中国科学技术大学
University of Science and Technology of China

概率统計思考與总結

程嘉杰　整理

USTC, 96 JinZhai Road, Hefei,
Anhui, P.R.China, 230026
中国安徽省合肥市金寨路 96 号
中国科学技术大学

文件网址：home.ustc.edu.cn/∼jiajie
编写日期：2021 年 12 月 23 日

目录 CONTENTS

概率统计思考与总结

程嘉杰

前言

本书总结了概率论数理统计的相关概念，为中科大概率论与数理统计习题给出了思考和解答。

第一章 概率论基础

第 1 节 事件

术语

1. 随机试验
   - ♣ 相同情况可以重复
   - ♣ 事先已知全部结果
   - ♣ 不确定产生何结果
2. 样本空间 (Ω)：随机试验所有可能结果集合
3. 基本事件 (ω)：随机试验可能结果
4. 事件 (A,B,$\cdots$): 关心的结果的集合, $A \in \Omega$ 但不是所有 Ω 子集都是事件
5. 必然事件 (Ω)/不可能事件 ($\varnothing$)

事件运算 (Venn 图)

事件的交并补、属于关系、或者关系、加减关系, 如何写至少发生几个，不同时发生

注 1.1.1. $A+B$ 代表着互斥事件的并 $AB=\varnothing$且$A\cup B$

性质 1.1.1. 事件的运算

- 交换律 $AB=BA, \quad A\cup B=B\cup A$
- 结合律 $A\cup B\cup C=A\cup(B\cup C), \quad A\cap B\cap C=A\cap(B\cap C)$
- 分配律 $(A\cup B)\cap C=(AC)\cup(BC), (A\cap B)\cup C=(A\cup C)(B\cup C)$
- 对偶原则 *(De Morgan)*:

$$\overline{\bigcup_{i=1}^{n} A_i}=\bigcap_{i=1}^{n}\overline{A_i}, \quad \overline{\bigcap_{i=1}^{n} A_i}=\bigcup_{i=1}^{n}\overline{A_i}$$

第 2 节 概率

概率定义

1. 古典概型

(有限性) 试验结果只有有限个, $\Omega=\{\omega_1,\cdots,\omega_n\}$
(等可能性) 每个基本事件发生的可能性相同. $P\{\omega_j\}=\dfrac{1}{n}, j=1,\cdots,n$

2. 统计定义:

   事件 A 的随机试验独立重复做 n 次 (Bernouli 试验), 设事件 A 发生了 n_A 次, 称比值 $f_n(A)=\dfrac{n_A}{n}$ 为事件 A 发生的频率, 当 n 越来越大时, 频率会在某个值 p 附近波动, 且波动越来越小, 这个值 p 就定义为事件 A 的概率, $\lim\limits_{n\to+\infty} f_n(A)=p$.

3. 主观概率: 合理的信念的测度

4. 公理化定义 (Колмогоров)

   称 $P(\cdot)$ 为一概率 (集函数定义域 $P(\cdot):\mathcal{F}\to[0,1]$), 如果

   (a) 设 A 是随机事件, 则 $0\leqslant P(A)\leqslant 1$

   (b) 设 Ω 为必然事件, 则 $P(\Omega)=1$

   (c) 若事件 $A_1,A_2,\ldots$ 为两两不相容的事件序列, 则

$$P\left(\bigcup_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P\left(A_i\right) \text{ 互斥可列可加性}$$

概率性质

I. 不可能事件 $\quad P\left(\sum_{j=1}^{\infty}\phi\right)=\sum_{j=1}^{\infty}P(\phi)\Rightarrow P(\phi)=0$

II. 有限可加性

$$\begin{aligned}
P\left(\sum_{j=1}^{n} A_j\right) &= P\left(\sum_{j=1}^{n} A_j+\sum_{j=n+1}^{+\infty}\phi\right)=P\left(\sum_{j=1}^{n} A_j\right), \\
&= \sum_{j=1}^{+\infty} P\left(A_j\right)=\sum_{j=1}^{n} P\left(A_j\right)+\sum_{n=1}^{\infty} P(\phi) \\
&= \sum_{j=1}^{n} P\left(A_j\right)
\end{aligned}$$

III. 互斥事件:

$$P(A)=1-P(\bar{A})$$
$$1=P(\Omega)=P(A+\bar{A})=P(A)+P(\bar{A})$$

IV. 两事件关系

$$
\begin{aligned}
&A \subset B \Rightarrow P(B-A)=P(B)-P(A)=P(B\cap A)\\
&P(A,B)=P(A\cup B)=P(A)+P(B)-P(AB)\\
&P(A\cap B)=P(AB);P(A\bar{B})+P(AB)=P(A)
\end{aligned}
$$

V. 容斥原理（进出公式）

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n}|A_i| - \sum_{1\leqslant i<j\leqslant n}|A_i\cap A_j| + \sum_{1\leqslant i<j<k\leqslant n}|A_i\cap A_j\cap A_k| - \cdots + (-1)^{n-1}|A_1\cap\cdots\cap A_n|$$

$$\Rightarrow P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1\leqslant i<j\leqslant n} P(A_iA_j) + \sum_{1\leqslant i<j<k\leqslant n} P(A_iA_jA_k) - \cdots + (-1)^{n-1}P(A_1A_2\cdots A_n)$$

VI. 次可加性：任意的事件 $A_1,\cdots,A_n,\cdots \in \mathcal{F}$, 有 $P\left(\sum_{n=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} P(A_n)$.

VII. 上、下连续性

若 $A_n \in \mathcal{F}$ 且 $A_n \subset A_{n+1}, n=1,2,\cdots$, 则

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \lim_{n\to+\infty} P(A_n)$$

若 $A_n \in \mathcal{F}$ 且 $A_n \supset A_{n+1}, n=1,2,\cdots$, 则

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \lim_{n\to+\infty} P(A_n)$$

一些概型

1、 古典概型

- 概率计算：$P(A)=\dfrac{m}{n}=\dfrac{|A|}{|\Omega|}$
- 计数原理：

排列数：$A_n^r = P_r^n = n(n-1)\cdots(n-r+1)$

组合数：$\dbinom{n}{r} = \dfrac{n(n-1)\cdots(n-r+1)}{r!} = \dfrac{n!}{r!(n-r)!} = \dfrac{A_n^r}{(n-r)!}$

有放回组合：$\dbinom{n+r-1}{r}$ 相当于 r 球和 (n-1) 个板一起排

多组组合：$\dfrac{n!}{n_1!\cdot n_2!\cdot\ldots\cdot n_k!}$

n 个质点分配到 N 个盒子情形

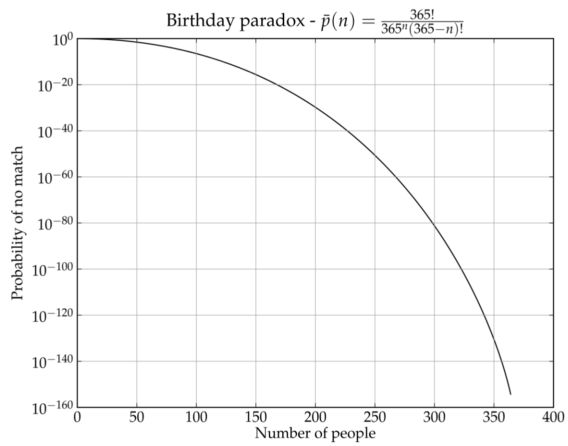
分配方式		不同分法的总数
每盒可容纳任意多个质点	质点可辨	N^n
	质点不可辨	$\binom{N+n-1}{n}$
每盒最多容纳一个质点	质点可辨	$A_N^n = N(N-1)\cdots\cdots(N-n+1)$
	质点不可辨	$\binom{N}{n} = \frac{A_N^n}{n!}$

N 个元素 n 次简单随机抽样情形

抽样方式		不同抽法的总数
有放回	有序	N^n
	无序	$\binom{N+n-1}{n}$
无放回	有序	$A_N^n = N(N-1)\cdots\cdots(N-n+1)$
	无序	$\binom{N}{n} = \frac{A_N^n}{n!}$

例 1.2.1. 生日同天问题

对立事件 (无人同天）概率：$\bar{p}(n) = 1 \times \left(1-\frac{1}{365}\right) \times \left(1-\frac{2}{365}\right) \times \cdots \times \left(1-\frac{n-1}{365}\right) = \frac{n!\cdot\binom{365}{n}}{365^n}$



2、 几何概型

$$P(A) = \frac{m(A)}{m(\Omega)}$$落在区域 A 的概率与区域 A 的测度成正比并且与其形状位置无关

3、 条件概率

★ 全概率公式：设 $\{B_1, B_2, \cdots B_n\}$ 是样本空间 Ω 的一个分割, 且 $P(B_i) > 0 (i = 1, \cdots, n)$, A 为 Ω 中的一个事件, 则

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i)$$

★ Bayes(逆概率) 公式：设 $\{B_1, B_2, \cdots B_n\}$ 是样本空间的一个分割, A 为 Ω 中的一个事件, $P(B_i) > 0, i = 1, 2, \cdots, n, P(A) > 0$, 则

$$P(B_i \mid A) = \frac{P(A \mid B_i) P(B_i)}{\sum_{j=1}^{n} P(A \mid B_j) P(B_j)}$$

4、 事件的独立性: 检验独立性：$P(AB) = P(A)P(B)$

(a) 相互独立和两两独立区别：

设 $A_1, A_2, \cdots A_n$ 是随机试验中的 n 个事件, 以 $\tilde{A}_i$ 表示 A_i 或 $\bar{A}_i$ 之一. 若满足

$$P\left(\tilde{A}_1 \tilde{A}_2 \cdots \tilde{A}_n\right) = P\left(\tilde{A}_1\right) P\left(\tilde{A}_2\right) \cdots P\left(\tilde{A}_n\right),$$

则称事件列 $A_1, A_2, \cdots A_n$ 相互独立. (上面有 2^n 个等式)

5、 小概率事件：一个事件如果发生的概率很小，那它在一次试验中是几乎不可能发生的，但在多次重复试验中几乎是必然发生的

第 3 节 思考题

1. 写出下列各试验的样本空间及指定事件的样本点. (1) 连续两次掷色子, $A = \{$ 第一次掷出的值比第二次大 $\}, B = \{$ 两次点数相等 $\}, C = \{$ 两次点数之和为 $10\}$. (2) 连续掷硬币 3 次, $A = \{$ 第一次为反面 $\}, B = \{$ 有两个正面 $\}, C = \{$ 三面都相同$\}$. (3) 以原点为圆心的单位圆内随机取一点, $A = \{$ 所取之点与原点的距离小于 $1/2\}, C = \{$ 所取之点与原点的距离小于 $1/2$ 大于 $1/3\}$.

2. 画出第 1 题中各事件的关系的维恩图 (Venn 图).

3. 某炮弹射击目标 3 次, 记 $A_i = \{$ 第 i 次集中目标 $\}(i = 1, 2, 3)$, 用 A_1, A_2, A_3 表示下列事件 (1) 仅有一次击中目标. (2) 至少有一次击中目标. (3) 第一次击中且第二次第三次至少有一次击中. (4) 最多击中一次.

4. 设一个试验的样本空间为 $[0, 2]$, 记事件 $A = \{1/2 < x \leqslant 1\}, B = \{1/4 < x \leqslant 3/2\}$, 写出下列各事件下列事件 (1) $A\bar{B}$, (2) $\bar{A} \cup B$, (3) $\overline{AB}$, (4) $\overline{\bar{A}B}$.

   解: 此题注意一切讨论在样本空间进行

$$A\bar{B} = \varnothing, \quad \bar{A} \cup B = \Omega, \quad \overline{AB} = \{0 \leqslant x \leqslant \frac{1}{2}\} \cup \{1 < x \leqslant 2\}, \quad \overline{\bar{A}B} = \{\frac{1}{4} < x \leqslant \frac{3}{2}\}$$

5. 设 A, B 是两事件且 $\mathbb{P}(A) = 0.7, \mathbb{P}(B) = 0.8$, 问: (1) 在什么条件下, $\mathbb{P}(AB)$ 取到最大值, 最大值多少? (2) 在什么条件下, $\mathbb{P}(AB)$ 取到最小值, 最小值多少?

6. 设 A, B, C 是三事件, 已知 $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1/3, \mathbb{P}(AB) = \mathbb{P}(BC) = 1/8, \mathbb{P}(AC) = 0$. 求 A, B, C 至少发生一个的概率.

7. 随机排列字母 $a, b, b, i, i, l, o, p, r, t, y$, 求恰好拼成单词 probability 的概率. 从这 11 个字母任意连续抽取 7 个字母, 恰好排列成 ability 的概率.

