密度区间上严格单调增加即可.

(2) 本题是一个重要结论, 即在满足 $F_X(x)$ 在 X 的正概率密度区间上严格单调增加时, 若 $X \sim F_X(x)$, 则 $Y = F_X(X) \sim U(0, 1)$. 这一结论考研中常用.

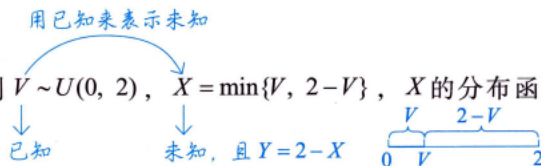
例 2.16 在区间 $(0, 2)$ 上随机取一点, 将该区间分成两段, 较短一段的长度记为 X , 较长一

段的长度记为 Y . 令 $Z = \frac{Y}{X}$.

(1) 求 X 的概率密度;

(2) 求 Z 的概率密度.

解 (1) 设随机取的点的坐标记为 V , 则 $V \sim U(0, 2)$, $X = \min\{V, 2-V\}$, X 的分布函数记为 $F_X(x)$. 由于 $P\{0 \leq X \leq 1\} = 1$, 故



当 $x < 0$ 时, $F_X(x) = 0$;

当 $x \geq 1$ 时, $F_X(x) = 1$;

当 $0 \leq x < 1$ 时,

$$F_X(x) = P\{X \leq x\}$$

$$= P\{\min\{V, 2-V\} \leq x\}$$

$$= 1 - P\{\min\{V, 2-V\} > x\} \quad V > x \text{ 且 } 2-V > x \Rightarrow x < V < 2-x$$

$$= 1 - P\{x < V < 2-x\}$$

$$= 1 - \frac{(2-x)-x}{2} = x \quad \xrightarrow{\text{数轴}} \quad \begin{array}{c} 0 \quad x \quad 2-x \quad 2 \\ \hline \end{array}$$

$$= 1 - \frac{2-2x}{2} = x \quad \xrightarrow{\text{数轴}} \quad \begin{array}{c} 0 \quad x \quad 2-x \quad 2 \\ \hline \end{array}$$

所以 X 的分布函数为 $F_X(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases}$

故 X 的概率密度为 $f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{其他} \end{cases}$

能用一元表示就不要用多元

(2) 由条件知, $Z = \frac{Y}{X} = \frac{2-X}{X}$. 由于函数 $z = \frac{2-x}{x}$ 在 $(0, 1)$ 内严格单调减少且可导, 反函数为

$x = \frac{2}{1+z}$, 且 $\frac{dx}{dz} = -\frac{2}{(1+z)^2}$, 故 Z 的概率密度为

$$f_Z(z) = \begin{cases} f_X\left(\frac{2}{1+z}\right) \left| -\frac{2}{(1+z)^2} \right|, & z > 1, \\ 0, & \text{其他} \end{cases} \quad \xrightarrow{\text{数轴}} \quad z(0) \rightarrow \infty, z(1) = 1$$

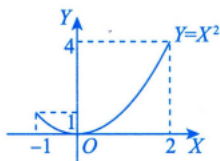
$$= \begin{cases} \frac{2}{(1+z)^2}, & z > 1, \\ 0, & \text{其他} \end{cases}$$

例 2.17 设随机变量 X 的概率密度为 $f_X(x) = \begin{cases} \frac{1}{2}, & -1 < x < 0, \\ \frac{1}{4}, & 0 \leq x < 2, \\ 0, & \text{其他} \end{cases}$ 令 $Y = X^2$, 求 Y 的概率密度 $f_Y(y)$.

分析 利用定义来做.

$$Y \sim F_Y(y) = P\{Y \leq y\} = P\{X^2 \leq y\},$$

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ P\{-\sqrt{y} \leq X \leq \sqrt{y}\}, & 0 \leq y < 1, \\ P\{-1 \leq X \leq \sqrt{y}\}, & 1 \leq y < 4, \\ 1, & y \geq 4. \end{cases}$$



解 Y 的分布函数为 $F_Y(y) = P\{Y \leq y\} = P\{X^2 \leq y\}$.

① 当 $y < 0$ 时, $F_Y(y) = 0$;

② 当 $0 \leq y < 1$ 时,

$$\begin{aligned} F_Y(y) &= P\{-\sqrt{y} \leq X \leq \sqrt{y}\} = P\{-\sqrt{y} \leq X < 0\} + P\{0 \leq X \leq \sqrt{y}\} \\ &= \frac{1}{2}\sqrt{y} + \frac{1}{4}\sqrt{y} = \frac{3}{4}\sqrt{y}; \end{aligned}$$

③ 当 $1 \leq y < 4$ 时,

$$F_Y(y) = P\{-1 \leq X < 0\} + P\{0 \leq X \leq \sqrt{y}\} = \frac{1}{2} + \frac{1}{4}\sqrt{y};$$

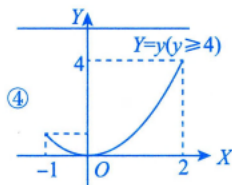
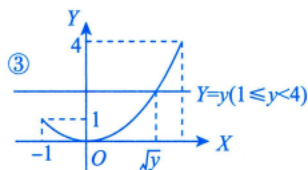
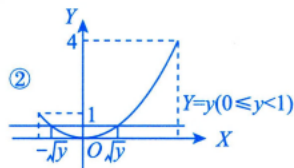
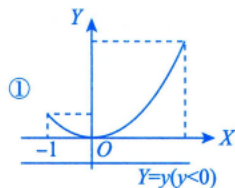
④ 当 $y \geq 4$ 时, $F_Y(y) = 1$.

所以 Y 的分布函数为

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ \frac{3}{4}\sqrt{y}, & 0 \leq y < 1, \\ \frac{1}{2} + \frac{\sqrt{y}}{4}, & 1 \leq y < 4, \\ 1, & y \geq 4. \end{cases}$$

故 Y 的概率密度为

$$f_Y(y) = \begin{cases} \frac{3}{8\sqrt{y}}, & 0 < y < 1, \\ \frac{1}{8\sqrt{y}}, & 1 < y < 4, \\ 0, & \text{其他.} \end{cases}$$



五 综合题举例

例 2.18 投保人的损失事件发生时, 保险公司的赔付额 Y 与投保人的损失额 X 的关系为

$$Y = \begin{cases} 0, & X \leq 100, \\ X - 100, & X > 100. \end{cases}$$

设损失事件发生时, 投保人的损失额 X 的概率密度为

$$f(x) = \begin{cases} \frac{2 \times 100^2}{(100 + x)^3}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

(1) 求 $P\{Y > 0\}$ 及 EY ;



(2) 这种损失事件在一年内发生的次数记为 N , 保险公司在一年内就这种损失事件产生的理赔次数记为 M . 假设 N 服从参数为 8 的泊松分布, 在 $N=n(n \geq 1)$ 的条件下, M 服从二项分布 $B(n, p)$, 其中 $p = P\{Y > 0\}$, 求 M 的概率分布.

解 (1) $P\{Y > 0\} = P\{X - 100 > 0\} = P\{X > 100\} = \int_{100}^{+\infty} \frac{2 \times 100^2}{(100+x)^3} dx = \frac{1}{4}.$

$$EY = \int_{100}^{+\infty} (x-100) \frac{2 \times 100^2}{(100+x)^3} dx = 50.$$

(2) 由题知, $N \sim P(8), P\{M = m | N = n\} = C_n^m \left(\frac{1}{4}\right)^m \left(\frac{3}{4}\right)^{n-m}$, 由乘法公式可得

$$\begin{aligned} P\{M = m, N = n\} &= P\{N = n\} P\{M = m | N = n\} \\ &= \frac{8^n}{n!} e^{-8} \cdot C_n^m \cdot \left(\frac{1}{4}\right)^m \cdot \left(\frac{3}{4}\right)^{n-m}, \end{aligned}$$

其中 $m = 0, 1, \dots, n; n = 1, 2, \dots$.

故 M 的概率分布为

$$\begin{aligned} P\{M = m\} &= \sum_{n=m}^{\infty} P\{M = m, N = n\} = \sum_{n=m}^{\infty} \frac{8^n}{n!} e^{-8} \cdot C_n^m \cdot \left(\frac{1}{4}\right)^m \cdot \left(\frac{3}{4}\right)^{n-m} \\ &= \sum_{n=m}^{\infty} \frac{8^n}{n!} e^{-8} \cdot \frac{n!}{m!(n-m)!} \left(\frac{1}{4}\right)^m \cdot \left(\frac{3}{4}\right)^{n-m} \\ &= \left(\frac{1}{4}\right)^m e^{-8} \frac{8^m}{m!} \sum_{n=m}^{\infty} \frac{8^{n-m}}{(n-m)!} \cdot \left(\frac{3}{4}\right)^{n-m} \\ &= \left(\frac{1}{4}\right)^m e^{-8} \frac{8^m}{m!} \sum_{n=m}^{\infty} \frac{6^{n-m}}{(n-m)!} = \frac{2^m}{m!} e^{-8} \cdot e^6 = \frac{2^m}{m!} e^{-2}, \end{aligned}$$

即 $P\{M = m\} = \frac{2^m}{m!} e^{-2}, m = 0, 1, 2, \dots$, 即 $M \sim P(2)$.

注 本题第(2)问在 2001 年考研真题中出现过其中一部分, 如下.

设某班车起点站上车人数 X 服从参数为 $\lambda (\lambda > 0)$ 的泊松分布, 每位乘客在中途下车的概率为 $p (0 < p < 1)$, 且中途下车与否相互独立. 以 Y 表示在中途下车的人数, 求:

(1) 在发车时有 n 个乘客的条件下, 中途有 m 人下车的概率;

(2) 二维随机变量 (X, Y) 的概率分布.

解 (1) $P\{Y = m | X = n\} = C_n^m p^m (1-p)^{n-m}, 0 \leq m \leq n, n = 0, 1, 2, \dots$.

(2) $P\{X = n, Y = m\} = P\{Y = m | X = n\} P\{X = n\}$

$$= C_n^m p^m (1-p)^{n-m} \cdot \frac{e^{-\lambda}}{n!} \lambda^n, 0 \leq m \leq n, n = 0, 1, 2, \dots$$

 基础习题精练

习题

2.1 已知 $F(x)$ 是分布函数, 下列函数中:

① $aF(x) (a > 0, a \neq 1)$; ② $F(x) + F(-x)$; ③ $F(x) - F(-x)$; ④ $F(x) \cdot F(-x)$.

不能作为分布函数的个数是 ().

(A) 1 (B) 2 (C) 3 (D) 4

2.2 设 $X \sim B(2, p)$, $Y \sim B(4, p)$, 且 $P\{X \geq 1\} = \frac{5}{9}$, 则 $P\{Y \geq 1\} = ()$.

(A) $\frac{65}{81}$ (B) $\frac{16}{81}$ (C) 1 (D) $\frac{4}{7}$

2.3 设随机变量 X 服从正态分布 $N(\mu_1, \sigma_1^2)$, 随机变量 Y 服从正态分布 $N(\mu_2, \sigma_2^2)$, $\sigma_1, \sigma_2 > 0$, 且

$P\{|X - \mu_1| < 1\} > P\{|Y - \mu_2| < 1\}$, 则必有 ().

(A) $\sigma_1 < \sigma_2$ (B) $\sigma_1 > \sigma_2$ (C) $\mu_1 < \mu_2$ (D) $\mu_1 > \mu_2$

2.4 设随机变量 X 的分布函数 $F(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{2}, & 0 \leq x < 1, \\ 1 - e^{-x}, & x \geq 1, \end{cases}$ 则 $P\{0 \leq X \leq 1\} = \underline{\hspace{2cm}}$, $P\{0 < X < 1\} = \underline{\hspace{2cm}}$.

2.5 离散型随机变量 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < -1, \\ 0.4, & -1 \leq x < 2, \\ 0.8, & 2 \leq x < 3, \\ 1, & x \geq 3, \end{cases}$$

则 X 的分布列为 $\underline{\hspace{2cm}}$, $P\{X \leq 3.2\} = \underline{\hspace{2cm}}$, 方程 $x^2 + Xx + 1 = 0$ 有实根的概率为 $\underline{\hspace{2cm}}$.

2.6 设连续型随机变量 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < -a, \\ A + B \arcsin \frac{x}{a}, & -a \leq x < a, \\ 1, & x \geq a, \end{cases}$$

则 $A = \underline{\hspace{2cm}}$, $B = \underline{\hspace{2cm}}$, X 的概率密度 $f(x) = \underline{\hspace{2cm}}$, $P\left\{-\frac{a}{2} < X < \frac{a}{2}\right\} = \underline{\hspace{2cm}}$.

2.7 设连续型随机变量 X 的概率密度为

$$f(x) = \begin{cases} \frac{A}{x^2}, & x > 100, \\ 0, & x \leq 100, \end{cases}$$

则 $A = \underline{\hspace{2cm}}$, $P\{100 < X \leq 150\} = \underline{\hspace{2cm}}$, $P\{X \geq 1000\} = \underline{\hspace{2cm}}$, $P\{X = 1000\} = \underline{\hspace{2cm}}$, X 的分布函数 $F(x) = \underline{\hspace{2cm}}$.

2.8 设随机变量 X 服从正态分布 $N(\mu, \sigma^2)$ ($\sigma > 0$), 且二次方程 $y^2 + 4y + X = 0$ 无实根的概率为 $\frac{1}{2}$, 则 $\mu = \underline{\hspace{2cm}}$.

2.9 设随机变量 X 服从区间 $(2, 5)$ 上的均匀分布, 则对 X 进行 3 次独立观测中, 至少有 2 次的观测值大于 3 的概率为 $\underline{\hspace{2cm}}$.

2.10 设随机变量 X 的概率密度为 $f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{其他,} \end{cases}$ Y 表示对 X 的 3 次独立重复试验中事件

$\left\{X \leq \frac{1}{2}\right\}$ 出现的次数, 则 $P\{Y = 2\} = \underline{\hspace{2cm}}$.

2.11 设随机变量 X 和 Y 同分布, X 的概率密度为

$$f(x) = \begin{cases} \frac{3}{8}x^2, & 0 < x < 2, \\ 0, & \text{其他.} \end{cases}$$

已知事件 $A = \{X > a\}$ 和 $B = \{Y > a\}$ 独立, 且 $P(A \cup B) = \frac{3}{4}$, 则常数 $a = \underline{\hspace{2cm}}$.

2.12 已知 X 为随机变量, $Y = X^2 + X + 1$, 已知 X 的概率分布为 $P\{X = -1\} = P\{X = 0\} = P\{X = 1\} = \frac{1}{3}$, 则 Y 的分布函数 $F_Y(y) = \underline{\hspace{2cm}}$.

2.13 已知离散型随机变量 X 的正概率点为 $-1, 0, 2$, 它们各自的概率互不相等且成等差数列.

(1) 求 X 的分布律;

(2) 求 X 的分布函数;

(3) 求 $P\{|X| \leq 1 | X \geq 0\}$.

2.14 随机变量 X 的概率密度为

$$f(x) = \begin{cases} x, & 0 \leq x < 1, \\ 2-x, & 1 \leq x < 2, \\ 0, & \text{其他.} \end{cases}$$

求 X 的分布函数 $F(x)$.

2.15 设连续型随机变量 X 的概率密度 $f(x)$ 是一个偶函数, $F(x)$ 为 X 的分布函数, 证明: 对任意实数 $a > 0$, 有

$$(1) F(-a) = 1 - F(a) = 0.5 - \int_0^a f(x) dx;$$

$$(2) P\{|X| < a\} = 2F(a) - 1;$$

$$(3) P\{|X| > a\} = 2[1 - F(a)].$$

2.16 设随机变量 X 的概率密度为 $f_X(x) = \frac{k}{1+x^2} (x \in \mathbf{R})$. 求:

(1) 常数 k ;

(2) 随机变量 $Y = 1 - \sqrt[3]{X}$ 的概率密度 $f_Y(y)$.

2.17 设随机变量 X 的概率密度为 $f(x) = \frac{e^x}{(1+e^x)^2}, -\infty < x < +\infty$, 令 $Y = e^x$.

(1) 求 X 的分布函数;

(2) 求 Y 的概率密度.

2.18 设随机变量 $X \sim U(-1, 2)$, 求 $Y = |X|$ 的概率密度.

解答

2.1 (D) 解 应用分布函数的充要条件求解. 由于

$$\textcircled{1} \quad aF(+\infty) = a \neq 1,$$

$$\textcircled{2} \quad F(-\infty) + F(+\infty) = 1 \neq 0,$$

$$\textcircled{3} \quad F(-\infty) - F(+\infty) = -1 \neq 0,$$

$$\textcircled{4} \quad F(+\infty) \cdot F(-\infty) = 0 \neq 1,$$

因此①, ②, ③, ④都不能作为分布函数, 故选 (D).

2.2 (A) 解 由题设 $P\{X \geq 1\} = \frac{5}{9} = 1 - P\{X = 0\} = 1 - C_2^0 p^0 (1-p)^2$, 即 $(1-p)^2 = \frac{4}{9}$, 解得 $p = \frac{1}{3}$, 从

而有

$$P\{Y \geq 1\} = 1 - P\{Y = 0\} = 1 - C_4^0 p^0 (1-p)^4 = 1 - \left(\frac{2}{3}\right)^4 = \frac{65}{81},$$

故选 (A).

2.3 (A) 解 所有有关正态分布不等式的判别都应先标准化, 即由

$$P\{|X - \mu_1| < 1\} > P\{|Y - \mu_2| < 1\},$$

有

$$P\left\{\left|\frac{X - \mu_1}{\sigma_1}\right| < \frac{1}{\sigma_1}\right\} > P\left\{\left|\frac{Y - \mu_2}{\sigma_2}\right| < \frac{1}{\sigma_2}\right\},$$

从而有

$$\Phi\left(\frac{1}{\sigma_1}\right) > \Phi\left(\frac{1}{\sigma_2}\right).$$

又由函数 $\Phi(x)$ 的单调增加性, 有 $\frac{1}{\sigma_1} > \frac{1}{\sigma_2} > 0$, 即 $\sigma_1 < \sigma_2$, 故选 (A).

2.4 $1 - e^{-1}$; 0 解 $P\{0 \leq X \leq 1\} = P\{X \leq 1\} - P\{X < 0\} = F(1) - F(0-0)$

$$= 1 - e^{-1} - 0 = 1 - e^{-1};$$

$$P\{0 < X < 1\} = P\{X < 1\} - P\{X \leq 0\} = F(1-0) - F(0) = \frac{1}{2} - \frac{1}{2} = 0.$$

2.5 $X \sim \begin{pmatrix} -1 & 2 & 3 \\ 0.4 & 0.4 & 0.2 \end{pmatrix}$; 1; 0.6 解 分布函数的分段点即为离散型随机变量的正概率点, 正概

率点的概率即为对应点的跃度, 即 $X = -1, 2, 3$, 分布列为

$$X \sim \begin{pmatrix} -1 & 2 & 3 \\ 0.4 & 0.4 & 0.2 \end{pmatrix},$$

$$P\{X \leq 3.2\} = F(3.2) = 1.$$

方程 $x^2 + Xx + 1 = 0$ 有实根, 即 $X^2 - 4 \geq 0$, 从而有

$$P\{X^2 - 4 \geq 0\} = P\{X \leq -2\} + P\{X \geq 2\} = F(-2) + 1 - F(2-0) = 1 - 0.4 = 0.6.$$

2.6 $\frac{1}{2}; \frac{1}{\pi}; \begin{cases} \frac{1}{\pi\sqrt{a^2-x^2}}, & -a < x < a, \\ 0, & \text{其他;} \end{cases} \frac{1}{3}$ 解 根据连续型随机变量分布函数的连续性, 有

$$F(-a) = F(-a-0), F(a) = F(a-0),$$

即有

$$\begin{cases} A + B \cdot \left(-\frac{\pi}{2}\right) = 0, \\ A + B \cdot \frac{\pi}{2} = 1, \end{cases}$$

解得 $A = \frac{1}{2}, B = \frac{1}{\pi}$.

于是

$$f(x) = \begin{cases} F'(x), & -a < x < a, \\ 0, & \text{其他} \end{cases} = \begin{cases} \frac{1}{\pi\sqrt{a^2-x^2}}, & -a < x < a, \\ 0, & \text{其他}. \end{cases}$$

$$P\left\{-\frac{a}{2} < X < \frac{a}{2}\right\} = \int_{-\frac{a}{2}}^{\frac{a}{2}} f(x) dx = \int_{-\frac{a}{2}}^{\frac{a}{2}} \frac{1}{\pi\sqrt{a^2-x^2}} dx = \frac{1}{3},$$

或

$$P\left\{-\frac{a}{2} < X < \frac{a}{2}\right\} = F\left(\frac{a}{2}-0\right) - F\left(-\frac{a}{2}\right) = \frac{1}{\pi} \left[\arcsin \frac{1}{2} - \arcsin \left(-\frac{1}{2}\right) \right] = \frac{1}{3}.$$

2.7 $100; \frac{1}{3}; \frac{1}{10}; 0;$ $\begin{cases} 0, & x < 100, \\ 1 - \frac{100}{x}, & x \geq 100 \end{cases}$ 解 根据连续型随机变量概率密度的性质, 有

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{100}^{+\infty} \frac{A}{x^2} dx = -\frac{A}{x} \Big|_{100}^{+\infty} = \frac{A}{100} = 1, \quad A = 100,$$

于是

$$P\{100 < X \leq 150\} = \int_{100}^{150} \frac{100}{x^2} dx = -\frac{100}{x} \Big|_{100}^{150} = \frac{1}{3},$$

$$P\{X \geq 1000\} = \int_{1000}^{+\infty} \frac{100}{x^2} dx = -\frac{100}{x} \Big|_{1000}^{+\infty} = \frac{1}{10},$$

$$P\{X = 1000\} = 0,$$

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0, & x < 100, \\ \int_{100}^x \frac{100}{t^2} dt, & x \geq 100 \end{cases} = \begin{cases} 0, & x < 100, \\ 1 - \frac{100}{x}, & x \geq 100. \end{cases}$$

2.8 4 解 二次方程 $y^2 + 4y + X = 0$ 无实根的充要条件是 $4 - X < 0$. 故由条件知

$$P\{X > 4\} = \frac{1}{2},$$

于是有 $\frac{1}{2} = P\{X > 4\} = 1 - P\{X \leq 4\} = 1 - P\left\{\frac{X - \mu}{\sigma} \leq \frac{4 - \mu}{\sigma}\right\} = 1 - P\left\{Y \leq \frac{4 - \mu}{\sigma}\right\} = 1 - \Phi\left(\frac{4 - \mu}{\sigma}\right),$

$$\Phi\left(\frac{4 - \mu}{\sigma}\right) = \frac{1}{2},$$

其中 $Y = \frac{X - \mu}{\sigma} \sim N(0, 1)$, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$. 又 $\Phi(0) = \frac{1}{2}$, 所以 $\frac{4 - \mu}{\sigma} = 0$, $\mu = 4$.

注 事实上, 根据正态分布的对称性, 可直接得 $\mu = 4$.

2.9 $\frac{20}{27}$ 解 在一次观测中, 观测值大于 3 的概率为

$$p = P\{X > 3\} = \int_3^5 \frac{1}{3} dx = \frac{2}{3}.$$

记 $Y = \{3 \text{ 次独立观测中 } X \text{ 的观测值大于 } 3 \text{ 的次数}\}$, 由 $Y \sim B\left(3, \frac{2}{3}\right)$ 得

$$P\{Y \geq 2\} = 1 - P\{Y = 0\} - P\{Y = 1\} = 1 - \left(\frac{1}{3}\right)^3 - 3 \times \frac{2}{3} \times \left(\frac{1}{3}\right)^2 = \frac{20}{27}.$$

2.10 $\frac{9}{64}$ 解 依题意得 $Y \sim B(3, p)$, 其中 $p = P\left\{X \leq \frac{1}{2}\right\} = \int_0^{\frac{1}{2}} 2x dx = \frac{1}{4}$, 故

$$P\{Y=2\}=C_3^2 p^2(1-p)=\frac{9}{64}.$$

2.11 $\sqrt[3]{4}$ 解 由 X 和 Y 同分布可得 $P(A)=P(B)$, 从而

$$\frac{3}{4}=P(A\cup B)=P(A)+P(B)-P(AB)=2P(A)-[P(A)]^2,$$

由此解得 $P(A)=\frac{1}{2}$, 则 $0 < a < 2$. 又

$$\frac{1}{2}=P(A)=P\{X>a\}=\int_a^2 \frac{3}{8}x^2 dx=\frac{1}{8}x^3\Big|_a^2=1-\frac{1}{8}a^3,$$

解得 $a=\sqrt[3]{4}$.

注 随机变量 X 与 Y 同分布, 并不意味着 $X=Y$, 但反之成立, 即若 $X=Y$, 则 X 与 Y 同分布.

2.12 $\begin{cases} 0, & y < 1, \\ \frac{2}{3}, & 1 \leq y < 3, \\ 1, & y \geq 3 \end{cases}$ 解 由题设知 Y 为离散型随机变量, 当 $X=-1$ 时, Y 取 1; 当 $X=0$ 时, Y 取 1;

当 $X=1$ 时, Y 取 3. 则 Y 的概率分布为

$$P\{Y=1\}=P\{X=-1\}+P\{X=0\}=\frac{2}{3}, \quad P\{Y=3\}=P\{X=1\}=\frac{1}{3}.$$

故 Y 的分布函数为

$$F_Y(y)=P\{Y \leq y\}=\begin{cases} 0, & y < 1, \\ \frac{2}{3}, & 1 \leq y < 3, \\ 1, & y \geq 3. \end{cases}$$

2.13 解 (1) 设 $P\{X=-1\}=a+d$, $P\{X=0\}=a$, $P\{X=2\}=a-d$, 其中 $0 < a+d < 1$, $0 < a < 1$, $0 < a-d < 1$, 由离散型随机变量 X 分布律的性质, 有

$$a+d+a+a-d=3a=1,$$

解得 $a=\frac{1}{3}$, 且 $-\frac{1}{3} < d < \frac{1}{3}$, $d \neq 0$, 因此 $X \sim \begin{pmatrix} -1 & 0 & 2 \\ \frac{1}{3}+d & \frac{1}{3} & \frac{1}{3}-d \end{pmatrix}$, 其中 $-\frac{1}{3} < d < \frac{1}{3}$, $d \neq 0$.

(2) 由 (1) 可知,

$$F(x)=\begin{cases} 0, & x < -1, \\ \frac{1}{3}+d, & -1 \leq x < 0, \\ \frac{2}{3}+d, & 0 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

(3) 由条件概率公式及 (1) 可知, $P\{|X| \leq 1 | X \geq 0\} = \frac{P\{0 \leq X \leq 1\}}{P\{X \geq 0\}} = \frac{1}{2-3d}$.

2.14 解 由题设知, 当 $x < 0$ 时, $F(x) = 0$;

当 $0 \leq x < 1$ 时,

$$F(x) = \int_{-\infty}^x f(t) dt = \int_0^x t dt = \frac{1}{2} t^2 \Big|_0^x = \frac{x^2}{2};$$

当 $1 \leq x < 2$ 时,

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt = \int_1^x (2-t) dt + \int_0^1 t dt \\ &= \left(2t - \frac{t^2}{2} \right) \Big|_1^x + \frac{t^2}{2} \Big|_0^1 = 2x - \frac{x^2}{2} - 1; \end{aligned}$$

当 $x \geq 2$ 时,

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt = \int_0^1 t dt + \int_1^2 (2-t) dt \\ &= \frac{t^2}{2} \Big|_0^1 + \left(2t - \frac{t^2}{2} \right) \Big|_1^2 = 1. \end{aligned}$$

所以

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{x^2}{2}, & 0 \leq x < 1, \\ 2x - \frac{x^2}{2} - 1, & 1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

2.15 证明 因为 $f(x)$ 是一个偶函数, 所以 $f(-x) = f(x)$, 且 $1 = \int_{-\infty}^{+\infty} f(x) dx = 2 \int_0^{+\infty} f(x) dx$, 可得

$$\int_0^{+\infty} f(x) dx = 0.5.$$

(1) $F(-a) = \int_{-\infty}^{-a} f(x) dx$, 令 $x = -t$, 则 $F(-a) = \int_{+\infty}^a f(t)(-dt) = \int_a^{+\infty} f(t) dt = 1 - F(a)$.

$$\begin{aligned} F(-a) &= \int_a^{+\infty} f(t) dt = \int_0^{+\infty} f(t) dt - \int_0^a f(t) dt \\ &= 0.5 - \int_0^a f(x) dx. \end{aligned}$$

$$\begin{aligned} (2) \quad P\{|X| < a\} &= P\{-a < X < a\} = F(a-0) - F(-a) \\ &= F(a) - [1 - F(a)] \\ &= 2F(a) - 1. \end{aligned}$$

$$(3) \quad P\{|X| > a\} = P\{X < -a\} + P\{X > a\} = F(-a-0) + 1 - F(a)$$

$$= 1 - F(a) + 1 - F(a)$$

$$= 2[1 - F(a)] .$$

2.16 解 (1) 由于

$$\int_{-\infty}^{+\infty} f_X(x) dx = \int_{-\infty}^{+\infty} \frac{k}{1+x^2} dx = k \arctan x \Big|_{-\infty}^{+\infty} = k\pi = 1,$$

因此 $k = \frac{1}{\pi}$.

(2) X 的概率密度为 $f_X(x) = \frac{1}{\pi(1+x^2)} (x \in \mathbf{R})$. 随机变量 $Y = 1 - \sqrt[3]{X}$ 的分布函数为

$$F_Y(y) = P\{Y \leq y\} = P\{1 - \sqrt[3]{X} \leq y\} = P\{X \geq (1-y)^3\}$$

$$= \int_{(1-y)^3}^{+\infty} f_X(x) dx = \int_{(1-y)^3}^{+\infty} \frac{1}{\pi(1+x^2)} dx$$

$$= \frac{1}{\pi} \arctan x \Big|_{(1-y)^3}^{+\infty} = \frac{1}{\pi} \left[\frac{\pi}{2} - \arctan(1-y)^3 \right] (y \in \mathbf{R}),$$

由此可得 $Y = 1 - \sqrt[3]{X}$ 的概率密度为 $f_Y(y) = F'_Y(y) = \frac{3(1-y)^2}{\pi[1+(1-y)^6]} (y \in \mathbf{R})$.

2.17 解 (1) X 的分布函数为

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{e^t}{(1+e^t)^2} dt = \int_{-\infty}^x \frac{d(1+e^t)}{(1+e^t)^2} .$$

$$= -\frac{1}{1+e^t} \Big|_{-\infty}^x = 1 - \frac{1}{1+e^x} (-\infty < x < +\infty) .$$

(2) 函数 $y = e^x$ 单调且反函数为 $x = \ln y (y > 0)$, 从而 Y 的概率密度为

$$f_Y(y) = \begin{cases} \frac{1}{y} f(\ln y), & y > 0, \\ 0, & \text{其他} \end{cases}$$

$$= \begin{cases} \frac{1}{(1+y)^2}, & y > 0, \\ 0, & \text{其他}. \end{cases}$$

2.18 解 X 的概率密度为

$$f_X(x) = \begin{cases} \frac{1}{3}, & -1 < x < 2, \\ 0, & \text{其他}. \end{cases}$$

记 X 的分布函数为 $F_X(x)$, 先求 Y 的分布函数 $F_Y(y)$.

当 $y < 0$ 时, $F_Y(y) = P\{Y \leq y\} = 0$;

当 $y \geq 0$ 时, $F_Y(y) = P\{|X| \leq y\} = P\{-y \leq X \leq y\} = F_X(y) - F_X(-y-0)$.

将 $F_Y(y)$ 关于 y 求导可得 Y 的概率密度为

$$f_Y(y) = \begin{cases} f_X(y) + f_X(-y), & y > 0, \\ 0, & \text{其他.} \end{cases}$$

当 $0 \leq y < 1$ 时, $-1 < -y \leq 0$, 因而 $f_X(y) = \frac{1}{3}$, $f_X(-y) = \frac{1}{3}$, 此时

$$f_Y(y) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3};$$

当 $1 \leq y < 2$ 时, $-2 < -y \leq -1$, 因而 $f_X(y) = \frac{1}{3}$, $f_X(-y) = 0$, 此时

$$f_Y(y) = \frac{1}{3};$$

当 $y \geq 2$ 时, $-y \leq -2$, 因而 $f_X(y) = 0$, $f_X(-y) = 0$, 此时 $f_Y(y) = 0$. 故 Y 的概率密度为

$$f_Y(y) = \begin{cases} \frac{2}{3}, & 0 \leq y < 1, \\ \frac{1}{3}, & 1 \leq y < 2, \\ 0, & \text{其他.} \end{cases}$$

第3讲

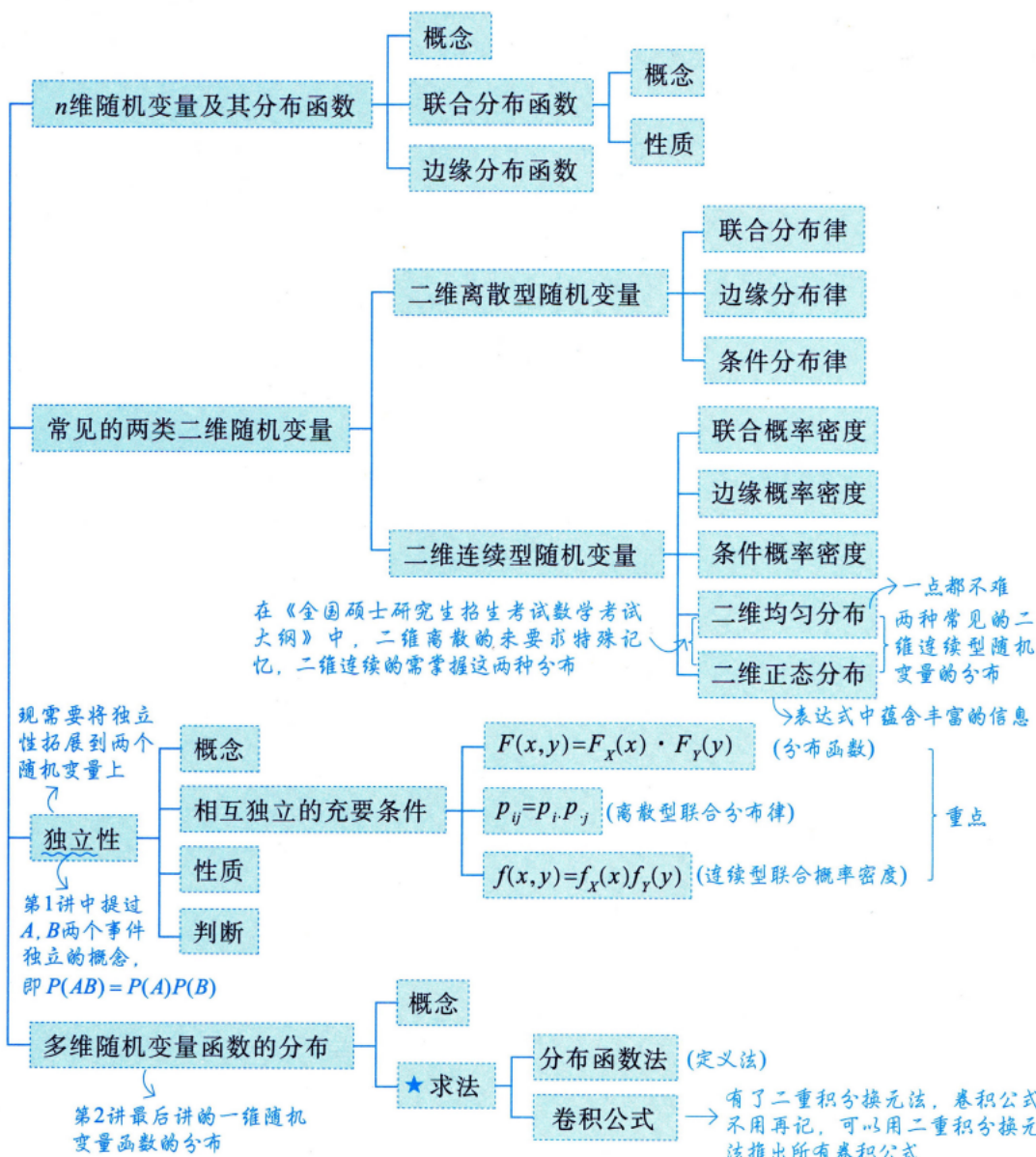
多维随机变量及其分布

也可称随机向量

在一维基础上, 引入两个或更多的随机变量



基础知识结构 → 结构与一维一致, 但需找到多维(二维)与一维的联系与区别



基础内容精讲

一 n 维随机变量及其分布函数

在一维中 $X = X(w)$, $w \in \Omega$, 将事件映射到实数轴上使其成为一个数. 数是有概率出现的, 随机而定的, 称之为随机变量

1 n 维随机变量的概念

要指出维数

习惯称 如果 $X_1, X_2, \dots, X_n$ 是定义在同一个样本空间 Ω 上的 n 个随机变量, 则称 $(X_1, X_2, \dots, X_n)$ 为 n 维随机变量或 n 维随机向量, $X_i (i=1, 2, \dots, n)$ 称为第 i 个分量.

将这些指标写成数组

n 个有序数组 $\rightarrow X_1, X_2, \dots, X_n$ 的取值都是有概率的

当 $n=2$ 时, 称 (X, Y) 为二维随机变量或二维随机向量.

主要考 $n=2$, 但 $n=3, n=4$ 也需要准备. 学习时, 主要是二维和 n 维

注 在实际问题中, 所研究问题的数量指标往往不止一个. 比如根据一个体检的结果来反映他的健康程度. 年龄越大, 体检的项目就越多, 这些数量指标显然可以构成一个多维向量.

比如身高、体重、血压等

2 n 维随机变量的分布函数的概念和性质

(1) 概念.

对任意的 n 个实数 $x_1, x_2, \dots, x_n$, 称 n 元函数

即 $P(A_1 A_2 \dots A_n)$: $A_1, A_2, \dots, A_n$ 同时发生的概率

$$F(x_1, x_2, \dots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}$$

多个随机变量

为 n 维随机变量 $(X_1, X_2, \dots, X_n)$ 的分布函数或随机变量 $X_1, X_2, \dots, X_n$ 的联合分布函数.

当 $n=2$ 时, 对任意的实数 x, y , 称二元函数

$F(x, y)$ 是事件 $A = \{X \leq x\}$ 与 $B = \{Y \leq y\}$ 同时发生的概率

$$F(x, y) = P\{X \leq x, Y \leq y\}$$

为二维随机变量 (X, Y) 的分布函数或随机变量 X 和 Y 的联合分布函数, 记为 $(X, Y) \sim F(x, y)$.

4个性质, 尤其④最重要

(2) 性质.

有两个自变量

①单调性 $F(x, y)$ 是 x, y 的单调不减函数:

类似“偏”导数, 固定一个字

母, 看另一个字母的变化

二元函数研究单调性的方法

如固定 $y = y_0$, 一刀切进去, 研究关于 x 的变化率

对任意固定的 y , 当 $x_1 < x_2$ 时, $F(x_1, y) \leq F(x_2, y)$; $\rightarrow$ 关于 x 单调不减

对任意固定的 x , 当 $y_1 < y_2$ 时, $F(x, y_1) \leq F(x, y_2)$. $\rightarrow$ 关于 y 单调不减

②右连续性 $F(x, y)$ 是 x, y 的右连续函数:

一维中 $F(x)$
也是右连续

→ 固定 y

$$\lim_{x \rightarrow x_0^+} F(x, y) = F(x_0 + 0, y) = F(x_0, y);$$

$$\lim_{y \rightarrow y_0^+} F(x, y) = F(x, y_0 + 0) = F(x, y_0).$$

右极限值 = 函数值, 故右连续

③有界性 $F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0, F(+\infty, +\infty) = 1$.

有一个负无穷, 为不可能事件, 不可能事件交任何事件都是不可能事件

$$P(X \leq +\infty, Y \leq +\infty)$$

必然事件 必然事件

必然事件的交集也是必然事件, 必然事件的概率为1

注 $F(x, y)$ 关于某一个自变量单调不减; $F(x, y)$ 关于某一个自变量右连续; $F(x, y)$ 的两个自变量都趋于正无穷时, 概率是1. 性质①, ②, ③与一维类似.

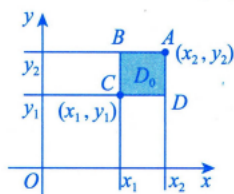
④非负性 对任意的 $x_1 < x_2, y_1 < y_2$, 有 → 点 (x_2, y_2) 在点 (x_1, y_1) 的右上方

$$P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0.$$

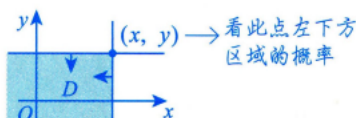
→ $F(A)$ → $F(D)$ → $F(B)$ → $F(C)$

①实质上是 $(x, y) \in D_0$ 的概率 ≥ 0 .

→ 非0区域, 有面积



$$② F(x, y) = P\{X \leq x, Y \leq y\}$$



→ 看此点左下方区域的概率

3 边缘分布函数

单独看

设二维随机变量 (X, Y) 的分布函数为 $F(x, y)$, 随机变量 X 与 Y 的分布函数 $F_X(x)$ 与 $F_Y(y)$ 分别称为 (X, Y) 关于 X 和关于 Y 的边缘分布函数. 由概率性质得

$$\text{即 } P(A) = P(A \cap \Omega), A = A \cap \Omega \leftarrow$$

$$\text{定义没有变} \leftarrow F_X(x) = P\{X \leq x\} = P\{X \leq x, Y < +\infty\}$$

$$= \lim_{y \rightarrow +\infty} P\{X \leq x, Y \leq y\}$$

$$= \lim_{y \rightarrow +\infty} F(x, y) = F(x, +\infty).$$

直接写, 不需要中间的过程

同理, 有 $F_Y(y) = F(+\infty, y)$.

随机变量 (r.v.)	
一维 r.v. $X \sim F(x)$	二维 r.v. $(X, Y) \sim F(x, y) \longrightarrow$ 联合分布函数
	$\left. \begin{array}{l} X \sim F_X(x) = F(x, +\infty) \\ Y \sim F_Y(y) = F(+\infty, y) \end{array} \right\}$ 多了边缘分布函数

二 常见的两类二维随机变量——离散型随机变量与连续型随机变量



1 二维离散型随机变量的概率分布、边缘分布和条件分布

(1) 概率分布.

① 如果二维随机变量 (X, Y) 的可能取值是有限对值或可列无限对值, 则称 (X, Y) 为二维离散型随机变量.

称

$$p_{ij} = P\{X = x_i, Y = y_j\}, \quad i, j = 1, 2, \dots$$

为 (X, Y) 的分布律或随机变量 X 和 Y 的联合分布律, 记为 $(X, Y) \sim p_{ij}$. 联合分布律常用表格形式表示.

		Y 的可能取值		
		y_1	$\dots$	y_j
X	x_1	p_{11}	$\dots$	p_{1j}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	x_i	p_{i1}	$\dots$	p_{ij}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$P\{Y = y_j\}$		$p_{\cdot 1}$	$\dots$	$p_{\cdot j}$
		$P\{Y = y_1\}$		$P\{X = x_1\}$

利用全集分解 $P\{X = x_1\} = P\{X = x_1, \Omega\}$. 比如 Y 只能取 y_1 和 y_2 , 则

$$上式 = P\{X = x_1, (Y = y_1 \cup Y = y_2)\} = P\{X = x_1, Y = y_1\} + P\{X = x_1, Y = y_2\}$$

② 数列 $\{p_{ij}\}, i, j = 1, 2, \dots$ 是某一二维离散型随机变量的概率分布的充分必要条件为

$$p_{ij} \geq 0, \quad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1.$$

$\xrightarrow{\text{归一性}}$
 $\xrightarrow{\text{所有 } p_{ij} \text{ 相加}}$

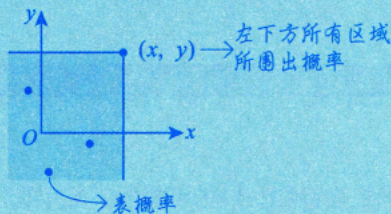
(2) 联合分布函数、边缘分布、条件分布.

① 联合分布函数.

设 (X, Y) 的概率分布为 p_{ij} , $i, j = 1, 2, \dots$, 则 (X, Y) 的分布函数或 X 和 Y 的联合分布函数为

$$F(x, y) = P\{X \leq x, Y \leq y\} = \sum_{x_i \leq x} \sum_{y_j \leq y} p_{ij}.$$

注 它是以 (x, y) 为顶点的左下方平面上 (X, Y) 取所有可能值的概率的和.

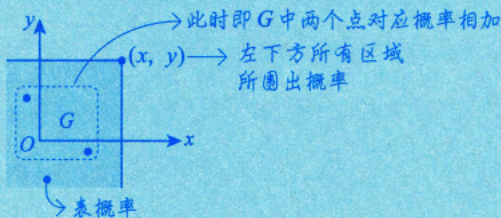


设 G 是平面上的某个区域, 则

$$P\{(X, Y) \in G\} = \sum_{(x_i, y_j) \in G} p_{ij}.$$

注 (X, Y) 落入 G 的概率等于 (X, Y) 在 G 内取所有可能值的概率的和.

离散型问题相对来讲不难, 考试为了有区分度, 出现的频率很低, 但要注意与连续型对比.



② 边缘分布.

X, Y 的边缘分布分别为

$$\left. \begin{aligned} p_{i \cdot} &= P\{X = x_i\} = \sum_{j=1}^{\infty} P\{X = x_i, Y = y_j\} = \sum_{j=1}^{\infty} p_{ij} \quad (i = 1, 2, \dots); \\ p_{\cdot j} &= P\{Y = y_j\} = \sum_{i=1}^{\infty} P\{X = x_i, Y = y_j\} = \sum_{i=1}^{\infty} p_{ij} \quad (j = 1, 2, \dots). \end{aligned} \right\} \text{结合表格计算}$$

固定 x_i , 把所有 y_j 的概率相加
 y_j 不动, 把所有 x_i 的概率相加

★③ 条件分布. $\rightarrow$ 不管是离散还是连续都经常考

如果 $(X, Y) \sim p_{ij} (i, j = 1, 2, \dots)$, 对固定的 j , 如果 $p_{\cdot j} = P\{Y = y_j\} > 0$, 则称

$\rightarrow$ 条件

$$P\{X=x_i | Y=y_j\} = \frac{P\{X=x_i, Y=y_j\}}{P\{Y=y_j\}} = \frac{p_{ij}}{p_{\cdot j}} \quad (i=1, 2, \dots)$$

看谁在帘子后边, 分母就是谁

为 X 在 “ $Y=y_j$ ” 条件下的条件分布.

同理, 对固定的 i , 如果 $p_{i\cdot} > 0$, 可定义 Y 在 “ $X=x_i$ ” 条件下的条件分布

$$P\{Y=y_j | X=x_i\} = \frac{p_{ij}}{p_{i\cdot}} \quad (j=1, 2, \dots).$$

→ 第1讲中条件概率的写法

条件分布 = $\frac{\text{联合分布}}{\text{边缘分布}}$;

联合分布 = 边缘分布 · 条件分布;
“三知二求一” 易考

$$P(A) > 0, P(B|A) = \frac{P(AB)}{P(A)}$$

$$P(AB) = P(A) \cdot P(B|A)$$

在二维表格中

注 二维离散型随机变量 $\sim p_{ij} \rightarrow$ 联合分布律.

$$\text{边缘分布} \begin{cases} X \sim p_{i\cdot} \text{ (把 } j \text{ 的所有情况相加),} \\ Y \sim p_{\cdot j} \text{ (把 } i \text{ 的所有情况相加).} \end{cases}$$

$$\text{条件分布} \begin{cases} P\{Y=y_j | X=x_i\} = \frac{p_{ij}}{p_{i\cdot}}, \\ P\{X=x_i | Y=y_j\} = \frac{p_{ij}}{p_{\cdot j}}. \end{cases}$$

2 二维连续型随机变量的概率密度、边缘概率密度和条件概率密度

(1) 概率密度.

① 概念.

如果二维随机变量 (X, Y) 的分布函数 $F(x, y)$ 可以表示为

$$F(x, y) = \int_{-\infty}^y dv \int_{-\infty}^x f(u, v) du, \quad (x, y) \in \mathbf{R}^2,$$

$= P\{X \leq x, Y \leq y\}$

$F(x) = \int_{-\infty}^x f(t) dt$

其中 $f(x, y)$ 是非负可积函数, 则称 (X, Y) 为二维连续型随机变量, 称 $f(x, y)$ 为 (X, Y) 的概率密度, 记为 $(X, Y) \sim f(x, y)$.

注 连续, 即点与点之间的距离是无穷小量.

② 二元函数 $f(x, y)$ 是概率密度的充分必要条件为

$$\underbrace{f(x, y) \geq 0}_{\text{非负}}, \underbrace{\int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} f(x, y) dx = 1}_{\text{在全平面积分结果是1}}.$$

注 改变 $f(x, y)$ 的有限个点的值 (仍取非负值), $f(x, y)$ 仍然是概率密度.

(2) 联合分布函数与概率密度、边缘概率密度、条件概率密度.

① 联合分布函数与概率密度.

设 (X, Y) 的分布函数为 $F(x, y)$, 概率密度为 $f(x, y)$, 则

a. $F(x, y)$ 为 (x, y) 的二元连续函数, 且

$$F(x, y) = P\{X \leq x, Y \leq y\} = \int_{-\infty}^y dv \int_{-\infty}^x f(u, v) du;$$

b. 设 G 为平面上某个区域, 则

$$P\{(X, Y) \in G\} = \iint_G f(x, y) dx dy;$$

→ 在某个区域上取值
体积对应概率

c. 若 $f(x, y)$ 在点 (x, y) 处连续, 则 $\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$;

以一维为例, $F'(x) = f(x)$. $F(x)$ 不一定处处可导, 但一定连续. 在不可导的地方给 $f(x)$ 赋值即可. 一般赋值为 0. 二维同理

d. 若 $F(x, y)$ 连续且可导, 则 (X, Y) 是连续型随机变量, 且 $\frac{\partial^2 F(x, y)}{\partial x \partial y}$ 是它的概率密度.

② 边缘概率密度.

设 $(X, Y) \sim f(x, y)$, 则 X 的边缘分布函数为

$$F_X(x) = F(x, +\infty) = \int_{-\infty}^x \left[\int_{-\infty}^{+\infty} f(u, v) dv \right] du, \rightarrow \int_{-\infty}^x g(u) du$$

$g(u) = \int_{-\infty}^{+\infty} f(u, v) dv$

所以 X 是连续型随机变量, 其概率密度为

→ 连续即离散得足够靠近, 故两者性质应类似 → x 不动, 累加 y

称 $f_X(x)$ 为 (X, Y) 关于 X 的边缘概率密度.

$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$, 曲面的面积

求谁不积谁

同理, Y 也是连续型随机变量, 其概率密度为 $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$.

③ 条件概率密度.

设 $(X, Y) \sim f(x, y)$, 边缘概率密度 $f_X(x) > 0$, 则称

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

为 Y 在 “ $X=x$ ” 条件下的条件概率密度.

$$\rightarrow \text{条件概率密度} = \frac{\text{联合概率密度}}{\text{边缘概率密度}}$$

同理, 可定义 X 在 “ $Y=y$ ” 条件下的条件概率密度

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} (f_Y(y) > 0).$$

由以上讨论可知, 若 $f_X(x) > 0$, $f_Y(y) > 0$, 则有概率密度乘法公式

$$f(x, y) = f_X(x)f_{Y|X}(y|x) = f_Y(y)f_{X|Y}(x|y).$$

注1 称

$$F_{Y|X}(y|x) = \int_{-\infty}^y f_{Y|X}(y|x) dy = \int_{-\infty}^y \frac{f(x, y)}{f_X(x)} dy$$

为 Y 在 “ $X=x$ ” 条件下的条件分布函数.

为了避免字母混淆, 可以写成

$$F_{Y|X}(y|x) = \int_{-\infty}^y f_{Y|X}(t|x) dt$$

同理, 可定义 X 在 “ $Y=y$ ” 条件下的条件分布函数

$$F_{X|Y}(x|y) = \int_{-\infty}^x f_{X|Y}(x|y) dx = \int_{-\infty}^x \frac{f(x, y)}{f_Y(y)} dx.$$

注2

	二维离散型随机变量	二维连续型随机变量
概率分布 / 概率密度	$(X, Y) \sim p_{ij}$	$(X, Y) \sim f(x, y)$
联合分布 $F(x, y)$	$\sum_{x_i \leq x} \sum_{y_j \leq y} p_{ij}$	$\int_{-\infty}^y dv \int_{-\infty}^x f(u, v) du$
边缘分布 $F_X(x), F_Y(y)$	$\begin{cases} X \sim p_{i\cdot} = \sum_{j=1}^{\infty} p_{ij}, \\ Y \sim p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij} \end{cases}$	$\begin{cases} X \sim F_X(x) = \int_{-\infty}^x f_X(u) du = \int_{-\infty}^x \int_{-\infty}^{+\infty} f(u, v) dv du, \\ Y \sim F_Y(y) = \int_{-\infty}^y f_Y(v) dv = \int_{-\infty}^y \int_{-\infty}^{+\infty} f(u, v) du dv \end{cases}$
条件分布 $F_{Y X}(y x), F_{X Y}(x y)$	$\begin{cases} P\{Y = y_j X = x_i\} = \frac{p_{ij}}{p_{i\cdot}}, \\ P\{X = x_i Y = y_j\} = \frac{p_{ij}}{p_{\cdot j}} \end{cases}$	$\begin{cases} F_{Y X}(y x) = \int_{-\infty}^y \frac{f(u, v)}{f_X(x)} dv, \\ F_{X Y}(x y) = \int_{-\infty}^x \frac{f(u, v)}{f_Y(y)} du \end{cases}$

(3) 常见的二维分布. → 《全国硕士研究生招生考试数学考试大纲》要求掌握两种

① 二维均匀分布.

如果 (X, Y) 的概率密度为

一维均匀分布: $X \sim f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & \text{其他} \end{cases}$ 区间长度的倒数

$$f(x, y) = \begin{cases} \frac{1}{S_D}, & (x, y) \in D, \\ 0, & \text{其他}, \end{cases}$$

$S_D \cdot h = 1 \Rightarrow h = \frac{1}{S_D}$
不画图直接算面积也可以

其中 S_D 为区域 D 的面积, 则称 (X, Y) 在平面有界区域 D 上服从均匀分布.

★★★ ② 二维正态分布. 记住以下结论即可, 不必看证明

如果 (X, Y) 的概率密度为

一维正态分布: $X \sim f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ → 记住! 考试会考

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\},$$

表示 X 与 Y 间的相关系数

其中 $\mu_1 \in \mathbf{R}, \mu_2 \in \mathbf{R}, \sigma_1 > 0, \sigma_2 > 0, -1 < \rho < 1$, 则称 (X, Y) 服从参数为 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$ 的二维正态分布, 记为 $(X, Y) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$. 此时有

5个参数

若 $\rho = 0, (X, Y) \sim f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]}$

如果 $\rho = 0$, 联合概率密度等于两个边缘概率密度的乘积. 是最简单的情况

$$= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} = f_X(x)f_Y(y)$$

a. 若 $(X_1, X_2) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$, 则

边缘分布一定是正态分布 $\rightarrow X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2)$.

→ 但两个正态分布联合未必是二维正态分布

b. 若 $X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2)$ 且 X_1, X_2 相互独立, 则

恰好说明:
联合分布 ≠ 边缘分布

相关性是指线性相关的程度. 不相关即不线性相关, 但可能有非线性关系, 故不相关未必独立

独立 $\Rightarrow$ 不相关, 故 $\rho = 0$ $(X_1, X_2) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; 0)$.

$\rightarrow X_2$ 取什么值, 不影响 X_1 取什么值的概率, 反之也是. 即 X_1 与 X_2 没有任何关系

c. $(X_1, X_2) \sim N \Rightarrow k_1 X_1 + k_2 X_2 \sim N(k_1, k_2 \text{ 是不全为 } 0 \text{ 的常数})$.

线性组合依然服从正态分布

此类题几乎不考大题, 如果考大题可学习本题的过程. 本题是基本问题, 在传统考研题中会出现, 第(1)问是随机事件与概率, 属于古典概型问题

例 3.1 将两封信投入3个信箱, 设 X_1, X_2 分别表示第一个和第二个信箱投进的信的数量, 求:

(1) (X_1, X_2) 的分布律、边缘分布;

(2) 在条件 $X_2 = 1$ 下, X_1 的条件分布.

① ② ③ 3个信箱

分析 不限定每个信箱放入信的数量, X_1 与 X_2 的可能取值为 0, 1, 2.

第一封信投入3个信箱有3种可能, 第二封信投入3个信箱也有3种可能

解 (1) X_1 与 X_2 可能取值为 0, 1, 2, 样本点总数为 $3^2 = 9$, 则

$$p_{00} = P\{X_1 = 0, X_2 = 0\} = \frac{1}{9},$$

第1个、第2个信箱没有信, 即信都在第3个信箱中

$$p_{01} = P\{X_1 = 0, X_2 = 1\} = \frac{2}{9},$$

第1个信箱中没有, 第2个信箱有1封, 第3个信箱有1封, $C_2^1 C_1^1$ 种可能

$$p_{02} = P\{X_1 = 0, X_2 = 2\} = \frac{1}{9},$$

$$p_{10} = P\{X_1 = 1, X_2 = 0\} = \frac{2}{9},$$

$$p_{11} = P\{X_1 = 1, X_2 = 1\} = \frac{2}{9},$$

$$p_{12} = P\{X_1 = 1, X_2 = 2\} = 0,$$

$$p_{20} = P\{X_1 = 2, X_2 = 0\} = \frac{1}{9},$$

$$p_{21} = p_{22} = 0,$$

于是 (X_1, X_2) 的分布律与边缘分布为

$X_1 \backslash X_2$	0	1	2		$p_{i\cdot}$
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	行相加 $\rightarrow$	$\frac{4}{9}$
1	$\frac{2}{9}$	$\frac{2}{9}$	0	$\rightarrow$	$\frac{4}{9}$
2	$\frac{1}{9}$	0	0	$\rightarrow$	$\frac{1}{9}$
$p_{\cdot j}$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$	列相加 $\downarrow$	1 $\rightarrow$ 归一性

(2)

$$P\{X_1=0|X_2=1\}=\frac{p_{01}}{p_{\cdot 1}}=\frac{\frac{2}{9}}{\frac{4}{9}}=\frac{1}{2},$$

按数字标

没有独立性前提，在一个事件发生的条件下，会影响另一个事件的概率

$$P\{X_1=1|X_2=1\}=\frac{p_{11}}{p_{\cdot 1}}=\frac{\frac{2}{9}}{\frac{4}{9}}=\frac{1}{2},$$

$$P\{X_1=2|X_2=1\}=0,$$

 所以在条件 $X_2=1$ 下， X_1 的条件分布为

$$X_1 \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

例 3.2 二维随机变量 (X, Y) 的概率分布为

$X \backslash Y$		0	1
0		0.4	a
1		b	0.1

 若随机事件 $\{X=0\}$ 与 $\{X+Y=1\}$ 相互独立，则 ().

 (A) $a=0.2, b=0.3$

 (B) $a=0.1, b=0.4$

 (C) $a=0.3, b=0.2$

 (D) $a=0.4, b=0.1$
解 应选 (D).

由联合分布、边缘分布及条件分布的性质，有

$$P\{X=0\}=p_{00}+p_{01}=0.4+a,$$

$$P\{X+Y=1\}=p_{01}+p_{10}=a+b,$$

$$p_{00}+p_{01}+p_{10}+p_{11}=a+b+0.5=1 \Rightarrow a+b=0.5,$$

归一性

$$P\{X=0, X+Y=1\}=P\{X=0, Y=1\}=a.$$

 又 $\{X=0\}$ 与 $\{X+Y=1\}$ 相互独立，故

$$\begin{aligned} P\{X=0, X+Y=1\} &= P\{X=0\} \cdot P\{X+Y=1\} \\ &= (0.4+a)(a+b) \\ &= 0.5(0.4+a). \end{aligned}$$

 因此 $0.5(0.4+a)=a$ ，解得 $a=0.4, b=0.5-0.4=0.1$ ，故选择 (D).

例 3.3 已知二维随机变量 (X, Y) 在 G 上服从均匀分布, G 由直线 $x-y=0$, $x+y=2$ 与 $y=0$ 围成, 求:

→ 连续型

(1) 边缘概率密度 $f_X(x)$ 和 $f_Y(y)$;

(2) 条件概率密度 $f_{X|Y}(x|y)$.

分析 先画图, 求面积, 再写出联合概率密度. 由联合概率密度求边缘概率密度几乎必考, 很重要.

解 (X, Y) 的正概率密度区域如图 3-1 所示.

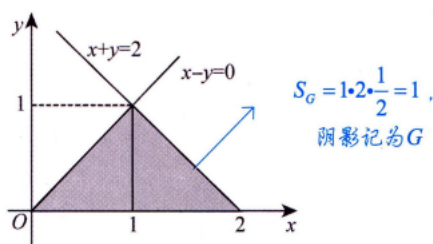


图 3-1

由题设, (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 1, & 0 \leq y \leq 1, y \leq x \leq 2-y, \\ 0, & \text{其他.} \end{cases}$$

(1) 求边缘概率密度 $f_X(x)$:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \begin{cases} \int_0^x 1 dy, & 0 < x < 1, \\ \int_0^{2-x} 1 dy, & 1 < x < 2, \\ 0, & \text{其他} \end{cases}$$

不积先定限

$$= \begin{cases} x, & 0 < x < 1, \\ 2-x, & 1 < x < 2, \\ 0, & \text{其他,} \end{cases} \quad \text{或} = \begin{cases} x, & 0 \leq x < 1, \\ 2-x, & 1 \leq x \leq 2, \\ 0, & \text{其他} \end{cases}$$

求边缘概率密度口诀: 求谁不积谁, 不积先定限, 限内画条线, 先交写下限, 后交写上限. 等号归哪里都可以.

求谁不积谁

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \begin{cases} \int_y^{2-y} 1 dx, & 0 < y < 1, \\ 0, & \text{其他} \end{cases} = \begin{cases} 2-2y, & 0 < y < 1, \\ 0, & \text{其他.} \end{cases}$$

不积先定限

(2) 在 $0 < y < 1$ 的条件下, 条件概率密度

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{1}{2-2y}, y \leq x \leq 2-y,$$

条件概率密度 = $\frac{\text{联合概率密度}}{\text{边缘概率密度}}$ → 帘子后边的

所以在 $Y=y(0 < y < 1)$ 的条件下, X 的条件概率密度为 $f_{X|Y}(x|y) = \begin{cases} \frac{1}{2-2y}, & y \leq x \leq 2-y, \\ 0, & \text{其他.} \end{cases}$ → x 的范围

无条件分布 ←

0到1内任意实数

例 3.4 已知随机变量 X 在区间 $[0, 1]$ 上服从均匀分布, 在 $X=x(0 < x < 1)$ 的条件下, 随机变量 Y 在区间 $[0, x]$ 上服从均匀分布, 求: 一般写 $(0, 1)$

(1) (X, Y) 的概率密度;

(2) Y 的概率密度;

(3) 概率 $P\{X+Y>1\}$. → 很重要, 典型问题

解 (1) 由题设, X 的边缘概率密度为

$$f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{其他.} \end{cases}$$

在 $X=x(0 < x < 1)$ 的条件下, 随机变量 Y 的条件概率密度为

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{x}, & 0 < y < x < 1, \\ 0, & \text{其他.} \end{cases}$$

← 常数
相同的表达式含义不同

因此, 当 $0 < y < x < 1$ 时, $f(x, y) = f_X(x)f_{Y|X}(y|x) = \frac{1}{x}$, 故 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} \frac{1}{x}, & 0 < y < x < 1, \\ 0, & \text{其他.} \end{cases}$$

↑ 统一 ↑ ↑ 变量

← 二元

(2) (X, Y) 的正概率密度区域如图 3-2 所示.

$$\text{当 } 0 < y < 1 \text{ 时, } f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_y^1 \frac{1}{x} dx = -\ln y;$$

$$\text{当 } y \leq 0 \text{ 或 } y \geq 1 \text{ 时, } f_Y(y) = 0.$$

从而有

$$f_Y(y) = \begin{cases} -\ln y, & 0 < y < 1, \\ 0, & \text{其他.} \end{cases}$$

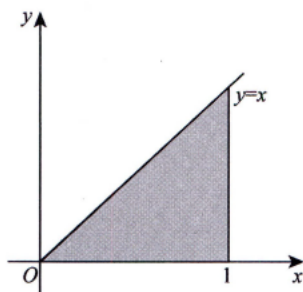


图 3-2

★★★ (3) $P\{X+Y>1\} = \iint_{x+y>1} f(x, y) dx dy = \int_{\frac{1}{2}}^1 dx \int_{1-x}^x \frac{1}{x} dy = \int_{\frac{1}{2}}^1 (2x-1) \cdot \frac{1}{x} dx = 1 - \ln 2.$

注 (1) 第 (1) 问中, 当 $f_X(x) > 0$ 时, 才有 $f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$, 所以 $f(x, y) = f_X(x) \cdot f_{Y|X}(y|x)$

是在 $f_X(x) > 0$ 时的表达式, 故最终结果要加上 $f_X(x) = 0$ 时的定义, 一般说来, 会归类到分段表达式的“其他”中, 这样就定义全面了.

(2) 连续型: $(X, Y) \sim f(x, y).$

$$\text{边缘概率密度} \begin{cases} f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy, \\ f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx. \end{cases}$$

$$\text{条件概率密度} \begin{cases} f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, \\ f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}. \end{cases}$$

→ 一个事件的发生, 不会促进, 也不会抑止另一个事件的发生

三 随机变量的相互独立性



1 概念

(1) 设二维随机变量 (X, Y) 的分布函数为 $F(x, y)$, 边缘分布函数分别为 $F_X(x)$, $F_Y(y)$, 如果对任意的实数 x, y , 都有

→ 联合分布函数 = 两边缘分布函数乘积

充要条件

$$F(x, y) = F_X(x) \cdot F_Y(y) \quad (\text{即事件 } \{X \leq x\} \text{ 与 } \{Y \leq y\} \text{ 相互独立}),$$

则称 X 与 Y 相互独立, 否则称 X 与 Y 不相互独立.

(2) 如果 n 维随机变量 $(X_1, X_2, \dots, X_n)$ 的分布函数等于边缘分布函数的乘积, 即

$$F(x_1, x_2, \dots, x_n) = F_1(x_1) \cdot F_2(x_2) \cdot \dots \cdot F_n(x_n),$$

其中 $F_i(x_i) (i=1, 2, \dots, n)$ 为 X_i 的边缘分布函数, x_i 为任意实数, 则称 $X_1, X_2, \dots, X_n$ 相互独立.

→ 以 $n=3$ 为例, $F(x_1, x_2, x_3) = F_1(x_1) \cdot F_2(x_2) \cdot F_3(x_3)$ ①

$F_1(x_1) = P\{X_1 \leq x_1\}$, x_1 是可以任意取值的.

当 $x_1 \rightarrow +\infty$, $1 = P\{\Omega\}$, 此时

$$F(x_1, x_2, x_3) = F(x_2, x_3) = F_2(x_2) \cdot F_3(x_3).$$

$$F(x_1, x_2) = P\{X_1 \leq x_1, X_2 \leq x_2\};$$

$$F_1(x_1) = P\{X_1 \leq x_1\};$$

$$F_1(x_1) \geq F(x_1, x_2). \quad \text{②}$$

这个式子其实包含了任意的两个变量独立的情况, 不需要像事件的独立性, 列多个式子

①, ②的不等关系和蕴含关系, 可以取特殊值, 或把定义写出来之后, 用不定式、单调性对式子进行处理

(3) 如果对任意实数 $x_i (i=1, 2, \dots, n)$ 与 $y_j (j=1, 2, \dots, m)$, 有

分别满足了两个事件的积再取交

$$\begin{aligned} P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n, Y_1 \leq y_1, Y_2 \leq y_2, \dots, Y_m \leq y_m\} \\ = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} \cdot P\{Y_1 \leq y_1, Y_2 \leq y_2, \dots, Y_m \leq y_m\}, \end{aligned}$$

即联合分布函数等于各自的分布函数相乘:

$$F(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = F_1(x_1, x_2, \dots, x_n) \cdot F_2(y_1, y_2, \dots, y_m),$$

则两个多维随机变量 $(X_1, X_2, \dots, X_n)$ 与 $(Y_1, Y_2, \dots, Y_m)$ 相互独立. 理论上了解, 不必验证

2 相互独立的充要条件

(1) n 个随机变量 $X_1, X_2, \dots, X_n$ 相互独立

$\Leftrightarrow$ 对任意的 n 个实数 $x_i (i=1, 2, \dots, n)$, n 个事件 $\{X_1 \leq x_1\}, \{X_2 \leq x_2\}, \dots, \{X_n \leq x_n\}$ 相互独立.

具有任意性, $x_i \in (-\infty, +\infty)$

★★★ (2) ① 设 (X, Y) 为二维离散型随机变量, 则

虽然简单, 但经常考 X 与 Y 相互独立 $\Leftrightarrow$ 联合分布律等于边缘分布相乘,

如可在例3.1中加(3) X_1, X_2 是否独立?

即

由于 $p_{00} \neq p_{0\cdot} p_{\cdot 0}$, 故 X_1 与 X_2 不独立.

$$P\{X = x_i, Y = y_j\} = P\{X = x_i\} \cdot P\{Y = y_j\} (i, j = 1, 2, \dots)$$

验证的方法很多, 找一个就行

实际考题一般只考二维

② n 个离散型随机变量 $X_1, X_2, \dots, X_n$ 相互独立

$\Leftrightarrow$ 对任意的 $x_i \in D_i = \{X_i \text{ 的一切可能值} \} (i=1, 2, \dots, n)$, 有

$$P\{X_1 = x_1, \dots, X_n = x_n\} = \prod_{i=1}^n P\{X_i = x_i\}.$$

★★★ (3) ① 设 (X, Y) 为二维连续型随机变量, 则

X 与 Y 相互独立 $\Leftrightarrow$ 概率密度等于边缘概率密度相乘,

即

几乎必考

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

② 设 $(X_1, X_2, \dots, X_n)$ 为 n 维连续型随机变量, 则

$X_1, X_2, \dots, X_n$ 相互独立 $\Leftrightarrow$ 概率密度等于边缘概率密度相乘,

即

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) \cdot f_2(x_2) \cdots f_n(x_n),$$

其中 $f_i(x_i)$ 为 X_i 的边缘概率密度.

3 相互独立的性质

类似线性代数中的向量组, 整体向量组线性无关, 则部分向量组线性无关

(1) 设 $X_1, X_2, \dots, X_n$ 相互独立, 则其中任意 $k (2 \leq k \leq n)$ 个随机变量也相互独立.

↑ 整体

↑ 部分

(2) ① 设 (X, Y) 为二维离散型随机变量, X 与 Y 独立, 则条件分布等于边缘分布:

$$P\{X = x_i | Y = y_j\} = P\{X = x_i\} (P\{Y = y_j\} > 0),$$

类似 $P(B|A) = \frac{P(B)}{P(A)} > 0$

Y 的取值不影响 X 的概率

前提

第1讲事件概率中推导的基本性质

$P\{Y=y_j | X=x_i\} = P\{Y=y_j\}(P\{X=x_i\} > 0)$.
 X 的取值不影响 Y 的概率

② 设 (X, Y) 为二维连续型随机变量, X 与 Y 独立, 则条件概率密度等于边缘概率密度:

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = f_X(x) \quad (f_Y(y) > 0),$$

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = f_Y(y) \quad (f_X(x) > 0).$$

前提
同理 $f_X(x)$ 约掉

③ 若 $X_1, X_2, \dots, X_n$ 相互独立, $g_1(x), g_2(x), \dots, g_n(x)$ 为一元连续函数, 则 $g_1(X_1), g_2(X_2), \dots, g_n(X_n)$ 相互独立.

分组(分成若干组, 变量不重叠)

一般地, 若 $X_{11}, \dots, X_{l_1}, X_{21}, \dots, X_{2l_2}, \dots, X_{n1}, \dots, X_{nl_n}$ 相互独立, g_i 是 $t_i (i=1, 2, \dots, n)$ 元连续函数, 则 $g_1(X_{11}, \dots, X_{l_1}), g_2(X_{21}, \dots, X_{2l_2}), \dots, g_n(X_{n1}, \dots, X_{nl_n})$ 也相互独立.

注 (1) 所有涉及的 n 维, 只是为了数理统计中应用这些原理、性质, 只需掌握二维中的性质即可.

(2) 二维随机变量 (X, Y) .

独立 $\Leftrightarrow F(x, y) = F_X(x) \cdot F_Y(y)$, 任意 x, y

离散型
 $\Leftrightarrow p_{ij} = p_{i \cdot} \cdot p_{\cdot j}$, 任意 i, j

连续型
 $\Leftrightarrow f(x, y) = f_X(x) \cdot f_Y(y)$.

考试重点

★★★ 4 X 与 Y 不独立的判断与证明

X 与 Y 不独立 $\Leftrightarrow$ 存在 x_0, y_0 , 使 $A = \{X \leq x_0\}$ 与 $B = \{Y \leq y_0\}$ 不独立, 即

从分布函数判断也常考 $F(x_0, y_0) \neq F_X(x_0) \cdot F_Y(y_0)$.

比如前边补充的例3.1(3)

也即取合适的 x_0, y_0 , 使 $P\{X \leq x_0\} \cdot P\{Y \leq y_0\} \neq P\{X \leq x_0, Y \leq y_0\}$. 如: 取 x_0, y_0 , 使 $0 < P\{X \leq x_0\} < 1$, $0 < P\{Y \leq y_0\} < 1$, 但 $\{X \leq x_0\} \subset \{Y \leq y_0\}$ 或 $\{Y \leq y_0\} \subset \{X \leq x_0\}$ 或 $\{X \leq x_0, Y \leq y_0\} = \emptyset$ 等.

三种取法应掌握. 应用见例3.6

注 正向会推, 反向也需要会写严格的证明过程.

例 3.5 设随机变量 $X \sim N(0, 1)$, 在 $X=x$ 的条件下, 随机变量 $Y \sim N(x, 1)$, 则 (X, Y) 的概率密度 $f(x, y) = \underline{\hspace{2cm}}$.

解 应填 $\frac{1}{2\pi} e^{-\frac{2x^2+y^2-2xy}{2}}$.

由题意知 $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}},$

故 $f(x, y) = f_X(x) \cdot f_{Y|X}(y|x) = \frac{1}{2\pi} e^{-\frac{2x^2+y^2-2xy}{2}}.$

注 考题中不会单独考查这一点, 会结合其他知识点一起命题.