8. 市场调查员报道了如下数据: 在被询问的 1000 名顾客中, 有 811 人喜欢巧克力糖, 752 人喜欢夹心糖, 418 人喜欢大白兔糖, 570 人喜欢巧克力糖和夹心糖, 356 人喜欢巧克力糖和大白兔糖, 348 人喜欢夹心糖和大白兔糖以及 297 人喜欢全部三种糖果. 证明这一消息有误.

9. 从 0 到 9 中不放回的任取三个数排好, 求恰好排成一个 3 位数偶数的概率.

10. 一个班有 50 个同学, 其中至少有 2 个人生日相同的概率是多少?

    解:

$$P = 1 - \frac{50!\binom{365}{50}}{365^{50}} \approx 0.97037$$

11. 袋中有 a 个白球, b 个黑球, 现任意不放回的一一摸出, 求第 k 次取出白球的概率 $(1 \leqslant k \leqslant a + b)$.

12. 设 100 件产品, 其中有 3 件是次品, 现从中不放回的随机取 2 件, 求抽到的两件都是次品的概率是多少? 抽到的两件都是合格品的概率是多少?

13. 设在某考卷上某一同学有 4 道选择题不会做, 每道题有 4 个可供选择的答案, 只许选择一个, 于是瞎猜随机选一个, 试问能猜对 m 道题的概率是多少, $m = 0, 1, 2, 3, 4$.

    解：此处情景符合 n 重 Bernoulli 试验，故采取公式：

    $$P = \binom{4}{m}(\frac{1}{4})^m(\frac{3}{4})^{4-m}$$

14. 平面上画有等距离为 1 的一些平行线, 现向此平面任意投掷一根长为 0.6 的针, 现投掷 3204 次针, 发现其中针与线相交次数为 1218 次, 由此求出 π 的近似值.

    解：可以采取如下方法说明：可能存在针长度比平行线距离短、长的情况

    短的时候

    $$P(x,\theta) = P(x)P(\theta) = \frac{2}{d} \cdot \frac{2}{\pi}$$
    $$P = \int_0^{\frac{\pi}{2}} \int_0^{\frac{l}{2}\sin\theta} \frac{4}{d\pi} d\theta dx = \frac{2l}{d\pi}$$

    长的时候

    $$P = \left(\int_0^{\arcsin(\frac{d}{l})} \int_0^{\frac{l}{2}\sin\theta} \frac{4}{d\pi} d\theta dx \right) + \left(\int_{\arcsin(\frac{d}{l})}^{\frac{\pi}{2}} \frac{2}{\pi} d\theta dx \right) = \frac{2l}{d\pi} - \frac{2}{d\pi} \left\{ \sqrt{l^2 - d^2} + d \arcsin\left(\frac{d}{l}\right) \right\} + 1$$

    本题中 π 近似值是 3.1566

15. 甲乙两人约定在下午 3 点和 4 点之间到某公交始发站乘公交车, 该公交始发车站每隔 15 分钟发出一辆公交车. 现约定见车就乘, 求甲乙同乘一辆车的概率. 现假定甲乙两人在这期间到达为等可能.

16. 在一次游戏过程中有两队需要合作, 甲队将于 12 点半到 1 点间到达某地问关过河, 乙队将于 12 点到 12 点半到达此地为甲队准备船只, 准备船只需要时间为 15 分钟，请问甲队到达即能过河的概率是多少?

17. 合肥市的电话号码由 8 位数组成, 除第一个数字为 6,8 中的一个外, 其他各位数字可以是 $0, 1, \cdots, 9$ 中的任意一个, 现随机抽查一户居民的电话号码, 问其后四位数是由不同数字组成的概率为多大.

18. 将一枚骰子连掷 12 次. 试求 $1, 2, 3, 4, 5, 6$ 各点均出现两次的概率.

19. 某小学一年级有 8 个班, 一年级有 6 个班, 三年级有 4 个班. 如果将所有班级随机分成 3 组, 每组班级数相同. 求每组都有三年级班的概率是多少?

20. 四人玩一副扑克牌 (54 张牌), 求大小王被一个人拿到的概率.

21. 从一副 52 张扑克牌中随机抽取 10 张, 求包含所有 4 种花色牌的概率.

22. 一小区居民订阅报纸的统计数字如下: 定甲种报纸的占 40%, 订乙种报纸的占 25%, 同时订上述两种报纸的占 15%. 求下列事件的概率：(1) 只订甲报的; (2) 只订一种报的; (3) 至少订一种报的; (4) 两种报都不订的.

23. 甲乙两选手进行兵乓球单打比赛, 已知在每局中甲胜的概率为 $p(p>1/2)$, 乙胜的概率为 $1-p$. 比赛可采用三局两胜制或五局三胜制, 问哪一种比赛制度对甲更有利?

24. * 设有 n 个人随机地坐到礼堂第一排的 N 个座位上, 试求下列事件的概率: (1) 任何人都没有邻座. (2) 每人恰有一个邻座. (3) 关于中央对称的两个座位至少有一个空着.

25. 考虑一元二次方程 $x^2+Bx+C=0$, 其中 B,C 分别是将一枚均匀骰子连掷两次先后出现的点数. 求该方程有实根的概率和有重根的概率.

26. 连续投掷骰子直到出现 6 点为止, 求最终投掷了偶数次的概率.

27. 设某袋子中有红白黑三种颜色的球, 红色球的数目是白球的 2 倍, 黑球为红球的 $1/3$, 试求从袋中随机摸出一球恰好是白球的概率.

28. 现投掷三枚均匀骰子, 试求恰好有两枚出现相同点数的概率.

29. 一栋 20 层楼中的一架电梯在底层 (第一层) 上来 8 位乘客. 电梯在每一层都停, 设每位乘客在每层离开是等可能的, 求没有两位乘客在同一层离开的概率.

30. 某路公共汽车共有 11 个停车站, 由始发站开车时车上共有 8 名乘客. 假设每人在各站 (始发站除外) 下车的概率相同. 试求下列各事件的概率: (1) 8 人在不同的车站下车. (2) 8 人在同一车站下车. (3) 8 人中恰有 3 人在终点站下车.

31. 在一种双骰子博弈中, 玩家投两枚骰子, 如果其和是 7 或 11 , 则玩家赢; 如果其和是 2,3 或者 12 , 玩家输; 若是其他结果时就继续玩, 直到玩家输或者赢为止. 计算玩家 的概率.

32. 从 $1,2,\ldots,9$ 这 10 个数中不放回的取 n 个, 求这 n 个数的乘积可以被 10 整除.

33. 从 n 个数 $1,2,\cdots,n$ 中随机地不放回取出两个数, 试求此两数一个小于 $k(1<k<n)$ 、另一个大于 k 的概率.

34. 在数字 $1\sim 9$ 中不放回的随机取两个数, 每次一个数, 则在第一次取出偶数的条件下, 第二次取出奇数的概率.

35. 在数字 $1\sim 9$ 中随机的取一个数 X, 然后再在 1 到 X 中随机的取一个数, 试求第二次取的数为 $m(1\leqslant m\leqslant 9)$ 的概率是多少?

36. 连续投掷一枚骰子两次, 已知其中一次是 6 点, 试问另一次也是 6 点的概率是多 ?

37. 某人忘了电话号码的最后一个数字, 因而随机拨号, 问拨号不超过 3 次而接通所需拨的电话号码的概率是多少? 若已知最后一位是偶数, 那么此概率是多少?

38. 掷两枚均匀的骰子, 已知点数只和为 5 , 试用两种方法, 求其中有一枚骰子的点数为 1 的概率.

39. 连续投掷一枚骰子两次, 已知这两次的点数和是 5 点, 求两次点数之差不大于 2 的概率.

40. 连续投掷一枚骰子两次, 已知这两次的点数差小于 2 , 求两次点数之和大于 6 的概率.

41. 一人的网上银行密码共有 8 位数字和字母组成, 每位数字都可从 $0 \sim 9$ 中任选一个. 某人在登陆网上银行时忘记了密码的最后一位数字, 求 (1) 任意尝试最后一位数字, 不超过 3 次就试对的概率. (2) 如果他记得最后一位是偶数, 不超过 3 次就成功登陆的概率.

42. 连续掷三颗骰子, 已知所得三个数都不相同, 试问含有 1 点的概率是多少? 三次中最大结果是 6 的概率是多少?

43. 投郑两枚均匀的骰子, 问至少有一个是 6 的概率是多少? 若这两个面不一样, 求至少有一个是 6 的概率.

44. 袋中有 6 个白球和 4 个黑球, 从中无放回的随机取出 3 个球, 已知其中之一为黑球, 试求其余两球均是白球的概率.

45. 袋中有一个球, 它为白球黑球的概率相等. 现从中放入一个白球, 再从中随机取出一球, 现发现是白球, 试求袋中所剩之球也是白球的概率.

46. 掷一枚均匀骰子, 如果得到点数为 N, 则连续抛掷一均匀硬币 N 次. 已知抛鄉硬币的过程中有 3 次正面向上, 计算 $N=4$ 的概率.

47. 有两積同种类型的零件. 第一箱装 50 只, 其中 10 只一等品; 第二昔装 30 只, 其中 18 只为一等品. 今从两箱中任挑出一箱, 然后从该箱中取零件两次, 每次任取一只, 做不放回抽样. 试求 (1). 第一次取到的零件是一等品的概率. (2). 第一次取到的零件是一等品的条件下, 第二次取到的也是一等品的概率.

48. 掷三枚硬币, 已知其中有一枚出现了正面, 求至少出现一枚反面的概率.

49. 为防止意外, 办公大楼楼道里同时装有两个报警装置 1 和 2 . 已知报警装置 1 单独使用时有效概率为 0.95; 报警装置 2 单独使用时有效概率为 0.90. 在报警装置 1 失效的条件下装置 2 失效的概率为 0.86. 求发生意外时至少有一个报警装置有效的概率.

50. 设笔袋中有 r 支红色铅笔, b 支黑色铅笔. 每次从袋中任取一支笔, 观察其颜色后放回并再放入 a 支同色的球. 求第一第二次取到红色并第三第四次取到黑色笔的概率.

51. 袋中有 10 个白球, 5 个黑球, 现从中无放回随机摸出 3 球, 求均为同色球的概率是多少?

52. 设事件 A 与自己独立, 证明 $\mathbb{P}(A)$ 等于 0 或者 1 .

53. 证明: 如果 $\mathbb{P}(B \mid A)=\mathbb{P}(B \mid \bar{A})$, 则事件 A 与 B 独立.

54. 如果 $\mathbb{P}(B \mid A)>\mathbb{P}(B)$, 则称 A 倾向于 B. 证明: 如果 A 倾向于 B, 那么 $\bar{A}$ 也倾向于 $\bar{B}$.

55. 事件 A 与 B 至少发生一个的概率是 0.12, 同时发生的概率是 0.1, 请问事件 A 与 B 相互独立吗?

56. 对于三个事件 A, B, C, 若
$$\mathbb{P}(AB \mid C)=\mathbb{P}(A \mid C)\mathbb{P}(B \mid C)$$
成立, 则称 A 与 B 关于 C 条件独立. 若已知 A 与 B 关于 C 与 $\bar{C}$ 条件独立, 且 $\mathbb{P}(C)=0.5, \mathbb{P}(A \mid C)=\mathbb{P}(B \mid C)=0.9, \mathbb{P}(A \mid \bar{C})=0.2, \mathbb{P}(B \mid \bar{C})=0.1$, 试求 $\mathbb{P}(A), \mathbb{P}(B), \mathbb{P}(AB)$ 并证明 A 与 B 不独立.

57. 有 4 个一年级男生, 6 个一年级女生, 6 个二年级男生共上一门课, 为了使在随机选取一个学生时性别与班级独立, 在这个班还需要出现多少个二年级女生?

58. 设敌机俯冲时被步枪击落的概率 $p < 0.05$, 求当 n 只步枪同时开火时, 击落敌机的概率. 若 $p = 0.008n = 25$, 敌机被击中的概率.

59. 对同一目标进行三次独立射击, 第一、二、三次射击的命中率分别为 $0.5, 0.6$ 和 0.8, 试求: (1) 在这三次射击中, 恰好有一次射中的概率. (2) 在这三次射击中, 至少射中一次的概率.

60. 某炮弹单发命中率为 p, 连续向某个目标发射 m 次, 试求 (1) 击中目标的概率是多少? (2) 恰好击中 10 次的概率是多少? (3) 至少击中 3 次的概率是多少?

61. 某人从家到学校要经过 4 个红绿灯口, 1⊠3 红绿灯口出现红灯的概率均为 0.4, 2、4 路口出现红灯的可能性为 0.6, 且相互独立. 求某人从家到学校至少碰到两个红灯的概率是多少?