→ 很重要, 但真正考试不会单独作为一个大题, 一般会作为选择题

★★★ **例 3.6** 设随机变量 X 的概率密度为

$$f(x) = \frac{1}{2} e^{-|x|}, x \in (-\infty, +\infty),$$

证明: X 与 $|X|$ 不独立.

分析 首先需要想到什么是独立、什么是不独立, 再看题目已给出 X 的概率密度, 需要找 $|X|$ 的概率密度,

可以根据前边学习的知识, 求 $|X|$ 的分布函数, 也可以求联合概率密度. 具体选什么方法, 需要结合具体问题确定,

本题最优解法是根据定义找 x_0 , 使 $P\{X \leq x_0, |X| \leq x_0\} \neq P\{X \leq x_0\} \cdot P\{|X| \leq x_0\}$.

证 取 $x_0 = 1$, 则

$$P\{X \leq 1\} = \frac{1}{2} \int_{-\infty}^1 e^{-|x|} dx = \frac{1}{2} \int_{-\infty}^0 e^x dx + \frac{1}{2} \int_0^1 e^{-x} dx = 1 - \frac{1}{2} e^{-1},$$

$$P\{|X| \leq 1\} = \frac{1}{2} \int_{-1}^1 e^{-|x|} dx = \int_0^1 e^{-x} dx = 1 - e^{-1}.$$

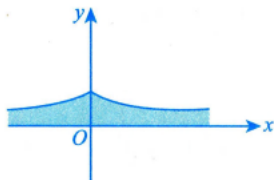
又事件 $\{|X| \leq 1\} \subset \{X \leq 1\}$, 即 $\{|X| \leq 1\} \cap \{X \leq 1\} = \{|X| \leq 1\}$, 故有

$$P(A \cap B) = P(A) \rightarrow P\{X \leq 1, |X| \leq 1\} = P\{|X| \leq 1\} = 1 - e^{-1},$$

从而有

$$A = \{X \leq 1\} \subset \{|X| \leq 1\} = B$$

$$P\{X \leq 1, |X| \leq 1\} \neq P\{X \leq 1\} \cdot P\{|X| \leq 1\}, \leftarrow \text{代入, 显然不等}$$



积分收敛的部分面积为1,

最好记住, 即 $\int_{-\infty}^0 e^x dx = 1$

因此, X 与 $|X|$ 不独立.

注 本题中 x_0 可取任意正数.

作为选择题时, 可以口算

$$P\{X \leq 1, |X| \leq 1\} = P\{|X| \leq 1\},$$

而 $P\{X \leq 1\} \cdot P\{|X| \leq 1\}$ 显然是 $P\{|X| \leq 1\}$ 乘上一个小于 1 的正数. 故必然有

$$P\{X \leq 1, |X| \leq 1\} \neq P\{X \leq 1\} \cdot P\{|X| \leq 1\}.$$

基础阶段先学好基础知识, 强化阶段慢慢总结做题技巧.

★★★ 四 多维随机变量函数的分布

随机事件和概率, 没有统一化的研究对象

对分布函数的考查达到顶峰, 既是知识的一脉相承, 也是对前边所有知识的总结. 此部分内容需要考生好好掌握



1 概念

设 X, Y 为随机变量, $g(x, y)$ 是二元函数, 则以随机变量 X, Y 作为变量的函数 $Z = g(X, Y)$ 也是随机变量, 称为随机变量 X, Y 的函数. 例如: $Z = X + Y, Z = \begin{cases} 1, & X < Y \\ 0, & X \geq Y \end{cases}$ 等. 输入 (X, Y) , 输出一个 Z 值

加减乘除.

有很好的实际背景, 很多问题都可以抽象成线性组合

常考问题: 已知 (X, Y) 的分布, 求 $Z = g(X, Y)$ 的分布; 又 $U = h(X, Y)$, 求 (Z, U) 的分布.

由于 Z 是一维的, 求 Z 的分布, 即求一维随机变量的分布

联合分布函数

2 求法

已知 (X, Y) 的分布, 求函数 $Z = g(X, Y)$ 的分布. 首先要确定 X, Y 的类型, 然后采用相应的公式计算.

(1) 如果 (X, Y) 是二维离散型随机变量, 则 $Z = g(X, Y)$ 也是离散型的, 先确定 Z 的值, 而后求其相应的概率, 即可求得 Z 的分布, 具体见例 3.7.

此类问题比较简单

★★★ (2) 如果 X, Y 其中一个离散型的, 另一个是非离散型的, 将事件对离散型的一切可能值进行全集分解, 而后应用全概率公式求得 Z 的分布, 具体见例 3.10.

很重要

考研重点考 $\begin{cases} \text{离散} \\ (X, Y) \sim p_{ij}, \\ \text{连续} \\ (X, Y) \sim f(x, y), \end{cases}$ 若 (X, Y) 是 $\begin{cases} \text{离散} \\ \text{连续} \end{cases}$, 便没有 p_{ij} , 也没有 $f(x, y)$, 问题会复杂一些

(3) 如果 (X, Y) 是二维连续型随机变量, 即 $(X, Y) \sim f(x, y)$, 则 $Z = g(X, Y)$ 的分布函数

$$F(z) = P\{g(X, Y) \leq z\} = \iint_{g(x, y) \leq z} f(x, y) dx dy,$$

其中 $-\infty < z < +\infty$, 讨论起来会复杂一些

要求概率密度 $f(z)$, 亦可直接用下面的“四、3.”中所讲公式, 具体见例 3.8, 例 3.9.

→ 求概率密度前一般先求分布函数 $F(z)$, 再求导得到 $f(z)$

$$\rightarrow F(z) = P\{Z \leq z\}$$

注 设 (X, Y) 是二维连续型随机变量, 且 $(X, Y) \sim f(x, y)$, 若 $\begin{cases} u = u(x, y), \\ v = v(x, y), \end{cases}$ 且 u, v 具有一阶连续偏

导数. $\begin{cases} x = x(u, v), \\ y = y(u, v) \end{cases}$ 是唯一反函数, 则 (X, Y) 的函数 $\begin{cases} U = U(X, Y), \\ V = V(X, Y) \end{cases}$ 的 $f_{U, V}(u, v) = f[x(u, v), y(u, v)] \cdot |J|$,
 已知 $f(x, y)$ 欲求的 f 雅可比行列式的绝对值

$$\text{其中 } J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \left[\frac{\partial(u, v)}{\partial(x, y)} \right]^{-1} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}^{-1} \neq 0, \text{ 那么}$$

$$\left. \begin{aligned} U \sim f_U(u) &= \int_{-\infty}^{+\infty} f[x(u, v), y(u, v)] |J| dv, \\ V \sim f_V(v) &= \int_{-\infty}^{+\infty} f[x(u, v), y(u, v)] |J| du. \end{aligned} \right\} \text{ 直接用这两个公式即可}$$

证

$$F_{U, V}(u, v) = P\{U \leq u, V \leq v\}$$

$$= P\{U(X, Y) \leq u, V(X, Y) \leq v\}$$

X, Y 的一个范围 $\leftarrow$ 即两个区域的交集 $\rightarrow X, Y$ 的另一个范围

$$= \iint_{D_{xy}} f(x, y) dx dy$$

$$\xrightarrow{\text{换元}} \begin{aligned} & \begin{matrix} x = x(u, v) \\ y = y(u, v) \end{matrix} \iint_{D_{uv}} f[x(u, v), y(u, v)] |J| du dv, \end{aligned}$$

换积分区域 $\leftarrow$ 换被积函数 $\rightarrow$ 换积分变量
 换雅可比行列式的绝对值, 表示从一个坐标系到另一个坐标系几何度量上的转换系数

$$\text{故 } f_{U, V}(u, v) = \frac{\partial^2 F_{U, V}(u, v)}{\partial u \partial v} = f[x(u, v), y(u, v)] |J|. \text{ 由 } U, V \text{ 的联合概率密度可得 } U, V \text{ 的边缘概率密度,}$$

如上所写.

在高等数学中, 我们已经熟练掌握了二重积分的换元积分法, 故在此给出换元法求 $f_{U, V}(u, v)$, 这是比卷积公式更为有用且极为简便的方法.

3 相互独立随机变量函数的分布及卷积公式

(1) 和的分布.

设 $(X, Y) \sim f(x, y)$, 则 $Z = X + Y$ 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx = \int_{-\infty}^{+\infty} f(z-y, y) dy$$

$$\stackrel{\text{独立}}{=} \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy.$$

(2) 差的分布.

设 $(X, Y) \sim f(x, y)$, 则 $Z = X - Y$ 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, x-z) dx = \int_{-\infty}^{+\infty} f(y+z, y) dy$$

$$\stackrel{\text{独立}}{=} \int_{-\infty}^{+\infty} f_X(x) f_Y(x-z) dx = \int_{-\infty}^{+\infty} f_X(y+z) f_Y(y) dy.$$

注1 令 $\begin{cases} u = x - y, \\ v = y, \end{cases}$ 则 $\begin{cases} x = u + v, \\ y = v, \end{cases}$ 雅可比行列式为 $J = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}$, 由

$$f_{u,v}(u, v) = f(u+v, v) \cdot 1,$$

得

$$f_u(u) = \int_{-\infty}^{+\infty} f_X(u+v) \cdot f_Y(v) dv.$$

注2 由此处可知, 卷积公式均可由换元法经简单计算得出, 无须再记卷积公式, 也省去了使用卷积公式后关于变量和区域的复杂讨论. 如 $X \sim f_X(x)$, $Y \sim f_Y(y)$, X, Y 相互独立. $Z = aX + bY (ab \neq 0)$.

$$\text{令 } \begin{cases} u = ax + by, \\ v = y, \end{cases} \text{ 则有 } \begin{cases} x = \frac{u-bv}{a}, \\ y = v. \end{cases} J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{a} & -\frac{b}{a} \\ 0 & 1 \end{vmatrix} = \frac{1}{a}.$$

$$\text{故 } f_{u,v}(u, v) = f\left(\frac{u-bv}{a}, v\right) \cdot \frac{1}{|a|} = f_X\left(\frac{u-bv}{a}\right) \cdot f_Y(v) \cdot \frac{1}{|a|}, \text{ 则 } f_u(u) = \int_{-\infty}^{+\infty} f_X\left(\frac{u-bv}{a}\right) f_Y(v) \cdot \frac{1}{|a|} dv,$$

将 u 换为 z , $f_z(z)$ 即为 $Z = aX + bY$ 的概率密度.

(3) 积的分布.

设 $(X, Y) \sim f(x, y)$, 则 $Z = XY$ 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{z}{x}\right) dx = \int_{-\infty}^{+\infty} \frac{1}{|y|} f\left(\frac{z}{y}, y\right) dy$$

$$\stackrel{\text{独立}}{=} \int_{-\infty}^{+\infty} \frac{1}{|x|} f_X(x) f_Y\left(\frac{z}{x}\right) dx = \int_{-\infty}^{+\infty} \frac{1}{|y|} f_X\left(\frac{z}{y}\right) f_Y(y) dy.$$

注 令 $\begin{cases} u=xy, \\ v=y, \end{cases}$ 则有 $\begin{cases} x=\frac{u}{v}, \\ y=v, \end{cases} J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{v} & -\frac{u}{v^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v} \neq 0.$

故 $f_{U,V}(u,v) = f\left(\frac{u}{v}, v\right) \cdot \frac{1}{|v|} \stackrel{\text{独立}}{=} f_X\left(\frac{u}{v}\right) \cdot f_Y(v) \cdot \frac{1}{|v|}$, 于是

$$f_U(u) = \int_{-\infty}^{+\infty} f_{U,V}(u,v) dv = \int_{-\infty}^{+\infty} f_X\left(\frac{u}{v}\right) \cdot f_Y(v) \cdot \frac{1}{|v|} dv,$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X\left(\frac{z}{v}\right) \cdot f_Y(v) \cdot \frac{1}{|v|} dv.$$

(4) 商的分布.

设 $(X, Y) \sim f(x, y)$, 则 $Z = \frac{X}{Y}$ 的概率密度为

$$f_Z(z) = \int_{-\infty}^{+\infty} |y| f(yz, y) dy \stackrel{\text{独立}}{=} \int_{-\infty}^{+\infty} |y| f_X(yz) f_Y(y) dy.$$

注 1 令 $\begin{cases} u=\frac{x}{y}, \\ v=y, \end{cases}$ 则 $\begin{cases} x=uv, \\ y=v, \end{cases}$ 雅可比行列式为 $J = \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} = v$, 由

$$f_{U,V}(u, v) = f(uv, v) \cdot |v|,$$

得

$$f_U(u) = \int_{-\infty}^{+\infty} f_X(uv) \cdot f_Y(v) |v| dv.$$

注 2 以上四组公式, 可用“口诀”来记忆: “积谁不换谁, 换完求偏导.”

如 $Z = X - Y$ 的 $f_Z(z) = \int_{-\infty}^{+\infty} f(x, x-z) dx$ 中.

先用第一句“积谁不换谁”. 对 x 积分, 则不换 x , 写成 $f(x, \square)$, 换 $y = x - z$, 为 $f(x, x - z)$.

再用第二句“换完求偏导”. $\frac{\partial(x-z)}{\partial z} = -1$, 因概率密度非负, 要加绝对值, 即为 $|-1| = 1$. 这样

便记住了公式.

会了二重积分换元,
将不再需要

(5) $\max\{X, Y\}$ 分布.

设 $(X, Y) \sim F(x, y)$, 则 $Z = \max\{X, Y\}$ 的分布函数为

$$F_{\max}(z) = P\{\max\{X, Y\} \leq z\} = P\{X \leq z, Y \leq z\} = F(z, z).$$

当 X 与 Y 独立时,

$$F_{\max}(z) = F_X(z) \cdot F_Y(z).$$

(6) $\min\{X, Y\}$ 分布.

设 $(X, Y) \sim F(x, y)$, 则 $Z = \min\{X, Y\}$ 的分布函数为

$$\begin{aligned} F_{\min}(z) &= P\{\min\{X, Y\} \leq z\} = P\{\{X \leq z\} \cup \{Y \leq z\}\} \\ &= P\{X \leq z\} + P\{Y \leq z\} - P\{X \leq z, Y \leq z\} \quad \text{加法公式} \\ &= F_X(z) + F_Y(z) - F(z, z). \end{aligned}$$

当 X 与 Y 独立时,

$$\begin{aligned} F_{\min}(z) &= F_X(z) + F_Y(z) - F_X(z)F_Y(z) \\ &= 1 - [1 - F_X(z)][1 - F_Y(z)]. \end{aligned}$$

$$\begin{aligned} F_{\min}(z) &= P\{\min\{X, Y\} \leq z\} \\ &= 1 - P\{\min\{X, Y\} > z\} \\ &= 1 - P\{X > z\}P\{Y > z\} \quad \text{生存函数} \\ &= 1 - [1 - F_X(z)][1 - F_Y(z)] \end{aligned}$$

一般不考, 或放数理统计里考

注 上述结果容易推广到 n 个相互独立的随机变量 $X_1, X_2, \dots, X_n$ 的情况, 即

$$F_{\max}(z) = F_{X_1}(z)F_{X_2}(z) \cdots F_{X_n}(z),$$

$$F_{\min}(z) = 1 - [1 - F_{X_1}(z)][1 - F_{X_2}(z)] \cdots [1 - F_{X_n}(z)].$$

特别地, 当 $X_i (i=1, 2, \dots, n)$ 相互独立且有相同的分布函数 $F(x)$ 与概率密度 $f(x)$ 时,

$$F_{\max}(x) = [F(x)]^n,$$

见例3.11应用

$$f_{\max}(x) = n[F(x)]^{n-1}f(x).$$

$$F_{\min}(x) = 1 - [1 - F(x)]^n,$$

$$f_{\min}(x) = n[1 - F(x)]^{n-1}f(x).$$

这些结果在数理统计部分是要用到的.

(7) 常见分布的可加性.

有些相互独立且服从同类型分布的随机变量, 其和的分布也是同类型的, 它们是二项分布、泊松分布、正态分布与 χ^2 分布.

设随机变量 X 与 Y 相互独立, 则

若 $X \sim B(n, p)$, $Y \sim B(m, p)$, 则 $X+Y \sim B(n+m, p)$ (注意 p 相同);

若 $X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$, 则 $X+Y \sim P(\lambda_1 + \lambda_2)$;

强度相加

若 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 则 $X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$;

若 $X \sim \chi^2(n)$, $Y \sim \chi^2(m)$, 则 $X+Y \sim \chi^2(n+m)$. → 第6讲再讲
自由度 自由度相加

注 上述结果对 n 个相互独立的随机变量也成立.

是对“四、2(1)”的应用.

(离散, 离散) → 离散

例 3.7 将两封信投入 3 个信箱, 设 X_1, X_2 分别表示第一个和第二个信箱投进的信的数量,

求随机变量 $Y_1 = X_1 + X_2$ 和 $Y_2 = X_1 - X_2$ 的分布. 分别对应“四、2(1)”中的 X 和 Y

分析 利用例 3.1 的联合分布律, 再计算 $Y_1 = X_1 + X_2$ 的所有可能取值: 0, 1, 2, 3, 4.

但概率为 0 的可以去掉, 即得答案上 Y_1 的分布律, Y_2 的分布律同理.

解 由例 3.1 知 (X_1, X_2) 的分布律, 则 $Y_1 = X_1 + X_2$ 的可能取值为 0, 1, 2, 于是

$$P\{Y_1=0\}=p_{00}=\frac{1}{9}, P\{Y_1=1\}=p_{01}+p_{10}=\frac{4}{9}, P\{Y_1=2\}=p_{02}+p_{11}+p_{20}=\frac{4}{9},$$

从而有

$$Y_1 \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{9} & \frac{4}{9} & \frac{4}{9} \end{pmatrix}.$$

又 $Y_2 = X_1 - X_2$ 的可能取值为 -2, -1, 0, 1, 2, 于是

$$P\{Y_2=-2\}=p_{02}=\frac{1}{9}, P\{Y_2=-1\}=p_{01}=\frac{2}{9}, P\{Y_2=0\}=p_{00}+p_{11}=\frac{1}{3},$$

$$P\{Y_2=1\}=p_{10}=\frac{2}{9}, P\{Y_2=2\}=p_{20}=\frac{1}{9},$$

从而有

$$Y_2 \sim \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ \frac{1}{9} & \frac{2}{9} & \frac{1}{3} & \frac{2}{9} & \frac{1}{9} \end{pmatrix}.$$

例 3.8 设随机变量 X, Y 相互独立, 它们的概率密度分别为

$$f_X(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & \text{其他}, \end{cases} f_Y(y) = \begin{cases} 2y, & 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

→ 引导大家走第(2)步

(1) 求 (X, Y) 的概率密度; 联合概率密度 = 边缘概率密度的乘积, 但现在考题不会这样送分

(2) 求 $Z = X + Y$ 的概率密度.

解 (1) 由于 X, Y 相互独立, 则 (X, Y) 的概率密度为

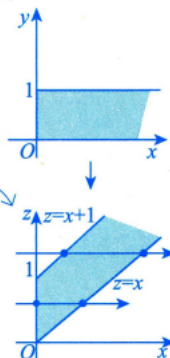
$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} 2ye^{-x}, & x > 0, 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

用卷积公式“三部曲”:

① 换字母. $x > 0, 0 < z - x < 1 \Rightarrow$

$x > 0, x < z < x + 1.$

② 换区域.



(2) 方法一 在 X, Y 相互独立的条件下, 用公式 $f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx$ 计算, 有

$$f_X(x)f_Y(z-x) = \begin{cases} 2(z-x)e^{-x}, & x > 0, 0 < z-x < 1, \\ 0, & \text{其他.} \end{cases}$$

当 $z < 0$ 时, $f_Z(z) = 0$;

当 $0 < z < 1$ 时, $f_Z(z) = \int_0^z 2(z-x)e^{-x} dx = 2z + 2e^{-z} - 2$;

当 $z > 1$ 时, $f_Z(z) = \int_{z-1}^z 2(z-x)e^{-x} dx = 2e^{-z}$.

③ 背口诀.
求 z 不积 z ,
不积先定限,
限内画条线,
先交写下限,
后交写上限

因此 $Z = X + Y$ 的概率密度为

$$f_Z(z) = \begin{cases} 2z + 2e^{-z} - 2, & 0 < z < 1, \\ 2e^{-z}, & z > 1, \\ 0, & \text{其他.} \end{cases}$$

方法二 先记 $U = X + Y \sim f_U(u)$, 令 $\begin{cases} u = x + y, \\ v = y. \end{cases}$ 换元法保证两边同维

保证二重积分换元, 二元到二元, 几元到几元, 元不够, 凑. 类似线性代数中二次型, 换元保证维数相同

一定有一阶连续偏导数, 且反函数是唯一的一个

$$\begin{cases} x = u - v, \\ y = v, \end{cases}$$

$$\text{再计算 } |J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$$

说明两个坐标系转换后, 测度相同

则 $f_{u,v}(u,v) = f(u-v,v) \cdot 1$, 故

$$f_U(u) = \int_{-\infty}^{+\infty} f(u-v, v) dv = \int_{-\infty}^{+\infty} f_X(u-v) \cdot f_Y(v) dv,$$

又因为 $\begin{cases} x > 0, \\ 0 < y < 1, \end{cases}$ 则 $\begin{cases} u-v > 0, \\ 0 < v < 1, \end{cases}$ 解得 $\begin{cases} u > v, \\ 0 < v < 1. \end{cases}$

故

$$f_U(u) = \begin{cases} \int_0^u e^{-(u-v)} \cdot 2v dv = 2u - 2 + 2e^{-u}, & 0 < u < 1, \\ 2e^{-u}, & u > 1, \\ 0, & \text{其他.} \end{cases}$$

将 u 换为 z 即为答案

$$2e^{-u} \int_0^u e^v \cdot v dv$$

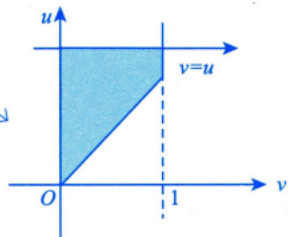
$$= 2e^{-u} \left[(ve^v - e^v) \Big|_0^u \right]$$

$$= 2e^{-u} \left[(v-1)e^v \Big|_0^u \right] = 2e^{-u} \left[(u-1)e^u + 1 \right]$$

$$= 2(u-1) + 2e^{-u}$$

即

$$f_Z(z) = \begin{cases} 2z + 2e^{-z} - 2, & 0 < z < 1, \\ 2e^{-z}, & z > 1, \\ 0, & \text{其他.} \end{cases}$$



注 本题亦可用分布函数法. (X, Y) 的正概率区域 D 与所求概率 $F_Z(z) = P\{X+Y \leq z\}$ 的积分区域的公共部分有三种不同的组合形式 (见图 3-3), 于是:

当 $z < 0$ 时, $F_Z(z) = 0$;

当 $0 \leq z < 1$ 时, $F_Z(z) = \int_0^z dx \int_0^{z-x} 2ye^{-x} dy = z^2 - 2z - 2e^{-z} + 2$;

当 $z \geq 1$ 时, $F_Z(z) = 1 - \int_0^1 dy \int_{z-y}^{+\infty} 2ye^{-x} dx = 1 - 2e^{-z}$.

因此 $Z = X + Y$ 的概率密度为

$$f_Z(z) = \begin{cases} 2z + 2e^{-z} - 2, & 0 < z < 1, \\ 2e^{-z}, & z > 1, \\ 0, & \text{其他.} \end{cases}$$

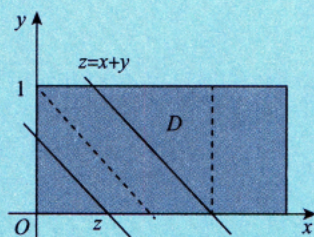


图 3-3

例 3.9 设二维随机变量 (X, Y) 在矩形区域 $D = \{(x, y) | 0 \leq x \leq 2, 0 \leq y \leq 1\}$ 上服从均匀分布, 求边长为 X 和 Y 的矩形面积 Z 的概率密度.

解 方法一 令 $\begin{cases} u = xy, \\ v = y, \end{cases}$ 则 $\begin{cases} x = \frac{u}{v}, \\ y = v, \end{cases}$ 则雅可比行列式为

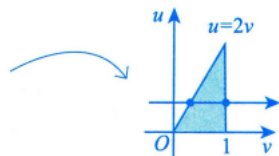
$$J = \begin{vmatrix} \frac{1}{v} & -\frac{u}{v^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v},$$

则

$$f(u, v) = \frac{1}{2} \cdot \frac{1}{|v|} = \frac{1}{2v},$$

所以

$$f_u(u) = \int_{-\infty}^{+\infty} \frac{1}{2v} dv = \int_{\frac{u}{2}}^1 \frac{1}{2v} dv = \frac{1}{2} \ln v \Big|_{\frac{u}{2}}^1 = \frac{1}{2} (\ln 2 - \ln u), 0 < u < 2,$$



即

$$f_Z(z) = \begin{cases} \frac{1}{2} (\ln 2 - \ln z), & 0 < z < 2, \\ 0, & \text{其他.} \end{cases}$$

方法二 由题设 $Z = XY$, (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} \frac{1}{S_D}, & 0 \leq x \leq 2, 0 \leq y \leq 1, \\ 0, & \text{其他.} \end{cases}$$

用卷积公式

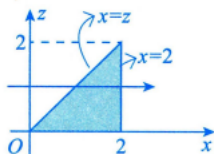
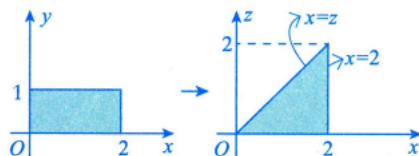
$$f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f\left(x, \frac{z}{x}\right) dx, \quad 0 \leq z \leq x \leq 2.$$

用卷积公式“三部曲”:

①换字母.

$$0 < x \leq 2, 0 \leq \frac{z}{x} \leq 1 \Rightarrow 0 < x \leq 2, 0 \leq z \leq x.$$

②换区域.



当 $z \leq 0$ 或 $z \geq 2$ 时, $f_Z(z) = 0$;

当 $0 < z < 2$ 时,

$$f_Z(z) = \frac{1}{2} \int_z^2 \frac{1}{x} dx = \frac{1}{2} (\ln 2 - \ln z).$$

③ 背口诀.
求 z 不积 z ,
不积先定限,
限内画条线,
先交写下限,
后交写上限

因此 $Z = XY$ 的概率密度为

$$f_Z(z) = \begin{cases} \frac{1}{2} (\ln 2 - \ln z), & 0 < z < 2, \\ 0, & \text{其他.} \end{cases}$$

注 本题亦可用分布函数法.

正概率密度区域 D 与所求概率 $F_Z(z) = P\{XY \leq z\}$ 的积分区域的公共部分有三种不同的组合形式 (见图 3-4), 于是

当 $z \leq 0$ 时, $F_Z(z) = 0$;

当 $0 < z < 2$ 时,

$$\begin{aligned} F_Z(z) &= \int_0^z dx \int_0^1 \frac{1}{2} dy + \int_z^2 dx \int_0^{\frac{z}{x}} \frac{1}{2} dy \\ &= \frac{1}{2} z (1 - \ln z + \ln 2); \end{aligned}$$

当 $z \geq 2$ 时, $F_Z(z) = 1$.

因此 $Z = XY$ 的概率密度为

$$f_Z(z) = \begin{cases} \frac{1}{2} (\ln 2 - \ln z), & 0 < z < 2, \\ 0, & \text{其他.} \end{cases}$$

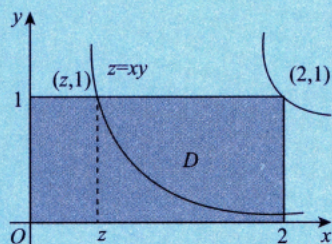


图 3-4

→ 此类型为考研最后一个大题, 但不难

例 3.10 设随机变量 X_1, X_2, X_3 相互独立, 其中 X_1 与 X_2 均服从标准正态分布, X_3 的概率

分布为 $P\{X_3 = 0\} = P\{X_3 = 1\} = \frac{1}{2}$, $Y = X_3 X_1 + (1 - X_3) X_2$. $\rightarrow \begin{cases} X_3 = 0 \text{ 时, } Y = X_2, \\ X_3 = 1 \text{ 时, } Y = X_1, \end{cases}$
所以 Y 有一半的概率取 X_1 ,
有一半的概率取 X_2 .

→ 离散的两点分布, 即 0-1 分布

(1) 求二维随机变量 (X_1, Y) 的分布函数, 结果用标准正态分布函数 $\Phi(x)$ 表示;

(2) 证明随机变量 Y 服从标准正态分布.

(1) **解** 记 (X_1, Y) 的分布函数为 $F(x, y)$, 则对任意实数 x 和 y , 有

$$\begin{aligned} F(x, y) &= P\{X_1 \leq x, Y \leq y\} \\ &= P\{X_1 \leq x, X_3 X_1 + (1 - X_3) X_2 \leq y\} \\ &= P\{X_1 \leq x, X_3 X_1 + (1 - X_3) X_2 \leq y, (X_3 = 0 \cup X_3 = 1)\} \end{aligned}$$

分解
完备事件组

$$\begin{aligned}
 &= P\{X_3=0\}P\{X_1 \leq x, X_3X_1+(1-X_3)X_2 \leq y | X_3=0\} + \\
 &\quad P\{X_3=1\}P\{X_1 \leq x, X_3X_1+(1-X_3)X_2 \leq y | X_3=1\} \quad \left. \vphantom{\begin{aligned} &= P\{X_3=0\}P\{X_1 \leq x, X_3X_1+(1-X_3)X_2 \leq y | X_3=0\} + \\ &\quad P\{X_3=1\}P\{X_1 \leq x, X_3X_1+(1-X_3)X_2 \leq y | X_3=1\} \end{aligned}} \right\} \text{全集分解} \\
 &= \frac{1}{2}P\{X_1 \leq x, X_2 \leq y | X_3=0\} + \frac{1}{2}P\{X_1 \leq x, X_1 \leq y | X_3=1\} \\
 &= \frac{1}{2}P\{X_1 \leq x, X_2 \leq y\} + \frac{1}{2}P\{X_1 \leq x, X_1 \leq y\} \quad \begin{array}{l} \text{独立性} \\ \text{独立性} \end{array} \\
 &= \frac{1}{2}P\{X_1 \leq x, X_2 \leq y\} + \frac{1}{2}P\{X_1 \leq \min\{x, y\}\} \quad \text{取交集} \\
 &= \frac{1}{2}\Phi(x)\Phi(y) + \frac{1}{2}\Phi(\min\{x, y\}) \quad \text{独立}
 \end{aligned}$$

(2) **证** 由(1)知, Y 的分布函数为

$$\begin{aligned}
 F_Y(y) &= \lim_{x \rightarrow +\infty} F(x, y) \quad \text{找边缘, 取极限} \\
 &= \lim_{x \rightarrow +\infty} \left[\frac{1}{2}\Phi(x)\Phi(y) + \frac{1}{2}\Phi(\min\{x, y\}) \right] \\
 &= \frac{1}{2}\Phi(y) + \frac{1}{2}\Phi(y) = \Phi(y),
 \end{aligned}$$

所以 Y 服从标准正态分布.

注 本题较简单, 考研题不会考这么单一, 但核心都是全集分解思想, 热点题型.

例 3.11 设随机变量 X, Y 相互独立, 且都在区间 $[a, b]$ 上服从均匀分布, 求 $Z = \max\{X, Y\}$ 的概率密度.

解 方法一 由题设, 有

$$\begin{aligned}
 f(x) &= \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{其他,} \end{cases} \\
 F(x) &= \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x < b, \\ 1, & x \geq b, \end{cases}
 \end{aligned}$$

则

故

$$f_z(z) = f_{\max}(z) = 2F(z)f(z) = \begin{cases} \frac{2(z-a)}{(b-a)^2}, & a \leq z \leq b, \\ 0, & \text{其他.} \end{cases}$$

方法二 由题意可知 X, Y 相互独立, 且

→很重要, 会写公式, 把握解题方法

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{其他}, \end{cases}$$

$$f(y) = \begin{cases} \frac{1}{b-a}, & a \leq y \leq b, \\ 0, & \text{其他}, \end{cases}$$

所以 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} \frac{1}{(b-a)^2}, & a \leq x \leq b, a \leq y \leq b, \\ 0, & \text{其他}. \end{cases}$$

由 (X, Y) 的取值区域 $\{(x, y) | a \leq x \leq b, a \leq y \leq b\}$, 可知 $Z = \max\{X, Y\}$ 的取值区间为 $[a, b]$.

当 $z < a$ 或 $z > b$ 时, 有 $f_z(z) = 0$.

当 $a \leq z \leq b$ 时, 有

$$F_z(z) = P\{Z \leq z\} = P\{\max\{X, Y\} \leq z\} = P\{X \leq z, Y \leq z\}$$

$$\begin{aligned} &= \iint_{x \leq z, y \leq z} f(x, y) dx dy = \int_a^z dx \int_a^z \frac{1}{(b-a)^2} dy \\ &= \frac{(z-a)^2}{(b-a)^2}, \end{aligned}$$

所以

$$f_z(z) = F'_z(z) = \frac{2(z-a)}{(b-a)^2}.$$

综上, $Z = \max\{X, Y\}$ 的概率密度为

$$f_z(z) = \begin{cases} \frac{2(z-a)}{(b-a)^2}, & a \leq z \leq b, \\ 0, & \text{其他}. \end{cases}$$

五 综合题举例

例 3.12 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} \frac{2}{\pi}(x^2 + y^2), & x^2 + y^2 \leq 1, \\ 0, & \text{其他}. \end{cases}$$

- (1) 求 X 与 Y 的协方差;
- (2) X 与 Y 是否相互独立?



(3) 求 $Z = X^2 + Y^2$ 的概率密度.

解 (1) 因为

$$EX = \iint_{x^2+y^2 \leq 1} x \cdot \frac{2}{\pi} (x^2 + y^2) dx dy = 0,$$

$$EY = \iint_{x^2+y^2 \leq 1} y \cdot \frac{2}{\pi} (x^2 + y^2) dx dy = 0,$$

$$E(XY) = \iint_{x^2+y^2 \leq 1} xy \cdot \frac{2}{\pi} (x^2 + y^2) dx dy = 0,$$

所以 X 与 Y 的协方差为

$$\text{Cov}(X, Y) = E(XY) - EXEY = 0.$$

(2) 当 $x \geq 1$ 或 $x \leq -1$ 时, X 的边缘概率密度 $f_X(x) = 0$.

当 $-1 < x < 1$ 时,

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{2}{\pi} (x^2 + y^2) dy = \frac{4}{3\pi} (1 + 2x^2) \sqrt{1-x^2},$$

所以

$$f_X(x) = \begin{cases} \frac{4}{3\pi} (1 + 2x^2) \sqrt{1-x^2}, & -1 < x < 1, \\ 0, & \text{其他.} \end{cases}$$

同理得

$$f_Y(y) = \begin{cases} \frac{4}{3\pi} (1 + 2y^2) \sqrt{1-y^2}, & -1 < y < 1, \\ 0, & \text{其他.} \end{cases}$$

由于 $f(x, y) \neq f_X(x)f_Y(y)$, 故 X 与 Y 不相互独立.

(3) 记 Z 的分布函数为 $F_Z(z)$.

因为 $P\{0 \leq Z \leq 1\} = 1$, 所以当 $z < 0$ 时, $F_Z(z) = 0$; 当 $z \geq 1$ 时, $F_Z(z) = 1$.

当 $0 \leq z < 1$,

$$\begin{aligned} F_Z(z) &= P\{Z \leq z\} \\ &= P\{X^2 + Y^2 \leq z\} \\ &= \iint_{x^2+y^2 \leq z} \frac{2}{\pi} (x^2 + y^2) dx dy \\ &= \int_0^{2\pi} d\theta \int_0^{\sqrt{z}} \frac{2}{\pi} r^3 dr \\ &= z^2. \end{aligned}$$

所以 Z 的概率密度为

$$f_z(z) = \begin{cases} 2z, & 0 < z < 1. \\ 0, & \text{其他.} \end{cases}$$

基础习题精练

习题

3.1 设二维随机变量 (X_1, Y_1) 和 (X_2, Y_2) 的概率密度分别为 $f_1(x, y)$ 与 $f_2(x, y)$, 令

$$f(x, y) = af_1(x, y) + bf_2(x, y),$$

若 $f(x, y)$ 是某二维连续型随机变量的概率密度, 则 a, b 满足条件 ().

(A) $a+b=1$

(B) $a>0$ 且 $b>0$

(C) $0 \leq a \leq 1, 0 \leq b \leq 1$

(D) $a \geq 0, b \geq 0$ 且 $a+b=1$

3.2 设随机变量 $X_i \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix} (i=1, 2)$, 且满足 $P\{X_1 X_2 = 0\} = 1$, 则 $P\{X_1 = X_2\}$ 等于 ().

(A) 0

(B) $\frac{1}{4}$

(C) $\frac{1}{2}$

(D) 1

3.3 设随机变量 X, Y 相互独立, 且均服从标准正态分布 $N(0, 1)$, 则 ().

(A) $P\{X+Y \geq 1\} = \frac{1}{2}$

(B) $P\{X-Y \geq 1\} = \frac{1}{2}$

(C) $P\{\max\{X, Y\} \geq 0\} = \frac{1}{4}$

(D) $P\{\min\{X, Y\} \geq 0\} = \frac{1}{4}$

3.4 设随机变量 X 和 Y 独立同分布, 且 X 的分布函数为 $F(x)$, 则 $Z = \max\{X, Y\}$ 的分布函数为 ().

(A) $F^2(z)$

(B) $F(x)F(y)$

(C) $1 - [1 - F(z)]^2$

(D) $[1 - F(x)][1 - F(y)]$

3.5 如果二维随机变量 (X, Y) 的分布函数为

$$F(x, y) = \begin{cases} 1 - e^{-\lambda_1 x} - e^{-\lambda_2 y} + e^{-\lambda_1 x - \lambda_2 y - \lambda_{12} \max\{x, y\}}, & x > 0, y > 0, \\ 0, & \text{其他,} \end{cases}$$

其中 $\lambda_1, \lambda_2, \lambda_{12} > 0$, 则 X 和 Y 的边缘分布函数分别为_____.

3.6 设随机变量 (X, Y) 的概率密度为

$$f(x, y) = \frac{1 + \sin x \sin y}{2\pi} e^{-\frac{1}{2}(x^2 + y^2)} \quad (-\infty < x, y < +\infty),$$

则两个边缘概率密度 $f_X(x)$, $f_Y(y)$ 分别为_____.

3.7 设随机变量 X_1, X_2, X_3, X_4 相互独立且同分布, $P\{X_i = 0\} = 0.6, P\{X_i = 1\} = 0.4, i = 1, 2, 3, 4$,

则行列式 $X = \begin{vmatrix} X_1 & X_2 \\ X_3 & X_4 \end{vmatrix}$ 的概率分布为_____.

3.8 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 2 - x - y, & 0 < x < 1, 0 < y < 1, \\ 0, & \text{其他,} \end{cases}$$

则 $Z = X + Y$ 的概率密度 $f_Z(z) =$ _____.

3.9 设 X 与 Y 相互独立, 下表是随机变量 X 与 Y 的联合分布律以及两个边缘分布律中的部分数值, 请将其余数值填入表中空白处.

$X \backslash Y$	y_1	y_2	y_3	$P\{X = x_i\}$
x_1		$\frac{1}{8}$		
x_2	$\frac{1}{8}$			
$P\{Y = y_j\}$	$\frac{1}{6}$			1

3.10 袋中有编号为 1, 1, 2, 3 的四个球, 现从中无放回地取两次, 每次任取一个, 设 X_1, X_2 分别为第一次、第二次取到的球的号码, 求:

- (1) (X_1, X_2) 的分布律, 并判断 X_1 与 X_2 的独立性;
- (2) 在 $X_2 = 2$ 的条件下, X_1 的条件分布;
- (3) 随机变量 $Y = X_1 X_2$ 的分布.

3.11 已知 (X, Y) 的概率分布为

$X \backslash Y$	-1	1
-1	$\frac{1}{6}$	$\frac{1}{6}$
0	$\frac{1}{3}$	$\frac{1}{6}$
1	0	$\frac{1}{6}$

(1) 求 $Z = X - Y$ 的概率分布;

(2) 设 $U_1 = XY$, $V_1 = \frac{X}{Y}$, 求 (U_1, V_1) 的概率分布;

(3) 设 $U_2 = \max\{X, Y\}$, $V_2 = \min\{X, Y\}$, 求 (U_2, V_2) 的概率分布, $U_2 V_2$ 的概率分布.

3.12 求以下给出概率密度的 (X, Y) 的边缘概率密度 $f_X(x)$ 和 $f_Y(y)$:

$$(1) f_1(x, y) = \begin{cases} e^{-y}, & 0 < x < y, \\ 0, & \text{其他}; \end{cases}$$

$$(2) f_2(x, y) = \begin{cases} \frac{5}{4}(x^2 + y), & 0 < y < 1 - x^2, \\ 0, & \text{其他}. \end{cases}$$

3.13 设二维随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} k, & 0 < x^2 < y < x < 1, \\ 0, & \text{其他}. \end{cases}$$

(1) 求常数 k ;

$$(2) \text{求 } P\left\{X > \frac{1}{2}\right\} \text{ 和 } P\left\{Y < \frac{1}{2}\right\}.$$

3.14 已知随机变量 (X, Y) 的概率密度

$$f(x, y) = \begin{cases} Ae^{-y}, & 0 < x < y, \\ 0, & \text{其他}, \end{cases}$$

求常数 A , 并计算概率 $P\{X + Y \geq 1\}$, $P\left\{\frac{X}{Y} \leq \frac{1}{2}\right\}$.

3.15 已知随机变量 X 与 Y 相互独立, X 服从参数为 λ 的指数分布, $P\{Y = -1\} = \frac{1}{4}$, $P\{Y = 1\} = \frac{3}{4}$,

求概率 $P\{X - Y \leq 1\}$, $P\{XY \leq 2\}$.

3.16 设随机变量 X, Y 相互独立, 且都服从指数分布, 它们的概率密度分别为

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & \text{其他}, \end{cases} \quad f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y \geq 0, \\ 0, & \text{其他}, \end{cases}$$

求 $Z = \frac{X}{Y}$ 的概率密度.

3.17 设随机变量 X, Y 相互独立, 且都在区间 $[a, b]$ 上服从均匀分布, 求 $Z = \min\{X, Y\}$ 的概率密度.

3.18 设随机变量 X 与 Y 相互独立, 且 X 服从标准正态分布 $N(0, 1)$, Y 的概率分布为 $P\{Y = 0\} =$

$P\{Y = 1\} = \frac{1}{2}$, 记 $F_Z(z)$ 为随机变量 $Z = XY$ 的分布函数, 求函数 $F_Z(z)$ 的间断点.

解答

3.1 (D) 解 函数 $f(x, y)$ 为某二维连续型随机变量的概率密度的充分必要条件是

$$f(x, y) \geq 0, \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f(x, y) dy = 1,$$

即同时有

$$f(x, y) = af_1(x, y) + bf_2(x, y) \geq 0,$$

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f(x, y) dy = a \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f_1(x, y) dy + b \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f_2(x, y) dy = a + b = 1,$$

故选择(D).

3.2 (A) 解 首先, 列出二维随机变量 (X_1, X_2) 的分布律及其边缘分布律中的部分数值.

$X_1 \backslash X_2$	-1	0	1	$P\{X_1 = x_{1i}\}$
-1	a	b	c	$\frac{1}{4}$
0	d	e	f	$\frac{1}{2}$
1	g	h	k	$\frac{1}{4}$
$P\{X_2 = x_{2j}\}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

由于 $P\{X_1 X_2 = 0\} = 1$, 故 $P\{X_1 X_2 \neq 0\} = 0$, 因此 $a = 0, c = 0, g = 0, k = 0$. 根据边缘分布的性质

$$b = \frac{1}{4}, h = \frac{1}{4}, d = \frac{1}{4}, f = \frac{1}{4}, e = \frac{1}{2} - (b + h) = \frac{1}{2} - \frac{1}{2} = 0,$$

故 (X_1, X_2) 的分布律为

$X_1 \backslash X_2$	-1	0	1	$P\{X_1 = x_{1i}\}$
-1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$P\{X_2 = x_{2j}\}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

可见 $P\{X_1 = X_2\} = P\{X_1 = -1, X_2 = -1\} + P\{X_1 = 0, X_2 = 0\} + P\{X_1 = 1, X_2 = 1\} = 0$.

故应选 (A).

3.3 (D) 解 因为 X, Y 相互独立, 且均服从标准正态分布 $N(0, 1)$, 所以 $U = X + Y \sim N(0, 2)$, $V = X - Y \sim N(0, 2)$, 排除 (A), (B). 又

$$P\{\max\{X, Y\} \geq 0\} = 1 - P\{\max\{X, Y\} < 0\} = 1 - P\{X < 0, Y < 0\} = \frac{3}{4},$$

$$P\{\min\{X, Y\} \geq 0\} = P\{X \geq 0, Y \geq 0\} = \frac{1}{4},$$

故应选择 (D).

3.4 (A) 解 由于 X 和 Y 独立同分布, 且 X 的分布函数为 $F(x)$, 则 Z 的分布函数为

$$F_Z(z) = P\{Z \leq z\} = P\{\max\{X, Y\} \leq z\} = P\{X \leq z, Y \leq z\} = P\{X \leq z\}P\{Y \leq z\} = F^2(z),$$

故选择 (A).

$$3.5 \quad F_X(x) = \begin{cases} 1 - e^{-\lambda_1 x}, & x > 0, \\ 0, & \text{其他}; \end{cases} \quad F_Y(y) = \begin{cases} 1 - e^{-\lambda_2 y}, & y > 0, \\ 0, & \text{其他} \end{cases} \quad \text{解 因为}$$

$$\lim_{y \rightarrow +\infty} (1 - e^{-\lambda_1 x} - e^{-\lambda_2 y} + e^{-\lambda_1 x - \lambda_2 y - \lambda_{12} \max\{x, y\}}) = 1 - e^{-\lambda_1 x},$$

$$\lim_{x \rightarrow +\infty} (1 - e^{-\lambda_1 x} - e^{-\lambda_2 y} + e^{-\lambda_1 x - \lambda_2 y - \lambda_{12} \max\{x, y\}}) = 1 - e^{-\lambda_2 y},$$

所以 X 和 Y 的边缘分布函数分别为

$$F_X(x) = F(x, +\infty) = \lim_{y \rightarrow +\infty} F(x, y) = \begin{cases} 1 - e^{-\lambda_1 x}, & x > 0, \\ 0, & \text{其他}, \end{cases}$$

$$F_Y(y) = F(+\infty, y) = \lim_{x \rightarrow +\infty} F(x, y) = \begin{cases} 1 - e^{-\lambda_2 y}, & y > 0, \\ 0, & \text{其他}. \end{cases}$$

$$3.6 \quad f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}; \quad f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \quad \text{解 由边缘概率密度的公式, 有}$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \frac{1}{2\pi} e^{-\frac{x^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy + \frac{\sin x}{2\pi} e^{-\frac{x^2}{2}} \int_{-\infty}^{+\infty} \sin y \cdot e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

即 $f_X(x)$ 是标准正态概率密度. 由对称性知 $f_Y(y)$ 也是标准正态概率密度, 所以 $f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$.

$$3.7 \quad X \sim \begin{pmatrix} -1 & 0 & 1 \\ 0.1344 & 0.7312 & 0.1344 \end{pmatrix} \quad \text{解 因为 } X = X_1 X_4 - X_2 X_3, \text{ 所以 } X \text{ 的可能取值为 } -1, 0, 1.$$

易知 $Z_1 = X_1 X_4$ 与 $Z_2 = X_2 X_3$ 独立同分布, 且

$$P\{Z_1 = 1\} = P\{X_1 = 1, X_4 = 1\} = 0.4^2 = 0.16,$$

$$P\{Z_1 = 0\} = 1 - P\{Z_1 = 1\} = 0.84,$$

同理, 得 $P\{Z_2 = 1\} = 0.16, P\{Z_2 = 0\} = 0.84$. 由此得

$$P\{X=-1\}=P\{Z_1=0, Z_2=1\}=P\{Z_1=0\}P\{Z_2=1\}=0.84\times 0.16=0.1344,$$

$$P\{X=1\}=P\{Z_1=1, Z_2=0\}=P\{Z_1=1\}P\{Z_2=0\}=0.16\times 0.84=0.1344,$$

$$P\{X=0\}=1-P\{X=-1\}-P\{X=1\}=1-2\times 0.1344=0.7312.$$

所以 X 的概率分布为

$$X \sim \begin{pmatrix} -1 & 0 & 1 \\ 0.1344 & 0.7312 & 0.1344 \end{pmatrix}.$$

$$3.8 \quad \begin{cases} z(2-z), & 0 < z < 1, \\ (2-z)^2, & 1 \leq z < 2, \\ 0, & \text{其他} \end{cases} \quad \text{解} \quad f_z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx, \text{ 其中}$$

$$f(x, z-x) = \begin{cases} 2-x-(z-x), & 0 < x < 1, 0 < z-x < 1, \\ 0, & \text{其他} \end{cases} = \begin{cases} 2-z, & 0 < x < 1, 0 < z-x < 1, \\ 0, & \text{其他} \end{cases}$$

当 $z \leq 0$ 或 $z \geq 2$ 时, $f_z(z) = 0$;

当 $0 < z < 1$ 时, $f_z(z) = \int_0^z (2-z) dx = z(2-z)$;

当 $1 \leq z < 2$ 时, $f_z(z) = \int_{z-1}^1 (2-z) dx = (2-z)^2$.

所以 Z 的概率密度为

$$f_z(z) = \begin{cases} z(2-z), & 0 < z < 1, \\ (2-z)^2, & 1 \leq z < 2, \\ 0, & \text{其他} \end{cases}$$

3.9 解 由于 X 与 Y 相互独立, 因此

$$p_{ij} = P\{X=x_i, Y=y_j\} = P\{X=x_i\}P\{Y=y_j\} = p_{i\cdot}p_{\cdot j} (i=1, 2; j=1, 2, 3).$$

由边缘分布的性质, 有

$$p_{i\cdot} = P\{X=x_i\} = \sum_{j=1}^3 p_{ij}, \quad p_{\cdot j} = P\{Y=y_j\} = \sum_{i=1}^2 p_{ij},$$

于是

$$p_{11} = p_{1\cdot} - p_{21} = \frac{1}{6} - \frac{1}{8} = \frac{1}{24},$$

$$p_{1\cdot} = \frac{p_{11}}{p_{\cdot 1}} = \frac{1}{24} \div \frac{1}{6} = \frac{1}{4}, \quad p_{2\cdot} = 1 - p_{1\cdot} = \frac{3}{4},$$

$$p_{\cdot 2} = \frac{p_{12}}{p_{1\cdot}} = \frac{1}{8} \div \frac{1}{4} = \frac{1}{2}, \quad p_{\cdot 3} = 1 - p_{\cdot 1} - p_{\cdot 2} = \frac{1}{3},$$

$$p_{22} = p_{2\cdot}p_{\cdot 2} = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8}, \quad p_{13} = p_{1\cdot}p_{\cdot 3} = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12},$$

$$p_{23} = p_{2\cdot}p_{\cdot 3} = \frac{3}{4} \times \frac{1}{3} = \frac{1}{4}.$$

因此所得数值填空如下表:

$\begin{matrix} Y \\ X \end{matrix}$	y_1	y_2	y_3	$P\{X=x_i\}$
x_1	$\frac{1}{24}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{4}$
x_2	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{3}{4}$
$P\{Y=y_j\}$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$	1

3.10 解 (1) X_1 与 X_2 的可能取值为 1, 2, 3, 则

$$p_{11} = P\{X_1=1, X_2=1\} = \frac{2}{4} \times \frac{1}{3} = \frac{1}{6}, \quad p_{12} = P\{X_1=1, X_2=2\} = \frac{2}{4} \times \frac{1}{3} = \frac{1}{6},$$

$$p_{13} = P\{X_1=1, X_2=3\} = \frac{2}{4} \times \frac{1}{3} = \frac{1}{6}, \quad p_{21} = P\{X_1=2, X_2=1\} = \frac{1}{4} \times \frac{2}{3} = \frac{1}{6},$$

$$p_{22} = 0, \quad p_{23} = P\{X_1=2, X_2=3\} = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}, \quad p_{31} = P\{X_1=3, X_2=1\} = \frac{1}{4} \times \frac{2}{3} = \frac{1}{6},$$

$$p_{32} = P\{X_1=3, X_2=2\} = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}, \quad p_{33} = 0.$$

于是 (X_1, X_2) 的分布律为

$\begin{matrix} X_2 \\ X_1 \end{matrix}$	1	2	3	$p_{i\cdot}$
1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$
2	$\frac{1}{6}$	0	$\frac{1}{12}$	$\frac{1}{4}$
3	$\frac{1}{6}$	$\frac{1}{12}$	0	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	1

因为 $p_{11} \neq p_{\cdot 1} \cdot p_{1\cdot}$, 所以 X_1 与 X_2 不相互独立.

(2) 由于

$$P\{X_1=1|X_2=2\} = \frac{p_{12}}{p_{\cdot 2}} = \frac{\frac{1}{6}}{\frac{1}{4}} = \frac{2}{3},$$

$$P\{X_1=2|X_2=2\} = 0, \quad P\{X_1=3|X_2=2\} = \frac{p_{32}}{p_{\cdot 2}} = \frac{\frac{1}{12}}{\frac{1}{4}} = \frac{1}{3},$$

因此在 $X_2=2$ 的条件下, X_1 的条件分布为 $X_1 \sim \begin{pmatrix} 1 & 3 \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$.

(3) $Y = X_1X_2$ 的可能取值为 1, 2, 3, 6, 于是

$$P\{Y=1\} = p_{11} = \frac{1}{6}, \quad P\{Y=2\} = p_{12} + p_{21} = \frac{1}{3},$$

$$P\{Y=3\} = p_{13} + p_{31} = \frac{1}{3}, \quad P\{Y=6\} = p_{23} + p_{32} = \frac{1}{6},$$

从而有

$$Y \sim \begin{pmatrix} 1 & 2 & 3 & 6 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}.$$

3.11 解 由题设得

p_{ij}	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	0	$\frac{1}{6}$
(X, Y)	$(-1, -1)$	$(-1, 1)$	$(0, -1)$	$(0, 1)$	$(1, -1)$	$(1, 1)$
$Z = X - Y$	0	-2	1	-1	2	0
$U_1 = XY$	1	-1	0	0	-1	1
$V_1 = X/Y$	1	-1	0	0	-1	1
$U_2 = \max\{X, Y\}$	-1	1	0	1	1	1
$V_2 = \min\{X, Y\}$	-1	-1	-1	0	-1	1
U_2V_2	1	-1	0	0	-1	1

(1) 由上表可得 $Z = X - Y$ 的概率分布为

$Z = X - Y$	-2	-1	0	1
$P\{Z=k\}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$

(2) (U_1, V_1) 的概率分布为

V_1	-1	0	1
U_1			
-1	$\frac{1}{6}$	0	0
0	0	$\frac{1}{2}$	0
1	0	0	$\frac{1}{3}$

(3) (U_2, V_2) 的概率分布为

$U_2 \backslash V_2$	-1	0	1
-1	$\frac{1}{6}$	0	0
0	$\frac{1}{3}$	0	0
1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

U_2V_2 的概率分布为

U_2V_2	-1	0	1
$P\{U_2V_2=k\}$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

3.12 解 (1) 当 $x > 0$ 时, 有 $f_X(x) = \int_x^{+\infty} e^{-y} dy = e^{-x}$, 所以 X 的边缘概率密度为

$$f_X(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & \text{其他}, \end{cases}$$

这是指数分布 $E(1)$.

当 $y > 0$ 时, 有 $f_Y(y) = \int_0^y xe^{-y} dx = ye^{-y}$, 所以 Y 的边缘概率密度为

$$f_Y(y) = \begin{cases} ye^{-y}, & y > 0, \\ 0, & \text{其他}. \end{cases}$$

(2) 因为 $f_2(x, y)$ 的非零区域如图 3-5 阴影部分所示, 所以当 $-1 < x < 1$ 时, 有

$$f_X(x) = \int_0^{1-x^2} \frac{5}{4}(x^2 + y) dy = \frac{5}{4}x^2(1-x^2) + \frac{5}{8}(1-x^2)^2 = \frac{5}{8}(1-x^4),$$

所以 X 的边缘概率密度为

$$f_X(x) = \begin{cases} \frac{5}{8}(1-x^4), & -1 < x < 1, \\ 0, & \text{其他}. \end{cases}$$

又因为当 $0 < y < 1$ 时, 有

$$\begin{aligned} f_Y(y) &= \int_{-\sqrt{1-y}}^{\sqrt{1-y}} \frac{5}{4}(x^2 + y) dx = \frac{5}{12}x^3 \Big|_{-\sqrt{1-y}}^{\sqrt{1-y}} + \frac{5}{2}y\sqrt{1-y} \\ &= \frac{5}{6}\sqrt{1-y}(1+2y), \end{aligned}$$

所以 Y 的边缘概率密度为 $f_Y(y) = \begin{cases} \frac{5}{6}\sqrt{1-y}(1+2y), & 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$

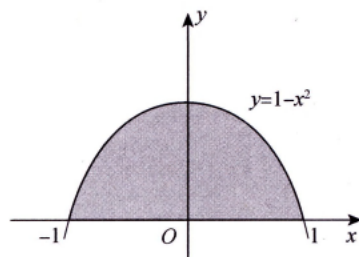


图 3-5

3.13 解 (1) $f(x, y)$ 的非零区域如图 3-6(a) 阴影部分所示.

由 $k \int_0^1 dx \int_{x^2}^x dy = k \int_0^1 (x - x^2) dx = \frac{k}{6} = 1$, 解得 $k = 6$.

(2) $f(x, y)$ 的非零区域与 $x > \frac{1}{2}$ 的交集如图 3-6(b) 阴影部分所示, 所以

$$P\left\{X > \frac{1}{2}\right\} = 6 \int_{\frac{1}{2}}^1 dx \int_{x^2}^x dy = 6 \int_{\frac{1}{2}}^1 (x - x^2) dx = 6 \left(\frac{1}{2} x^2 - \frac{1}{3} x^3 \right) \Big|_{\frac{1}{2}}^1 = \frac{1}{2}.$$

又因为 $f(x, y)$ 的非零区域与 $y < \frac{1}{2}$ 的交集如图 3-6(c) 阴影部分所示, 所以

$$P\left\{Y < \frac{1}{2}\right\} = 6 \int_0^{\frac{1}{2}} dy \int_y^{\sqrt{y}} dx = 6 \int_0^{\frac{1}{2}} (\sqrt{y} - y) dy = \sqrt{2} - \frac{3}{4}.$$

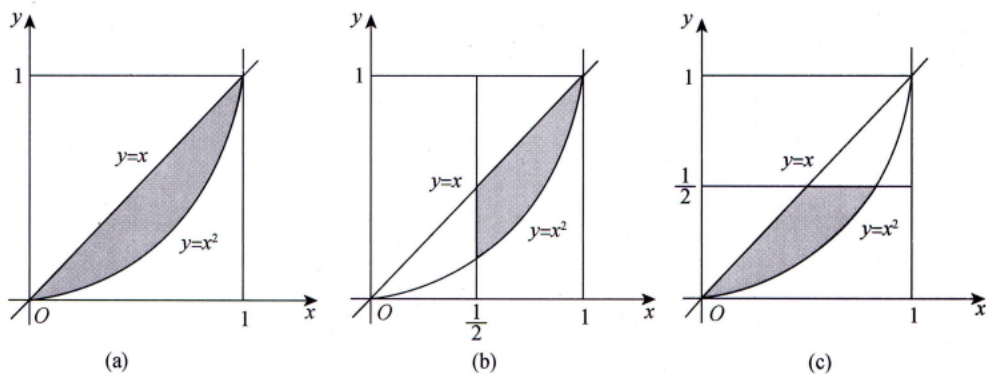


图 3-6

3.14 解 由于

$$1 = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^{+\infty} dx \int_x^{+\infty} A e^{-y} dy = A \int_0^{+\infty} e^{-x} dx = A,$$

因此 $A = 1$.

$$\begin{aligned} P\{X+Y \geq 1\} &= 1 - P\{X+Y < 1\} = 1 - \iint_{x+y < 1} f(x, y) dx dy \\ &= 1 - \int_0^{\frac{1}{2}} dx \int_x^{1-x} e^{-y} dy = 1 - \int_0^{\frac{1}{2}} (e^{-x} - e^{-x-1}) dx \\ &= 1 - \left(1 + e^{-1} - 2e^{-\frac{1}{2}} \right) = 2e^{-\frac{1}{2}} - e^{-1}, \end{aligned}$$

积分区域如图 3-7 中阴影部分所示.

$$P\left\{\frac{X}{Y} \leq \frac{1}{2}\right\} = \iint_{\frac{x}{y} \leq \frac{1}{2}} f(x, y) dx dy = \int_0^{+\infty} dx \int_{2x}^{+\infty} e^{-y} dy = \frac{1}{2},$$

积分区域如图 3-8 中阴影部分所示.

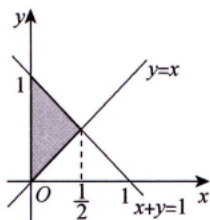


图 3-7

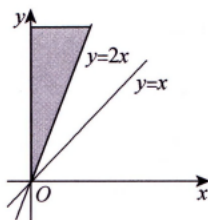


图 3-8

3.15 解 由于 Y 是离散型随机变量, X 与 Y 相互独立, 因此计算与 X, Y 有关的事件 A 的概率时, 自然想到将 A 对 Y 的可能值作全集分解, 应用全概率公式计算 $P(A)$.

已知 $X \sim f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0, \end{cases}$ $P\{Y = -1\} = \frac{1}{4}$, $P\{Y = 1\} = \frac{3}{4}$, X 与 Y 独立, 所以由全概率公式得

$$\begin{aligned} P\{X - Y \leq 1\} &= P\{X - Y \leq 1, Y = 1\} + P\{X - Y \leq 1, Y = -1\} \\ &= P\{X - 1 \leq 1, Y = 1\} + P\{X + 1 \leq 1, Y = -1\} \\ &= P\{Y = 1\}P\{X \leq 2\} + P\{Y = -1\}P\{X \leq 0\} \\ &= \frac{3}{4} \int_0^2 \lambda e^{-\lambda x} dx + \frac{1}{4} (1 - e^{-2\lambda}), \end{aligned}$$

$$\begin{aligned} P\{XY \leq 2\} &= P\{XY \leq 2, Y = 1\} + P\{XY \leq 2, Y = -1\} \\ &= P\{X \leq 2, Y = 1\} + P\{-X \leq 2, Y = -1\} \\ &= P\{Y = 1\}P\{X \leq 2\} + P\{Y = -1\}P\{X \geq -2\} \\ &= \frac{3}{4} \int_0^2 \lambda e^{-\lambda x} dx + \frac{1}{4} \int_0^{+\infty} \lambda e^{-\lambda x} dx = 1 - \frac{3}{4} e^{-2\lambda}. \end{aligned}$$

3.16 解 由于 X, Y 相互独立, (X, Y) 的概率密度为

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} \lambda^2 e^{-\lambda(x+y)}, & x \geq 0, y \geq 0, \\ 0, & \text{其他.} \end{cases}$$

用卷积公式

$$f_Z(z) = \int_{-\infty}^{+\infty} |y| f(yz, y) dy, \quad y \geq 0, z \geq 0.$$

当 $z \leq 0$ 时, $f_Z(z) = 0$;

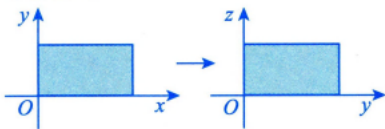
当 $z > 0$ 时,

$$\begin{aligned} f_Z(z) &= \int_0^{+\infty} y \cdot \lambda^2 e^{-\lambda(yz+y)} dy \\ &= \frac{1}{(1+z)^2} \int_0^{+\infty} \lambda(1+z) y e^{-\lambda(1+z)y} d[\lambda(1+z)y] \\ &\stackrel{t=\lambda(1+z)y}{=} \frac{1}{(1+z)^2} \int_0^{+\infty} t e^{-t} dt \\ &= \frac{1}{(1+z)^2}. \end{aligned}$$

用卷积公式“三部曲”:

① 换字母. $yz \geq 0, y \geq 0 \Rightarrow y \geq 0, z \geq 0$.

② 换区域.



③ 背口诀.
求 z 不积 z
不积先定限
限内画条线
先交写下限
后交写上限

因此 $Z = \frac{X}{Y}$ 的概率密度为

$$f_z(z) = \begin{cases} \frac{1}{(1+z)^2}, & z > 0, \\ 0, & \text{其他.} \end{cases}$$

3.17 解 由例 3.11 可知 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} \frac{1}{(b-a)^2}, & a \leq x \leq b, a \leq y \leq b, \\ 0, & \text{其他.} \end{cases}$$

又由 (X, Y) 的取值区域 $\{(x, y) | a \leq x \leq b, a \leq y \leq b\}$, 可知 $Z = \min\{X, Y\}$ 的取值区间为 $[a, b]$.

当 $z < a$ 或 $z > b$ 时, 有 $f_z(z) = 0$.

当 $a \leq z \leq b$ 时, 有

$$\begin{aligned} F_z(z) &= P\{Z \leq z\} = P\{\min\{X, Y\} \leq z\} = 1 - P\{X > z, Y > z\} \\ &= 1 - \iint_{x>z, y>z} f(x, y) dx dy = 1 - \int_z^b dx \int_z^b \frac{1}{(b-a)^2} dy \\ &= 1 - \frac{(b-z)^2}{(b-a)^2}, \end{aligned}$$

所以

$$f_z(z) = F'_z(z) = \frac{2(b-z)}{(b-a)^2}.$$

综上, $Z = \min\{X, Y\}$ 的概率密度为

$$f_z(z) = \begin{cases} \frac{2(b-z)}{(b-a)^2}, & a \leq z \leq b, \\ 0, & \text{其他.} \end{cases}$$

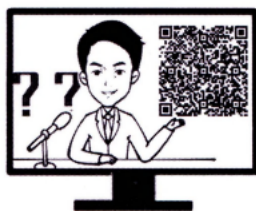
3.18 解 $F_z(z) = P\{XY \leq z\} = P\{XY \leq z | Y = 0\}P\{Y = 0\} + P\{XY \leq z | Y = 1\}P\{Y = 1\}$

$$\begin{aligned} &= \frac{1}{2} (P\{XY \leq z | Y = 0\} + P\{XY \leq z | Y = 1\}) \\ &= \frac{1}{2} (P\{0 \leq z\} + P\{X \leq z\}) \\ &= \begin{cases} \frac{1}{2} \Phi(z), & z < 0, \\ \frac{1}{2} [1 + \Phi(z)], & z \geq 0. \end{cases} \end{aligned}$$

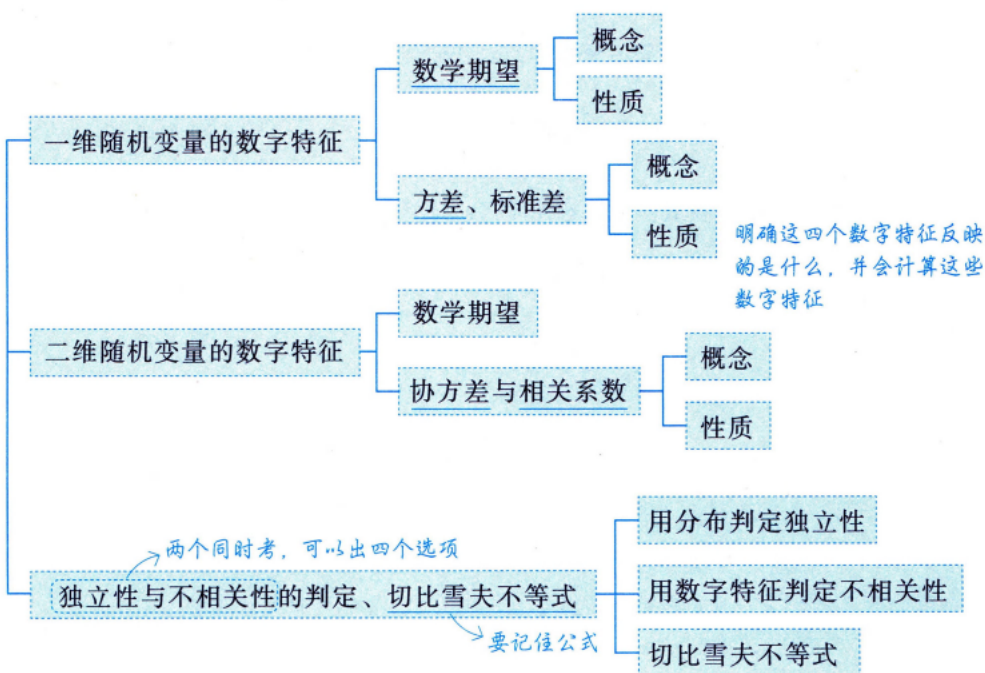
所以 $z = 0$ 是唯一间断点.

第4讲

随机变量的数字特征



基础知识结构



基础内容精讲

一 一维随机变量的数字特征

1 随机变量的数学期望

(1) 概念.

设 X 是随机变量, Y 是 X 的函数, $Y = g(X)$ (g 为连续函数).



①如果 X 是离散型随机变量, 其分布列为 $p_i = P\{X = x_i\} (i=1, 2, \dots)$. 若级数 $\sum_{i=1}^{\infty} x_i p_i$ 绝对收敛, 则

称随机变量 X 的数学期望存在, 并将级数 $\sum_{i=1}^{\infty} x_i p_i$ 的和称为随机变量 X 的数学期望, 记为 EX , 即

$$EX = \sum_{i=1}^{\infty} x_i p_i,$$

数 权重

否则称 X 的数学期望不存在.

若级数 $\sum_{i=1}^{\infty} g(x_i) p_i$ 绝对收敛, 则称 $Y = g(X)$ 的数学期望 $E[g(X)]$ 存在, 且 $E[g(X)] = \sum_{i=1}^{\infty} g(x_i) p_i$, 否

则称 $g(X)$ 的数学期望不存在.

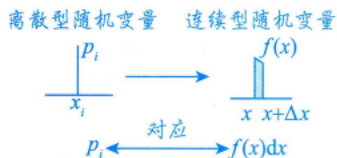
$$g(X) = Y \sim \begin{pmatrix} g(x_1) & g(x_2) & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$$

②如果 X 是连续型随机变量, 其概率密度为 $f(x)$. 若积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛, 则称 X 的数学期望

存在, 且 $EX = \int_{-\infty}^{+\infty} xf(x)dx$, 否则称 X 的数学期望不存在.

若积分 $\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛, 则称 $g(X)$ 的数学期望存在, 且 $E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$, 否则

称 $g(X)$ 的数学期望不存在.



注 (1) 数学期望又称为概率平均值, 常常简称期望或均值. 数学期望是描述随机变量平均取值状况特征的指标, 它刻画随机变量的一切可能值的集中位置. $\rightarrow$ 中心位置

(2) 在数学期望的定义中要求级数 (或积分) 绝对收敛, 否则称期望不存在.

(2) 性质.

$$\left. \begin{array}{l} \text{若 } X \sim p_i \Rightarrow EX = \sum_i x_i p_i \\ X \sim f(x) \Rightarrow EX = \int_{-\infty}^{+\infty} xf(x)dx \end{array} \right\} \xRightarrow{Y=g(X)} EY = \begin{cases} \sum_i g(x_i) p_i, & X \sim p_i, \\ \int_{-\infty}^{+\infty} g(x)f(x)dx, & X \sim f(x) \end{cases}$$

① $Ea = a$, $E(EX) = EX$. $E[E(EX)] = EX$

② $E(aX + bY) = aEX + bEY$, $E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i EX_i$.

$\rightarrow$ 线性运算

$\rightarrow$ 无条件打开

③若 X, Y 相互独立, 则 $E(XY) = EXEY$.

2 随机变量的方差、标准差

现有A,B两名射击运动员,打靶环数平均水平都为9.6,那要选择哪名运动员参加比赛呢?由此,需要引入一个衡量数据波动情况的特征量——方差

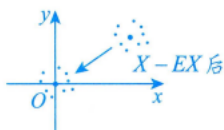
(1) 概念.

设 X 是随机变量,如果 $E[(X-EX)^2]$ 存在,则称 $E[(X-EX)^2]$ 为 X 的方差,记为 DX ,即

$$\begin{aligned} \star DX &= E[(X-EX)^2] = E(X^2) - (EX)^2, & DX &= E[(X-EX)^2] = E[X^2 - 2X \cdot EX + (EX)^2] \\ & & &= E(X^2) - 2EX \cdot EX + (EX)^2 \\ & & &= E(X^2) - (EX)^2 \end{aligned}$$

称 $\sqrt{DX}$ 为 X 的标准差或均方差,记为 $\sigma(X)$,称随机变量 $X^* = \frac{X-EX}{\sqrt{DX}}$ 为 X 的标准化随机变量,此时

$$EX^* = 0, DX^* = 1.$$



$$\begin{aligned} &\Downarrow \text{去中心化, 以二维为例} \\ DX^* &= D\left(\frac{X-EX}{\sqrt{DX}}\right) = \frac{1}{DX} D(X-EX) \\ &= \frac{DX-0}{DX} = 1 \end{aligned}$$

最早研究波动情况时,采用的是 $|X-EX|$,但由于绝对值不方便计算,因此选择平方这种计算方法.

(2) 性质.

$$\textcircled{1} DX \geq 0, E(X^2) = DX + (EX)^2 \geq (EX)^2. \quad \begin{aligned} DX &= E[(X-EX)^2] \\ &= E(X^2) - (EX)^2 \geq 0 \end{aligned}$$

$$\textcircled{2} Dc = 0 \quad (c \text{ 为常数}).$$

$$DX = 0 \Leftrightarrow X \text{ 几乎处处为某个常数 } a, \text{ 即 } P\{X=a\}=1.$$

$$\textcircled{3} D(aX+b) = a^2 DX. \quad \text{注意这里是 } a^2$$

$$\textcircled{4} D(X \pm Y) = DX + DY \pm 2\text{Cov}(X, Y), \quad D\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 DX_i + 2 \sum_{1 \leq i < j \leq n} a_i a_j \text{Cov}(X_i, X_j).$$

⑤如果 X 与 Y 相互独立,则

★每两个随机变量之间
都要求协方差

$$D(aX+bY) = a^2 DX + b^2 DY,$$

$$D(XY) = DXDY + DX(EY)^2 + DY(EX)^2 \geq DXDY.$$

注 证 因为 X, Y 相互独立,故 X^2, Y^2 也相互独立,即有

$$E(XY) = EXEY,$$

$$E(X^2 Y^2) = E(X^2)E(Y^2),$$

$$D(XY) = E(X^2 Y^2) - [E(XY)]^2 = E(X^2)E(Y^2) - (EX)^2(EY)^2$$

$$= [DX + (EX)^2][DY + (EY)^2] - (EX)^2(EY)^2$$

$$= DXDY + (EX)^2 DY + (EY)^2 DX.$$

又 $(EX)^2 DY + (EY)^2 DX \geq 0$, 所以 $D(XY) \geq DXDY$.