62. 一个电路共有 3 个继电器, 当第一个继电器断开, 或者第二、第三个同时断开, 电路断开. 现设三个继电器断开的概率依次为 $0.3, 0.4, 0.6$, 且三个继电器断开与否相互独立. 求电路断开的概率.

63. 连续投掷一均匀硬币三次, 事件 $A = \{$ 至多有一次正面 $\}$, 事件 $B = \{$ 三次中有正面有反面 $\}$, 求证两事件 A 与 B 相互独立.

64. 投郑两枚均匀的骰子, 记事件 $A = \{$ 两次点数和为奇数 $\}, B = \{$ 第一枚的点数为奇数 $\}, C = \{$ 第二枚的点数为奇数 $\}$. 证明这三个事件两两独立但不相互独立.

65. 求下列各系统能正常工作的概率, 其中框图中的字母代表元件, 字母相同但下标不同的都是同一种元件, 只是装配在不同的位置上, A, B, C, D 类元件能正常工作的概率分别为 p_A, p_B, p_C, p_D.

66. 某系统由四个部件 I, II, III, IV 构成, 设四个部件之间相互独立, 可靠性 (正常工作的概率) 均为 p. 求 (1) (2) 两个系统的可靠性, 哪个系统更可靠?

67. 某炮弹单发命中率为 p, 连续向某个目标发射 m 次, 试求 (1) 击中目标的概率是多少? (2) 恰好击中 10 次的概率是多少? (3) 至少击中 3 次的概率是多少?

68. 设事件 $A_1, \cdots, A_n$ 相互独立, 记 $\mathbb{P}(A_i) = p_i > 0, i = 1, 2, \cdots n$, 假设 $\sum_{i=1}^{n} p_i = 1$. 求 (1) 这些事件至少有一件不发生的概率. (2) 这些事件均不发生的概率. (3) 这些事件恰好发生一件的概率.

69. 桌子上有两个盒子, 第一个盒子里装有 3 个橙色乒乓球和 2 个白色乒乓球, 第二个盒子是 2 个橙色的和 5 个白色的. 随机的抽出一个盒子中的一个乒乓球, 发现是橙色的, 问来自第 1 个盒子的概率大还是来自第二个盒子的概率大?

70. 在某个社区, 60% 的家庭拥有汽车, 30% 的家庭拥有房产, 而 20% 的家庭既有汽车又有房产, 随机选取一个家庭, 求此家庭或有汽车或有房产但不是两者都有的概率.

71. 假设镬子中有 1 个白球和 1 个黑球, 每次从罐子中随机取一个球, 放回后再放入 1 个同色的球. 计算当镬子中有 $m(m \geqslant 2)$ 个球时, 黑球个数为 $k(1 \leqslant k \leqslant m$ 的概率.

72. 假设 A 罐子中有 10 个红球和 6 个白球, B 镬子中是 6 个红球和 10 个白球. 从中随机选取一个罐子, 抽出一个红球后放回, 问再次抽取仍然抽到红球的概率是多少.

73. 从北京到达拉斯有两个航班, 从达拉斯到艺加哥有 3 个航班, 从北京直飞艺加哥有 2 个航班. 这些航班的票务之间完全独立, 头到票的概率都是 $p(0<p<1)$. 假设从北京飞芝加哥没有其他途径可选, 且某人从北京出发到达了巴黎, 计算他乘坐直飞航班完成的旅程的概率.

74. 某工厂的第一、二、三号车间生产同一种产品, 产量各占总产量的 $1/2, 1/3, 1/6$, 次品率分别为 $1\%, 1\%$ 和 2%. 现从该厂产品中随机抽取一件产品 (1) 求该产品是次品的概率. (2) 若发现该产品是次品, 求它是一号车间生产的概率.

75. 考卷中的某选择题有四个答案, 其中只有一个是正确的. 某考生可能知道哪个是正确的, 也可能是乱猜一个. 假设此考生知道正确答案的概率为 p, 而且在不知答案的情况时是随机地选择一个答案. 如果已知他答对了这道题, 问他确实知道正确答案的概率是多少?

76. 设有来自三个地区的考生报名表共 50 份, 三个地区分别有 10,15 和 25 份, 其中女生的报名表分别为 3 份, 7 份和 5 份, 现随机地选一个地区, 从该地区的报名表中先后抽出 2 份. (1) 求先抽到的 1 份是女生报名表的概率. (2) 已知后抽到的 1 份是男生报名表, 求先抽到的 1 份是女生报名表的概率.

77. 镬子中有 25 个球, 20 个红球, 5 个黑球. 从中不放回去出 5 个球, 不看颜色直接扔掉. 然后再从剩下的球中随机摸出一球发现是红球, 问扔掉的球里至少有两个红球的概率.

78. 装有 $m(m>3)$ 个白球和 n 个黑球的罐子中失去一球, 但不知是什么颜色的球. 为猜测它是什么颜色, 随机地从罐中摸出两个球, 结果都得到的是白球, 试求失去的球是白球的概率.

79. 假设患乙肝的人通过检查能被诊断出来的概率为 0.98, 而正常人经检查被误诊为有乙肝的概率为 0.05, 设某城市乙肝患病率为 0.05. 现从该城市居民中随机抽出一人进行检查, 如果其被诊断为乙肝患者, 求该人确实患有乙肝的概率.

80. 盒中有三枚硬币, 一枚是双正面的硬币, 另外两枚是正反面硬币 (其中一枚是均匀的硬币, 一枚是正面出现概率为 75% 的不均匀硬币). 当从这三枚硬币中随机选取一枚抛掷时, 它出现正面. 问它是双正面硬币的概率是多少?

81. 假定某种病菌在群体中的带菌率为 10%. 在检测时, 带菌者和不带菌者被检测出阳性的概率分别为 0.95 和 0.01. (1) 现有某人被测出呈阳性反应, 该人确为带菌者的概率是多少? (2)* 该人又独立地做了一次检测, 检测结果依然是阳性, 问在两次检测均呈阳性的情况下, 该人确为带菌者的概率是多少.

82. 桌上有三个笔筒, 第 1 个笔筒装有 2 支红芯圆珠笔, 4 支蓝芯圆珠笔, 第 2 个笔筒装有 4 支红芯圆珠笔, 2 支蓝芯圆珠笔; 第三个笔筒装有 3 支红芯圆珠笔, 3 支蓝芯圆珠笔. 外表看起来一模一样, 先随机取一个笔筒, 任取一支笔出来. 试求 (1). 取得红笔的概率. (2). 在已知取得红笔的条件下, 问笔从哪个盒子中取出的概率最大?

83. 计算机信号“0”和“1”传递出去, 信息站接收的时候, “0”被误收为“1”的概率为 0.02, “1”被误收为“0”的概率为 0.01. 信号“0”和“1”传输的频繁程度为 $2:1$. 若接收到的信号是“0”, 真实信号是“0”的概率是多少?

84. 假设有 4 个镬子, 其中第 k 个镬子里有 $k-1$ 个红球和 $4-k$ 个蓝球, $k=1,2,3,4$. 现随机取出一个罐子, 然后不放回的从中取两球. 求 (1) 取出的两个球颜色相同; (2) 若已知其中一个球为红球, 则另外一个球也是红球的概率为多少?

85. 有甲乙两只口袋, 甲袋中有 5 只白球和 2 只黑球, 乙袋中有 4 只白球 5 只黑球. 先从甲袋中任取两球放入乙袋, 然后再从乙袋中任取一球. 试 (1) 求从乙袋中取出的球为白球的概率. (2) 若已知从乙袋中取出的球为白球, 求从甲袋中取的两只球中有白色球的概率.

86. * 甲、乙二人约定了这样一个赌博规则: 有无穷个盒子, 编号为 n 的盒子中有 n 红球 1 个白球, $n=1,2,\cdots$. 甲拿一个均匀硬币掷到出现正面为止, 若到这时甲掷了 n 次, 则甲在编号为 n 的盒子中抽出一个球, 若抽到白球算甲胜, 否则乙胜. 你认为这规则对谁更有利?

87. 设摸奖箱里有 n 张奖券, 其中 m 张可以中奖. 每人每次摸奖一次, 试问摸奖的顺序会影响中奖概率吗?

第二章 随机变量及其分布

第 1 节　标准正态分布表

$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t=P\{X\leqslant x\}$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500 0	0.504 0	0.508 0	0.512 0	0.516 0	0.519 9	0.523 9	0.527 9	0.531 9	0.535 9
0.1	0.539 8	0.543 8	0.547 8	0.551 7	0.555 7	0.559 6	0.563 6	0.567 5	0.571 4	0.575 3
0.2	0.579 3	0.583 2	0.587 1	0.591 0	0.594 8	0.598 7	0.602 6	0.606 4	0.610 3	0.614 1
0.3	0.617 9	0.621 7	0.625 5	0.629 3	0.633 1	0.636 8	0.640 4	0.644 3	0.648 0	0.651 7
0.4	0.655 4	0.659 1	0.662 8	0.666 4	0.670 0	0.673 6	0.677 2	0.680 8	0.684 4	0.687 9
0.5	0.691 5	0.695 0	0.698 5	0.701 9	0.705 4	0.708 8	0.712 3	0.715 7	0.719 0	0.722 4
0.6	0.725 7	0.729 1	0.732 4	0.735 7	0.738 9	0.742 2	0.745 4	0.748 6	0.751 7	0.754 9
0.7	0.758 0	0.761 1	0.764 2	0.767 3	0.770 3	0.773 4	0.776 4	0.779 4	0.782 3	0.785 2
0.8	0.788 1	0.791 0	0.793 9	0.796 7	0.799 5	0.802 3	0.805 1	0.807 8	0.810 6	0.813 3
0.9	0.815 9	0.818 6	0.821 2	0.823 8	0.826 4	0.828 9	0.835 5	0.834 0	0.836 5	0.838 9
1.0	0.841 3	0.843 8	0.846 1	0.848 5	0.850 8	0.853 1	0.855 4	0.857 7	0.859 9	0.862 1
1.1	0.864 3	0.866 5	0.868 6	0.870 8	0.872 9	0.874 9	0.877 0	0.879 0	0.881 0	0.883 0
1.2	0.884 9	0.886 9	0.888 8	0.890 7	0.892 5	0.894 4	0.896 2	0.898 0	0.899 7	0.901 5
1.3	0.903 2	0.904 9	0.906 6	0.908 2	0.909 9	0.911 5	0.913 1	0.914 7	0.916 2	0.917 7
1.4	0.919 2	0.920 7	0.922 2	0.923 6	0.925 1	0.926 5	0.927 9	0.929 2	0.930 6	0.931 9
1.5	0.933 2	0.934 5	0.935 7	0.937 0	0.938 2	0.939 4	0.940 6	0.941 8	0.943 0	0.944 1
1.6	0.945 2	0.946 3	0.947 4	0.948 4	0.949 5	0.950 5	0.951 5	0.952 5	0.953 5	0.953 5
1.7	0.955 4	0.956 4	0.957 3	0.958 2	0.959 1	0.959 9	0.960 8	0.961 6	0.962 5	0.963 3
1.8	0.964 1	0.964 8	0.965 6	0.966 4	0.967 2	0.967 8	0.968 6	0.969 3	0.970 0	0.970 6
1.9	0.971 3	0.971 9	0.972 6	0.973 2	0.973 8	0.974 4	0.975 0	0.975 6	0.976 2	0.976 7
2.0	0.977 2	0.977 8	0.978 3	0.978 8	0.979 3	0.979 8	0.980 3	0.980 8	0.981 2	0.981 7
2.1	0.982 1	0.982 6	0.983 0	0.983 4	0.983 8	0.984 2	0.984 6	0.985 0	0.985 4	0.985 7
2.2	0.986 1	0.986 4	0.986 8	0.987 1	0.987 4	0.987 8	0.988 1	0.988 4	0.988 7	0.989 0
2.3	0.989 3	0.989 6	0.989 8	0.990 1	0.990 4	0.990 6	0.990 9	0.991 1	0.991 3	0.991 6
2.4	0.991 8	0.992 0	0.992 2	0.992 5	0.992 7	0.992 9	0.993 1	0.993 2	0.993 4	0.993 6
2.5	0.993 8	0.994 0	0.994 1	0.994 3	0.994 5	0.994 6	0.994 8	0.994 9	0.995 1	0.995 2
2.6	0.995 3	0.995 5	0.995 6	0.995 7	0.995 9	0.996 0	0.996 1	0.996 2	0.996 3	0.996 4
2.7	0.996 5	0.996 6	0.996 7	0.996 8	0.996 9	0.997 0	0.997 1	0.997 2	0.997 3	0.997 4
2.8	0.997 4	0.997 5	0.997 6	0.997 7	0.997 7	0.997 8	0.997 9	0.997 9	0.998 0	0.998 1
2.9	0.998 1	0.998 2	0.998 2	0.998 3	0.998 4	0.998 4	0.998 5	0.998 5	0.998 6	0.998 6
3	0.998 7	0.999 0	0.999 3	0.999 5	0.999 7	0.999 8	0.999 8	0.999 9	0.999 9	1.000 0