一般地, 如果 $X_1, X_2, \dots, X_n$ 相互独立, $g_i(x)(i=1, 2, \dots, n)$ 为 x 的连续函数, 则

$$D\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 D X_i,$$

$$D\left[\sum_{i=1}^n g_i(X_i)\right] = \sum_{i=1}^n D[g_i(X_i)].$$

⑥对任意常数 c , 有 $DX = E[(X - EX)^2] \leq E[(X - c)^2]$. $\rightarrow$ 关于 c 的函数

证

$$\begin{aligned} E[(X - c)^2] &= E(X^2 - 2cX + c^2) \\ &= E(X^2) - 2cEX + c^2 \\ &\stackrel{\text{令}}{=} g(c). \end{aligned}$$

由

$$g'(c) = -2EX + 2c \stackrel{\text{令}}{=} 0,$$

得 $c = EX$, 又 $g''(c) = 2 > 0$, 故 $c = EX$ 是 $g(c)$ 的最小值点.

我们将常用分布的数学期望和方差列表如下. **要记住**

分布	分布列 p_k 或概率密度 $f(x)$	数学期望	方差
0—1 分布	$P\{X = k\} = p^k(1-p)^{1-k}, k = 0, 1$	p	$p(1-p)$
二项分布 $B(n, p)$	$P\{X = k\} = C_n^k p^k(1-p)^{n-k}, k = 0, 1, \dots, n$	np	$np(1-p)$
泊松分布 $P(\lambda)$ $\rightarrow$ 泊松分布与指数分布要分清楚	$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, \dots$	λ	λ
几何分布 $G(p)$	$P\{X = k\} = (1-p)^{k-1} p, k = 1, 2, \dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
正态分布 $N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, -\infty < x < +\infty$	μ	σ^2
均匀分布 $U(a, b)$	$f(x) = \frac{1}{b-a}, a < x < b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $E(\lambda)$	$f(x) = \lambda e^{-\lambda x}, x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

注: 表中仅列出各分布概率密度的非零区域.

例 4.1 设一次试验成功的概率为 p ，进行 100 次独立重复试验，则成功次数的标准差的最大值为_____.

解 应填 5.

设成功次数为 X ，则 $X \sim B(100, p)$ ， X 的标准差为 $\sqrt{DX} = \sqrt{100p(1-p)}$.

由于 $p(1-p) \leq \frac{1}{4}$ ，当且仅当 $p = \frac{1}{2}$ 时等号成立，故当 $p = \frac{1}{2}$ 时成功次数的标准差最大，且其最大值为 $\sqrt{100 \times \frac{1}{4}} = 5$.

例 4.2 设随机变量 X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{2} \cos \frac{x}{2}, & 0 \leq x < \pi, \\ 0, & \text{其他,} \end{cases}$$

对 X 独立观察 4 次，用 Y 表示观察值大于 $\frac{\pi}{3}$ 的次数，求 Y^2 的数学期望.

解 由题意知 $P\left\{X > \frac{\pi}{3}\right\} = \int_{\frac{\pi}{3}}^{+\infty} f(x) dx = \int_{\frac{\pi}{3}}^{\pi} \frac{1}{2} \cos \frac{x}{2} dx = \frac{1}{2}$ ，所以 $Y \sim B\left(4, \frac{1}{2}\right)$ ，于是

$$EY = 4 \times \frac{1}{2} = 2, \quad DY = 4 \times \frac{1}{2} \times \left(1 - \frac{1}{2}\right) = 1,$$

所以

$$E(Y^2) = DY + (EY)^2 = 1 + 4 = 5.$$

例 4.3 设连续型随机变量 X 的概率密度为

$$f(x) = \frac{1}{\pi(1+x^2)} \quad (-\infty < x < +\infty),$$

则 EX ().

- (A) 等于 0 (B) 等于 1 (C) 等于 π (D) 不存在

解 应选 (D).

由连续型随机变量数学期望的定义，有

$$\int_{-\infty}^{+\infty} \frac{x}{\pi(1+x^2)} dx = \frac{1}{2\pi} \ln(1+x^2) \Big|_{-\infty}^{+\infty},$$

由于该积分发散，因此 EX 不存在.

例 4.4 设连续型随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sqrt{\pi}} e^{-x^2+2x-1} \quad (-\infty < x < +\infty),$$

则 $EX \cdot \sqrt{DX} =$ _____.

解 应填 $\frac{1}{\sqrt{2}}$.

形如 Ae^{-ax^2+bx+c} ($a > 0$) 的概率密度

要凑成 $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ 的形式

$$\Rightarrow \mu=1, \sigma^2=\frac{1}{2}$$

由

$$f(x) = \frac{1}{\sqrt{\pi}} e^{-x^2+2x-1} = \frac{1}{\sqrt{2\pi} \cdot \frac{1}{\sqrt{2}}} \exp \left\{ -\frac{(x-1)^2}{2 \left(\frac{1}{\sqrt{2}} \right)^2} \right\},$$

知 $EX=1$, $\sqrt{DX}=\frac{1}{\sqrt{2}}$, 故 $EX \cdot \sqrt{DX} = \frac{1}{\sqrt{2}}$.

例 4.5 设随机变量 X 的概率分布为 $P\{X=k\} = \frac{C}{k!}$, $k=0, 1, 2, \dots$, 则 $E(|X-EX|) = \underline{\hspace{2cm}}$.

解 应填 $\frac{2}{e}$.

形似泊松分布 $P\{X=k\} = \frac{\lambda^k}{k!} e^{-\lambda}$, $k=0, 1, 2, \dots$, 且 $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, 结合起来解题

利用性质 $\sum_{k=0}^{\infty} P\{X=k\} = 1$. 因为 $\sum_{k=0}^{\infty} P\{X=k\} = \sum_{k=0}^{\infty} \frac{C}{k!} = Ce = 1$, 所以 $C = e^{-1}$. 可知随机变量 X 服从参数为 1 的泊松分布. 于是

$$EX=1, E(X-EX) = E(X-1) = EX-1=0,$$

而

$$E(|X-EX|) = E(|X-1|)$$

$$= P\{X=0\} \cdot |0-1| + \sum_{k=1}^{\infty} P\{X=k\} \cdot |k-1|$$

$$= \frac{e^{-1}}{0!} + \sum_{k=1}^{\infty} P\{X=k\} \cdot (k-1)$$

$$= \frac{1}{e} + \left[\sum_{k=0}^{\infty} P\{X=k\} \cdot (k-1) - P\{X=0\} \cdot (0-1) \right]$$

$\Downarrow$

$$E(X-1) = 0$$

$$= \frac{1}{e} + \left(0 + \frac{1}{e} \right)$$

$$= \frac{2}{e}.$$

注 由泊松分布的分布列 $P\{X=k\} = \frac{\lambda^k}{k!} e^{-\lambda}$, $k=0, 1, 2, \dots$, 与本题给出的分布列 $P\{X=k\} =$

$\frac{C}{k!}$, $k=0, 1, 2, \dots$, 比较可以看出题目中的随机变量 X 就是服从参数 $\lambda=1$ 的泊松分布.

例 4.6 设随机变量 X 的概率密度为 $f(x) = \begin{cases} \frac{x}{2}, & 0 < x < 2, \\ 0, & \text{其他}, \end{cases}$ $F(x)$ 为 X 的分布函数, EX 为 X 的

数学期望, 则 $P\{F(X) > EX - 1\} = \underline{\hspace{2cm}}$.

解 应填 $\frac{2}{3}$.

方法一 令 $Y = F(X)$, 因为 $F(x)$ 严格单调增加, 所以 $Y \sim U(0, 1)$, 又 $EX = \int_{-\infty}^{+\infty} xf(x)dx =$

$$\int_0^2 \frac{x^2}{2} dx = \frac{4}{3}, \text{ 则}$$

由例2.15可得

$$P\{F(X) > EX - 1\} = P\left\{Y > \frac{1}{3}\right\} = \frac{2}{3}.$$

方法二 由题意知,

$$EX = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^2 \frac{x^2}{2} dx = \frac{4}{3}.$$

由 $F(x) = \int_{-\infty}^x f(t)dt$, 得

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{x^2}{4}, & 0 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

从而,

$$P\{F(X) > EX - 1\} = P\{F(X) > EX - 1, X < 0\} + P\{F(X) > EX - 1, 0 \leq X < 2\} +$$

$$P\{F(X) > EX - 1, X \geq 2\} = P\left\{\frac{X^2}{4} > \frac{1}{3}, 0 \leq X < 2\right\}$$

$$= P\left\{\frac{2}{\sqrt{3}} < X < 2\right\} = \int_{\frac{2}{\sqrt{3}}}^2 \frac{x}{2} dx = \frac{2}{3}.$$

注 显然方法二解法比较复杂.

二维随机变量的数字特征

1 二维随机变量函数的数学期望

$Z = g(X, Y)$ 仍然是一维, 期望, 方差都是

→ 对一维变量进行研究

设 X, Y 为随机变量, $g(X, Y)$ 为 X, Y 的函数 (g 是连续函数).

(1) 如果 (X, Y) 为离散型随机变量, 其联合分布律为



$$p_{ij} = P\{X = x_i, Y = y_j\} (i, j = 1, 2, \dots),$$

若级数 $\sum_i \sum_j g(x_i, y_j) p_{ij}$ 绝对收敛, 则定义

$$E[g(X, Y)] \stackrel{Z=g(X, Y)}{=} EZ = \sum_i \sum_j g(x_i, y_j) p_{ij}.$$

概率取值没变

(2) 如果 (X, Y) 为连续型随机变量, 其概率密度为 $f(x, y)$, 若积分 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$ 绝对收敛, 则定义

$$E[g(X, Y)] \stackrel{Z=g(X, Y)}{=} EZ = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy.$$

二重积分

$$DZ = E(Z^2) - (EZ)^2.$$

$$E(Z^2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g^2(x, y) f(x, y) dx dy$$

2 两个随机变量的协方差与相关系数

(1) 概念.

描述两个随机变量之间的关系

如果随机变量 X 与 Y 的方差存在且 $DX > 0, DY > 0$, 则称 $E[(X - EX)(Y - EY)]$ 为随机变量 X 与 Y 的协方差, 并记为 $\text{Cov}(X, Y)$, 即

$$\text{Cov}(X, Y) = E[(X - EX)(Y - EY)] = E(XY) - EX \cdot EY,$$

若 $X=Y$, 则 $\text{Cov}(X, Y) = E[(X - EX)^2] = DX$

要想求出 $\text{Cov}(X, Y)$, 需要先求 $EX, EY, E(XY)$

其中

$$E(XY) = \begin{cases} \sum_i \sum_j x_i y_j P\{X = x_i, Y = y_j\} & (\text{离散型}), \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f(x, y) dx dy & (\text{连续型}). \end{cases}$$

称 $\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{DX} \sqrt{DY}}$ 为随机变量 X 与 Y 的相关系数. 如果 $\rho_{XY} = 0$, 则称 X 与 Y 不相关; 如果 $\rho_{XY} \neq 0$,

则称 X 与 Y 相关.

若有随机变量 X , 数字特征为 EX, DX , 有 $X^* = \frac{X - EX}{\sqrt{DX}}$ 为 X 的标准化; 类似的, 有随机变量 Y ,

Y^* 为 Y 的标准化, 则

$EX^* = 0$ (去中心化), $DX^* = 1$ (去扰动, 去量纲)

$$\begin{aligned} \text{Cov}(X^*, Y^*) &= \text{Cov}\left(\frac{X - EX}{\sqrt{DX}}, \frac{Y - EY}{\sqrt{DY}}\right) \\ &= \frac{\text{Cov}(X - EX, Y - EY)}{\sqrt{DX} \sqrt{DY}} = \frac{\text{Cov}(X, Y)}{\sqrt{DX} \sqrt{DY}} = \rho_{XY}. \end{aligned}$$

注 (1) 协方差 $\text{Cov}(X, Y)$ 是描述随机变量 X 与 Y 之间偏差的关联程度的特征, 比如研究父亲身高 X 与孩子身高 Y 之间的偏差程度, 便可用 $\text{Cov}(X, Y)$ 来刻画. 相关系数 ρ_{XY} 描述随机变量 X 与 Y 之间的线性相依性, $|\rho_{XY}|$ 是刻画 X 与 Y 之间线性相关程度的一种度量. $\rho_{XY} = 0$ 表示 X 与 Y 之间不存在线性关系, 故称 X 与 Y 不相关, 但这并不意味着 X 与 Y 之间不存在相依关系, 它们之间还可能存在着某种非线性关系.

如 $X \sim U(-1, 1)$, $Y = \cos X$, 则

$$EX = 0, E(XY) = E(X \cos X) = \int_{-1}^1 x \cos x \cdot \frac{1}{2} dx = 0.$$

故 $\rho_{XY} = 0$, 说明 X 与 Y 之间没有线性关系, 但它们有函数关系.

(2) 随机变量的矩.

设 (X, Y) 是二维随机变量. 如果 $E(X^k Y^l)$ 存在, 则称 $E(X^k) (k=1, 2, \dots)$ 为 X 的 k 阶原点矩; 称 $E[(X - EX)^k] (k=2, 3, \dots)$ 为 X 的 k 阶中心矩; 称 $E(X^k Y^l) (k, l=1, 2, \dots)$ 为 X 与 Y 的 $k+l$ 阶混合原点矩; 称 $E[(X - EX)^k (Y - EY)^l] (k, l=1, 2, \dots)$ 为 X 与 Y 的 $k+l$ 阶混合中心矩.

从上述定义看出: 数学期望 EX 是 X 的一阶原点矩, 方差 DX 是 X 的二阶中心矩, 协方差 $\text{Cov}(X, Y)$ 是 X 与 Y 的二阶混合中心矩.

(2) 性质.

① $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.

② $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$.

③ $\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$.

④ $|\rho_{XY}| \leq 1$.

⑤ $\rho_{XY} = 1 \Leftrightarrow P\{Y = aX + b\} = 1 (a > 0)$.

$\rho_{XY} = -1 \Leftrightarrow P\{Y = aX + b\} = 1 (a < 0)$.

注 $Y = aX + b, a > 0 \Rightarrow \rho_{XY} = 1$.

$Y = aX + b, a < 0 \Rightarrow \rho_{XY} = -1$. $\longrightarrow$ 作为结论记住

例 4.7 设随机变量 X 与 Y 的概率分布分别为

X	0	1
P	$\frac{1}{3}$	$\frac{2}{3}$

Y	-1	0	1
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

且 $P\{X^2 = Y^2\} = 1$.

- (1) 求随机变量 (X, Y) 的概率分布;
- (2) 求 $Z = XY$ 的概率分布;
- (3) 求 X 与 Y 的相关系数 ρ_{XY} .

解 (1) 由 $P\{X^2 = Y^2\} = 1$, 知 $P\{X^2 \neq Y^2\} = 0$, 即

$$P\{X=0, Y=-1\} = P\{X=0, Y=1\} = P\{X=1, Y=0\} = 0,$$

从而有

$$P\{X=0, Y=0\} = P\{X=0\} = \frac{1}{3},$$

$$P\{X=1, Y=-1\} = P\{Y=-1\} = \frac{1}{3},$$

$$P\{X=1, Y=1\} = P\{Y=1\} = \frac{1}{3},$$

因此, (X, Y) 的概率分布为

$X \backslash Y$	-1	0	1	$p_{i\cdot}$
0	0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
$p_{\cdot j}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

- (2) $Z = XY$ 的可能取值为 -1, 0, 1.

$$X=1, Y=-1 \Rightarrow Z=-1,$$

$$X=0, Y=-1, 0, 1 \Rightarrow Z=0,$$

$$X=1, Y=0 \Rightarrow Z=0,$$

$$X=1, Y=1 \Rightarrow Z=1$$

$$P\{Z=-1\} = P\{X=1, Y=-1\} = \frac{1}{3},$$

$$P\{Z=1\} = P\{X=1, Y=1\} = \frac{1}{3},$$

$$P\{Z=0\} = 1 - \frac{1}{3} - \frac{1}{3} = \frac{1}{3},$$

所以

$$Z \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

- (3) 由于

$$EX = 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3}, \quad E(X^2) = \frac{2}{3}, \quad DX = \frac{2}{3} - \left(\frac{2}{3}\right)^2 = \frac{2}{9},$$

$$EY = -1 \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} = 0, \quad E(Y^2) = \frac{2}{3}, \quad DY = \frac{2}{3},$$

$$E(XY) = -1 \times 1 \times \frac{1}{3} + 1 \times 1 \times \frac{1}{3} = 0,$$

因此

$$\rho_{XY} = \frac{E(XY) - EXEY}{\sqrt{DX}\sqrt{DY}} = 0.$$

例 4.8 甲、乙两人相约于某地在 12:00~13:00 会面, 设 X, Y 分别是甲、乙到达超过 12:00 的时间 (单位: 小时), 且假设 X 和 Y 相互独立, 已知 X, Y 的概率密度分别为

$$f_X(x) = \begin{cases} 3x^2, & 0 < x < 1, \\ 0, & \text{其他}, \end{cases} \quad f_Y(y) = \begin{cases} 2y, & 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

求先到达者需要等待的时间的数学期望.

→ $|X-Y|$

解 由题意可得 $f(x, y) = \begin{cases} 6x^2y, & 0 < x < 1, 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$

由于 X, Y 相互独立

根据题意可知, 所求即为 $|X-Y|$ 的数学期望, 如图 4-1 所示, 有

$$\begin{aligned} E(|X-Y|) &= \int_0^1 \int_0^1 |x-y| 6x^2y \, dx \, dy \\ &= \iint_{D_1} [-(x-y) 6x^2y] \, dx \, dy + \iint_{D_2} (x-y) 6x^2y \, dx \, dy \\ &= \frac{1}{12} + \frac{1}{6} = \frac{1}{4} \text{ (小时)}. \end{aligned}$$

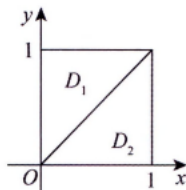


图 4-1

三 独立性与不相关性的判定、切比雪夫不等式

1 独立性与不相关性的判定



(1) 随机变量 X 与 Y 相互独立, 意指对任意实数 x, y , 事件 $\{X \leq x\}$ 与 $\{Y \leq y\}$ 相互独立, 即 (X, Y) 的分布等于边缘分布相乘: $F(x, y) = F_X(x) \cdot F_Y(y)$.

若 (X, Y) 是离散型的, 则 X 与 Y 独立的充要条件是

$$P\{X = x_i, Y = y_j\} = P\{X = x_i\} \cdot P\{Y = y_j\}; \quad \text{存在 } x_0, y_0 \text{ 使得}$$

若 (X, Y) 是连续型的, 则 X 与 Y 独立的充要条件是

$$F(x_0, y_0) \neq F_X(x_0) \cdot F_Y(y_0)$$

↓

X, Y 不独立

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

一般通过分布判定独立性.

(2) 随机变量 X 与 Y 不相关, 意指 X 与 Y 之间不存在线性相依性, 即 $\rho_{XY} = 0$, 其充要条件是

$$\rho_{XY} = 0 \Leftrightarrow \text{Cov}(X, Y) = 0 \Leftrightarrow E(XY) = EX \cdot EY \Leftrightarrow D(X \pm Y) = DX + DY.$$

一般通过数字特征判定不相关性.

→ $D(X \pm Y) = DX + DY \pm 2\text{Cov}(X, Y)$

(3) 几个重要结论: $\nearrow$ 没有任何关系

① 如果 X 与 Y 独立, 则 X, Y 不相关, 反之不然;

② 如果 (X, Y) 服从二维正态分布, 则 X, Y 独立 $\Leftrightarrow X, Y$ 不相关;

③ 由①知, 如果 X 与 Y 相关, 则 X, Y 不独立.

综上所述, 我们在讨论随机变量 X 与 Y 的不相关性、独立性时, 总是先计算 $\text{Cov}(X, Y)$, 而后按下列程序进行判断或再计算:

$$\star \text{Cov}(X, Y) = E(XY) - EXEY \begin{cases} \neq 0 \Leftrightarrow X \text{ 与 } Y \text{ 相关} \Rightarrow X \text{ 与 } Y \text{ 不独立,} \\ = 0 \Leftrightarrow X \text{ 与 } Y \text{ 不相关, 通过分布推断} \begin{cases} X, Y \text{ 独立,} \\ X, Y \text{ 不独立.} \end{cases} \end{cases}$$

注 上述讨论均假设方差存在且不为零.

2 切比雪夫不等式

如果随机变量 X 的期望 EX 和方差 DX 存在, 则对任意 $\varepsilon > 0$, 有

$$P\{|X - EX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2} \text{ 或 } P\{|X - EX| < \varepsilon\} \geq 1 - \frac{DX}{\varepsilon^2}.$$

随机变量离中心点
应当相差不远

注 由切比雪夫不等式知, 当 DX 越小时, 概率 $P\{|X - EX| < \varepsilon\}$ 越大, 这表明方差是刻画随机变量与其期望值偏离程度的量, 是描述随机变量 X “分散程度”特征的指标.

例 4.9 设随机变量 (X, Y) 服从二维正态分布, 则随机变量 $U = X + Y, V = X - Y$ 不相关的充分必要条件为 ().

(A) $EX = EY$

(B) $E(X^2) - (EX)^2 = E(Y^2) - (EY)^2$

(C) $E(X^2) = E(Y^2)$

(D) $E(X^2) + (EX)^2 = E(Y^2) + (EY)^2$

分析 $\text{Cov}(X + Y, X - Y) = 0 \Leftrightarrow E[(X + Y)(X - Y)] = E(X + Y) \cdot E(X - Y) \Leftrightarrow E(X^2) - E(Y^2) = (EX)^2 - (EY)^2$.

解 应选 (B).

随机变量 U, V 不相关的充分必要条件是协方差 $\text{Cov}(U, V) = 0$, 由

$$E(UV) = E(X^2) - E(Y^2), \quad EU = EX + EY, \quad EV = EX - EY,$$

得 $\text{Cov}(U, V) = E(UV) - EUEV = E(X^2) - E(Y^2) - (EX + EY)(EX - EY)$

$$= E(X^2) - (EX)^2 - [E(Y^2) - (EY)^2],$$

因此, 随机变量 U, V 不相关的充分必要条件为 $E(X^2) - (EX)^2 - [E(Y^2) - (EY)^2] = 0$, 即 $E(X^2) - (EX)^2 = E(Y^2) - (EY)^2$, 故选择 (B).

例 4.10 设随机变量 X 的概率密度为常考题, 研究 X , $|X|$ 的独立性、相关性

$$f(x) = \frac{1}{2} e^{-|x|}, \quad x \in (-\infty, +\infty),$$

则 ().

(A) X 与 $|X|$ 不相关, X 与 $|X|$ 不独立(B) X 与 $|X|$ 不相关, X 与 $|X|$ 独立(C) X 与 $|X|$ 相关, X 与 $|X|$ 不独立(D) X 与 $|X|$ 相关, X 与 $|X|$ 独立**解** 应选 (A).

①由对称性, 有

$$EX = \int_{-\infty}^{+\infty} xf(x)dx = 0, \quad E(X|X|) = \int_{-\infty}^{+\infty} x|x|f(x)dx = 0,$$

奇函数

因此, $\text{Cov}(X, |X|) = E(X|X|) - EXE(|X|) = 0$, 知 X 与 $|X|$ 不相关.②由例 3.6 知, X 与 $|X|$ 不独立, 故选 (A).**例 4.11** 设 X, Y 为随机变量, 数学期望都是 2, 方差分别为 1 和 4, 相关系数为 0.5, 则由切比雪夫不等式, $P\{|X - Y| \geq 6\} \leq$ _____.

$$P\{|U - EU| \geq 6\}$$

$$P\{|X - EX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2}$$

解 应填 $\frac{1}{12}$.

常数

已知 $EX = EY = 2$, $DX = 1$, $DY = 4$, $\rho_{XY} = 0.5 = \frac{\text{Cov}(X, Y)}{\sqrt{DX} \cdot \sqrt{DY}} = \frac{\text{Cov}(X, Y)}{2}$. 记 $U = X - Y$, 则

$$EU = 0, \quad DU = D(X - Y) = DX + DY - 2\text{Cov}(X, Y) = 1 + 4 - 2 \times 0.5 \times 2 = 3.$$

取 $\varepsilon = 6$, 由切比雪夫不等式, 得

$$P\{|X - Y| \geq 6\} \leq \frac{D(X - Y)}{6^2} = \frac{3}{36} = \frac{1}{12}.$$

**四 综合题举例****例 4.12** 在区间 $(0, 2)$ 上随机取一点, 将该区间分成两段, 较短一段的长度记为 X , 较长一段的长度记为 Y , 令 $Z = \frac{Y}{X}$.(1) 求 X 的概率密度;(2) 求 Z 的概率密度;

(3) 求 $E\left(\frac{X}{Y}\right)$.

解 (1) 设随机取的点的坐标记为 V , 则 $V \sim U(0, 2)$, $X = \min\{V, 2-V\}$. X 的分布函数记为 $F_X(x)$.

由于 $P\{0 \leq X \leq 1\} = 1$, 故

当 $x < 0$ 时, $F_X(x) = 0$;

当 $x \geq 1$ 时, $F_X(x) = 1$;

当 $0 \leq x < 1$ 时,

$$\begin{aligned} F_X(x) &= P\{X \leq x\} \\ &= P\{\min\{V, 2-V\} \leq x\} \\ &= 1 - P\{\min\{V, 2-V\} > x\} \\ &= 1 - P\{x < V < 2-x\} \\ &= 1 - \frac{(2-x)-x}{2} = x. \end{aligned}$$

所以 X 的分布函数为

$$F_X(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases}$$

故 X 的概率密度为

$$f_X(x) = \begin{cases} 1, & 0 < x < 1. \\ 0, & \text{其他.} \end{cases}$$

(2) 由条件, 知 $Z = \frac{Y}{X} = \frac{2-X}{X}$. 由于函数 $z = \frac{2-x}{x}$ 在 $(0, 1)$ 内严格单调且可导, 反函数为 $x = \frac{2}{1+z}$,

$\frac{dx}{dz} = -\frac{2}{(1+z)^2}$, 故 Z 的概率密度为

$$\begin{aligned} f_Z(z) &= \begin{cases} f_X\left(\frac{2}{1+z}\right) \left| -\frac{2}{(1+z)^2} \right|, & z > 1, \\ 0, & \text{其他} \end{cases} \\ &= \begin{cases} \frac{2}{(1+z)^2}, & z > 1, \\ 0, & \text{其他.} \end{cases} \end{aligned}$$

$$(3) E\left(\frac{X}{Y}\right) = E\left(\frac{X}{2-X}\right) = \int_0^1 \frac{x}{2-x} dx = 2 \ln 2 - 1.$$



基础习题精练

习题

4.1 已知连续型随机变量 X 与 Y 有相同的概率密度, 且

$$X \sim f(x) = \begin{cases} 2x\theta^2, & 0 < x < \frac{1}{\theta}, \\ 0, & \text{其他} \end{cases} \quad (\theta > 0), \quad E[a(X+2Y)] = \frac{1}{\theta},$$

则 $a = (\quad)$.

- (A) $\frac{2}{3}$ (B) $\frac{1}{2}$ (C) $\frac{1}{3}$ (D) $\frac{1}{6}$

4.2 设随机变量 (X, Y) 服从二维正态分布, X 与 Y 的期望值均为 μ , 方差均为 σ^2 , X, Y 的相关系数为 $\rho_{XY} = 0$, 记 $Z_1 = 2X + Y$, $Z_2 = 2X - Y$, 则 Z_1, Z_2 的相关系数为 $(\quad)$.

- (A) 0 (B) $\frac{3}{5}$ (C) $\frac{3}{\sqrt{15}}$ (D) $\frac{3}{\sqrt{10}}$

4.3 设二维随机变量 (X, Y) 的概率分布为

X \ Y	Y		
	0	1	2
-1	0.1	0.1	b
1	a	0.1	0.1

若事件 $\{\max\{X, Y\} = 2\}$ 与事件 $\{\min\{X, Y\} = 1\}$ 相互独立, 则 $\text{Cov}(X, Y) = (\quad)$.

- (A) -0.6 (B) -0.36 (C) 0 (D) 0.48

4.4 设随机变量 X 在 $[-1, 1]$ 上服从均匀分布, $Y = X^3$, 则 X 与 Y $(\quad)$.

- (A) 不相关且相互独立 (B) 不相关且不相互独立
(C) 相关且相互独立 (D) 相关且不相互独立

4.5 设随机变量 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < -1, \\ 0.2, & -1 \leq x < 0, \\ 0.8, & 0 \leq x < 1, \\ 1, & x \geq 1, \end{cases}$$

则 $D(X^2) = \underline{\hspace{2cm}}$.

4.6 设连续型随机变量 X 服从参数为 1 的指数分布, 则 $D(e^{-2X}) = \underline{\hspace{2cm}}$.

4.7 已知连续型随机变量 Y 的概率密度为

$$f(y) = \begin{cases} \frac{y}{a^2} e^{-\frac{y^2}{2a^2}}, & y > 0, \\ 0, & y \leq 0, \end{cases} \quad a > 0,$$

则随机变量 $Z = \frac{1}{Y}$ 的数学期望 $EZ =$ _____.

4.8 设随机变量 X 服从参数为 λ 的泊松分布, 且已知 $E[(X-1)(X-2)] = 1$, 则 $\lambda =$ _____.

4.9 设连续型随机变量 X 的概率密度为

$$f(x) = \frac{1}{\pi(1+x^2)} \quad (-\infty < x < +\infty),$$

则 $E(\min\{|X|, 1\}) =$ _____.

4.10 已知随机变量 $(X, Y) \sim N(0, 0; 1, 1; 0.5)$, $U = \max\{X, Y\}$, $V = \min\{X, Y\}$, 则 $E(U+V) =$ _____, $E(UV) =$ _____.

4.11 设 X 为随机变量, $P\{|X-EX| < \varepsilon\} \geq 0.9$, $DX = 0.009$, 用切比雪夫不等式估计 ε 的取值范围为 _____.

4.12 设随机变量

$$X \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

且 $\text{Cov}(X, Y) = \frac{1}{8}$, X 与 Y 的联合分布律为 _____.

4.13 设 A, B 为两个随机事件, 且 $P(A) = \frac{1}{4}$, $P(B|A) = \frac{1}{3}$, $P(A|B) = \frac{1}{2}$. 令

$$X = \begin{cases} 1, & A \text{ 发生,} \\ 0, & A \text{ 不发生,} \end{cases} \quad Y = \begin{cases} 1, & B \text{ 发生,} \\ 0, & B \text{ 不发生,} \end{cases}$$

求: (1) 二维随机变量 (X, Y) 的概率分布;

(2) X, Y 的相关系数 ρ_{XY} ;

(3) $Z = X^2 + Y^2$ 的概率分布.

4.14 设 $X \sim N(0, \sigma^2) (\sigma > 0)$, $Y = X^3$.

(1) 求 EY ;

(2) 判断 X 与 Y 是否相关? 是否独立? 说明理由.

解答

4.1 (B) 解 由于 X 与 Y 有相同的概率密度, 因此 $EX = EY$, 于是

$$\begin{aligned} E[a(X+2Y)] &= a(EX+2EY) = 3aEX \\ &= 3a \int_{-\infty}^{+\infty} xf(x)dx = 3a \int_0^{\frac{1}{\theta}} 2x^2\theta^2 dx \\ &= 3a \cdot \frac{2}{3} x^3 \theta^2 \Big|_0^{\frac{1}{\theta}} = \frac{2a}{\theta} = \frac{1}{\theta}, \end{aligned}$$

解得 $a = \frac{1}{2}$, 故选择 (B).

4.2 (B) 解 利用已有的关系计算, 不必从 (Z_1, Z_2) 的分布去考虑.

由 $\rho_{XY} = 0$, 知 X 与 Y 不相关且相互独立. 于是

$$\begin{aligned} DZ_1 &= 4DX + DY = 5\sigma^2, \quad DZ_2 = 4DX + DY = 5\sigma^2, \\ \text{Cov}(Z_1, Z_2) &= E[(2X+Y)(2X-Y)] - E(2X+Y)E(2X-Y) \\ &= 4E(X^2) - 4(EX)^2 - E(Y^2) + (EY)^2 \\ &= 4DX - DY = 3\sigma^2, \end{aligned}$$

因此, Z_1, Z_2 的相关系数 $\rho = \frac{\text{Cov}(Z_1, Z_2)}{\sqrt{DZ_1}\sqrt{DZ_2}} = \frac{3}{5}$, 故选择 (B).

4.3 (B) 解 $P\{\max\{X, Y\} = 2\} = P\{Y = 2\} = b + 0.1$,

$$P\{\min\{X, Y\} = 1\} = P\{X = 1, Y = 1\} + P\{X = 1, Y = 2\} = 0.2,$$

$$P\{\max\{X, Y\} = 2, \min\{X, Y\} = 1\} = P\{X = 1, Y = 2\} = 0.1,$$

由于事件 $\{\max\{X, Y\} = 2\}$ 与事件 $\{\min\{X, Y\} = 1\}$ 相互独立, 从而有

$$P\{\max\{X, Y\} = 2, \min\{X, Y\} = 1\} = P\{\max\{X, Y\} = 2\}P\{\min\{X, Y\} = 1\},$$

即 $0.1 = (b + 0.1) \times 0.2$, 解得 $b = 0.4$.

由分布律的性质知 $0.1 + 0.1 + b + a + 0.1 + 0.1 = 1$, 解得 $a = 0.2$. 从而有

$$EX = -1 \times (0.1 + 0.1 + 0.4) + 1 \times (0.2 + 0.1 + 0.1) = -0.2,$$

$$EY = 0 + 1 \times (0.1 + 0.1) + 2 \times (0.4 + 0.1) = 1.2,$$

$$E(XY) = -1 \times 1 \times 0.1 + 1 \times 1 \times 0.1 + (-1) \times 2 \times 0.4 + 1 \times 2 \times 0.1 = -0.6,$$

所以

$$\text{Cov}(X, Y) = E(XY) - EXEY = -0.6 - (-0.2) \times 1.2 = -0.36.$$

故选 (B).

4.4 (D) 解 显然, 结论 (C) 对任意两个随机变量都不正确. 又 X 与 Y 之间存在函数关系, 因而

不可能独立, 结论 (A) 不正确. 由题设, $EX=0$, $E(XY)=\int_{-1}^1 x^4 \cdot \frac{1}{2} dx \neq 0$, 从而知 $\text{Cov}(X, Y) \neq 0$, 则 X 与 Y 必相关, 也必不相互独立, 故选择 (D).

4.5 0.24 解 先求 X 的分布列, 即有

$$X \sim \begin{pmatrix} -1 & 0 & 1 \\ 0.2 & 0.6 & 0.2 \end{pmatrix}.$$

因此

$$E(X^2) = (-1)^2 \times 0.2 + 0^2 \times 0.6 + 1^2 \times 0.2 = 0.4,$$

$$E(X^4) = (-1)^4 \times 0.2 + 0^4 \times 0.6 + 1^4 \times 0.2 = 0.4,$$

$$D(X^2) = E(X^4) - [E(X^2)]^2 = 0.4 - 0.4^2 = 0.24.$$

4.6 $\frac{4}{45}$ 解 因为 $E(e^{-2X}) = \int_0^{+\infty} e^{-2x} e^{-x} dx = \int_0^{+\infty} e^{-3x} dx = \frac{1}{3}$, 且

$$\begin{aligned} E[(e^{-2X})^2] &= E(e^{-4X}) = \int_0^{+\infty} e^{-4x} e^{-x} dx \\ &= \int_0^{+\infty} e^{-5x} dx = \frac{1}{5}, \end{aligned}$$

$$\text{所以 } D(e^{-2X}) = E[(e^{-2X})^2] - [E(e^{-2X})]^2 = \frac{1}{5} - \left(\frac{1}{3}\right)^2 = \frac{4}{45}.$$

4.7 $\frac{\sqrt{2\pi}}{2a}$ 解
$$\begin{aligned} EZ &= E\left(\frac{1}{Y}\right) = \int_{-\infty}^{+\infty} \frac{1}{y} \cdot f(y) dy = \int_0^{+\infty} \frac{1}{a^2} e^{-\frac{y^2}{2a^2}} dy \\ &= \frac{1}{2a^2} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2a^2}} dy = \frac{\sqrt{2\pi}a}{2a^2} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}a} e^{-\frac{y^2}{2a^2}} dy = \frac{\sqrt{2\pi}}{2a}. \end{aligned}$$

注 上面是凑正态分布来处理 $\int_0^{+\infty} \frac{1}{a^2} e^{-\frac{y^2}{2a^2}} dy$ 的, 也可利用泊松积分 $\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$ 来完成.

4.8 1 解 由于随机变量 X 服从参数为 λ 的泊松分布, 因此 $EX = DX = \lambda$, 从而有

$$\begin{aligned} E[(X-1)(X-2)] &= E(X^2) - 3EX + 2 \\ &= DX + (EX)^2 - 3EX + 2 = \lambda^2 - 2\lambda + 2 = 1, \end{aligned}$$

解得 $\lambda = 1$.

4.9 $\frac{1}{\pi} \ln 2 + \frac{1}{2}$ 解 由对称性知,

$$\begin{aligned} E(\min\{|X|, 1\}) &= \int_{-\infty}^{+\infty} \min\{|x|, 1\} f(x) dx \\ &= 2 \int_0^{+\infty} \min\{x, 1\} f(x) dx \end{aligned}$$

$$\begin{aligned}
&= 2 \left[\int_0^1 x f(x) dx + \int_1^{+\infty} f(x) dx \right] \\
&= \frac{2}{\pi} \left(\int_0^1 \frac{x}{1+x^2} dx + \int_1^{+\infty} \frac{1}{1+x^2} dx \right) \\
&= \frac{1}{\pi} \ln(1+x^2) \Big|_0^1 + \frac{2}{\pi} \arctan x \Big|_1^{+\infty} \\
&= \frac{1}{\pi} \ln 2 + \frac{1}{2}.
\end{aligned}$$

4.10 0; 0.5 解 由于

$$U = \max\{X, Y\} = \frac{1}{2}[(X+Y) + |X-Y|],$$

$$V = \min\{X, Y\} = \frac{1}{2}[(X+Y) - |X-Y|],$$

则

$$U+V = X+Y, \quad UV = XY.$$

所以

$$E(U+V) = EX + EY = 0,$$

$$E(UV) = E(XY) = \text{Cov}(X, Y) + EXEY$$

$$= 0.5 + 0 = 0.5.$$

4.11 $\varepsilon \geq 0.3$ 解 应用切比雪夫不等式 $P\{|X-EX| < \varepsilon\} \geq 1 - \frac{DX}{\varepsilon^2}$ 求解. 由题设得

$$P\{|X-EX| < \varepsilon\} \geq 1 - \frac{0.009}{\varepsilon^2} \geq 0.9,$$

$$\varepsilon^2 \geq 0.09,$$

$$\varepsilon \geq 0.3.$$

4.12

$X \backslash Y$	0	1	$p_{i \cdot}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
$p_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{2}$	1

解 依题设, $EX = \frac{3}{4}$, $EY = \frac{1}{2}$, 于是

$$\text{Cov}(X, Y) = E(XY) - EXEY = E(XY) - \frac{3}{8} = \frac{1}{8},$$

得 $E(XY) = \frac{1}{2}$, 即有

$$E(XY) = \frac{1}{2} = P\{X=1, Y=1\},$$

$$P\{X=0, Y=1\} = P\{Y=1\} - P\{X=1, Y=1\} = 0,$$

$$P\{X=0, Y=0\} = P\{X=0\} - P\{X=0, Y=1\} = \frac{1}{4},$$

$$P\{X=1, Y=0\} = 1 - \frac{1}{4} - \frac{1}{2} - 0 = \frac{1}{4}.$$

因此, X 与 Y 的联合分布律为

$X \backslash Y$			$P_{i \cdot}$
	0	1	
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
$P_{\cdot j}$	$\frac{1}{2}$	$\frac{1}{2}$	1

4.13 解 (1) (X, Y) 的所有取值为 $(0, 0), (0, 1), (1, 0), (1, 1)$, 并有

$$\begin{aligned} P\{X=0, Y=0\} &= P(\overline{AB}) = P(\overline{A \cup B}) \\ &= 1 - P(A) - P(B) + P(AB), \end{aligned}$$

其中, $P(AB) = P(A)P(B|A) = \frac{1}{12}$, $P(B) = \frac{P(AB)}{P(A|B)} = \frac{\frac{1}{12}}{\frac{1}{2}} = \frac{1}{6}$, 因此

$$P\{X=0, Y=0\} = 1 - \frac{1}{4} - \frac{1}{6} + \frac{1}{12} = \frac{2}{3},$$

$$\begin{aligned} P\{X=0, Y=1\} &= P(\overline{AB}) = P(B) - P(AB) \\ &= \frac{1}{6} - \frac{1}{12} = \frac{1}{12}, \end{aligned}$$

$$\begin{aligned} P\{X=1, Y=0\} &= P(\overline{AB}) = P(A) - P(AB) \\ &= \frac{1}{4} - \frac{1}{12} = \frac{1}{6}, \end{aligned}$$

$$P\{X=1, Y=1\} = P(AB) = \frac{1}{12},$$

从而 (X, Y) 的概率分布为

$X \backslash Y$	0	1	$p_{i\cdot}$
0	$\frac{2}{3}$	$\frac{1}{12}$	$\frac{3}{4}$
1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{4}$
$p_{\cdot j}$	$\frac{5}{6}$	$\frac{1}{6}$	1

(2) 由 X 与 Y 的联合分布律, 得 X 与 Y 的边缘分布律分别为

$$X \sim \begin{pmatrix} 0 & 1 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}, Y \sim \begin{pmatrix} 0 & 1 \\ \frac{5}{6} & \frac{1}{6} \end{pmatrix},$$

从而

$$EX = \frac{1}{4}, E(X^2) = \frac{1}{4},$$

$$DX = \frac{1}{4} - \left(\frac{1}{4}\right)^2 = \frac{3}{16},$$

$$EY = \frac{1}{6}, E(Y^2) = \frac{1}{6},$$

$$DY = \frac{1}{6} - \left(\frac{1}{6}\right)^2 = \frac{5}{36},$$

$$E(XY) = \frac{2}{3} \times 0 \times 0 + \frac{1}{12} \times 0 \times 1 + \frac{1}{6} \times 1 \times 0 + \frac{1}{12} \times 1 \times 1 = \frac{1}{12},$$

$$\text{Cov}(X, Y) = E(XY) - EXEY = \frac{1}{24},$$

因此

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{DX}\sqrt{DY}} = \frac{1}{\sqrt{15}} = \frac{\sqrt{15}}{15}.$$

(3) $Z = X^2 + Y^2$ 的可能取值为 0, 1, 2, 因此

$$P\{Z = 0\} = P\{X = 0, Y = 0\} = \frac{2}{3},$$

$$P\{Z = 1\} = P\{X = 0, Y = 1\} + P\{X = 1, Y = 0\} = \frac{1}{4},$$

$$P\{Z = 2\} = P\{X = 1, Y = 1\} = \frac{1}{12},$$

从而

$$Z \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{2}{3} & \frac{1}{4} & \frac{1}{12} \end{pmatrix}.$$

4.14 解 (1) 已知 X 的概率密度为 $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}$, $-\infty < x < +\infty$, $f(x)$ 为 x 的偶函数, 故

$$EY = E(X^3) = \int_{-\infty}^{+\infty} x^3 f(x) dx = 0.$$

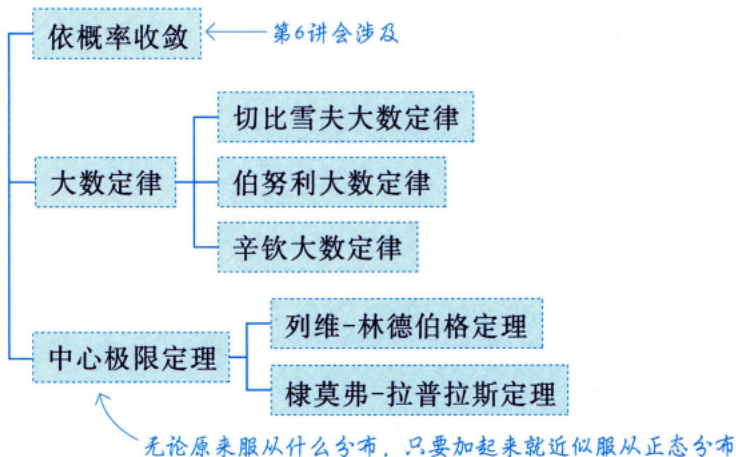
$$\begin{aligned} (2) \quad E(XY) &= E(X^4) = \int_{-\infty}^{+\infty} x^4 \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} dx \\ &= \frac{2}{\sqrt{2\pi}\sigma} \int_0^{+\infty} x^4 e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} dx \neq 0, \end{aligned}$$

故 $EX \cdot EY \neq E(XY)$, 因此 X 与 Y 相关, 从而可知 X 与 Y 不独立.

第5讲 大数定律与中心极限定理



基础知识结构



基础内容精讲

一 依概率收敛

若 $x_n \rightarrow a$, 则对于任意的 $\varepsilon > 0$, 当 $n \rightarrow \infty$ 时, 有 $|x_n - a| < \varepsilon$.

若 $X_n \xrightarrow{P} a$, 则对于任意的 $\varepsilon > 0$, 当 $n \rightarrow \infty$ 时, 有 $P\{|X_n - a| < \varepsilon\} \rightarrow 1$.

依据概率的



设随机变量 X 与随机变量序列 $\{X_n\} (n=1, 2, \dots)$, 如果对任意的 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\{|X_n - X| \geq \varepsilon\} = 0 \text{ 或 } \lim_{n \rightarrow \infty} P\{|X_n - X| < \varepsilon\} = 1,$$

则称随机变量序列 $\{X_n\}$ 依概率收敛于随机变量 X , 记为

$$\lim_{n \rightarrow \infty} X_n = X(P) \text{ 或 } X_n \xrightarrow{P} X (n \rightarrow \infty).$$

注 (1) 以上定义中将随机变量 X 写成数 a 也成立.

(2) 设 $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, $g(x, y)$ 是二元连续函数, 则 $g(X_n, Y_n) \xrightarrow{P} g(X, Y)$.

一般地, 对 m 元连续函数 $g(x_1, x_2, \dots, x_m)$, 上述结论亦成立.

二 大数定律


 总而言之: $\bar{X} \xrightarrow{P} EX (\bar{X} \text{ 依概率收敛到 } EX)$

定理 1 (切比雪夫大数定律) 假设 $\{X_n\} (n=1, 2, \dots)$ 是相互独立的随机变量序列, 如果方差 DX_i ($i \geq 1$) 存在且一致有上界, 即存在常数 C , 使 $DX_i \leq C$ 对一切 $i \geq 1$ 均成立, 则 $\{X_n\}$ 服从大数定律:

$$\text{随机变量的平均值} \rightarrow \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \frac{1}{n} \sum_{i=1}^n EX_i \leftarrow \text{期望的均值}$$

定理 2 (伯努利大数定律) 假设 μ_n 是 n 重伯努利试验中事件 A 发生的次数, 在每次试验中事件 A 发生的概率为 p ($0 < p < 1$), 则 $\frac{\mu_n}{n} \xrightarrow{P} p$, 即对任意 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = 1.$$

若 μ_n 表示投篮投中的次数, n 表示投篮总次数, 则 $\frac{\mu_n}{n}$ 表示投中的频率. 当 n 足够大时, 频率依概率收敛到概率本身

注 (仅数学三) 在数理统计中, 若 $(x_1, x_2, \dots, x_n)$ 为总体样本 $(X_1, X_2, \dots, X_n)$ 的一个观测值, 按大小顺序排列为 $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$. 对任意实数 x , 称函数

$$F_n(x) = \frac{x_1, x_2, \dots, x_n \text{ 中小于等于 } x \text{ 的样本值个数}}{n}$$

$$= \begin{cases} 0, & x < x_{(1)}, \\ \frac{k}{n}, & x_{(k)} \leq x < x_{(k+1)} (k=1, 2, \dots, n-1), \\ 1, & x \geq x_{(n)} \end{cases}$$

为样本 $(X_1, X_2, \dots, X_n)$ 的经验分布函数.

事实上, $F_n(x)$ 就是事件 $\{X \leq x\}$ 在 n 次试验中出现的频率, 而 $P\{X \leq x\} = F(x)$ 是事件 $\{X \leq x\}$ 出现的概率, 由伯努利大数定律 (即频率收敛于概率) 可知, 当 n 充分大时, $F_n(x)$ 可作为未知分布函数 $F(x)$ 的一个近似, n 越大, 近似效果越好.

★ **定理 3 (辛钦大数定律)** 假设 $\{X_n\} (n=1, 2, \dots)$ 是独立同分布的随机变量序列, 如果数学期望

$EX_i = \mu (i=1, 2, \dots)$ 存在, 则 $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \mu$, 即对任意 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P \left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| < \varepsilon \right\} = 1. \quad \bar{X} \xrightarrow{P} EX = \mu$$

注 (仅数学一) 在考查未知参数估计量是否具有一致性(相合性)时, 常常要用到依概率收敛和大数定律.

例 5.1 ① 设 $X_1, X_2, \dots, X_n, \dots$ 是相互独立的随机变量序列, X_n 服从参数为 $n(n \geq 1)$ 的指数分布, 则下列随机变量序列中不服从切比雪夫大数定律的是 ().

- (A) $X_1, \frac{1}{2}X_2, \dots, \frac{1}{n}X_n, \dots$ (B) $X_1, X_2, \dots, X_n, \dots$
 (C) $X_1, 2X_2, \dots, nX_n, \dots$ (D) $X_1, 2^2X_2, \dots, n^2X_n, \dots$

解 应选 (D).

切比雪夫大数定律要求 $\{X_n\}$ 相互独立, 方差存在且一致有界, 即 $DX_n \leq C$. 逐一验证各选项是否满足这一条件, 从而确定正确选项.

由题设知 $\{X_n\}$ 相互独立, 且 $DX_n = \frac{1}{n^2} \leq 1$, 所以选项 (B) 满足切比雪夫大数定律的条件.

又

$$D\left(\frac{1}{n}X_n\right) = \frac{1}{n^2}DX_n = \frac{1}{n^4} \leq 1, \quad D(nX_n) = n^2DX_n = 1 \leq 2,$$

由此可知, 选项 (A), (C) 均满足切比雪夫大数定律的条件, 然而 $D(n^2X_n) = n^4DX_n = n^2$, 选项 (D) 不满足切比雪夫大数定律的条件, 故选择 (D).

例 5.2 (仅数学三) 设 $(2, 1, 5, 2, 1, 3, 1)$ 是来自总体 X 的简单随机样本值, 求总体 X 的经验分布函数 $F_7(x)$.

解 将各观测值按从小到大的顺序排列, 得 $1, 1, 1, 2, 2, 3, 5$, 则经验分布函数为

$$F_7(x) = \begin{cases} 0, & x < 1, \\ \frac{3}{7}, & 1 \leq x < 2, \\ \frac{5}{7}, & 2 \leq x < 3, \\ \frac{6}{7}, & 3 \leq x < 5, \\ 1, & x \geq 5. \end{cases}$$

例 5.3 设随机变量序列 $X_1, X_2, \dots, X_n, \dots$ 独立同分布, 且 X_1 的概率密度为

$$f(x) = \begin{cases} 1-|x|, & |x| < 1, \\ 0, & \text{其他,} \end{cases}$$

满足辛钦大数定律: ①独立;
②同分布; ③ EX 存在,
则 $\bar{X} \xrightarrow{P} EX$, 即

则当 $n \rightarrow \infty$ 时, $\frac{1}{n} \sum_{i=1}^n X_i^2$ 依概率收敛于 ().

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} E\left(\frac{1}{n} \sum_{i=1}^n X_i^2\right)$$

(A) $\frac{1}{8}$

(B) $\frac{1}{6}$

(C) $\frac{1}{3}$

(D) $\frac{1}{2}$

解 应选 (B).

$$E(X_1^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_{-1}^1 x^2 (1-|x|) dx = 2 \int_0^1 x^2 (1-x) dx = \frac{1}{6}.$$

依题意知随机变量序列 $X_1^2, X_2^2, \dots, X_n^2, \dots$ 独立同分布, 从而由辛钦大数定律知, 当 $n \rightarrow \infty$ 时,

$\frac{1}{n} \sum_{i=1}^n X_i^2$ 依概率收敛于 $E(X_1^2)$, 即依概率收敛于 $\frac{1}{6}$. 故选 (B).



三 中心极限定理

定理 4 (列维 - 林德伯格定理) 假设 $\{X_n\} (n=1, 2, \dots)$ 是独立同分布的随机变量序列, 如果

$$EX_i = \mu, DX_i = \sigma^2 > 0 (i=1, 2, \dots)$$

存在, 则对任意的实数 x , 有

$$\lim_{n \rightarrow \infty} P \left\{ \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \Phi(x).$$

注 (1) 定理的三个条件“独立、同分布、期望和方差存在”, 缺一不可.

(2) 只要 X_n 满足定理条件, 那么当 n 很大时, 独立同分布随机变量的和 $\sum_{i=1}^n X_i$ 近似服从正态分布 $N(n\mu, n\sigma^2)$, 由此可知, 当 n 很大时, 有

$$P \left\{ a < \sum_{i=1}^n X_i < b \right\} \approx \Phi \left(\frac{b-n\mu}{\sqrt{n}\sigma} \right) - \Phi \left(\frac{a-n\mu}{\sqrt{n}\sigma} \right),$$

这常常是解题的依据. 只要题目涉及独立同分布随机变量的和 $\sum_{i=1}^n X_i$, 我们就要考虑独立同分布中心极限定理.

定理 5 (棣莫弗 - 拉普拉斯定理) 假设随机变量 $Y_n \sim B(n, p) (0 < p < 1, n \geq 1)$, 则对任意实数 x , 有

$$\lim_{n \rightarrow \infty} P \left\{ \frac{Y_n - np}{\sqrt{np(1-p)}} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \Phi(x).$$

注 (1) 如果记 $X_i \sim B(1, p) (0 < p < 1, i = 1, 2, \dots)$, 即 $X_i \sim \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$ 且相互独立, 则

$$Y_n = \sum_{i=1}^n X_i \sim B(n, p),$$

由列维-林德伯格定理推出棣莫弗-拉普拉斯定理.

(2) 二项分布概率计算的三种方法.

设 $X \sim B(n, p)$.

① 当 n 不太大时 ($n \leq 10$), 直接计算

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n;$$

② 当 n 较大且 p 较小时 ($n > 10, p < 0.1$), $\lambda = np$ 适中, 根据泊松定理有近似公式

$$P\{X = k\} = C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots, n;$$

③ 当 n 较大而 p 不太大时 ($p < 0.1, np \geq 10$), 根据中心极限定理, 有近似公式

$$P\{a < X < b\} \approx \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right).$$

例 5.4 设随机变量 $X_1, X_2, \dots, X_n$ 相互独立, $S_n = X_1 + X_2 + \dots + X_n$, 则根据列维-林德伯格定理, 当 n 充分大时, S_n 近似服从正态分布, 只要 $X_1, X_2, \dots, X_n$ ().

(A) 有相同的期望和方差 $\leftarrow$ 不一定同分布

(B) 服从同一离散型分布 $\leftarrow$ 同分布但期望或方差

(C) 服从同一指数分布

(D) 服从同一连续型分布 $\leftarrow$ 不一定存在

解 应选 (C).

列维-林德伯格定理的条件: 随机变量 $X_1, X_2, \dots, X_n$ 独立同分布, 并且其数学期望和方差均存在. 由于有相同的数学期望和方差未必服从相同的分布, 可见选项 (A) 不满足定理条件. 满足选项 (B) 和选项 (D) 的随机变量的数学期望或方差未必存在, 故选项 (B) 和选项 (D) 也不满足定理条件. 于是, 只有选项 (C) 成立 (指数分布的数学期望和方差都存在).

例 5.5 设 X_n 表示将一枚硬币随意投掷 n 次出现“正面”的次数, 则 ().

$$(A) \lim_{n \rightarrow \infty} P\left\{\frac{X_n - n}{\sqrt{n}} \leq x\right\} = \Phi(x)$$

$$(B) \lim_{n \rightarrow \infty} P\left\{\frac{2X_n - n}{\sqrt{n}} \leq x\right\} = \Phi(x)$$

$$(C) \lim_{n \rightarrow \infty} P\left\{\frac{X_n - 2n}{\sqrt{n}} \leq x\right\} = \Phi(x)$$

$$(D) \lim_{n \rightarrow \infty} P\left\{\frac{2X_n - 2n}{\sqrt{n}} \leq x\right\} = \Phi(x)$$

解 应选 (B).

由题设知 $X_n \sim B\left(n, \frac{1}{2}\right)$, 根据棣莫弗-拉普拉斯定理, 得

$$\lim_{n \rightarrow \infty} P\left\{\frac{X_n - \frac{1}{2}n}{\sqrt{\frac{1}{4}n}} \leq x\right\} = \lim_{n \rightarrow \infty} P\left\{\frac{2X_n - n}{\sqrt{n}} \leq x\right\} = \Phi(x),$$

故选择 (B).

例 5.6 生产线生产的产品成箱包装, 每箱质量是随机的. 假设平均每箱重 50 千克, 标准差为 5 千克, 若用载重为 5 吨的汽车承运, 试用中心极限定理说明每辆汽车最多可以装多少箱, 才能保证不超载的概率大于 0.977. ($\Phi(2) = 0.977$)

解 设 $X_i (i=1, 2, \dots, n)$ 为“第 i 箱产品的质量 (千克)”, 由题意, X_i 独立同分布, 且 $EX_i = 50$, $\sqrt{DX_i} = 5$, n 箱总质量为 $T_n = \sum_{i=1}^n X_i$, 从而有

平均每箱重 50 千克

$$ET_n = 50n, \sqrt{DT_n} = 5\sqrt{n}.$$

由列维-林德伯格定理, 知 $T_n \stackrel{\text{近似}}{\sim} N(50n, 25n)$, 因此有

$$P\{T_n \leq 5000\} = P\left\{\frac{T_n - 50n}{5\sqrt{n}} \leq \frac{5000 - 50n}{5\sqrt{n}}\right\} \approx \Phi\left(\frac{1000 - 10n}{\sqrt{n}}\right) > 0.977,$$

即 $\frac{1000 - 10n}{\sqrt{n}} > 2$, $n < 98.0199$, 所以每辆汽车最多可以装 98 箱, 才能保证不超载的概率大于 0.977.

基础习题精练

习题

5.1 设 $X_1, X_2, \dots, X_n, \dots$ 为独立同分布随机变量序列, 且均服从参数为 $\lambda (\lambda > 1)$ 的指数分布, 记 $\Phi(x)$ 为标准正态分布函数, 则 ().

$$(A) \lim_{n \rightarrow \infty} P\left\{\frac{\sum_{i=1}^n X_i - n\lambda}{\lambda\sqrt{n}} \leq x\right\} = \Phi(x)$$

$$(B) \lim_{n \rightarrow \infty} P\left\{\frac{\sum_{i=1}^n X_i - n\lambda}{\sqrt{\lambda n}} \leq x\right\} = \Phi(x)$$

$$(C) \lim_{n \rightarrow \infty} P\left\{\frac{\lambda \sum_{i=1}^n X_i - n}{\sqrt{n}} \leq x\right\} = \Phi(x)$$

$$(D) \lim_{n \rightarrow \infty} P\left\{\frac{\sum_{i=1}^n X_i - \lambda}{\sqrt{n\lambda}} \leq x\right\} = \Phi(x)$$

5.2 设 $X_1, X_2, \dots, X_{100}$ 为来自总体 X 的简单随机样本, 其中 $P\{X=0\}=P\{X=1\}=\frac{1}{2}$. $\Phi(x)$ 表示

标准正态分布函数, 则利用中心极限定理可得 $P\left\{\sum_{i=1}^{100} X_i \leq 55\right\}$ 的近似值为 ().

(A) $1-\Phi(1)$

(B) $\Phi(1)$

(C) $1-\Phi(0.2)$

(D) $\Phi(0.2)$

5.3 设某种电气元件不能承受超负荷试验的概率为 0.05. 现在对 100 个这样的元件进行超负荷试验, 以 X 表示“不能承受试验而烧毁的元件数”, 则根据中心极限定理, $P\{5 \leq X \leq 10\} \approx$ _____.

($\Phi(2.29) = 0.989$)

5.4 将一枚骰子重复掷 n 次, 则当 $n \rightarrow \infty$ 时, n 次掷出点数的算术平均值 $\bar{X}_n$ 依概率收敛于 _____.

5.5 假设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的简单随机样本, 已知 $E(X^k) = \alpha_k (k=1, 2, 3, 4)$.

证明: 当 n 充分大时, 随机变量 $Z_n = \frac{1}{n} \sum_{i=1}^n X_i^2$ 近似服从正态分布, 并指出其分布参数.

解答

5.1 (C) 解 由于 $E\left(\sum_{i=1}^n X_i\right) = \frac{n}{\lambda}$, $D\left(\sum_{i=1}^n X_i\right) = \frac{n}{\lambda^2}$, 因此 $\lim_{n \rightarrow \infty} P\left\{\frac{\sum_{i=1}^n X_i - \frac{n}{\lambda}}{\frac{\sqrt{n}}{\lambda}} \leq x\right\} = \Phi(x)$, 即

$$\lim_{n \rightarrow \infty} P\left\{\frac{\lambda \sum_{i=1}^n X_i - n}{\sqrt{n}} \leq x\right\} = \Phi(x),$$

故选择 (C).

5.2 (B) 解 由总体的概率分布得总体的期望、方差分别为 $EX = \frac{1}{2}$, $DX = \frac{1}{4}$, 故

$$E\left(\sum_{i=1}^{100} X_i\right) = 50, D\left(\sum_{i=1}^{100} X_i\right) = 25.$$

由中心极限定理, $\sum_{i=1}^{100} X_i$ 近似服从正态分布 $N(50, 25)$, 从而有

$$P\left\{\sum_{i=1}^{100} X_i \leq 55\right\} = P\left\{\frac{\sum_{i=1}^{100} X_i - 50}{5} \leq \frac{55 - 50}{5}\right\} \approx \Phi\left(\frac{55 - 50}{\sqrt{25}}\right) = \Phi(1),$$

故选 (B).

5.3 0.489 解 不能承受试验而烧毁的元件数 $X \sim B(n, p)$. 根据棣莫弗-拉普拉斯定理, X 近似服从正态分布 $N(np, npq)$, 其中 $n=100$, $p=0.05$, $q=0.95$. 因此

$$\begin{aligned} P\{5 \leq X \leq 10\} &= P\left\{\frac{5-np}{\sqrt{npq}} \leq \frac{X-np}{\sqrt{npq}} \leq \frac{10-np}{\sqrt{npq}}\right\} = P\left\{0 \leq \frac{X-5}{\sqrt{4.75}} \leq \frac{10-5}{\sqrt{4.75}}\right\} \\ &= P\left\{0 \leq \frac{X-5}{\sqrt{4.75}} \leq 2.29\right\} \approx \Phi(2.29) - \Phi(0) = 0.489. \end{aligned}$$

5.4 $\frac{7}{2}$ 解 设 $X_1, X_2, \dots, X_n$ 是各次掷出的点数, 显然它们独立同分布, 每次掷出点数的数学期望等于 $\frac{7}{2}$. 因此, 根据辛钦大数定律, $\bar{X}_n$ 依概率收敛于 $\frac{7}{2}$.

5.5 证明 依题意 $X_1, X_2, \dots, X_n$ 独立同分布, 可知 $X_1^2, X_2^2, \dots, X_n^2$ 也独立同分布且有

$$E(X_i^2) = \alpha_2, D(X_i^2) = E(X_i^4) - [E(X_i^2)]^2 = \alpha_4 - \alpha_2^2.$$

由列维-林德伯格定理, $V_n = \frac{\sum_{i=1}^n X_i^2 - n\alpha_2}{\sqrt{n(\alpha_4 - \alpha_2^2)}} = \frac{\frac{1}{n} \sum_{i=1}^n X_i^2 - \alpha_2}{\sqrt{(\alpha_4 - \alpha_2^2)/n}} = \frac{Z_n - \alpha_2}{\sqrt{(\alpha_4 - \alpha_2^2)/n}}$ 的极限分布是标准正

态分布, 所以当 n 充分大时, V_n 近似服从标准正态分布, 从而 $Z_n = \sqrt{\frac{\alpha_4 - \alpha_2^2}{n}} \cdot V_n + \alpha_2$ 近似服从参数为

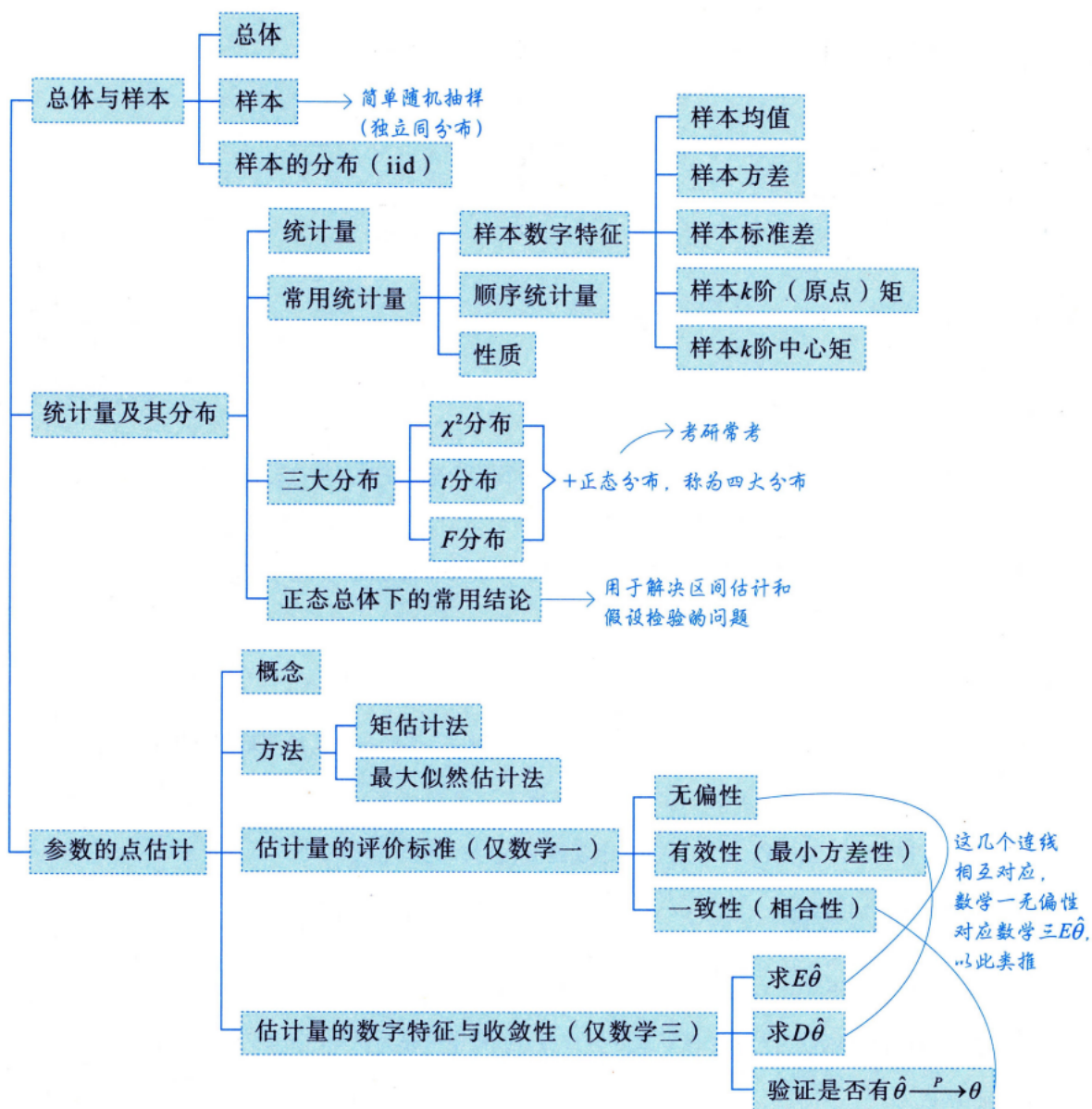
$\mu = \alpha_2, \sigma^2 = \frac{\alpha_4 - \alpha_2^2}{n}$ 的正态分布.

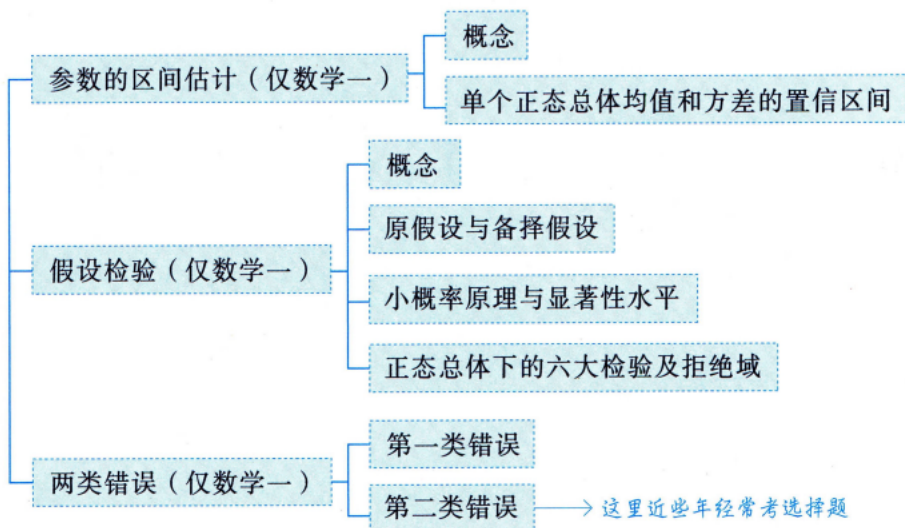
第6讲

数理统计



基础知识结构





基础内容精讲

一 总体与样本



1 总体

研究对象的全体称为总体, 组成总体的每一个元素称为个体. 在对总体进行统计研究时, 我们所关心的是表征总体状况的某个 (或某几个) 数量指标 X (可以是向量) 和该指标在总体中的分布情况. 我们把总体与随机变量 X 等同起来, 说“总体 X ”, 所谓总体的分布就是随机变量 X 的分布.

→ 大样本下, 其他分布收敛于正态分布, 所以常考的是正态分布

2 样本

n 个相互独立且与总体 X 具有相同概率分布的随机变量 $X_1, X_2, \dots, X_n$ 所组成的整体 $(X_1, X_2, \dots, X_n)$ 称为来自总体 X 、容量为 n 的一个简单随机样本, 简称样本. 一次抽样结果的 n 个具体数值 $(x_1, x_2, \dots, x_n)$ 称为样本 $X_1, X_2, \dots, X_n$ 的一个观测值 (或样本值).

→ 简单随机样本都是独立同分布的, 数学写法 $X_i \stackrel{\text{iid}}{\sim} F(x)$

3 样本的分布 (iid)

对于容量为 n 的样本 $X_1, X_2, \dots, X_n$, 有如下定理:

假设总体 X 的分布函数为 $F(x)$ (概率密度为 $f(x)$, 或概率分布为 $p_i = P\{X = x_i\}$), 则 $(X_1, X_2, \dots, X_n)$ 的分布函数为

$$F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F(x_i).$$

相应地:

→ 因为独立, 所以可以拆成连乘形式

(1) 对于离散型随机变量的样本 $X_1, X_2, \dots, X_n$, 联合分布为

$$P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\} = \prod_{i=1}^n P\{X_i = x_i\}.$$

→ 因为独立, 所以可以拆成连乘形式

(2) 对于连续型随机变量的样本 $X_1, X_2, \dots, X_n$, 联合概率密度为

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i).$$

→ 因为独立, 所以可以拆成连乘形式

二 统计量及其分布



1 统计量

设 $X_1, X_2, \dots, X_n$ 为来自总体 X 的一个样本, $g(x_1, x_2, \dots, x_n)$ 为 n 元函数, 如果 g 中不含任何未知参数, 则称 $g(X_1, X_2, \dots, X_n)$ 为样本 $X_1, X_2, \dots, X_n$ 的一个统计量. 若 $(x_1, x_2, \dots, x_n)$ 为样本值, 则称 $g(x_1, x_2, \dots, x_n)$ 为 $g(X_1, X_2, \dots, X_n)$ 的观测值.

→ 注意统计量是大写的 X

→ 注意是小写的 x

注 (1) 直观上, 统计量就是由统计数据计算得来的量. 数学上, 统计量是样本 $X_1, X_2, \dots, X_n$

的函数: $T = g(X_1, X_2, \dots, X_n)$. 统计量不依赖于任何未知参数.

(2) 统计量也是随机变量, 且是样本的函数.

→ 只能有样本 $X_1, X_2, \dots, X_n$

2 常用统计量

样本数字特征和顺序统计量都是常用的统计量. 统计量是统计分析和统计推断的重要工具.

(1) 样本数字特征.

数学期望 (均值)、方差和标准差、矩等都是总体 X 最重要的数字特征. 设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的简单随机样本, 则相应的样本数字特征定义如下.

① 样本均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$; (将所有的随机变量加起来再除以 n)

就是样本一阶(原点)矩

② 样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$;

是样本二阶中心矩的修正值

样本均值

$E(X - EX)^2 = DX$ 随机变量的方差
取得的值 自己的波动中心

注意是“ $n-1$ ”,而不是“ n ”.因 $\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ 不满足无偏性

样本标准差 $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$;

③ 样本 k 阶(原点)矩 $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k (k=1, 2, \dots)$; $\frac{1}{n} \sum_{i=1}^n (X_i - 0)^k$

④ 样本 k 阶中心矩 $B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k (k=2, 3, \dots)$.

减去中心值

★ (2) 顺序统计量.

将样本 $X_1, X_2, \dots, X_n$ 的 n 个观测值按其取值从小到大的顺序排列, 得

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}.$$

随机变量 $X_{(k)} (k=1, 2, \dots, n)$ 称作第 k 顺序统计量, 其中 $X_{(1)}$ 是最小顺序统计量, 而 $X_{(n)}$ 是最大顺序统计量:

$$X_{(1)} = \min\{X_1, X_2, \dots, X_n\}, X_{(n)} = \max\{X_1, X_2, \dots, X_n\}.$$

注 (1) $X_{(n)} = \max\{X_1, X_2, \dots, X_n\}$ 的分布函数为 $X_i \stackrel{\text{iid}}{\sim} F$

$$F_{(n)}(x) = [F(x)]^n$$

(概率密度为 $f_{(n)}(x) = n[F(x)]^{n-1}f(x)$).