注：最后一行为 3.1～3.9

第 2 节 基础知识

1. 离散型随机变量: 取有限个值；连续型随机变量：取值连续且密度存在

2. 分布律/概率质量函数（probability mass function）

$$p_i = P\left(X = a_i\right), \quad i = 1, 2, \ldots \tag{2.1}$$

3. pmf 满足的条件：

$$\begin{aligned} &p_i \geqslant 0, \quad i = 1, 2, \ldots \\ &\sum_i p_i = 1 \end{aligned} \tag{2.2}$$

4. Bernoulli 试验：一个随机试验只有两个可能结果 A 和 $\bar{A}$，
独立重复 n 次: n 重 Bernoulli 试验
$n \to \infty$：可列重 Bernoulli 试验

5. 几何分布的无记忆性

$$\begin{aligned} P(\xi > k) &= \sum_{j=k+1}^{\infty} P(\xi = j) = p \sum_{j=k+1}^{\infty} q^{j-1} = q^k \\ P(\xi > m+n \mid \xi > m) &= \frac{P(\xi > m+n, \xi > m)}{P(\xi > m)} \\ &= \frac{P(\xi > m+n)}{P(\xi > m)} = \frac{q^{m+n}}{q^n} = q^n = P(\xi > n) \\ P(\xi > m+n \mid \xi > m) &= P(\xi > n) \end{aligned} \tag{2.3}$$

6. n 重 Bernoulli 试验和 Poisson 分布
如果 $np_n \to \lambda$, 则当 $n \to \infty$ 时

$$\begin{pmatrix} n \\ k \end{pmatrix} p_n^k \left(1 - p_n\right)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}. \tag{2.4}$$

7. 概率密度函数 (probability density function)：

♣ 性质 1： $f(x) \geqslant 0, \quad \forall x \in (-\infty, +\infty)$

♣ 性质 2： $\displaystyle\int_{-\infty}^{+\infty} f(x) dx = 1$

♣ 性质 3： $\displaystyle P\{a \leqslant X \leqslant b\} = \int_a^b f(x) dx, -\infty < a \leqslant b < +\infty$

8. (累积) 分布函数 $F(x)$

$$F(x) = \int_{-\infty}^{x} f(u) du, \quad -\infty < x < +\infty \tag{2.5}$$

♣ 性质 1： F 是非减的函数; 对任何 $x_1 < x_2$ 都有, $F\left(x_2\right) - F\left(x_1\right) = P\left(x_1 < X \leqslant x_2\right) \geqslant 0$

♣ 性质 2： $0 \leqslant F(x) \leqslant 1, x \in R$, 且 $\lim\limits_{x \to -\infty} F(x) = 0$; $\lim\limits_{x \to +\infty} F(x) = 1$.

♣ 性质 3： $F(x)$ 右连续

其他分布

1. 多维分布:

$$P\left(a_1 \leqslant X_1 \leqslant b_1, \ldots, a_n \leqslant X_n \leqslant b_n\right)=\int_{a_n}^{b_n} \ldots \int_{a_1}^{b_1} f\left(x_1, \ldots, x_n\right) d x_1 \cdots d x_n \tag{2.6}$$

- $F\left(x_1, \cdots, x_n\right)=P(X_1 \leqslant x_1, \cdots, X_n \leqslant x_n)$ 对每个变元单调非降;
- 对任意的 $1 \leqslant j \leqslant n$ 有, $\lim\limits_{x_j \rightarrow-\infty} F\left(x_1, \cdots, x_n\right)=0$;
- $\lim\limits_{x_1 \rightarrow \infty, \cdots, x_n \rightarrow \infty} F\left(x_1, \cdots, x_n\right)=1$.

2. 边缘分布: 边际分布律不决定联合分布律

$Y \backslash X$	x_1	x_2	$\ldots$	x_n	行和
y_1	p_{11}	p_{21}	$\cdots$	p_{n1}	$p_{\cdot 1}$
y_2	p_{12}	p_{22}	$\cdots$	p_{n2}	$p_{\cdot 2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
y_m	p_{1m}	p_{2m}	$\vdots$	p_{nm}	$p_{\cdot m}$
列和	$p_{1\cdot}$	$p_{2\cdot}$	$\cdots$	$p_{n\cdot}$	1

条件分布、联合分布、边缘分布关系

1. 已知 $(X, Y) \sim F(x, y)$, 则 $F_X(x)=F(x,+\infty), F_Y(y)=F(+\infty, y)$.
2. 已知 $(X, Y) \sim p_{ij}=P\left\{X=x_i, Y=y_j\right\}$, 则

$$F(x, y)=\sum_{x_i \leqslant x} \sum_{y_j \leqslant y} p_{ij},$$

$$\begin{gathered}
p_{i\cdot}=P\left\{X=x_i\right\}=\sum_{j=1}^{\infty} p_{ij}, \\
p_{\cdot j}=P\left\{Y=y_j\right\}=\sum_{i=1}^{\infty} p_{ij}. \\
\text{条件分布:} \\
P\left\{X=x_i \mid Y=y_j\right\}=\frac{p_{ij}}{p \cdot j}(p_{\cdot j}>0), \\
P\left\{Y=y_j \mid X=x_i\right\}=\frac{p_{ij}}{p_{i\cdot}}\left(p_{i\cdot}>0\right).
\end{gathered}$$

3. 已知 $(X, Y) \sim f(x, y)$, 则 $F(x, y)=\displaystyle\int_{-\infty}^x \int_{-\infty}^y f(u, v) d u d v$,

$$\begin{aligned}
f_X(x) & =\int_{-\infty}^{+\infty} f(x, y) d y \\
f_Y(y) & =\int_{-\infty}^{+\infty} f(x, y) d x
\end{aligned}$$

若 $f(x, y)$ 在点 (x, y) 处连续, 则

$$\frac{\partial^2 F(x, y)}{\partial x \partial y}=f(x, y).$$

条件概率密度:

$$f_{X|Y}(x \mid y)=\frac{f(x,y)}{f_Y(y)}\left(f_Y(y)>0\right),$$
$$f_{Y|X}(y \mid x)=\frac{f(x,y)}{f_X(x)}\left(f_X(x)>0\right).$$

第 3 节 随机变量分布的函数的概率分布

离散型

$$\text{离散卷积公式：}P(\xi+\eta=n)=\sum_{k=0}^{n}a_k b_{n-k}$$

连续型

★ **定理 2.3.1.** 密度变换公式：*(保证 Y=g(X) 严格单调；X=h(y) 存在且可微)*

$$p(y)=f(h(y))\left|h'(y)\right|,\quad y\in(\alpha,\beta) \tag{2.7}$$

例 2.3.1.

$$X\sim U\left(-\frac{\pi}{2},\frac{\pi}{2}\right),Y=\tan X,X=\arctan Y$$
$$f_Y(y)=\frac{1}{\pi}I_{\left(-\frac{\pi}{2}<\arctan y<\frac{\pi}{2}\right)}|(\arctan(y))'|=\frac{1}{\pi\left(1+y^2\right)},y\in(-\infty,+\infty)(Y-Cauchy\text{分布})$$

一般的方法：

$$\begin{aligned}F(y)&=P(Y\leqslant y)=P(\tan(X)\leqslant y)\\&=P(X\leqslant\arctan(y))=\int_{-\frac{\pi}{2}}^{\arctan(y)}\frac{1}{\pi}dy=\frac{1}{\pi}\arctan(y)+\frac{1}{2}\end{aligned}$$
$$f(y)=F'(y)=\frac{1}{\pi\left(1+y^2\right)},y\in(-\infty,+\infty)$$

例 2.3.2. $Y=X^2$

$$\begin{aligned}P(Y\leqslant y)=P\left(x^2\leqslant y\right)&=P(-\sqrt{y}\leqslant x\leqslant\sqrt{y})\\&=F(\sqrt{y})-F(-\sqrt{y})\\f_Y(y)=f_Y(\sqrt{y})\tfrac{1}{2}y^{-\frac{1}{2}}&+f_Y(-\sqrt{y})\tfrac{1}{2}y^{-\frac{1}{2}}\end{aligned}$$

例 2.3.3. $Y=\sqrt{X}$

$$f_Y(y)=f_X(y^2)\cdot 2y$$

★ **定理 2.3.2.** *(多维的随机变量分布的函数密度变换)* 如果 $(\xi_1,\cdots,\xi_n)$ 是 n 维连续型随机向量，具有联合密度函数 $p(x_1,\cdots,x_n)$. 假设存在 n 个 n 元函数

$$y_j=f_j\left(x_1,\cdots,x_n\right),\quad j=1,\cdots,n,$$

使得

$$\zeta_j=f_j\left(\xi_1,\cdots,\xi_n\right),\quad j=1,\cdots,n,$$

若 $(\xi_1,\cdots,\xi_n)$ 与 $(\zeta_1,\cdots,\zeta_n)$ 之间一一对应, 逆映射为 $\xi_j=h_j(\zeta_1,\cdots,\zeta_n), j=1,\cdots,n$. 其中每个 $h_j(y_1,\cdots,y_n)$ 都有一阶连续偏导数, 那么随机向量 $(\zeta_1,\cdots,\zeta_n)$ 是连续型的, 且具有联合密度函数

$$q(y_1,\cdots,y_n)=\begin{cases} p(h_1(y_1,\cdots,y_n),\cdots,h_n(y_1,\cdots,y_n))\,|J|, & (y_1,\cdots,y_n)\in\mathbb{D}\\ 0, & (y_1,\cdots,y_n)\notin\mathbb{D}\end{cases}$$

$$\text{Jacob 行列式: } J=\begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \cdots & \frac{\partial h_1}{\partial y_n}\\ \vdots & \vdots & \vdots\\ \frac{\partial h_n}{\partial y_1} & \cdots & \frac{\partial h_n}{\partial y_n}\end{vmatrix}$$

例 2.3.4. $(x,y)\sim N(0,0,1,1,0)\begin{cases} x=r\cos\theta\\ y=r\sin\theta\end{cases}, J=\begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial t}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial t}\end{vmatrix}=\begin{vmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix}=r$

$$p(x,y)=\frac{1}{2\pi}\exp\left\{-\frac{x^2+y^2}{2}\right\}$$

$$q(r,\theta)=\frac{1}{2\pi}r\exp\left\{-\frac{r^2}{2}\right\}=q_1(r)q_2(\theta),\quad r>0,\quad \theta\in[0,2\pi).$$

结果表明: θ 与 ρ 相互独立, 其中 $\theta\sim U[0,2\pi)$; ρ 则服从 *Weibull* 分布 (参数 $\lambda=1/2,\alpha=2$).

第 4 节 特殊组合的分布

画出概率密度的区域十分重要!!!

二重积分：先确定函数的限定，再在一定范围内积分（后积先定限）

和的分布

★ **定理 2.4.1.** *Z=X+Y* 分布的概率密度函数：

$$p(z)=\int_{-\infty}^{\infty}f(x,z-x)dx=\int_{-\infty}^{\infty}f(z-y,y)dy \tag{2.8}$$

特别地，X，Y 独立则有卷积：

$$p(z)=\int_{-\infty}^{\infty}f_1(x)f_2(z-x)dx=\int_{-\infty}^{\infty}f_1(z-y)f_2(y)dy\triangleq f_1*f_2(z)=f_2*f_1(z) \tag{2.9}$$

例 2.4.1. $X\sim Exp(\frac{1}{2}), Y\sim U[0,1]$, 求 $X-Y$ 的概率密度和 $P(X\leqslant Y)$

解一: 由题设知 $-Y\sim U(-1,0)$, 并记 X 和 $-Y$ 的密度分别为 f_1 和 f_2, 从而由卷积公式有

$$\begin{aligned} f_{X-Y}(z)&=\int_{-\infty}^{\infty}f_1(x)f_2(z-x)dx\\ &=\begin{cases} e^{-\frac{z}{2}}\left(1-e^{-\frac{1}{2}}\right), & z\geqslant 0\\ 1-e^{-\frac{z+1}{2}}, & -1<z<0\\ 0, & z\leqslant -1\end{cases}\end{aligned}$$

所以 $P(X\leqslant Y)=P(X-Y\leqslant 0)=2e^{-\frac{1}{2}}-1$ 。

解二：由于

$$
\begin{aligned}
P(X-Y\leqslant z)&=\{P(X\leqslant z+y\mid Y=y)f(y)dy\\
&=\begin{cases}\int_0^1 P(X\leqslant z+y)dy & z\geqslant 0\\ \int_{-z}^1 P(X\leqslant z+y)dy & -1<z<0\\ 0 & z\leqslant -1\end{cases}\\
&=\begin{cases}1-2e^{-z/2}\left(1-e^{-1/2}\right), & z\geqslant 0\\ z+2e^{-(z+1)/2}-1, & -1<z<0\\ 0, & z\leqslant -1\end{cases}
\end{aligned}
$$

再对分布函数求导数即得所求.

乘积分布

设 $X,Y)\sim f(x,y)$, 则 $Z=XY$ 的概率密度为

$$f_Z(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)dx=\int_{-\infty}^{+\infty}\frac{1}{|y|}f\left(\frac{z}{y},y\right)dy.$$

$$F_7(z)=P\{Z\leqslant z\}=P\{XY\leqslant z\}=\iint_{D:xy\leqslant z}f(x,y)d\sigma=\int_{-\infty}^{0}dx\int_{\frac{\pi}{x}}^{+\infty}f(x,y)dy+\int_0^{+\infty}dx\int_{-\infty}^{\frac{z}{x}}f(x,y)dy$$

$$
\begin{aligned}
f_z(z)=F_z'(z)&=\int_{-1}^{0}\left[\int_{\frac{z}{x}}f(x,y)dy\right]_z^{+\infty}dx+\int_0^{+\infty}\left[\int_x^{\frac{z}{x}}f(x,y)dy\right]_x'dx\\
&=\int_{-\infty}^{0}\left[-f\left(x,\frac{z}{x}\right)\cdot\frac{1}{x}\right]dx+\int_0^{+\infty}\left[f\left(x,\frac{z}{x}\right)\cdot\frac{1}{x}\right]dx\\
&=\int_{-\infty}^{+\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)dx
\end{aligned}
$$

例 2.4.2. 设 $X_1,\ldots,X_n i.i.d\sim N(0,1)$, 试求 $Y_n=\sum_{i=1}^n X_i^2$ 的分布

解: 由前例知 $Y_1=X_1^2$ 的概率密度函数为

$$f_1(y)=\frac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2}I(y>0)=\frac{1}{2^{1/2}\Gamma\left(\frac{1}{2}\right)}y^{1/2-1}e^{-y/2}I(y>0)$$

由卷积公式知 $Y_2=X_1^2+X_2^2$ 的概率密度函数为

$$
\begin{aligned}
f_2(z)&=\int_R f_1(y)f_2(z-y)dy\\
&=\frac{1}{2\Gamma^2\left(\frac{1}{2}\right)}e^{-z/2}I(z>0)\int_0^1 t^{-1/2}(1-t)^{-1/2}dt\\
&=\frac{1}{2^{2/2}\Gamma\left(\frac{2}{2}\right)}z^{2/2-1}e^{-z/2}I(z>0)
\end{aligned}
$$

从而由归纳法, 假设 $Y_{n-1}=X_1^2+\cdots+X_{n-1}^2$ 的概率密度函数为

$$f_{n-1}(z)=\frac{1}{2^{(n-1)/2}\Gamma\left(\frac{n-1}{2}\right)}z^{(n-1)/2-1}e^{-z/2}I(z>0)$$

则 $Y_n = Y_{n-1} + X_n^2$ 的概率密度函数可由卷积公式得

$$
\begin{aligned}
f_n(z) &= \int_R f_{n-1}(y) f_1(z-y) dy \\
&= \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{1}{2}\right)} z^{\frac{n}{2}-1} e^{-z/2} I(z>0) \int_0^1 t^{(n-1)/2-1}(1-t)^{1/2-1} dt \\
&= \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} z^{\frac{n}{2}-1} e^{-z/2} I(z>0)
\end{aligned}
$$

这就是 χ_n^2- 分布的 *pdf*（具有再生性）

商的分布

设 $(X,Y) \sim f(x,y)$, 则 $Z = \dfrac{X}{Y}$ 的概率密度为

$$
f_Z(z) = \int_{-\infty}^{+\infty} |y| f(yz, y) dy.
$$

$$
\begin{aligned}
\because f_Z(z) = F_Z'(z) &= \int_{-\infty}^{0} \left[\int_{zy}^{+\infty} f(x,y) dx \right]_z' dy + \int_0^{+\infty} \left[\int_{-\infty}^{zy} f(x,y) dx \right]_z' dy \\
&= \int_{-\infty}^{0} [-f(zy,y) \cdot y] dy + \int_0^{+\infty} [f(zy,y) \cdot y] dy \\
&= \int_{-\infty}^{+\infty} |y| f(zy,y) dy
\end{aligned}
$$

例 2.4.3. 设随机变量 ξ 与 η 相互独立, 同服从参数 $\lambda = 1$ 的指数分布, 试求 ξ/η 的密度函数.

$$
p(u,v) = e^{-u-v}, \quad u>0, v>0
$$

$$
p_{\frac{\xi}{\eta}}(x) = \begin{cases} \int_0^\infty t e^{-xt-t} dt = \frac{1}{(1+x)^2} & x>0 \\ 0 & x \leqslant 0 \end{cases}
$$

例 2.4.4. 设 $X_1 \sim N(0,1), X_2 \sim \chi_n^2$, 且 X_1 与 X_2 独立。求 $Y = \frac{X_1}{\sqrt{X_2/n}}$ 的概率密度函数. (记 $Y \sim t_n$, 称为自由度为 n 的 t 分布)。

解: 先求 $Z = \sqrt{X_2/n}$ 的密度 $g(z)$:

$$
g(z) = 2nz f_{X_2}\left(nz^2\right) I(z>0)
$$

其中 f_{X_2} 为 X_2 的密度函数. 利用商的密度变换公式, 可得

$$
\begin{aligned}
f_Y(y) = f_{X_1/Z}(y) &= \int_R |t| \phi(yt) g(t) dt \\
&= \int_R |t| \frac{1}{\sqrt{2\pi}} e^{-(yt)^2/2} \frac{2nt}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \left(nt^2\right)^{\frac{n}{2}-1} e^{-nt^2/2} I(t>0) dt \\
&= \frac{1}{\sqrt{2\pi}} \frac{2n^{\frac{n}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^\infty t^n e^{-\left(y^2+n\right) t^2/2} dt
\end{aligned}
$$

令$x=\left(n+y^2\right)t^2/2$, 则上述积分为

$$
\begin{aligned}
f_Y(y)&=\frac{1}{\sqrt{2\pi}}\frac{2n^{\frac{n}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^{\infty}x^{(n-1)/2}e^{-x}dx\cdot\frac{1}{2}\left(\frac{2}{n+y^2}\right)^{(n+1)/2}\\
&=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{y^2}{n}\right)^{-\frac{n+1}{2}}
\end{aligned}
$$

max/min 最大值最小值分布

注 2.4.1. 这两种分布对应着：最大值（串联正常工作系统），最小值（并联正常工作系统）

★ **设** $(X,Y)\sim F(x,y)$, **则** $Z=\max\{X,Y\}$ **的分布函数为** $F_{\max}(z)=P\{\max\{X,Y\}\leqslant z\}=P\{X\leqslant z,Y\leqslant z\}=F(z,z)$.

当 X **与** Y **独立时**, $F_{\max}(z)=F_X(z)\cdot F_Y(z)$.

★ **设** $(X,Y)\sim F(x,y)$, **则** $Z=\min\{X,Y\}$ **的分布函数为** $F_{\min}(z)=P\{\min\{X,Y\}\leqslant z\}=P\{X\leqslant z,Y\leqslant z\}=F(z,z)$.

$$
\begin{aligned}
F_{\min}(z)&=P\{\min\{X,Y\}\leqslant z\}=P\{\{X\leqslant z\}\cup\{Y\leqslant z\}\}\\
&=P\{X\leqslant z\}+P\{Y\leqslant z\}-P\{X\leqslant z,Y\leqslant z\}\\
&=F_X(z)+F_Y(z)-F(z,z).
\end{aligned}
$$

当 X **与** Y **独立时**,

$$
F_{\min}(z)=F_X(z)+F_Y(z)-F_X(z)F_Y(z)=1-\left[1-F_X(z)\right]\left[1-F_Y(z)\right].
$$

注 2.4.2. 上述可以推广到 n 个相互独立的随机变量 $X_1,X_2,\cdots,X_n$ 的情况, 即

$$
\begin{gathered}
F_{\max}(z)=F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z)\\
F_{\min}(z)=1-\left[1-F_{X_1}(z)\right]\left[1-F_{X_2}(z)\right]\cdots\left[1-F_{X_n}(z)\right]
\end{gathered}
$$

特别地, 当 $X_i(i=1,2,\cdots,n)$ 相互独立且有相同的分布函数 $F(x)$ 与概率密度 $f(x)$ 时,

$$
\begin{gathered}
F_{\max}(x)=[F(x)]^n,\quad f_{\max}(x)=n[F(x)]^{n-1}f(x)\\
F_{\min}(x)=1-[1-F(x)]^n,\quad f_{\min}(x)=n[1-F(x)]^{n-1}f(x)
\end{gathered}
$$

例 2.4.5. 设 $X_1,\ldots,X_n i.i.d\sim U(0,\theta),\theta>0$, 求 $X_{(n)}=\max_{1\leqslant i\leqslant n}X_i$ 的密度函数.

解：由于对任意 $0<x<\theta$,

$$
\begin{aligned}
F_{X_{(n)}}(x)&=P\left(X_{(n)}\leqslant x\right)=P\left(\max\left\{X_1,\cdots,X_n\right\}\leqslant x\right)=\prod_{k=1}^{n}F_k(x)\\
&=\left(\frac{x}{\theta}\right)^n
\end{aligned}
$$

从而所求密度为

$$
g(x)=F'_{X_{(n)}}(x)=\frac{nx^{n-1}}{\theta^n}I(0<x<\theta).
$$

12.

X	-1	1	2
P	0.4	0.4	0.2

13.

X	1	2	$\cdots$	n	$\cdots$
P	p	$(1-p)p$	$\cdots$	$(1-p)^n p$	$\cdots$

14.

$$\mathbb{P}(X>2)=1-\mathbb{P}(X\leqslant 2)=1-\frac{2}{5}-\frac{2}{5}\times\frac{3}{5}=\frac{17}{25}$$

42.

$$a=1,\quad b=-1$$

$$f(x)=F'(x)=(x^2(\ln x-1)+1)'=2x\ln x-x$$

第 5 节 思考题

1. 假设 X 是一个实值随机变量, 下列说法正确的是 ( ). (A) 至少存在一个常数 $x\in\mathbb{R}$, 使得 $\mathbb{P}(X=x)>0$ (B) 若 X 表示某个随机试验的结果, 则该随机试验结束后它的值是确定的 (C) 若 X 对任意实数都有可能取值, 则它一定是连续型的随机变量 (D) 若对一实数 x 有 $\mathbb{P}(X=x)=0$, 则 X 不可能取值为 x

2. 双色球是目前中国福利彩票中最受欢迎的一种玩法. 投注区分为红色球号码区和蓝色球号码区, 红色球号码区由 1 到 33 共三十三个号码组成, 蓝色球号码区由 1 到 16 共十六个号码组成. 投注时选择 6 个不同的红色球号码和 1 个蓝色球号码组成一注进行单式投注, 每注金额人民币 2 元. 设开奖时, 由系统随机指定 6 个不同的红色球号码和 1 个蓝色球号码. 若某单式投注中分别有 $i(0\leqslant i\leqslant 6)$ 个红色球号码和 $j(j=0,1)$ 个蓝色球号码与指定号码相同, 则称该投注的形式为“$i+j$”. 最后所有奖项规则如下

等级	一等奖	二等奖	三等奖	四等奖	五等奖	六等奖
形式	$6+1$	$6+0$	$5+1$	$5+0,4+1$	$4+0,3+1$	$2+1,1+1,0+1$

(1) 试引入一个随机变量 X 来描述随机购买的一注单式投注的各种获奖等级情况, 并求它的分布律; (2) 试用 X 取值的方式来表示某人花 2 元头一注后的下述事件

$$A=\{\text{获奖}\},B=\{\text{获得一等奖或二等奖}\},$$

并求出它们发生的概率.

3. 设随机变量 X 的分布律为 $\begin{pmatrix} 0 & 1 & 2 \\ 4p & p & 0.5 \end{pmatrix}$, 则 $p=$

4. 在一闯关游戏中一共有 4 道关卡, 若每道关卡某选手能以 $2/3$ 的概率顺利通过然后进入下一关, 或以 $1/3$ 的概率被淘汰出局 (设每道关卡的通过情况相互独立), 以 X 表示该选手能顺利通过关卡的数目, 试求 X 的分布律.

5. 同时掷两枚均匀的骰子, 以 X 记它们的点数之和. 式求 X 的分布律.

6. 同时郑三枚均讯的骰子, 以 X 记它们中最大的点数. 试求 X 的分布律和数学期望.