(2) $X_{(1)} = \min\{X_1, X_2, \dots, X_n\}$ 的分布函数为

$$F_{(1)}(x) = 1 - [1 - F(x)]^n$$

(概率密度为 $f_{(1)}(x) = n[1 - F(x)]^{n-1}f(x)$).

(3) 性质.

设总体 X 的期望 $EX = \mu$, 方差 $DX = \sigma^2$, $X_1, X_2, \dots, X_n$ 是取自总体 X 、容量为 n 的一个样本, $\bar{X}, S^2$ 分别为样本均值和样本方差, 则

$$EX_i = \mu, DX_i = \sigma^2 (i=1, 2, \dots, n), E\bar{X} = EX = \mu, D\bar{X} = \frac{1}{n}DX = \frac{\sigma^2}{n}, E(S^2) = DX = \sigma^2.$$

因 X_i 独立同分布,
故有相同的数字特征

举例: 若 $\bar{X} = \frac{X_1 + X_2}{2}$, 则

举例: 若 $\bar{X} = \frac{X_1 + X_2}{2}$, 则

$$\begin{aligned} E\bar{X} &= E\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{2}(EX_1 + EX_2) \\ &= \frac{1}{2} \cdot 2\mu = \mu \end{aligned}$$

$$\begin{aligned} D\bar{X} &= D\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{2^2}(DX_1 + DX_2) \\ &= \frac{2}{2^2}\sigma^2 \\ &= \frac{1}{2}\sigma^2 \end{aligned}$$

★★★ ③ 三大分布—— χ^2 分布、 t 分布和 F 分布

χ^2 分布、 t 分布和 F 分布是统计推断中最常用的抽样分布. 考生不必记忆 χ^2 分布、 t 分布和 F 分布的概率密度, 只需了解相应统计量的概念, 以及它们的分布曲线的示意图和分位数, 会查相应分位数的数值表.

(1) χ^2 分布.

① 概念.

$X_i \stackrel{iid}{\sim} N(0,1)$

若随机变量 $X_1, X_2, \dots, X_n$ 相互独立, 且都服从标准正态分布, 则随机变量 $X = \sum_{i=1}^n X_i^2$ 服从自由度

为 n 的 χ^2 分布, 记为 $X \sim \chi^2(n)$. $X = X_1^2 + X_2^2 + \dots + X_n^2 \sim \chi^2(n)$

和式中独立变量的个数为自由度

② 上 α 分位数.

对于给定的 $\alpha (0 < \alpha < 1)$, 称满足

$$P\{\chi^2 > \chi_{\alpha}^2(n)\} = \int_{\chi_{\alpha}^2(n)}^{+\infty} f(x)dx = \alpha$$

的 $\chi_{\alpha}^2(n)$ 为 $\chi^2(n)$ 分布的上 α 分位数 (见图 6-1). 对于不同的 α, n ,

$\chi^2(n)$ 分布上 α 分位数可通过查表求得.

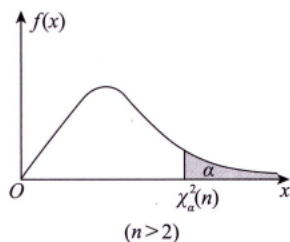
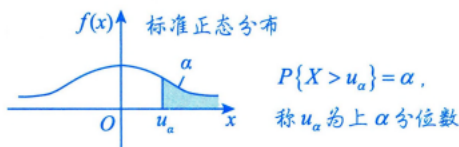


图 6-1

③ 性质.

a. 若 $X_1 \sim \chi^2(n_1)$, $X_2 \sim \chi^2(n_2)$, X_1 与 X_2 相互独立, 则

$$X_1 + X_2 \sim \chi^2(n_1 + n_2).$$

此结论可推广至有限多个的和.

b. 若 $X \sim \chi^2(n)$, 则 $EX = n$, $DX = 2n$.

期望是
自由度

方差是2倍
的自由度

(2) t 分布.

① 概念.

设随机变量 $X \sim N(0, 1)$, $Y \sim \chi^2(n)$, 且 X 与 Y 相互独立, 则随机变量 $t = \frac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布, 记为 $t \sim t(n)$.

$\frac{X}{\sqrt{Y/n}}$ 保证上下“公平”

 ② 上 α 分位数.

对于给定的 $\alpha (0 < \alpha < 1)$, 称满足

$$P\{t > t_\alpha(n)\} = \alpha$$

的 $t_\alpha(n)$ 为 $t(n)$ 分布的上 α 分位数 (见图 6-2).

③ 性质.

a. t 分布的概率密度 $f(x)$ 的图形关于 $x=0$ 对称 (见图 6-3), 因此

$$Et = 0 (n \geq 2).$$

因为 $f(x)$ 为偶函数.

b. 由 t 分布的概率密度 $f(x)$ 图形的对称性, 知 $P\{t > -t_\alpha(n)\} = P\{t > t_{1-\alpha}(n)\}$, 故 $t_{1-\alpha}(n) = -t_\alpha(n)$.
 当 α 值在表中没有时, 可用此式求得上 α 分位数.

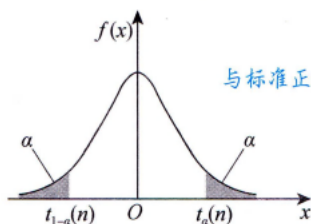


图 6-2

与标准正态分布类似

需要查2个数: α, n

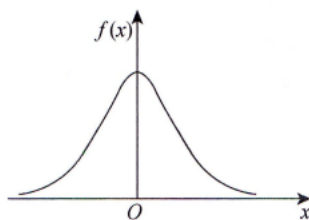


图 6-3

 (3) F 分布.

① 概念.

设随机变量 $X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, 且 X 与 Y 相互独立, 则 $F = \frac{X/n_1}{Y/n_2}$ 服从自由度为 (n_1, n_2) 的 F

分布, 记为 $F \sim F(n_1, n_2)$, 其中 n_1 称为第一自由度, n_2 称为第二自由度. F 分布的概率密度 $f(x)$ 的图形如图 6-4 所示.

n_1 分子的自由度 n_2 分母的自由度

 ② 上 α 分位数.

对于给定的 $\alpha (0 < \alpha < 1)$, 称满足

$$P\{F > F_\alpha(n_1, n_2)\} = \alpha$$

的 $F_\alpha(n_1, n_2)$ 为 $F(n_1, n_2)$ 分布的上 α 分位数 (见图 6-5).

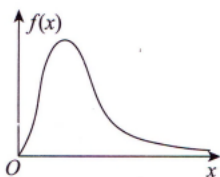


图 6-4

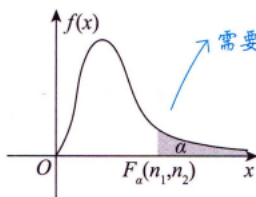


图 6-5

证明: 设 $F \sim F(n_2, n_1)$, 则

$$P\{F > F_\alpha(n_2, n_1)\} = \alpha,$$

$$P\{F \leq F_\alpha(n_2, n_1)\} = 1 - \alpha,$$

$$P\left\{\frac{1}{F} \geq \frac{1}{F_\alpha(n_2, n_1)}\right\} = 1 - \alpha,$$

因为 $\frac{1}{F} \sim F(n_1, n_2)$, 结合上 α 分位数

的定义, 可知 $\frac{1}{F_\alpha(n_2, n_1)} = F_{1-\alpha}(n_1, n_2)$

③性质.

a. 若 $F \sim F(n_1, n_2)$, 则 $\frac{1}{F} \sim F(n_2, n_1)$.

b. $F_{1-\alpha}(n_1, n_2) = \frac{1}{F_\alpha(n_2, n_1)}$, 常用来求 F 分布表中未列出的上 α 分位数.

c. 若 $t \sim t(n)$, 则 $t^2 \sim F(1, n)$.

自由度位置调换了

$$t = \frac{X}{\sqrt{Y/n}}, \quad t^2 = \frac{X^2}{Y/n} \sim F(1, n)$$

4 正态总体条件下的常用结论 (将 4 个结论当成 4 个例题)

设 $X_1, X_2, \dots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, $\bar{X}, S^2$ 分别是样本均值和样本方差, 则

$$\textcircled{1} \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \text{ 即 } \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1);$$

在正态总体下, $\frac{\bar{X} - X \text{ 的期望}}{X \text{ 的标准差} / \sqrt{n}} \sim N(0, 1)$

$$\textcircled{2} \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 \sim \chi^2(n);$$

客观, $\mu = EX = \int_{-\infty}^{+\infty} xf(x)dx$.

$$X \sim N(\mu, \sigma^2),$$

$$\frac{X_i - \mu}{\sigma} \sim N(0, 1),$$

$$\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(1),$$

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 \sim \chi^2(n)$$

$$\textcircled{3} \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1) (\mu \text{ 未知时, 在 } \textcircled{2} \text{ 中用 } \bar{X} \text{ 替代 } \mu);$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \Rightarrow \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \frac{(X_i - \bar{X})^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$$

数据

④ $\bar{X}$ 与 S^2 相互独立, $\frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1)$ (σ 未知时, 在 ① 中用 S 替代 σ). 进一步有

★ 正态总体下才有此结论

$$\frac{n(\bar{X} - \mu)^2}{S^2} \sim F(1, n-1).$$

$\sim N(0, 1)$
 $\frac{(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}})^2}{1} = \frac{n(\bar{X} - \mu)^2}{S^2}$
 $\sim \chi^2(n-1)$ (乘 $(n-1)$, 再除 $(n-1)$)

例 6.1 设 $X_1, X_2, \dots, X_n (n > 2)$ 为独立同分布的随机变量, 且均服从正态分布 $N(0, 1)$, 记

$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, Y_i = X_i - \bar{X}, i = 1, 2, \dots, n$. 则 Y_i 的方差 $DY_i = \underline{\hspace{2cm}}$. 考点: 结合第4讲 随机变量的数字特征

解 应填 $\frac{n-1}{n}$.

$$\begin{aligned}
 DY_i &= D(X_i - \bar{X}) = D\left(X_i - \frac{X_1 + X_2 + \dots + X_n}{n}\right) \\
 &= D\left[\left(1 - \frac{1}{n}\right)X_i - \frac{1}{n} \sum_{k \neq i} X_k\right] = \left(1 - \frac{1}{n}\right)^2 + \frac{n-1}{n^2} \\
 &= \frac{n-1}{n}, i = 1, 2, \dots, n.
 \end{aligned}$$

会有一个相同的 X_i

例 6.2 已知总体 X 的期望 $EX = 0$, 方差 $DX = \sigma^2$, 从总体 X 中抽取容量为 n 的简单随机样本,

其样本均值、样本方差分别为 $\bar{X}, S^2$. 记 $S_k^2 = \frac{n}{k} \bar{X}^2 + \frac{1}{k} S^2 (k = 1, 2, 3, 4)$, 则 (). $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} X$

(A) $E(S_1^2) = \sigma^2$

(B) $E(S_2^2) = \sigma^2$

(C) $E(S_3^2) = \sigma^2$

(D) $E(S_4^2) = \sigma^2$

解 应选 (B).

由 $E\bar{X} = EX = 0, D\bar{X} = \frac{DX}{n} = \frac{\sigma^2}{n}, E(S^2) = \sigma^2$, 且 $E(\bar{X}^2) = D\bar{X} + (E\bar{X})^2 = D\bar{X} = \frac{\sigma^2}{n}$, 故

$$E(S_k^2) = \frac{n}{k} E(\bar{X}^2) + \frac{1}{k} E(S^2) = \frac{n}{k} \cdot \frac{\sigma^2}{n} + \frac{1}{k} \sigma^2 = \frac{2}{k} \sigma^2.$$

当 $k = 2$ 时, $E(S_2^2) = \sigma^2$, 选择 (B).

例 6.3 设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, 2^2)$ 的简单随机样本, 记

$$X = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2,$$

若统计量 X 服从自由度为 2 的 χ^2 分布, 则 $ab = \underline{\hspace{2cm}}$.

分析 若 $X_i \stackrel{\text{iid}}{\sim} N(0, 1)$, 则 $X_1^2 + \dots + X_n^2 \sim \chi^2(n)$, 故

$$X = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2$$

自变量无重叠, 故独立

$$= \left(\frac{X_1 - 2X_2}{\frac{1}{\sqrt{a}}} \right)^2 + \left(\frac{3X_3 - 4X_4}{\frac{1}{\sqrt{b}}} \right)^2 \sim \chi^2(2).$$

解 应填 $\frac{1}{2000}$.

$$D(X_1 - 2X_2) = DX_1 + 4DX_2 = 20$$

由条件知 X_1, X_2, X_3, X_4 相互独立且同服从正态分布 $N(0, 2^2)$. 因此 $X_1 - 2X_2$ 服从正态分布 $N(0, 20)$, 而 $3X_3 - 4X_4$ 服从正态分布 $N(0, 100)$, 并且相互独立. 由 χ^2 分布的概念, 知

$$D(3X_3 - 4X_4) = 9DX_3 + 16DX_4 = 100 \quad T = \frac{(X_1 - 2X_2)^2}{20} + \frac{(3X_3 - 4X_4)^2}{100}$$

服从自由度为 2 的 χ^2 分布, 从而 $a = \frac{1}{20}, b = \frac{1}{100}$, 于是 $ab = \frac{1}{2000}$.

例 6.4 设总体 $X \sim N(0, 3^2)$, $X_1, X_2, \dots, X_8$ 是来自总体 X 的简单随机样本, 则统计量

$$Y = \frac{X_1 + X_2 + X_3 + X_4}{\sqrt{X_5^2 + X_6^2 + X_7^2 + X_8^2}}$$

若 $X \sim N(0, 1), Y \sim \chi^2(n)$.

则 $\frac{X}{\sqrt{Y/n}} \sim t(n)$

服从 ().

(A) $\chi^2(4)$

(B) $t(4)$

(C) $F(4, 4)$

(D) $N(0, 4^2)$

解 应选 (B).

由于相互独立且服从正态分布的随机变量的线性组合仍然服从正态分布, 易见

$$\begin{aligned} U &= \frac{X_1 + X_2 + X_3 + X_4}{\sqrt{D(X_1 + X_2 + X_3 + X_4)}} \\ &= \frac{X_1 + X_2 + X_3 + X_4}{6} \sim N(0, 1). \end{aligned}$$

作为相互独立且服从标准正态分布的随机变量的平方和,

$$V = \frac{X_5^2}{9} + \frac{X_6^2}{9} + \frac{X_7^2}{9} + \frac{X_8^2}{9}$$

服从 χ^2 分布, 自由度为 4. 随机变量 U 和 V 显然相互独立, 且随机变量 Y 可以表示为

$$Y = \frac{(X_1 + X_2 + X_3 + X_4)/6}{\sqrt{\frac{X_5^2 + X_6^2 + X_7^2 + X_8^2}{9}}/4} = \frac{U}{\sqrt{V/4}},$$

由 t 分布的概念, 可知随机变量 Y 服从自由度为 4 的 t 分布, 故选 (B).

例 6.5 设 $X_1, X_2, \dots, X_{15}$ 是来自正态总体 $N(0, 9)$ 的简单随机样本, 则统计量

$$\frac{X_i - 0}{3} \sim N(0, 1), \left(\frac{X_i}{3} \right)^2 \sim \chi^2(1) \quad Y = \frac{1}{2} \frac{X_1^2 + X_2^2 + \dots + X_{10}^2}{X_{11}^2 + X_{12}^2 + \dots + X_{15}^2}$$

服从 ().

(A) $t(10)$

(B) $t(15)$

(C) $F(10, 5)$

(D) $F(5, 10)$

解 应选 (C).

由 χ^2 分布的概念, 知

$$\chi_1^2 = \frac{X_1^2 + X_2^2 + \cdots + X_{10}^2}{9} \text{ 和 } \chi_2^2 = \frac{X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2}{9}$$

分别服从自由度为 10 和 5 的 χ^2 分布, 且相互独立. 从而, 由 F 分布的概念, 知

$$Y = \frac{\frac{1}{2} \frac{X_1^2 + X_2^2 + \cdots + X_{10}^2}{9}}{\frac{X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2}{5}} = \frac{\chi_1^2}{\chi_2^2}$$

服从自由度为 (10, 5) 的 F 分布, 故选 (C).

例 6.6 设随机变量 $X \sim t(n)$, $Y \sim F(1, n)$, 给定 $\alpha (0 < \alpha < 0.5)$, 常数 c 满足 $P\{X > c\} = \alpha$, 则

$P\{Y > c^2\} = (\quad)$.

(A) α

(B) $1 - \alpha$

(C) 2α

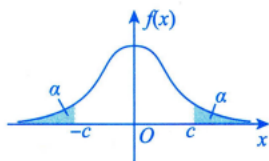
(D) $1 - 2\alpha$

解 应选 (C).

由 $X \sim t(n)$, 可得 $X^2 \sim F(1, n)$, 从而

$$P\{Y > c^2\} = P\{X^2 > c^2\} = P\{X > c\} + P\{X < -c\} = 2\alpha,$$

故正确选项为 (C).



注 此题只能有概率 $P\{Y > c^2\}$ 与 $P\{X^2 > c^2\}$ 相等, 而不能说 $Y = X^2$, 因为服从相同分布的随机变量并不一定相同.

例如: $X \sim U(0, 1)$,
 $Y = 1 - X \sim U(0, 1)$,
 X 和 Y 服从相同的分布但 $X \neq Y$

★★★ 三 参数的点估计

属于必考知识点

估计客观存在的未知参数, 接近实际参数值



设总体 X 的分布函数为 $F(x; \theta)$ (可以是多维的), 其中 θ 是一个未知参数, $X_1, X_2, \dots, X_n$ 是取自总体 X 的一个样本. 由样本构造一个适当的统计量 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 作为参数 θ 的估计, 则称统计量 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 为 θ 的估计量.

大写的样本 X

如果 $x_1, x_2, \dots, x_n$ 是样本的一个观测值, 将其代入估计量 $\hat{\theta}$ 中得值 $\hat{\theta}(x_1, x_2, \dots, x_n)$, 统计学中称这个值为未知参数 θ 的估计值.

小写的样本 x

小写时为样值

$\hat{\theta}$ 为估计值

具体的数

建立一个适当的统计量作为未知参数 θ 的估计量, 并以相应的观测值作为未知参数估计值的问题, 称为参数 θ 的点估计问题.

“以点估点”, 即由样本求出参数 θ 的估计值.

$\hat{\theta} \rightarrow \theta$ 两者是不等价的, 因为当 $\hat{\theta}$ 为连续型变量时, $P\{\hat{\theta} = \theta\} = 0$
估计值 $\leftarrow$ $\hat{\theta}$ $\rightarrow$ 客观参数 θ

均值 $\bar{X}$ 与期望 EX 是不相等的, 若令

$\bar{X} \xrightarrow[\text{数据}]{\text{强行}} EX$
相等 (客观), 从而建立方程, 可求出参数 θ

1 矩估计

- (1) 对于一个参数
- ① 用一阶矩建立方程: 令 $\bar{X} = EX$,
一阶原点矩
 - ② 若“①”不能用, 用二阶矩建立方程: 令 $\frac{1}{n} \sum_{i=1}^n X_i^2 = E(X^2)$.

① 中等式没有参数 θ 或解出 θ 很困难

求解方法: ① $Y = X^2$, 求 EY ;

② $E(X^2) = DX + (EX)^2$

一个方程解出一个参数即可作为矩估计.

- (2) 对于两个参数, 用一阶矩与二阶矩建立两个方程, 即① $\bar{X} = EX$ 与② $\frac{1}{n} \sum_{i=1}^n X_i^2 = E(X^2)$, 两个方

程解出两个参数即可作为矩估计.

k 不是随便取的, 优先用低阶的, 能用一阶绝不用二阶, 一阶不能用时才考虑二阶

注 显然, 一般情况下, 将 $\frac{1}{n} \sum_{i=1}^n X_i^k = E(X^k) (k=1, 2, \dots, m)$ 联立, 即可求 m 个参数的矩估计. 那

么, 若求一个参数的矩估计, 是否 k 可任取一个正整数呢? 答案是否定的, 在统计的问题中, 一般以低阶为原则, 即取能解出含参数的方程的最小 k 值.

2 最大似然估计

对未知参数 θ 进行估计时, 在该参数可能的取值范围内选取, 使“样本获此观测值 $x_1, x_2, \dots, x_n$ ”的概率最大的参数值 $\hat{\theta}$ 作为 θ 的估计, 这样选定的 $\hat{\theta}$ 最有利于 $x_1, x_2, \dots, x_n$ 的出现, 即“参数 $\theta = ?$ 时, 观测值出现的概率最大”.

即 $P\{x_1, x_2, \dots, x_n | \theta = ?\}$ 为最大.

$X_i \stackrel{\text{iid}}{\sim} F (i=1, 2, 3)$ 为连续型随机变量时, (x_1, x_2, x_3) 出现的概率 $= f(x_1, x_2, x_3; \theta) dx_1 dx_2 dx_3$

- (1) 写似然函数 $L(x_1, x_2, \dots, x_n; \theta) = \begin{cases} \prod_{i=1}^n p(x_i; \theta) & \text{(这是离散型总体 } X \text{ 取 } x_1, x_2, \dots, x_n \text{ 的概率),} \\ \prod_{i=1}^n f(x_i; \theta) & \text{(这是连续型总体 } X \text{ 取 } x_1, x_2, \dots, x_n \text{ 的联合概率密度).} \end{cases}$
- 为方便求解 $\hat{\theta}$, 也可取对数, 即
- $\ln L(\theta) = \begin{cases} \sum_{i=1}^n \ln p(x_i; \theta), \\ \sum_{i=1}^n \ln f(x_i; \theta) \end{cases}$
- 独立 $\rightarrow$ $p(x_1, x_2, \dots, x_n; \theta) = p_1(x_1; \theta) p_2(x_2; \theta) \cdots p_n(x_n; \theta) = \prod_{i=1}^n p(x_i; \theta)$
- 独立 $\rightarrow$ $f(x_1, x_2, \dots, x_n; \theta) = f_1(x_1; \theta) f_2(x_2; \theta) \cdots f_n(x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$
- 同分布 $\rightarrow$

- (2) 求参数 $\begin{cases} \text{若似然函数有驻点, 则令 } \frac{dL}{d\theta} = 0 \text{ 或 } \frac{d(\ln L)}{d\theta} = 0, \text{ 解出 } \hat{\theta}, \\ \text{若似然函数无驻点(单调), 则用定义求 } \hat{\theta}, \\ \text{若似然函数为常数, 则用定义求 } \hat{\theta}, \text{ 此时 } \hat{\theta} \text{ 不唯一.} \end{cases}$

(3) 最大似然估计量的不变性原则.

设 $\hat{\theta}$ 是总体分布中未知参数 θ 的最大似然估计, 函数 $u = u(\theta)$ 具有反函数, 则 $\hat{u} = u(\hat{\theta})$ 是 $u(\theta)$ 的最大似然估计, 如 $\hat{\theta}$ 是 θ 的最大似然估计, 则 $e^{\hat{\theta}}$ 是 e^{θ} 的最大似然估计.

3 估计量的评价标准 (仅数学一)

(1) 无偏性.

对于估计量 $\hat{\theta}$, 若 $E\hat{\theta} = \theta$, 则称 $\hat{\theta}$ 为 θ 的无偏估计量.

(2) 有效性 (最小方差性).

若 $E\hat{\theta}_1 = \theta$, $E\hat{\theta}_2 = \theta$, 即 $\hat{\theta}_1, \hat{\theta}_2$ 均是 θ 的无偏估计量, 当 $D\hat{\theta}_1 < D\hat{\theta}_2$ 时, 称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

(3) 一致性 (相合性) (只针对大样本 $n \rightarrow \infty$).

若 $\hat{\theta}$ 为 θ 的估计量, 对任意 $\varepsilon > 0$, 有

$$\star \lim_{n \rightarrow \infty} P\{|\hat{\theta} - \theta| \geq \varepsilon\} = 0,$$

或

$$\star \lim_{n \rightarrow \infty} P\{|\hat{\theta} - \theta| < \varepsilon\} = 1,$$

即当 $\hat{\theta} \xrightarrow{P} \theta$ 时, 称 $\hat{\theta}$ 为 θ 的一致 (或相合) 估计.

$\hat{\theta}$ 依概率收敛到 θ

一般用以下两种方法:

① 切比雪夫不等式 $P\{|X - EX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2}$.

② 辛钦大数定律 (独立同分布、 EX 存在)

$$\Rightarrow \bar{X} \xrightarrow{P} EX$$

均值依概率收敛到数学期望

4 估计量的数字特征与收敛性 (仅数学三)

(1) 求 $E\hat{\theta}$.

(2) 求 $D\hat{\theta}$.

(3) 验证 $\hat{\theta}$ 是否依概率收敛到 θ , 即是否有 $\hat{\theta} \xrightarrow{P} \theta$, 即对任意 $\varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\{|\hat{\theta} - \theta| \geq \varepsilon\} = 0 \text{ 或 } \lim_{n \rightarrow \infty} P\{|\hat{\theta} - \theta| < \varepsilon\} = 1.$$

例 6.7 设总体 X 的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $\theta \left(0 < \theta < \frac{1}{2}\right)$ 是未知参数, 从总体 X 中抽取容量为 8 的一组样本, 其样本值为 3, 1, 3, 0, 3, 1, 2, 3.

求 θ 的矩估计值和最大似然估计值.

解 由于总体分布仅含一个未知参数, 因此矩法方程为 $EX = \bar{X}$, 其中

$$\bar{x} = \frac{1}{8} \sum_{i=1}^8 x_i,$$

概率论知识去计算 $EX = 0 \times \theta^2 + 1 \times 2\theta(1-\theta) + 2\theta^2 + 3(1-2\theta) = 3 - 4\theta$,

令 $3 - 4\theta = \bar{X}$, 解得 θ 的矩估计量为 $\hat{\theta} = \frac{3 - \bar{X}}{4}$, 由样本值算得

$$\bar{x} = \frac{1}{8} \times (3 + 1 + 3 + 0 + 3 + 1 + 2 + 3) = 2,$$

所以 θ 的矩估计值为 $\hat{\theta} = \frac{3-2}{4} = \frac{1}{4}$.

已知总体 X 的概率分布为 $p_i = P\{X = x_i\} = p(x_i; \theta)$, 故样本值 $x_i (1 \leq i \leq 8)$ 的似然函数为

$$\begin{aligned} L(\theta) &= L(x_1, \dots, x_8; \theta) = \prod_{i=1}^8 p(x_i; \theta) = \theta^2 [2\theta(1-\theta)]^2 \theta^2 (1-2\theta)^4 \\ &= 4\theta^6 (1-\theta)^2 (1-2\theta)^4, \end{aligned}$$

取对数,

$$\ln L(\theta) = \ln 4 + 6 \ln \theta + 2 \ln(1-\theta) + 4 \ln(1-2\theta),$$

令

$$\frac{d[\ln L(\theta)]}{d\theta} = \frac{6}{\theta} - \frac{2}{1-\theta} - \frac{8}{1-2\theta} = \frac{24\theta^2 - 28\theta + 6}{\theta(1-\theta)(1-2\theta)} = 0,$$

即 $12\theta^2 - 14\theta + 3 = 0$, 解得 $\theta = \frac{7 \pm \sqrt{13}}{12}$, 因为 $\frac{7 + \sqrt{13}}{12} > \frac{1}{2}$ 不合题意, 故 θ 最大似然估计值为 $\hat{\theta} = \frac{7 - \sqrt{13}}{12}$.

注 似然方程满足条件的解 $\hat{\theta}$ 一般就认为是最大似然估计而不加以验证. 原则上是需要证明的, 然而有些问题证明是不容易的, 甚至是不可能的. 如果能断言 $L(\theta)$ 有最大值点, 而且似然方程只有唯一解 $\hat{\theta}$, 则 $\hat{\theta}$ 为最大似然估计. 此外有些问题可用微积分方法来验证, 如本例. 事实上, 由于

$$\frac{d^2[\ln L(\theta)]}{d\theta^2} = \frac{-6}{\theta^2} - \frac{2}{(1-\theta)^2} - \frac{16}{(1-2\theta)^2} < 0,$$

因此 $\hat{\theta} = \frac{7 - \sqrt{13}}{12}$ 使 $L(\theta)$ 达到最大, 因而满足条件的解为 $\hat{\theta} = \frac{7 - \sqrt{13}}{12}$.

例 6.8

设某种产品的使用寿命 T 的分布函数 $F(t)$ 满足方程 $F'(t) + \frac{2t}{\theta^2}[F(t)-1] = 0, t \geq 0$,

$F(0)=0$, 其中 θ 为大于 1 的参数, 且产品性能 $Q(\theta) = \frac{\theta^2}{2} \ln \theta - \frac{3}{4}\theta^2 + \theta$.

求解该微分方程, 得到 $F(t)$

(1) 求概率 $P\{T > t\}$ 与 $P\{T > s+t | T > s\}$, 其中 $s > 0, t > 0$;

(2) 任取 n 个这种产品做寿命试验, 测得它们的寿命分别为 $t_1, t_2, \dots, t_n$, 求 θ 的最大似然估计值 $\hat{\theta}$;

(3) 求产品性能 Q 的最大似然估计值 $\hat{Q}$.

解 首先求 $F(t)$, 由 $F'(t) + \frac{2t}{\theta^2} F(t) = \frac{2t}{\theta^2}$, 知

一阶线性微分方程, 直接用公式法求解

$$\begin{aligned} F(t) &= e^{-\int \frac{2t}{\theta^2} dt} \left(\int e^{\int \frac{2t}{\theta^2} dt} \frac{2t}{\theta^2} dt + C \right) \\ &= e^{-\frac{t^2}{\theta^2}} \left[\int e^{\frac{t^2}{\theta^2}} d\left(\frac{t^2}{\theta^2}\right) + C \right] \\ &= e^{-\frac{t^2}{\theta^2}} \left(e^{\frac{t^2}{\theta^2}} + C \right) = 1 + Ce^{-\frac{t^2}{\theta^2}}, t \geq 0. \end{aligned}$$

由 $F(0)=0$, 知 $C=-1$, 故

$$F(t) = \begin{cases} 1 - e^{-\frac{t^2}{\theta^2}}, & t \geq 0, \\ 0, & \text{其他.} \end{cases}$$

(1) 由条件知

$$\begin{aligned} P\{T > t\} &= e^{-\frac{t^2}{\theta^2}}, \\ P\{T > s+t | T > s\} &= \frac{P\{T > s+t, T > s\}}{P\{T > s\}} \\ &= \frac{P\{T > s+t\}}{P\{T > s\}} \\ &= e^{-\frac{t^2+2ts}{\theta^2}}. \end{aligned}$$

(2) 总体 T 的概率密度为

$$f(t; \theta) = \begin{cases} \frac{2t}{\theta^2} e^{-\frac{t^2}{\theta^2}}, & t > 0, \\ 0, & \text{其他.} \end{cases}$$

似然函数为

$$L(\theta) = \begin{cases} 2^n \theta^{-2n} t_1 t_2 \cdots t_n e^{-\frac{1}{\theta^2} \sum_{i=1}^n t_i^2}, & t_1, t_2, \dots, t_n > 0, \\ 0, & \text{其他.} \end{cases}$$

两边同时取对数

当 $t_1, t_2, \dots, t_n > 0$ 时, $\ln L(\theta) = n \ln 2 - 2n \ln \theta + \sum_{i=1}^n \ln t_i - \frac{1}{\theta^2} \sum_{i=1}^n t_i^2$.

令 $\frac{d[\ln L(\theta)]}{d\theta} = 0$, 得

$$-\frac{2n}{\theta} + \frac{2}{\theta^3} \sum_{i=1}^n t_i^2 = 0,$$

从而 θ 的最大似然估计值为 $\hat{\theta} = \left(\frac{1}{n} \sum_{i=1}^n t_i^2 \right)^{\frac{1}{2}}$.

(3) 由

$$Q'(\theta) = \theta \ln \theta + \frac{\theta}{2} - \frac{3}{2} \theta + 1 = \theta \ln \theta - \theta + 1$$

$$\underline{\underline{\text{令 } \theta - 1 = u}} (u+1) \ln(u+1) - u,$$

根据不等式 $\frac{u}{u+1} < \ln(u+1)$, $u > 0$, 知 $Q'(\theta) > 0$. 于是 $Q = Q(\theta)$ 有反函数, 且由 (2) 可知, Q 的最大似然估计值

$Q(\theta)$ 单调增加

$$\hat{Q} = \left(\frac{1}{2n} \sum_{i=1}^n t_i^2 \right) \ln \left(\frac{1}{n} \sum_{i=1}^n t_i^2 \right)^{\frac{1}{2}} - \frac{3}{4n} \sum_{i=1}^n t_i^2 + \left(\frac{1}{n} \sum_{i=1}^n t_i^2 \right)^{\frac{1}{2}}.$$

 例 6.9 设总体 X 的概率密度为

$$f(x; \theta) = \begin{cases} \frac{1}{1-\theta}, & \theta \leq x \leq 1, \\ 0, & \text{其他,} \end{cases}$$

其中 θ 为未知参数, $X_1, X_2, \dots, X_n$ 为来自该总体的简单随机样本.

(1) 求 θ 的矩估计量;

(2) 求 θ 的最大似然估计量.

 解 (1) 由于总体 X 服从区间 $[\theta, 1]$ 上的均匀分布, 因此 $EX = \frac{1+\theta}{2}$.

由 $\frac{1+\theta}{2} = \bar{X}$, 其中 $\bar{X}$ 为样本均值, 得 θ 的矩估计量 $\hat{\theta} = 2\bar{X} - 1$.

(2) 设 $x_1, x_2, \dots, x_n$ 为样本 $X_1, X_2, \dots, X_n$ 的观测值, 则似然函数为

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta) = \begin{cases} \frac{1}{(1-\theta)^n}, & \theta \leq x_i \leq 1 (i=1, 2, \dots, n), \\ 0, & \text{其他} \end{cases}$$

$$= \begin{cases} \frac{1}{(1-\theta)^n}, & \theta \leq \min\{x_1, x_2, \dots, x_n\}, x_i \leq 1, \\ 0, & \text{其他.} \end{cases}$$

由此可知, 当 $\theta = \min\{x_1, x_2, \dots, x_n\}$ 时, $L(\theta)$ 达到最大, 故 θ 的最大似然估计量为

$$\hat{\theta} = \min\{X_1, X_2, \dots, X_n\}.$$

例 6.10 (仅数学一) 设 $X_1, X_2, \dots, X_n$ 是来自正态总体 $X \sim N(\mu, \sigma^2)$ 的简单随机样本, 为使

$$D = k \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$$

成为总体方差 σ^2 的无偏估计量, 则 $k = (\quad)$.

- (A) $\frac{1}{n-1}$ (B) $\frac{1}{n}$ (C) $\frac{1}{2(n-1)}$ (D) $\frac{1}{2n}$

解 应选 (C).

由条件知 $E(X^2) = \sigma^2 + \mu^2$. 假如统计量 D 是总体方差 σ^2 的无偏估计量, 则

$$\begin{aligned} ED &= k \sum_{i=1}^{n-1} E[(X_{i+1} - X_i)^2] = k \sum_{i=1}^{n-1} E(X_{i+1}^2 + X_i^2 - 2X_i X_{i+1}) \\ &= k \sum_{i=1}^{n-1} [2E(X_i^2) - 2EX_i EX_{i+1}] = k \sum_{i=1}^{n-1} (2\sigma^2 + 2\mu^2 - 2\mu^2) \\ &= 2k(n-1)\sigma^2 = \sigma^2, \end{aligned}$$

于是 $k = \frac{1}{2(n-1)}$, 故选 (C).

(仅数学三) 设 $X_1, X_2, \dots, X_n$ 是来自正态总体 $X \sim N(\mu, \sigma^2)$ 的简单随机样本, 记

$$D = k \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2,$$

若 $ED = \sigma^2$, 则 $k = (\quad)$.

- (A) $\frac{1}{n-1}$ (B) $\frac{1}{n}$ (C) $\frac{1}{2(n-1)}$ (D) $\frac{1}{2n}$

解 应选 (C).

由条件知 $E(X^2) = \sigma^2 + \mu^2$. 又 $ED = \sigma^2$, 则

$$\begin{aligned} ED &= k \sum_{i=1}^{n-1} E[(X_{i+1} - X_i)^2] = k \sum_{i=1}^{n-1} E(X_{i+1}^2 + X_i^2 - 2X_i X_{i+1}) \\ &= k \sum_{i=1}^{n-1} [2E(X_i^2) - 2EX_i EX_{i+1}] = k \sum_{i=1}^{n-1} (2\sigma^2 + 2\mu^2 - 2\mu^2) \\ &= 2k(n-1)\sigma^2 = \sigma^2, \end{aligned}$$

$E(X_i^2) = DX_i + (EX_i)^2$

于是 $k = \frac{1}{2(n-1)}$, 故选 (C).

例 6.11 设总体 $X \sim N(\mu, \sigma^2)$, X_1, X_2, X_3 是来自总体 X 的简单随机样本, 记

$$\hat{\mu}_1 = \frac{1}{5}X_1 + \frac{3}{10}X_2 + \frac{1}{2}X_3,$$

$$\hat{\mu}_2 = \frac{1}{3}X_1 + \frac{1}{4}X_2 + \frac{5}{12}X_3,$$

$$\hat{\mu}_3 = \frac{1}{3}X_1 + \frac{1}{6}X_2 + \frac{1}{2}X_3.$$

(仅数学一) 证明估计量 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 都是 μ 的无偏估计, 并指出它们中哪一个最有效.

(仅数学三) 求估计量 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 的数学期望与方差.