7. 某物流公司和某工厂约定用车将一 货物按期无损地运到目的地可得佣金 100 元，但若不按期则扣 20 元 (即得佣金 80 元); 若货物有损坏则扣 50 元; 若货物不按期又有损坏则扣 160 元. 该物流公司按以往经验认为一箱货物按期无损地运到目的地有 60% 的把握, 不按期到达的占 20%, 货物有损坏的占 10%, 货物不按期又有损坏的占 10%. 以 X 记该物流公司用车将一箱货物运到目的地后的毛收益. 试求 X 的分布律.

8. 设某游乐场的一部设备在一天内发生故障的概率为 20%, 设备一旦发生故障则全天无法工作。若一周五个工作日内无故障可以获利 10 万元, 只发生一次故障可以获利 5 万元, 发生两次故障获利 0 元, 发生三次或三次以上故障则言损两万元. 试求一周内该游乐场在这台设备上的毛利润的分布律.

9. 有一种街头游戏的道具为一个不透明的布袋, 其中装有 20 颗两种不同颜色的弹珠, 红色和蓝色各 10 颗. 任何人都可以花 2 元钱从袋子中随机摸出 10 颗弹珠, 如果这 10 颗弹珠都是同一种颜色, 则奖励 100 元; 如果是 9 红 1 蓝或者 9 蓝 1 红, 则奖励 20 元; 如果是 8 红 2 蓝或者 8 蓝. 2 红, 则奖励 5 元; 其余情况则没有奖励. 试用随机变量 X 来表示在一局游戏中玩家获得各种奖励的情况, 并求它的分布律.

10. 设随机变量 X 的分布律见下表

X	-1	1	2
p	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

(1) 试求 X 的分布函数 $F(x)$; (2) 试求概率 $\mathbb{P}(X \leqslant 0)$; $\mathbb{P}\left(\frac{1}{2} < X \leqslant \frac{3}{2}\right)$; $\mathbb{P}(1 \leqslant X \leqslant 2)$ 和 $\mathbb{P}(1 < X \leqslant 2)$.

11. 设 10 件产品中有 8 件是正品, 2 件是次品. 现每次不放回地抽取一件产品直到取到正品为止. 以 X 记抽取的次数, 试求 X 的分布律和分布函数.

12. 设随机变量 X 的分布函数为

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

试求 X 的分布律.

13. 假设某种定期发行的奖券的中奖率为 $p(0<p<1)$. 某人每次购买一张, 若没中奖则接着再买一张, 直到中奖为止. 以 X 表示该人购买奖券的次数, 试求 X 的分布律.

14. 在一串独立试验中观察某事件 A 是否发生, 且假设每次 A 发生的概率都是 $\frac{2}{5}$. 若以 X 表示 A 发生时累计的试验次数, 试求概率 $\mathbb{P}(X$ 为偶数 $)$ 和 $\mathbb{P}(X>2)$.

15. 一篮球运动员连续投篮四次, 且假设每次的结果相互独立. 若至少投进一球的概率为 $80/81$, 则该运动员投篮的命中率为多少?

16. 若随机变量 X 和 Y 分别服从二项分布 $B(2,p)$ 和 $B(3,2p)$, 且 $\mathbb{P}(X\geqslant 1)=0.51$, 试求 $\mathbb{P}(Y\geqslant 1)$.

17. 向目标进行 20 次独立射击, 且假设每次射击的命中率为 0.2. 若以 X 记命中的次数, 试求概率 $\mathbb{P}(X\geqslant 1)$ 及 X 最有可能的取值.

18. 进行 4 次独立试验, 在每次试验中结果 A 出现的概率均为 0.3. 如果 A 不出现, 则 B 也不出现; 如果 A 只出现一次, 则 B 出现的概率是 0.6; 如果 A 出现至少两次, 则 B 出现的概率为 1 . 试求: (1) B 会出现的概率; (2) 若已知 B 出现, 求 A 恰出现一次的概率.

19. 某公司经理拟将一提案交董事代表会批准, 规定如提案获多数代表赞成则通过. 经理估计各代表对此提案投晚成票的概率是 0.6, 并且各代表投票情况相互独立. 为以较大的概率通过提案, 试问经理请 3 名董事代表好还是 5 名好?

20. 有两支篮球队进行友谊杯赛, 假定每一场甲乙两队获胜的概率分别是 0.6 和 0.4, 且各场胜负情况相互独立. 如果规定先胜 4 场者为冠军, 求甲队经过 $i(i=4,5,6,7)$ 场比赛而成为冠军的概率 p_i. 再问与“三场两胜”制比较, 采取哪种赛制对乙队更有利?

21. 甲乙两名运动员进行比赛, 规则采用“五局三胜制”, 即首先赢得三局者获胜. 假定各局比赛结果相互独立且两人的实力相当 (即在每局中各人获胜的概率均为 $\frac{1}{2}$), 试求: (1) 甲正好在四局比赛中就获胜的概率; (2) 若甲已经赢了第一局, 他最终获胜的概率.

22. 有一种赌博, 规则如下: 赌徒先在 1 到 6 中押一个数字, 然后掷三个豇子, 若赌徒所押的数字出现 i 次, $i=1,2,3$, 则署徒赢 i 元; 若其所押的数字没出现则输 1 元. 以随机变量 X 表示赌徒赌完一局后的收益, 试求它的分布律 (假设这些骰子都是均匀的且掷出的点数相互独立).

23. 设随机变量 X 服从参数为 λ 的 Poisson 分布, 且已知 $\mathbb{P}(X=1)=\mathbb{P}(X=2)$, 试求 X 等于 3 的概率.

24. 假设一个人在一年内患感冒的次数服从参数为 5 的 Poisson 分布. 设正在销售的一种新特效药对 75% 的人来说可以将上述 Poisson 分布的参数减小到 3 , 而对另外 25% 的人来说则是无效的. 现某人使用此药一年, 而在期间患了两次感冒, 问此药对他有效的可能性是多少?

25. 设某种昆虫单只每次产卵的数量服从参数为 λ 的 Poisson 分布, 而每个虫卵能孵出幼虫的概率均为 $p(0<p<1)$ 且相互独立. 分别以 Y 和 Z 记一只昆虫一次产卵后幼虫和末能紂出幼虫的虫卵的个数. 试问 Y 和 Z 分别服从什么分布? 它们是否相互独立?

26. 一个系统包含了 1000 个零件, 各个零件出故障是相互独立的并且在一个月内出故障的概率为 0.001. 试利用 Poisson 分布求系统在一个月内正常运转 (即没有零件出故障) 的概率.

27. 保险公司的资料表明持某种人寿保险单的人在保险期内死亡的概率为 2%. 利用 Poisson 分布, 试求在 400 份保单中最终至少赔付两份保单的概率 (精确到小数点后三位).

28. 某种数码传输系统每秒传送 5.12×10^5 个字符 (0 或 1), 由于会受到干扰, 传送中会出现误码, 即将 0 (或 1) 传送为 1 (或 0). 若误码率为 10^{-7}, 求在 10 秒内至少出现一个误码的概率. 在 100 秒内呢? 结果精确到小数点后三位.

29. 某航空公司知道预定航班的乘客有 5% 的概率最终不会来搭乘, 为了更多的盈利他们的政策是接受比实际座位更多的预定. 若一个恰有 50 个座位的航班一共被预定了 52 张票, 问最终出现无法满足所有乘客乘坐要求的情识的概率大约是多少? 结果精确到小数点后两位.

30. 假定有 100 万注彩票出售, 其中有 100 注有奖. (1) 若一个人头了 100 注, 求其中奖的概率; (2) 一个人买多少注才能保证有 95% 的概率中奖?

31. 一位篮球运动员练习投篮 100 次, 且已知他前两次只投进了一次. 从第 3 球开始, 假设他每次投篮的命中率为其前面所投进球的比率 (比如他前五次投进了四个球, 则第六次他的投篮命中率为 $\frac{4}{5}$). 求他最终在这 100 次投篮中投进次数的分布律.

32. 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0\\ \frac{x}{4}, & 0\leqslant x\leqslant 1\\ \frac{1}{2}+\frac{x-1}{4}, & 1\leqslant x<2\\ \frac{5}{6}, & 2\leqslant x<3\\ 1, & x\geqslant 3\end{cases}$$

试求: (1) $\mathbb{P}(X=k), k=1,2,3$; (2)$\mathbb{P}\left(\frac{1}{2}<x<\frac{3}{2}\right)$.

33. 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0\\ a\sin x, & 0\leqslant x\leqslant \frac{\pi}{2},\\ 1, & x>\frac{\pi}{2}.\end{cases}$$

试求常数 a 的值及概率 $\mathbb{P}\left(X>\frac{\pi}{6}\right)$.

34. 若函数 $F(x)=\frac{2}{2+x^2}$ 为随机变量 X 的分布函数, 则 X 可能的取值范围为 ( ). (A) $(-\infty,\infty)$ (B) $(0,\infty)$ (C) $(-\infty,0)$ (D) $(0,1)$

解: F(x) 单调不减且趋于无穷，C

35. (2010 年全国考研试题) 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0\\ \frac{1}{2}, & 0\leqslant x<1\\ 1-\mathrm{e}^{-x}, & x\geqslant 1\end{cases}$$

则 $\mathbb{P}(X=1)$ 等于 (C). (A) 0 (B) 1 (C) $\frac{1}{2}-\mathrm{e}^{-1}$ (D) $1-\mathrm{e}^{-1}$

解: $P(X=1)=1-e^{-1}-\frac{1}{2}=\frac{1}{2}-e^{-1}$

36. 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ \frac{1}{8}, & x=-1,\\ ax+b, & -1<x<1,\\ 1, & x\geqslant 1,\end{cases}$$

且. $\mathbb{P}(X=1)=\frac{1}{4}$, 试求常数 a 和 b 的值.

解: $a=\dfrac{\frac{3}{4}-\frac{1}{8}}{1-(-1)}=\dfrac{5}{16}, b=\dfrac{\frac{3}{4}+\frac{1}{8}}{2}=\dfrac{7}{16}$

37. 设随机变量 X 只在区间 $(0,1)$ 内取值, 且其分布函数 $F(x)$ 满足: 对任意 $0\leqslant a<b\leqslant 1, F(b)-F(a)$ 的值仅与差 $b-a$ 有关. 试证明 X 服从 $(0,1)$ 上的均匀分布.

38. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases} ax, 1<x<2, \\ b, 2\leqslant x<3, \\ 0, \text{其他}. \end{cases}$$

若又知 $\mathbb{P}(1<X<2)=\mathbb{P}(2<X<3)$, 试求常数 a 和 b 的值.

39. 设随机变量 X 的密度函数为

$$f(x)=\frac{a}{1+x^2}, \quad -\infty<x<\infty.$$

试求: (1) 常数 a; (2) 分布函数 $F(x)$; (3) 概率 $\mathbb{P}(|X|<1)$.

40. 设随机变量 X 的密度函数为

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

试求: (1) 常数 a; (2) 概率 $\mathbb{P}\left(|X|<\frac{1}{2}\right)$.

41. 在曲线 $y=2x-x^2$ 与 x 轴所围成的区域中随机取一点, 以 X 表示它与 y 轴之间的距离. 试求 X 的密度函数 $f(x)$ 和分布函数 $F(x)$.

42. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases} 0, & x<1 \\ ax^2\ln x+bx^2+1, & 1\leqslant x\leqslant \mathrm{e} \\ 1, & x>\mathrm{e} \end{cases}$$

试求: (1) 常数 a,b; (2) 随机变量 X 的密度函数 $f(x)$.

43. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases} a\mathrm{e}^{x}, & x<0 \\ b, & 0\leqslant x<1 \\ 1-a\mathrm{e}^{-(x-1)}, & x\geqslant 1 \end{cases}$$

试求: (1) 常数 a,b; (2) 随机变量 X 的密度函数 $f(x)$.

44. (2011 年全国考研试题) 设 $F_1(x), F_2(x)$ 为两个分布函数, 其相应密度 $f_1(x), f_2(x)$ 是连续函数, 则必为概率密度的是 ( ) (A) $f_1(x)f_2(x)$ (B) $2f_2(x)F_1(x)$ (C) $f_1(x)F_2(x)$ (D) $f_1(x)F_2(x)+f_2(x)F_1(x)$

解: $f_1(x)F_2(x)+f_2(x)F_1(x)=(F_1(x)F_2(x))'$ 相当于是乘积分布函数的导数

45. 若随机变量 X 服从区间 $(-5,5)$ 上的均匀分布, 求方程 $x^2+Xx+1=0$ 有实根的概率.

46. 某城际列车从早上六点整开始每 15 分钟发出一趟列车, 假设某乘客达到车站的时间服从七点到七点半的均匀分布, 若忽略买票等其它时间, 试求该乘客等车少于 5 分钟的概率.

47. 设随机变量 X 服从区间 $(1,4)$ 上的均匀分布, 现对 X 进行三次独立观测, 试求至少两次观测值大于 2 的概率.

48. (2013 年全国考研试题) 设随机变量 Y 服从参数为 1 的指数分布, a 为常数且大于零, 则 $\mathbb{P}(Y\leqslant a+1\mid Y>a)=$

49. 假定一机器的检修时间服从参数为 $\lambda=1$ 的指数分布 (单位: 小时). 试求: (1) 检修时间会超过 2 小时的概率; (2) 若已经检修了 2 小时, 总检修时间会超过 4 小时的概率.

50. 设顾客在某银行的窗口等待服务的时间 X 服从参数为 $\lambda=1/5$ 的指数分布 (单位: 分钟). 假设某顾客一旦等待时间超过 10 分钟他就立即离开, 且一个月内要到该银行 5 次, 试求他在一个月内至少有一次未接受服务而离开的概率.

51. (1) 设 X 为正值连续型随机变量, 试证明它服从指数分布的充分必要条件是对任意的常数 $t,x>0$, 均有

$$\mathbb{P}(X\leqslant t+x\mid X>t)=\mathbb{P}(X\leqslant x)$$

(2) 设 X 为取值为正整数的离散型随机变量, 试证明它服从几何分布的充分必要条件是对任意的正整数 m,n, 均有

$$\mathbb{P}(X\leqslant m+n\mid X>n)=\mathbb{P}(X\leqslant m)$$

52. 设随机变量 X 服从标准正态分布. (1) 试求概率 $\mathbb{P}(X<2)$ 和 $\mathbb{P}(|X|\leqslant 2)$; (2) 若常数 a 满足 $\mathbb{P}(|X|>a)<0.1$, 试求 a 的取值范围.

53. 设随机变量 $X\sim N(1,4)$. (1) 试求概率 $\mathbb{P}(0\leqslant X\leqslant 4),\mathbb{P}(X>2.4)$ 和 $\mathbb{P}(|X|>2)$; (2) 试求常数 c, 使得 $\mathbb{P}(X>c)=2\mathbb{P}(X\leqslant c)$.

54. 在一个流水线上我们测量每个电阻器的电阻值 R, 只有电阻值介于 96 和 104 欧姆之间的电阻器才是合格的. 对下列情形试求合格电阻器的比例: (1) 若 R 服从在 95 和 105 之间的均匀分布; (2) 若 R 服从正态分布 $N(100,4)$.

55. 由学校到飞机场有两条路线可供选择: 第一条要穿过市区, 路程短但塞车现象严重, 所需时间 (单位: 分钟) 服从正态分布 $N(30,100)$; 另一条是环城公路, 路程长但很少塞车, 所需时间服从正态分布 $N(40,16)$. 如果要求 (1) 在 50 分钟内达到机场; (2) 在 45 分钟内达到机场. 试问各应该选择哪条路线?

56. 假设在电源电压不超过 200 V, $200\sim 240$ V 和超过 240 V 三种情况下某电子元件损坏的概率分别为 10%, 0% 和 30%. 若电源电压服从 $N(220,225)$ 分布, 试求电子元件会损坏的概率.

57. 根据统计资料, 人的身高 (单位: 厘米) 服从正态分布 $N(\mu,\sigma^2)$. 对于中国成年男子, 可查得 $\mu=170,\sigma=6$. 已知公共汽车门的高度是按成年男子上下车需要低头的比例不超过 0.5% 来设计的, 请问车门的最低高度应该是多少 (保留到小数点后一位)?

58. (2006 年全国考研试题) 设随机变量 X 服从正态分布 $N(\mu_1,\sigma_1^2)$,Y 服从正态分布 $N(\mu_2,\sigma_2^2)$, 且 $\mathbb{P}(|X-\mu_1|<1)>\mathbb{P}(|Y-\mu_2|<1)$, 则 ( ). (A) $\sigma_1<\sigma_2$ (B) $\sigma_1>\sigma_2$ (C) $\mu_1<\mu_2$ (D) $\mu_1>\mu_2$

59. (2010 年全国考研试题) 设 $f_1(x)$ 为标准正态分布的概率密度, $f_2(x)$ 为 $[-1,3]$ 上的均匀分布的概率密度, 若 $f(x)=\begin{cases} af_1(x), & x\leqslant 0,\\ bf_2(x), & x>0,\end{cases}$ $(a>0,b>0)$ 为概率密度, 则 a,b 应满足 ( ). (A) $2a+3b=4$ (B) $3a+2b=4$ (C) $a+b=1$ (D) $a+b=2$

60. (2013 年全国考研试题) 设 X_1,X_2,X_3 是随机变量, 且 $X_1\sim N(0,1),X_2\sim N(0,2^2),X_3\sim N(5,3^2),P_i=\mathbb{P}(-2\leqslant X_i\leqslant 2)\,(i=1,2,3)$, 则 ( ). (A) $P_1>P_2>P_3$ (B) $P_2>P_1>P_3$ (C) $P_3>P_1>P_2$ (D) $P_1>P_3>P_2$

61. (2016 年全国考研试题) 设随机变量 $X\sim N(\mu,\sigma^2)\,(\sigma>0)$, 记 $p=\mathbb{P}(X\leqslant\mu+\sigma^2)$, 则 ( ) . (A) p 随 μ 的增加而增加 (B) p 随 σ 的增加而增加 (C) p 随 μ 的增加而减少 (D) p 随 σ 的增加而械少

62. 设随机变量 X 的分布律为

X	-1	0	1	2
P	0.2	0.3	0.1	0.4

试求下列随机变量的分布律: (1) $Y_1=-2X+1$; (2) $Y_2=|X|$; (3) $Y_3=(X-1)^2$.

63. 设随机变量 X 的分布律为

X	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$
P	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

试求随机变量 $Y=\cos(2X-\pi)$ 和 $Z=\left|X-\frac{\pi}{2}\right|$ 的分布律.

64. 设连续型随机变量 X 的分布函数为

$$F(x)=a+b\arctan x,\quad -\infty<x<\infty.$$

(1) 试求常数 a,b 的值. (2) 问 $\mathbb{E}X=0$ 是否成立? 为什么? (3) 试求随机变量 $Y=3-\sqrt[3]{X}$ 的密度函数 $p(y)$. (4) 试证明 X 与 $\frac{1}{X}$ 具有相同的分布.

65. 设随机变量 $X\sim U(0,1)$, 试求下列随机变量的密度函数. (1) $Y_1=\mathrm{e}^X$; (2) $Y_2=X^{-1}$; (3) $Y_3=-\frac{1}{\lambda}\ln X$, 其中 $\lambda>0$ 为常数.

66. 设随机变量 $X\sim U\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$, 试分别求 $Y_1=\tan X$ 和 $Y_2=\cos X$ 的密度函数.

67. 设随机变量 X 的分布函数 $F(x)$ 为严格单调函数, 证明: 随机变量 $Y=F(X)$ 服从区间 $(0,1)$ 上的均匀分布.

68. 设随机变量 X 服从参数为 1 的指数分布, 试分别求 $Y_1=X^2$ 和 $Y_2=1-\mathrm{e}^{-X}$ 的密度函数.

69. 设随机变量 X 的概率密度为 $f(x)=2(1-x),0<x<1$. 试构造区间 $(0,1)$ 上的一个单调递增函数 $g(x)$, 使得 $g(X)$ 恰好服从参数为 1 的指数分布.

70. 设随机变量 X 服从参数为 λ 的指数分布, 以 $Y = \lfloor X \rfloor$ 表示它的整数部分, 即不超过 X 的最大整数, 而以 Z 表示它的小数部分, 即 $Z = X - \lfloor X \rfloor$. 试求随机变量 Y 和 Z 各自的分布, 且它们是否相互独立?

71. 设随机变量 X 服从参数为 λ 的指数分布, 且随机变量 Y 定义为

$$Y = \begin{cases} X, & \text{若} X \geqslant 1, \\ -X^2, & \text{若} X < 1. \end{cases}$$

试求 Y 的密度函数 $p(y)$.

72. 设随机变量 X 服从标准正态分布, 试求下列随机变量的密度函数. (1) $Y_1 = \mathrm{e}^X$; (2) $Y_2 = |X|$; (3) $Y_3 = 2X^2 + 1$.

73. (2013 年全国考研试题) 设随机变量 X 的密度函数为 $f(x) = \frac{1}{a}x^2, 0 < x < 3$, 令随机变量

$$Y = \begin{cases} 2, & X \leqslant 1 \\ X, & 1 < X < 2 \\ 1, & X > 2 \end{cases}$$

(1) 求随机变量 Y 的分布函数. (2) 求概率 $\mathbb{P}(X \leqslant Y)$.

第三章 多维随机变量及其分布

本章知识点总结见上一章

第 1 节 思考题

1. (2013 年全国考研试题) 设随机变量 X 和 Y 相互独立, 且 X 和 Y 的概率分布分别为

X	0	1	2	3
$\mathbb{P}$	1/2	1/4	1/8	1/8

Y	-1	0	1
$\mathbb{P}$	1/3	1/3	1/3

则 $\mathbb{P}(X+Y=2)=($ $)$. (A) 1/12 (B) 1/8 (C) 1/6 (D) 1/2

2. 从 $1,2,3,4$ 四个数中任取一个数, 记为 X, 再从 1 到 X 中任取一个数, 记为 Y, 求 $\{Y=2\}$ 这个事件发生的概率是多少?

3. (2011 年全国考研试题) 设随机变量 X,Y 的分布律分别为

X	0	1
$\mathbb{P}$	1/3	2/3

Y	-1	0	1
$\mathbb{P}$	1/3	1/3	1/3

且 $\mathbb{P}(X^2=Y^2)=1$. 求: (1) (X,Y) 的分布; (2) $Z=XY$ 的分布.

4. (2010 年全国考研试题) 箱中装有 6 个球, 其中红、白、黑球的个数分别为 $1,2,3$ 个. 现从箱中随机地取出 2 个球, 记 X 为取出的红球个数, Y 为取出的白球个数. 求随机变量 (X,Y) 的概率分布.

5. (2009 年全国考研试题) 袋中有一个红球, 两个黑球, 三个白球, 现有放回地从袋中取两次, 每次取一球, 以 X,Y,Z 分别表示两次取球的红、黑、白球的个数. (1) 求 $\mathbb{P}(X=1\mid Z=0)$; (2) 求二维随机变量 (X,Y) 的概率分布.

6. 将 n 封信随机地投入到两个邮筒, 用 X_1,X_2 分别表示第 1,2 两个邮筒内信的数目, 求 (X_1,X_2) 的联合分布以及 X_1,X_2 的边缘分布.

7. 将同一硬币连续掷三次, 以 X 表示在三次中出现正面的次数, 以 Y 表示三次中出现的正面次数和出现的反面次数之差的绝对值. 试写出 X 和 Y 的联合分布.

8. 现有某种产品 100 个, 其中一、二、三等品分别为 $80,10,10$ 个, 现从中随机抽取一个产品, 记

$$X_1=\begin{cases}1, & \text{抽到一等品},\\ 0, & \text{其他},\end{cases}\qquad X_2=\begin{cases}1, & \text{抽到二等品},\\ 0, & \text{其他}.\end{cases}$$

求 (X_1,X_2) 的联合分布.

9. 从一副扑克牌 (共 52 张) 中任取 13 张, 以 X 和 Y 分别记其中的黑桃和红桃张数. 试求: (1) (X,Y) 的联合概率分布; (2) 已知取出的只有一张黑桃, 求此时 Y 的条件概率分布.

10. 设二维随机向量的概率分布为

X	-1	1
-1	0.2	b
1	a	0.3

已知事件 $\{X=-1\}$ 和 $\{X+Y=0\}$ 相互独立, 求 a,b.

11. 设某射手每次射中目标的概率为 $p(0<p<1)$, 射击进行到第二次射中目标为止, X 表示第一次射中目标所进行的射击次数, Y 表示第二次射中目标时所进行的射击次数. (1) 求二维随机变量 (X,Y) 的分布律; (2) 求 X 和 Y 的边缘分布.