(仅数学一) **证** 因

$$E\hat{\mu}_1 = \frac{1}{5}EX_1 + \frac{3}{10}EX_2 + \frac{1}{2}EX_3 = \left(\frac{1}{5} + \frac{3}{10} + \frac{1}{2}\right)\mu = \mu, \quad \text{EX}$$

直接线性拆分

$$E\hat{\mu}_2 = \left(\frac{1}{3} + \frac{1}{4} + \frac{5}{12}\right)\mu = \mu,$$

$$E\hat{\mu}_3 = \left(\frac{1}{3} + \frac{1}{6} + \frac{1}{2}\right)\mu = \mu,$$

故 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 都是 μ 的无偏估计.

又

$$D\hat{\mu}_1 = \frac{1}{25}DX_1 + \frac{9}{100}DX_2 + \frac{1}{4}DX_3 = \frac{19}{50}\sigma^2,$$

$$D\hat{\mu}_2 = \left(\frac{1}{9} + \frac{1}{16} + \frac{25}{144}\right)\sigma^2 = \frac{25}{72}\sigma^2,$$

$$D\hat{\mu}_3 = \left(\frac{1}{9} + \frac{1}{36} + \frac{1}{4}\right)\sigma^2 = \frac{7}{18}\sigma^2,$$

$$D\hat{\mu}_2 < D\hat{\mu}_1 < D\hat{\mu}_3,$$

所以 $\hat{\mu}_2$ 最有效.

(仅数学三) **解**

$$E\hat{\mu}_1 = \frac{1}{5}EX_1 + \frac{3}{10}EX_2 + \frac{1}{2}EX_3 = \left(\frac{1}{5} + \frac{3}{10} + \frac{1}{2}\right)\mu = \mu,$$

$$E\hat{\mu}_2 = \left(\frac{1}{3} + \frac{1}{4} + \frac{5}{12}\right)\mu = \mu,$$

$$E\hat{\mu}_3 = \left(\frac{1}{3} + \frac{1}{6} + \frac{1}{2}\right)\mu = \mu,$$

$$D\hat{\mu}_1 = \frac{1}{25}DX_1 + \frac{9}{100}DX_2 + \frac{1}{4}DX_3 = \frac{19}{50}\sigma^2,$$

$$D\hat{\mu}_2 = \left(\frac{1}{9} + \frac{1}{16} + \frac{25}{144}\right)\sigma^2 = \frac{25}{72}\sigma^2,$$

$$D\hat{\mu}_3 = \left(\frac{1}{9} + \frac{1}{36} + \frac{1}{4}\right)\sigma^2 = \frac{7}{18}\sigma^2.$$

例 6.12 设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的一个简单随机样本, $EX = \mu, DX = \sigma^2$.

(仅数学一) 证明统计量 $Y = \frac{2}{n(n+1)} \sum_{i=1}^n iX_i$ 是 μ 的无偏相合估计量.

(仅数学三) 求统计量 $Y = \frac{2}{n(n+1)} \sum_{i=1}^n iX_i$ 的数学期望, 并证明 Y 依概率收敛到 μ .

(仅数学一) **证** 由于

$$\begin{aligned} EY &= \frac{2}{n(n+1)} \sum_{i=1}^n E(iX_i) = \frac{2}{n(n+1)} \sum_{i=1}^n iEX_i \\ &= \frac{2}{n(n+1)} (\mu + 2\mu + \dots + n\mu) \\ &= \mu, \end{aligned}$$

从而知, Y 是 μ 的无偏估计量.

又因

$$\begin{aligned} DY &= \left[\frac{2}{n(n+1)} \right]^2 \sum_{i=1}^n i^2 DX_i = \frac{4\sigma^2}{n^2(n+1)^2} (1^2 + 2^2 + \dots + n^2) \\ &= \frac{4\sigma^2}{n^2(n+1)^2} \cdot \frac{1}{6} n(n+1)(2n+1) \\ &= \frac{2}{3} \cdot \frac{2n+1}{n(n+1)} \sigma^2, \end{aligned}$$

于是由切比雪夫不等式, 对任意实数 $\varepsilon > 0$, 有

$$1 \geq P\{|Y - \mu| < \varepsilon\} \geq 1 - \frac{DY}{\varepsilon^2} = 1 - \frac{2}{3} \cdot \frac{(2n+1)\sigma^2}{n(n+1)\varepsilon^2},$$

两边取极限, 由 $\lim_{n \rightarrow \infty} 1 \geq \lim_{n \rightarrow \infty} \left[1 - \frac{2}{3} \cdot \frac{(2n+1)\sigma^2}{n(n+1)\varepsilon^2} \right] = 1$, 有

$$\lim_{n \rightarrow \infty} P\{|Y - \mu| < \varepsilon\} = 1, \quad \leftarrow \text{这是我们的目标}$$

所以 Y 是 μ 的无偏相合估计量.

(仅数学三) **解**

$$\begin{aligned} EY &= \frac{2}{n(n+1)} \sum_{i=1}^n E(iX_i) = \frac{2}{n(n+1)} \sum_{i=1}^n iEX_i \\ &= \frac{2}{n(n+1)} (\mu + 2\mu + \dots + n\mu) \\ &= \mu, \end{aligned}$$

$$DY = \left[\frac{2}{n(n+1)} \right]^2 \sum_{i=1}^n i^2 DX_i = \frac{4\sigma^2}{n^2(n+1)^2} (1^2 + 2^2 + \dots + n^2)$$

$$= \frac{4\sigma^2}{n^2(n+1)^2} \cdot \frac{1}{6} n(n+1)(2n+1)$$

$$= \frac{2}{3} \cdot \frac{2n+1}{n(n+1)} \sigma^2,$$

于是由切比雪夫不等式, 对任意实数 $\varepsilon > 0$, 有

$$1 \geq P\{|Y - \mu| < \varepsilon\} \geq 1 - \frac{DY}{\varepsilon^2} = 1 - \frac{2}{3} \cdot \frac{(2n+1)\sigma^2}{n(n+1)\varepsilon^2},$$

两边取极限, 由 $\lim_{n \rightarrow \infty} 1 \geq \lim_{n \rightarrow \infty} \left[1 - \frac{2}{3} \cdot \frac{(2n+1)\sigma^2}{n(n+1)\varepsilon^2} \right] = 1$, 有

$$\lim_{n \rightarrow \infty} P\{|Y - \mu| < \varepsilon\} = 1,$$

所以 Y 依概率收敛到 μ .



★★★ 四 参数的区间估计 (仅数学一)

区别于点估计, $\hat{\theta}_1 \rightarrow \hat{\theta}_2$, 使参数 θ 落在区间里面, 故比点估计更加科学

1 概念

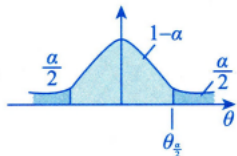
设 θ 是总体 X 的分布函数的一个未知参数, 对于给定 $\alpha (0 < \alpha < 1)$, 如果由样本 $X_1, X_2, \dots, X_n$ 确定的两个统计量 $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$, $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$, 使

$$P\{\hat{\theta}_1(X_1, X_2, \dots, X_n) < \theta < \hat{\theta}_2(X_1, X_2, \dots, X_n)\} = 1 - \alpha,$$

即 θ 落在区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 上的可能性为 $1 - \alpha$

则称随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 是 θ 的置信度为 $1 - \alpha$ 的置信区间, $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 分别称为 θ 的置信度为 $1 - \alpha$ 的双侧置信区间的置信下限和置信上限, $1 - \alpha$ 称为置信度或置信水平, α 称为显著性水平. 如果

$P\{\theta < \hat{\theta}_1\} = P\{\theta > \hat{\theta}_2\} = \frac{\alpha}{2}$, 则称这种置信区间为等尾置信区间.



2 单个正态总体均值和方差的置信区间

考前记一记, 考前摇一摇, 即可

设 $X \sim N(\mu, \sigma^2)$, 从总体 X 中抽取样本 $X_1, X_2, \dots, X_n$, 样本均值为 $\bar{X}$, 样本方差为 S^2 .

(1) σ^2 已知, μ 的置信水平是 $1 - \alpha$ 的置信区间为

也称置信度

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}} \right)$$

记为 I_1 $\rightarrow P\{\mu \in I_1\} = 1 - \alpha = P\{\bar{X} - \Delta < \mu < \bar{X} + \Delta\}$

Δ Δ

(2) σ^2 未知, μ 的置信水平是 $1-\alpha$ 的置信区间为

用样本方差 S^2 代替方差 σ^2 , S 是可以继续抽样数据

$$\left(\underbrace{\bar{X} - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)}_{\Delta}, \underbrace{\bar{X} + \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)}_{\Delta} \right) \xrightarrow{\text{记为 } I_2} P\{\mu \in I_2\} = 1-\alpha = P\{\bar{X} - \Delta < \mu < \bar{X} + \Delta\}$$

(3) μ 已知, σ^2 的置信水平是 $1-\alpha$ 的置信区间为

$$\left(\underbrace{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\frac{\alpha}{2}}^2(n)}}_{\Delta_1}, \underbrace{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\frac{\alpha}{2}}^2(n)}}_{\Delta_2} \right) \xrightarrow{\text{记为 } I_3} P\{\sigma^2 \in I_3\} = 1-\alpha = P\{\Delta_1 < \sigma^2 < \Delta_2\},$$

此种情况一般不出现

(4) μ 未知, σ^2 的置信水平是 $1-\alpha$ 的置信区间为

$$\left(\underbrace{\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}}^2(n-1)}}_{\Delta_1}, \underbrace{\frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}}^2(n-1)}}_{\Delta_2} \right) \xrightarrow{\text{记为 } I_4} P\{\sigma^2 \in I_4\} = 1-\alpha$$

$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$

注 (1) σ^2 已知时, 估 μ .

令 $P\left\{\left|\bar{X} - EX\right| < \Delta\right\} = 1-\alpha$ (人为给定的), 即令 $\bar{X}$ 到 μ 的距离小于某个正数的概率为 $1-\alpha$.

抽样数据 某个正数

由于 $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, 记 $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$.

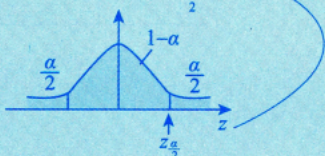
则 $P\left\{\left|\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right| < \frac{\Delta}{\sigma/\sqrt{n}}\right\} = 1-\alpha$, 即 $\frac{\Delta}{\sigma/\sqrt{n}} = z_{\frac{\alpha}{2}}$, 解得 $\Delta = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$. $\Rightarrow P\{\bar{X} - \Delta < \mu < \bar{X} + \Delta\} = 1-\alpha$

标准差 $\frac{\sigma}{\sqrt{n}}$ 上 α 分位数 $z_{\frac{\alpha}{2}}$ 样本容量 n

$\mu \in \left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$

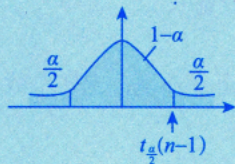
置信区间

故称 $\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$ 为置信度为 $1-\alpha$ 的置信区间.



(2) σ^2 未知时, 估 μ .

令 $P\left\{\left|\frac{\bar{X}-\mu}{S/\sqrt{n}}\right| < \frac{\Delta}{S/\sqrt{n}}\right\} = 1-\alpha$, 则 $P\left\{|T| < \frac{\Delta}{S/\sqrt{n}}\right\} = 1-\alpha$, 即 $\frac{\Delta}{S/\sqrt{n}} = t_{\frac{\alpha}{2}}(n-1)$, 解得 $\Delta = t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}}$.



(3) μ 已知时, 估 σ^2 .

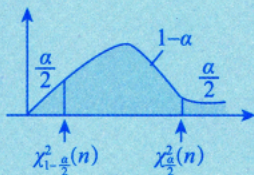
若 $P\{\Delta_1 < \sigma^2 < \Delta_2\} = 1-\alpha$, 则

$$P\left\{\frac{1}{\Delta_2} < \frac{1}{\sigma^2} < \frac{1}{\Delta_1}\right\} = 1-\alpha$$

$$= P\left\{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\Delta_2} < \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} < \frac{\sum_{i=1}^n (X_i - \mu)^2}{\Delta_1}\right\},$$

$\chi^2_{1-\frac{\alpha}{2}}(n)$ $\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$ $\chi^2_{\frac{\alpha}{2}}(n)$

故 $\Delta_1 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}$, $\Delta_2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}$.



(4) μ 未知时, 估 σ^2 .

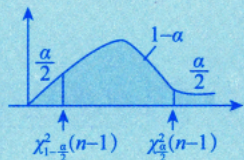
若 $P\{\Delta_1 < \sigma^2 < \Delta_2\} = 1-\alpha$, 则

$$P\left\{\frac{1}{\Delta_2} < \frac{1}{\sigma^2} < \frac{1}{\Delta_1}\right\} = 1-\alpha$$

$$= P\left\{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\Delta_2} < \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} < \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\Delta_1}\right\},$$

$\chi^2_{1-\frac{\alpha}{2}}(n-1)$ 服从 $\chi^2(n-1)$ $\chi^2_{\frac{\alpha}{2}}(n-1)$

故 $\Delta_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}$, $\Delta_2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$.



例 6.13

设总体 $X \sim N(\mu, \sigma^2)$ (σ^2 已知), 则在给定样本容量 n 及置信度 $1-\alpha$ 的情况下, 未知参数 μ 的置信区间长度随着样本均值 $\bar{X}$ 的增加而 ().

(A) 增加

(B) 减少

(C) 不变

(D) 不能确定增或减

分析 μ $\xrightarrow{\quad}$ $\overline{X}-\Delta$ $\overline{X}+\Delta$, $\Delta = \begin{cases} z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, & \sigma^2 \text{ 已知,} \\ t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}}, & \sigma^2 \text{ 未知,} \end{cases}$ 故区间长度 $L = 2\Delta = 2z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$. 与 $\overline{X}$ 无关

解 应选 (C).

在给定条件下, 未知参数 μ 的置信区间为 $\left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)$, 其中 $\Phi\left(z_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2}$, 可见该

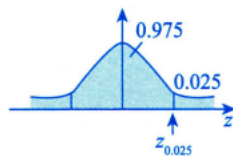
区间长度 $2 \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}$ 与 $\overline{X}$ 无关, 故选择 (C).

例 6.14 设总体 X 的方差为 1, 根据来自 X 的容量为 100 的简单随机样本测得样本均值为 5, 又 $\Phi(1.96) = 0.975$, 则 X 的数学期望的置信度近似等于 0.95 的置信区间为_____.

分析 由中心极限定理, 若 $X_i \stackrel{\text{i.i.d.}}{\sim} F$, 则当 $n \rightarrow \infty$ 时, $\sum_{i=1}^n X_i$ 服从 $N(n\mu, n\sigma^2)$, 故

$$\sum_{i=1}^{100} X_i \sim N(100\mu, 100), \overline{X} = \frac{1}{100} \sum_{i=1}^{100} X_i \sim N\left(\mu, \frac{1}{100}\right).$$

μ $\xrightarrow{\quad}$ $\overline{X}-\Delta$ $\overline{X}+\Delta$, 已知 $\sigma^2 = 1$, 故 $\Delta = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = z_{\frac{\alpha}{2}} \cdot \frac{1}{10} = 1.96 \times 0.1 = 0.196$.
 $(5 - 0.196, 5 + 0.196)$



$$\left. \begin{aligned} 1 - \alpha = 0.95, \alpha = 0.05, \frac{\alpha}{2} = 0.025 \\ 1 - \Phi(1.96) = 0.025 \end{aligned} \right\} \Rightarrow z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$$

解 应填 (4.804, 5.196).

尽管未说明 X 是否为正态总体, 但是由于样本容量 $n = 100$ 属于大样本, 因此由中心极限定理易得

$$Z = \frac{\overline{X} - \mu}{1/\sqrt{100}} \stackrel{\text{近似}}{\sim} N(0, 1),$$

又由 $\Phi(1.96) = 0.975$, 故 $P\{-1.96 < Z < 1.96\} = 0.95$, 即 μ 的置信度近似为 0.95 的置信区间为

$$\left(\overline{X} - \frac{1}{10} \times 1.96, \overline{X} + \frac{1}{10} \times 1.96\right),$$

将 $\overline{x} = 5$ 代入上式得置信区间为 (4.804, 5.196).

★★★ 五 假设检验 (仅数学一)

与估计的区别: 估计是通过实测数据 (样本值) 来反映背后客观存在的未知参数; 假设检验是对已有的说法 (命题) 进行检验, 判断是否正确

1 概念

关于总体 (分布中的未知参数, 分布的类型、特征、不相关性、独立性……) 的每一种论断 (“看法”) 称为统计假设, 然后根据样本观察数据或试验结果所提供的信息去推断 (检验) 这个 “看法” (即假设) 是否成立, 这类统计推断问题称为假设检验.



2 原假设与备择假设

常常把没有充分理由不能轻易否定的假设取为原假设(基本假设或零假设), 记为 H_0 , 将其否定的陈述(假设)称为对立假设或备择假设, 记为 H_1 .

3 小概率原理与显著性水平

(1) 小概率原理.

对假设进行检验的基本思想是采用某种带有概率性质的反证法. 这种方法的依据是小概率原理——概率很接近于 0 的事件在一次试验或观察中认为不会发生. 若小概率事件发生了, 则拒绝原假设.

(2) 显著性水平 α . 人为给定的

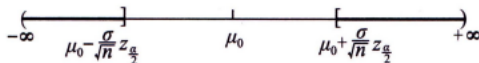
小概率事件中的“小概率”的值没有统一规定, 通常是根据实际问题的要求, 规定一个界限 $\alpha (0 < \alpha < 1)$, 当一个事件的概率不大于 α 时, 即认为它是小概率事件. 在假设检验问题中, α 称为显著性水平, 通常取 $\alpha = 0.1, 0.05, 0.01$ 等.

4 正态总体下的六大检验及拒绝域

考前记一记, 考前摇一摇, 即可

(1) σ^2 已知, μ 未知. $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$, 则拒绝域为 $\left(-\infty, \mu_0 - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right] \cup \left[\mu_0 + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, +\infty\right)$.

$$\Delta = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$



注 应熟知四、2 中 μ 的区间估计的相关知识. 以 σ^2 已知, 估 μ 为例.

$$P\left\{\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| < \frac{\Delta}{\sigma/\sqrt{n}}\right\} = 1 - \alpha, \quad \Delta = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}},$$

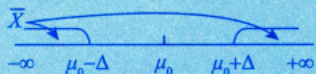
则

$$P\left\{\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \geq \frac{\Delta}{\sigma/\sqrt{n}}\right\} = \alpha,$$

故令

$$|\bar{X} - \mu_0| \geq z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}},$$

解得 $\bar{X} \geq \mu_0 + \Delta$ 或 $\bar{X} \leq \mu_0 - \Delta$.



则当小概率事件发生, 即 $\bar{X}$ 落到 $(-\infty, \mu_0 - \Delta] \cup [\mu_0 + \Delta, +\infty)$ 时, 拒绝原假设.

$$\left(-\infty, \mu_0 - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right] \cup \left[\mu_0 + \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1), +\infty\right),$$

代入数值, 计算得 $(-\infty, 9.4868] \cup [9.9132, +\infty)$, 由 $\bar{x}=10$ 知, 落入拒绝域, 故否定 H_0 , 即认为 $\mu \neq 9.7$.

六 两类错误 (仅数学一)



第一类错误 (“弃真”) : 若 H_0 为真, 按检验法则否定 H_0 , 此时犯了 “弃真” 的错误, 这种错误称为第一类错误, 犯第一类错误的概率为 $\alpha = P\{\text{拒绝 } H_0 | H_0 \text{ 为真}\}$. → 小概率事件的概率, 也称为显著性水平

第二类错误 (“取伪”) : 若 H_0 不真, 按检验法则接受 H_0 , 此时犯了 “取伪” 的错误, 这种错误称为第二类错误, 犯第二类错误的概率为 $\beta = P\{\text{接受 } H_0 | H_0 \text{ 为假}\} = P\{\text{接受 } H_0 | H_1 \text{ 为真}\}$.

→ 即两类错误不是对立事件

注 犯两类错误的概率 α 与 β , 并不满足 $\beta=1-\alpha$, 在固定样本容量 n 的条件下, α 小, β 就大; β 小, α 就大. 在实际应用中, 我们总是在控制 α 的条件下, 尽量使 β 小, 这是因为人们常常把拒绝 H_0 比错误地接受 H_0 看得更重要些.

例 6.16 假定 X 是连续型随机变量, U 是对 X 的一次观测值. 关于其概率密度 $f(x)$ 有如下假设:

$$H_0: f(x) = \begin{cases} \frac{1}{2}, & 0 \leq x \leq 2, \\ 0, & \text{其他}, \end{cases} \quad H_1: f(x) = \begin{cases} \frac{x}{2}, & 0 \leq x \leq 2, \\ 0, & \text{其他}, \end{cases}$$

检验规则: 当事件 $V = \left\{U > \frac{3}{2}\right\}$ 出现时, 否定假设 H_0 , 接受 H_1 . 则犯第一类错误的概率 α 与犯第二类错误的概率 β 分别为 ().

(A) $\frac{3}{4}, \frac{7}{16}$

(B) $\frac{7}{16}, \frac{3}{4}$

(C) $\frac{9}{16}, \frac{1}{4}$

(D) $\frac{1}{4}, \frac{9}{16}$

解 应选 (D).

由 α 和 β 的意义, 知

$$\alpha = P\left\{U > \frac{3}{2} \mid H_0\right\} = \int_{\frac{3}{2}}^2 \frac{1}{2} dx = \frac{1}{4},$$

$$\beta = P\left\{U \leq \frac{3}{2} \mid H_1\right\} = \int_0^{\frac{3}{2}} \frac{x}{2} dx = \frac{9}{16}.$$

故选 (D).

例 6.17 设 $X_1, X_2, \dots, X_{16}$ 是来自总体 $N(\mu, 4)$ 的简单随机样本, 考虑假设检验问题:

$H_0: \mu \leq 10, H_1: \mu > 10$. $\Phi(x)$ 表示标准正态分布函数. 若该检验问题的拒绝域为 $W = \{\bar{X} > 11\}$, 其中

$\bar{X} = \frac{1}{16} \sum_{i=1}^{16} X_i$, 则 $\mu = 11.5$ 时, 该检验犯第二类错误的概率为 ().

(A) $1 - \Phi(1)$

(B) $1 - \Phi(0.5)$

(C) $1 - \Phi(1.5)$

(D) $1 - \Phi(2)$

解 应选 (A).

检验犯第二类错误的概率为落在接受域的概率, 即

$$P = \{\bar{X} \leq 11\} = P\left\{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{11 - 11.5}{2/4}\right\} = \Phi(-1) = 1 - \Phi(1).$$

故选 (A).

基础习题精练

习题

6.1 设总体 X 服从正态分布 $N(\mu, \sigma^2)$, 其中 μ 已知, σ^2 未知. $X_1, X_2, \dots, X_n$ 是取自总体 X 的简单随机样本, 则下列样本函数中不是统计量的是 ().

(A) $\frac{1}{n} \sum_{i=1}^n X_i$

(B) $\max_{1 \leq i \leq n} \{X_i\}$

(C) $\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2$

(D) $\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$

6.2 设总体 X 与 Y 相互独立且都服从正态分布 $N(\mu, \sigma^2)$, $\bar{X}, \bar{Y}$ 是分别来自总体 X, Y , 容量都为 n 的样本的样本均值, 则当 n 固定时, 概率 $P\{|\bar{X} - \bar{Y}| > \sigma\}$ 的值随 σ 的增大而 ().

(A) 单调增大

(B) 单调减小

(C) 保持不变

(D) 增减不定

6.3 设 X_1, X_2, X_3, X_4 为来自总体 $N\left(0, \frac{1}{2}\right)$ 的简单随机样本, 则统计量 $Y = \frac{X_1 - X_2}{\sqrt{X_3^2 + X_4^2}}$ 服从 ().

(A) $N(0, 1)$

(B) $\chi^2(2)$

(C) $t(2)$

(D) $F(2, 2)$

6.4 设总体 $X \sim N(\mu_1, 4)$, $Y \sim N(\mu_2, 5)$, X 与 Y 相互独立, $X_1, \dots, X_8$ 和 $Y_1, \dots, Y_{10}$ 是分别来自总体 X 和 Y 的两个样本, S_X^2 与 S_Y^2 分别为两个样本的样本方差, 则 ().

- (A) $\frac{2S_x^2}{5S_y^2} \sim F(7, 9)$ (B) $\frac{5S_x^2}{2S_y^2} \sim F(7, 9)$ (C) $\frac{4S_x^2}{5S_y^2} \sim F(7, 9)$ (D) $\frac{5S_x^2}{4S_y^2} \sim F(7, 9)$

6.5 设总体 X 的概率分布为 $P\{X=1\}=\frac{1-\theta}{2}$, $P\{X=2\}=P\{X=3\}=\frac{1+\theta}{4}$. 利用来自总体 X 的样本值 1, 3, 2, 2, 1, 3, 1, 2, 可得 θ 的最大似然估计值为 ().

- (A) $\frac{1}{4}$ (B) $\frac{3}{8}$ (C) $\frac{1}{2}$ (D) $\frac{5}{8}$

6.6 (仅数学一) 一批零件的长度服从正态分布 $N(\mu, \sigma^2)$, 其中 μ, σ^2 均未知. 现从中随机抽取 16 个零件, 测得样本均值 $\bar{x}=20$ cm, 样本标准差 $s=1$ cm, 则 μ 的置信水平为 0.90 的置信区间为 ().

- (A) $\left(20 - \frac{1}{4}t_{0.05}(16), 20 + \frac{1}{4}t_{0.05}(16)\right)$ (B) $\left(20 - \frac{1}{4}t_{0.1}(16), 20 + \frac{1}{4}t_{0.1}(16)\right)$
 (C) $\left(20 - \frac{1}{4}t_{0.05}(15), 20 + \frac{1}{4}t_{0.05}(15)\right)$ (D) $\left(20 - \frac{1}{4}t_{0.1}(15), 20 + \frac{1}{4}t_{0.1}(15)\right)$

6.7 (仅数学一) 在假设检验中, 显著性水平 α 的意义是 ().

- (A) 原假设 H_0 成立, 经检验被拒绝的概率
 (B) 原假设 H_0 成立, 经检验被接受的概率
 (C) 原假设 H_0 不成立, 经检验被拒绝的概率
 (D) 原假设 H_0 不成立, 经检验被接受的概率

6.8 设总体 X 的概率密度为 $f(x)=\frac{1}{2}e^{-|x|}$ ($-\infty < x < +\infty$), $X_1, X_2, \dots, X_n$ 为来自总体 X 的简单随机样本, 其样本方差为 S^2 , 则 $E(S^2)=$ _____.

6.9 设总体 X 的概率密度为

$$f(x)=\begin{cases} \frac{6x}{\theta^3}(\theta-x), & 0 < x < \theta, \\ 0, & \text{其他,} \end{cases}$$

$X_1, X_2, \dots, X_n$ 是取自总体 X 的简单随机样本. 求:

- (1) θ 的矩估计量 $\hat{\theta}$;
 (2) $\hat{\theta}$ 的方差 $D\hat{\theta}$.

6.10 已知总体 X 的概率密度为

$$f(x)=\begin{cases} (1+\theta)x^\theta, & 0 < x < 1, \\ 0, & \text{其他,} \end{cases}$$

其中 $\theta > -1$ 是未知参数, 设 $X_1, X_2, \dots, X_n$ 为来自总体 X 的简单随机样本, 求 θ 的矩估计量和最大似然估计量.

6.11 设总体 X 的概率密度为 $f(x) = \begin{cases} \frac{2x}{\alpha^2}, & 0 \leq x \leq \alpha, \\ 0, & \text{其他,} \end{cases}$ 其中 $\alpha > 1$ 未知, $X_1, X_2, \dots, X_n$ 是来自总体

X 的简单随机样本.

(1) 求 $p = P\{0 < X < \sqrt{\alpha}\}$;

(2) 求 p 的最大似然估计量 $\hat{p}$.

6.12 设 $X_1, X_2, \dots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的一个简单随机样本, 记

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, T = \bar{X}^2 - \frac{1}{n} S^2.$$

(仅数学一) 证明: T 是 μ^2 的无偏估计量.

(仅数学三) 求 ET .

解答

6.1 (C) 解 由统计量的定义“不含任何未知参数的样本函数”, 即知不是统计量的选项应该是 (C), 因为选项 (C) 中含有未知参数 σ , 故选择 (C).

6.2 (C) 解 要计算概率需要知道 $\bar{X} - \bar{Y}$ 的分布. 依题意 $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, $\bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ 且相互独立, 故 $\bar{X} - \bar{Y} \sim N\left(0, \frac{2\sigma^2}{n}\right)$, 由此可知当 n 固定时, $P\{|\bar{X} - \bar{Y}| > \sigma\} = P\left\{\left|\frac{\bar{X} - \bar{Y}}{\sqrt{2\sigma/\sqrt{n}}}\right| > \sqrt{\frac{n}{2}}\right\}$ 与 σ 无关, 故所求概率值保持不变, 选择 (C). 事实上,

$$\begin{aligned} P\{|\bar{X} - \bar{Y}| > \sigma\} &= 1 - P\left\{\left|\frac{\bar{X} - \bar{Y}}{\sqrt{2\sigma/\sqrt{n}}}\right| \leq \sqrt{\frac{n}{2}}\right\} = 1 - \left[\Phi\left(\sqrt{\frac{n}{2}}\right) - \Phi\left(-\sqrt{\frac{n}{2}}\right)\right] \\ &= 2\left[1 - \Phi\left(\sqrt{\frac{n}{2}}\right)\right]. \end{aligned}$$

6.3 (C) 解 取 $X = X_1 - X_2$, 则 $EX = 0$, $DX = DX_1 + DX_2 = 1$, $X \sim N(0, 1)$.

又 $E(\sqrt{2}X_1) = 0$, $D(\sqrt{2}X_1) = 2DX_1 = 1$, 知 $\sqrt{2}X_1 \sim N(0, 1)$, 从而

$$Y_1 = (\sqrt{2}X_3)^2 + (\sqrt{2}X_4)^2 \sim \chi^2(2),$$

又 X 与 Y_1 相互独立, 因此 $Y = \frac{X}{\sqrt{\frac{Y_1}{2}}} \sim t(2)$, 即 Y 服从自由度为 2 的 t 分布.

6.4 (D) 解 如果 $aQ \sim \chi^2(m)$, $bT \sim \chi^2(n)$, Q 与 T 独立, 则 $\frac{aQ/m}{bT/n} = \frac{naQ}{mbT} \sim F(m, n)$. 依题意

$$\frac{7S_x^2}{4} \sim \chi^2(7), \frac{9S_y^2}{5} \sim \chi^2(9).$$

S_x^2 与 S_y^2 相互独立, 因此服从 $F(7, 9)$ 分布的统计量 $A \frac{S_x^2}{S_y^2}$, 其系数应是 $A = \frac{\frac{7}{4}/7}{\frac{9}{5}/9} = \frac{5}{4}$, 选择 (D).

注 通过 F 分布系数与自由度的关系确定选项.

6.5 (A) 解 由题意可知, 似然函数为

$$\begin{aligned} L(\theta) &= P\{X_1=1, X_2=3, X_3=2, X_4=2, X_5=1, X_6=3, X_7=1, X_8=2\} \\ &= \left(\frac{1-\theta}{2}\right)^3 \left(\frac{1+\theta}{4}\right)^3 \left(\frac{1+\theta}{4}\right)^2 \\ &= \frac{1}{2^{13}} (1-\theta)^3 (1+\theta)^5, \end{aligned}$$

对数似然函数为

$$\ln L(\theta) = -13 \ln 2 + 3 \ln(1-\theta) + 5 \ln(1+\theta),$$

令

$$\frac{d[\ln L(\theta)]}{d\theta} = -\frac{3}{1-\theta} + \frac{5}{1+\theta} = 0,$$

解得 $\theta = \frac{1}{4}$, 从而 θ 的最大似然估计值为 $\hat{\theta} = \frac{1}{4}$. 故选 (A).

6.6 (C) 解 σ^2 未知, 均值 μ 的置信水平为 $1-\alpha$ 的置信区间为

$$\left(\bar{X} - \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1), \bar{X} + \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1) \right),$$

其中, 在给定条件下, $\bar{x} = 20, s = 1, t_{\frac{\alpha}{2}}(n-1) = t_{0.05}(15)$, 则所求参数 μ 的置信水平为 0.90 的置信区间为

$$\left(20 - \frac{1}{4} t_{0.05}(15), 20 + \frac{1}{4} t_{0.05}(15) \right),$$

故选择 (C).

6.7 (A) 解 显著性水平 α 是确定小概率事件的一个界限, 由检验法则知, $\alpha = P\{\text{拒绝 } H_0 | H_0 \text{ 为真}\}$, 所以正确选项是 (A). 选项 (B) 是 $P\{\text{接受 } H_0 | H_0 \text{ 为真}\} = 1-\alpha$; 选项 (D) 是 $P\{\text{接受 } H_0 | H_0 \text{ 不成立}\} = P\{\text{犯第二类错误}\} = \beta$; 选项 (C) 是 $P\{\text{拒绝 } H_0 | H_0 \text{ 不成立}\} = 1-\beta$.

6.8 2 解 因为样本方差的数学期望等于总体的方差, 即 $E(S^2) = DX$, 而

$$EX = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^{+\infty} x \cdot \frac{1}{2} e^{-|x|} dx = 0,$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_0^{+\infty} x^2 e^{-x} dx = 2,$$

故

$$E(S^2) = DX = E(X^2) - (EX)^2 = 2 - 0 = 2.$$

$$6.9 \text{ 解 } (1) EX = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^\theta \frac{6x^2}{\theta^3}(\theta-x)dx = \frac{\theta}{2}.$$

记 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, 由 $\frac{\theta}{2} = \bar{X}$, 得 θ 的矩估计量为 $\hat{\theta} = 2\bar{X}$.

(2) 由于

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{+\infty} x^2 f(x)dx \\ &= \int_0^\theta \frac{6x^3}{\theta^3}(\theta-x)dx = \frac{3\theta^2}{10}, \end{aligned}$$

$$\begin{aligned} DX &= E(X^2) - (EX)^2 \\ &= \frac{3\theta^2}{10} - \left(\frac{\theta}{2}\right)^2 = \frac{\theta^2}{20}, \end{aligned}$$

因此 $\hat{\theta} = 2\bar{X}$ 的方差为 $D\hat{\theta} = D(2\bar{X}) = 4D\bar{X} = \frac{4}{n}DX = \frac{\theta^2}{5n}.$

$$6.10 \text{ 解 } \mu_1 = EX = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^1 (1+\theta)x^{\theta+1}dx = \frac{1+\theta}{2+\theta}, \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

令 $\mu_1 = \bar{X}$, 即 $\frac{1+\theta}{2+\theta} = \bar{X}$, 解得 θ 的矩估计量为 $\hat{\theta} = \frac{2\bar{X}-1}{1-\bar{X}}.$

设 $x_1, x_2, \dots, x_n$ 为样本值, 则似然函数为

$$L(\theta) = (1+\theta)^n \prod_{i=1}^n x_i^\theta \quad (0 < x_i < 1, i=1, 2, \dots, n),$$

取对数得

$$\ln L(\theta) = n \ln(1+\theta) + \theta \sum_{i=1}^n \ln x_i,$$

令

$$\frac{d[\ln L(\theta)]}{d\theta} = \frac{n}{1+\theta} + \sum_{i=1}^n \ln x_i = 0,$$

解得唯一解 $\hat{\theta} = -\frac{n}{\sum_{i=1}^n \ln x_i} - 1$, 则 θ 的最大似然估计量为 $\hat{\theta} = -\frac{n}{\sum_{i=1}^n \ln X_i} - 1.$

$$6.11 \text{ 解 } (1) p = P\{0 < X < \sqrt{\alpha}\} = \int_0^{\sqrt{\alpha}} f(x)dx = \frac{1}{\alpha}.$$

(2) 当 $0 \leq x_1 \leq \alpha, 0 \leq x_2 \leq \alpha, \dots, 0 \leq x_n \leq \alpha$ 时, 似然函数为

$$\begin{aligned} L(\alpha) &= f(x_1)f(x_2)\cdots f(x_n) \\ &= \frac{2^n}{\alpha^{2n}} x_1 x_2 \cdots x_n, \end{aligned}$$

显然 $L(\alpha)$ 关于 α 单调减少, 且 $\alpha \geq \max\{x_1, x_2, \dots, x_n\}$, 则 α 的最大似然估计量为

$$\hat{\alpha} = \max\{X_1, X_2, \dots, X_n\}.$$

又由 (1) 知 $p = \frac{1}{\alpha}$ 关于 α 是单调函数, 根据最大似然估计的不变性, 有 p 的最大似然估计量为

$$\hat{p} = \frac{1}{\max\{X_1, X_2, \dots, X_n\}}.$$

6.12 (仅数学一) 证明 由于

$$\begin{aligned} ET &= E\left(\overline{X}^2 - \frac{1}{n}S^2\right) = E(\overline{X}^2) - \frac{1}{n}E(S^2) \\ &= D\overline{X} + (E\overline{X})^2 - \frac{1}{n}E(S^2) = \frac{\sigma^2}{n} + \mu^2 - \frac{\sigma^2}{n} \\ &= \mu^2, \end{aligned}$$

从而知, T 是 μ^2 的无偏估计量.

(仅数学三) 解

$$\begin{aligned} ET &= E\left(\overline{X}^2 - \frac{1}{n}S^2\right) = E(\overline{X}^2) - \frac{1}{n}E(S^2) \\ &= D\overline{X} + (E\overline{X})^2 - \frac{1}{n}E(S^2) = \frac{\sigma^2}{n} + \mu^2 - \frac{\sigma^2}{n} \\ &= \mu^2. \end{aligned}$$