12. 设 $F_1(x),F_2(y)$ 为分布函数, $f_1(x),f_2(y)$ 为相应的密度函数, 则下述二元函数中哪个不是密度函数: ( ) (A) $f_1(x)f_2(y)$ (B) $f_1(x)f_2(y)\left(1+\left(2F_1(x)-1\right)\left(2F_2(y)-1\right)\right)$ (C) $0.5\left(f_1(x)+f_2(y)\right)$ (D) $0.5\left(f_1(x)f_1(y)+f_2(x)f_2(y)\right)$

13. (2011 年全国考研试题) 设二维随机变量 (X,Y) 服从 $N\left(\mu,\mu,\sigma^2,\sigma^2,0\right)$, 则 $\mathbb{E}\left(XY^2\right)=$

14. (2015 年全国考研试题) 设二维随机变量 X,Y 服从正态分布 $N(1,0,1,1,0)$, 则 $\mathbb{P}(XY-Y<0)=$

15. 设随机变量 X,Y 均服从正态分布 $N(0,1)$, 且已知 $\mathbb{P}(X\leqslant 2,Y\leqslant -2)=0.25$, 则 $\mathbb{P}(X>2,Y>-2)=$

16. 试举例说明: (X,Y) 的边缘分布相同时联合分布不一定相同.

17. 试举例说明: 存在标准正态随机变量 X,Y, 它们的联合分布 (X,Y) 不是二元止态的.

18. 设 $(X,Y)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$. (1) 求 X,Y 的边缘分布; (2) 求 $f_{X\mid Y}(x\mid y)$ 和 $f_{Y\mid X}(y\mid x)$.

19. 设二维随机变量 (X,Y) 的分布函数为

$$F(x,y)=a(b+\arctan x)(c+\arctan y),\quad x,y\in\mathbb{R}.$$

(1) 确定常数 a,b,c; (2) 求 $\mathbb{P}(X>0,Y>0)$; (3) 求 X 和 Y 的边缘密度函数.

20. 设二维随机变量 (X,Y) 的分布函数为

$$F(x,y)=\begin{cases}\left(1-e^{-\lambda x}\right)\left(1-e^{-\mu y}\right), & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$$

求 (X,Y) 的联合密度函数.

21. 设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases}\cos x\cos y, & 0<x<\pi/2,0<y<\pi/2,\\ 0, & \text{其他}.\end{cases}$$

(1) 试求 (X,Y) 的分布函数; (2) 试求概率 $\mathbb{P}(0<X<\pi/4,\pi/4<Y<\pi/2)$.

22. (2010 年全国考研试题) 设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=Ae^{-2x^2+2xy-y^2},\quad x,y\in\mathbb{R}.$$

求常数 A 及条件密度函数 $f_{Y\mid X}(y\mid x)$.

23. 设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases} Ax^2, & 0<|x|<y<1, \\ 0, & \text{其他}. \end{cases}$$

求常数 A 及条件概率 $\mathbb{P}(X\leqslant 0.25 \mid Y=0.5)$.

24. 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 $\mathbb{P}(Y=0)=\mathbb{P}(Y=1)=1/2$, 求随机变量 $Z=XY$ 的分布函数.

25. 设 X 和 Y 是相互独立的随机变量, $X\sim N\left(0,\sigma_1^2\right)$, $Y\sim N\left(0,\sigma_2^2\right)$, 其中 $\sigma_1,\sigma_2>0$ 为常数. 引入随机变量

$$Z=\begin{cases} 1, & X\leqslant Y \\ 0, & X>Y \end{cases}$$

求 Z 的分布律.

26. (2009 年全国考研试题) 设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases} e^{-x}, & 0<y<x, \\ 0, & \text{其他}. \end{cases}$$

(1) 求条件密度函数 $f_{Y|X}(y\mid x)$. (2) 求条件概率 $\mathbb{P}(X\leqslant 1\mid Y\leqslant 1)$.

27. 设二维随机变量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases} c\left(R-\sqrt{x^2+y^2}\right), & x^2+y^2<R^2, \\ 0, & x^2+y^2\geqslant R^2. \end{cases}$$

(1) 求 c 的值; (2) 求 (X,Y) 落在圆 $\{(x,y):x^2+y^2\leqslant r^2\}\,(r<R)$ 内的概率.

28. 设 (X,Y) 的联合密度函数为:

$$f(x,y)=\begin{cases} Ae^{-(3x+4y)}, & x>0,y>0, \\ 0, & \text{其他}. \end{cases}$$

(1) 求系数 A; (2) X 与 Y 是否独立; (3) 求 $Z=X+Y$ 的密度函数 $f_Z(z)$; (4) 试求 $\mathbb{E}(X\mid X+Y=1)$.

29. (2013 年全国考研试题) 设 (X,Y) 是二维随机变量, X 的边缘密度函数为

$$f_X(x)=\begin{cases} 3x^2, & 0<x<1 \\ 0, & \text{其他} \end{cases},$$

在给定 $X=x(0<x<1)$ 的条件下, Y 的条件密度函数为

$$f_{Y|X}(y\mid x)=\begin{cases} \frac{3y^2}{x^3}, & 0<y<x, \\ 0, & \text{其他}. \end{cases}$$

(1) 求 (X,Y) 的联合密度函数 $f(x,y)$; (2) Y 的边缘密度函数 $f_Y(y)$.

30. 设随机变量 X 的密度函数为 $f_X(x)=\begin{cases} xe^{-x}, & x>0 \\ 0, & x\leqslant 0 \end{cases}$, 而随机变量 Y 服从 $(0,X)$ 上的均匀分布, 求 (1) (X,Y) 的联合分布; (2) 随机变量 Y 的分布.

31. 设在 $X=x$ 的条件下, Y 服从参数为 x 的 Poisson 分布, 已知 X 服从标准指数分布, 求 Y 的分布律.

32. 在 n 重 Bernoulli 试验中, 以 X_i 记第 i 次试验中成功的次数, 求 $(X_1,X_2,\cdots,X_n)$ 的联合分布.

33. 设 (X,Y) 服从 $\{(x,y):0<x<1,|y|<x\}$ 上的均匀分布, 求 X 的边缘密度函数.

34. 设 (X,Y) 是矩形 $\{(x,y):0\leqslant x\leqslant 2,0\leqslant y\leqslant 2\}$ 上的均匀分布, 求随机变量 $Z=|X-Y|$ 的密度函数.

35. (2011 年全国考研试题) 设 (X,Y) 在 G 上服从均匀分布, G 由 $x-y=0,x+y=2$ 与 $y=0$ 围成. (1) 求边缘密度函数 $f_X(x)$; (2) 求 $f_{X|Y}(x\mid y)$.

36. 设随机向量 (X,Y) 服从区域 D 内的均匀分布, 其中 D 是由直线 $y=x,x=0,y=1$ 所围成的区域, 试求: (1) (X,Y) 的联合密度函数 $f(x,y)$; (2) (X,Y) 的边缘密度函数 $f_1(x)$ 和 $f_2(y)$; (3) 条件密度 $f_{X|Y}(x\mid y)$; (4) $\mathbb{E}(X\mid Y=y)$.

37. 设随机向量 (X,Y) 服从 $\{(x,y):|x+y|\leqslant 1,|x-y|\leqslant 1\}$ 内的均匀分布, (1) 试求出 X 和 Y 的边缘分布; (2) X 和 Y 是否相互独立? (3) 求在 $X=x(0<x<1)$ 时 Y 的条件密度函数.

38. 设 (X,Y) 服从矩形 $\{(x,y):0\leqslant x\leqslant 2,0\leqslant y\leqslant 1\}$ 内的均匀分布, 记

$$U=\begin{cases} 0, & X\leqslant Y, \\ 1, & X>Y, \end{cases}\quad V=\begin{cases} 0, & X\leqslant 2Y, \\ 1, & X>2Y, \end{cases}$$

求 (U,V) 的联合概率分布.

39. 假设有 $n(n\geqslant 3)$ 个不同的盒子与 m 个相同的小球, 每个小球独立地以概率 p_k 落入第 k 个盒子 $(k=1,2,\cdots,n)$. 分别以 $X_1,X_2,\cdots,X_n$ 表示落入各个盒子的球数. 试求 (1) $(X_1,X_2,\cdots,X_n)$ 的联合分布; (2) X_k 的边缘分布, 其中 $k=1,2,\cdots,n$; (3) (X_1,X_2) 的边缘分布; (4) 在 $X_1=m_1$ 的条件下 $(X_2,\cdots,X_n)$ 的条件分存.

40. 设 (X,Y,Z) 服从单位球 $\{(x,y,z):x^2+y^2+z^2\leqslant 1\}$ 内的均匀分布, 试求 X 的边缘分布.

41. (2012 年全国考研试题) 设随机变量 X 与 Y 相互独立, 且都服从区间 $(0,1)$ 上的均匀分存, 则 $\mathbb{P}\left(X^2+Y^2\leqslant 1\right)=(\quad)$. (A) $1/4$ (B) $1/2$ (C) $\pi/8$ (D) $\pi/4$

42. 设随机变量 X 与 Y 独立且具有相同的分布, 设

$$\mathbb{P}(X=0)=\mathbb{P}(X=1)=1/2.$$

则下列各式中正确的是 () . (A) $\mathbb{P}(X=Y)=1/2$ (B) $\mathbb{P}(X=Y)=1$ (C) $\mathbb{P}(XY=0)=1/4$ (D) $\mathbb{P}(X+Y=1)=1/4$

43. 设 X 和 Y 相互独立且分别服从均值为 1 和 $1/4$ 的指数分布, 则 $\mathbb{P}(X<Y)=$

44. 在区间 $[0,1]$ 中随机独立地取两个数 X 和 Y⊠ 则这两个数之差的绝对值小于 $1/2$ 的概率为

45. 设随机变量 X 与 Y 相互独立分别服从参数为 μ 和 λ 的 Possion 分布, 则 $\mathbb{P}(X=k\mid X+Y=n)=$ 即在给定 $X+Y=n$ 的条件下, X 的条件分布为

46. 设随机变量 X,Y 的分布律分别为

X	0	1
$\mathbb{P}$	$1/2$	$1/2$

Y	-1	0	1
$\mathbb{P}$	$1/4$	$1/2$	$1/4$

且. $\mathbb{P}(XY=0)=1$. (1) 求 (X,Y) 的联合分布; (2) 问 X,Y 是否独立.

47. 设某班车起点站上客人数 X 服从参数为 $\lambda(\lambda>0)$ 的 Poisson 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 个人下车的概率; (2) 二维随机变量 (X,Y) 的概率分布.

48. 连续地掷一颗均匀的骰子, 直到出现点数大于 2 为止, 以 X 表示掷骰子的次数, 以 Y 表示最后一次掷出的点数. (1) 求二维随机变量 (X,Y) 的联合分布以及边缘分布; (2) 问 X 和 Y 是不相互独立.

49. (2008 年全国考研试题) 设随机变量 X 与 Y 相互独立, X 的概率分布为

$$\mathbb{P}(X=i)=1/3,\quad i=-1,0,1,$$

Y 的密度函数为 $f_Y(y)=\begin{cases}1, & 0\leqslant y\leqslant 1\\ 0, & \text{其他}\end{cases}$, 记 $Z=X+Y$. (1) 求 $\mathbb{P}(Z\leqslant 1/2\mid X=0)$; (2) 求 Z 的密度函数.

50. 设随机变量 X 与 Y 相互独立, X 服从标准指数分布, Y 服从标准正态分布, 求 $(X,|Y|)$ 的联合密度函数.

51. 设 (X,Y) 的联合密度函数为:

$$f(x,y)=\begin{cases}0.25(1+xy), & |x|<1,|y|<1,\\ 0, & \text{其他}.\end{cases}$$

(1) 求给定 $X=1/2$ 时 Y 的条件密度函数; (2) 证明 X^2 和 Y^2 相互独立.

52. 设 (X,Y) 服从矩形 $\{(x,y):0<x<1,0<y<1\}$ 内的均匀分布, (1) 求 X 与 Y 的边缘分布; (2) 问 X,Y 是否相互独立.

53. 设 (X,Y) 服从单位圆 $\{(x,y):x^2+y^2\leqslant 1\}$ 内的均匀分布, (1) 求 X 与 Y 的边缘分布; (2) 问 X,Y 是否相互独立.

54. 设随机向量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases}3x, & 0<x<1,0<y<x,\\ 0, & \text{其他}.\end{cases}$$

(1) 求 X 与 Y 的边缘密度函数; (2) 问 X 与 Y 是否相互独立? (3) 计算 $\mathbb{P}(X+Y\leqslant 1)$.

55. 设 (X,Y) 的联合分布函数为 $F(x,y)=\begin{cases}\frac{\left(1-(x+1)e^{-x}\right)y}{1+y}, & x>0,y>0,\\ 0, & \text{其他}.\end{cases}$ (1) 求 X,Y 的边缘分布函数 $F_X(x)$ 和 $F_Y(y)$; (2) 求 (X,Y) 的联合密度函数 $f(x,y)$ 以及边缘密度函数 $f_X(x),f_Y(y)$; (3) 验证 X,Y 是否相互独立.

56. 设 X,Y 是两个相互独立的随机变量, X 在 $(0,1)$ 上服从均匀分布, Y 的密度函数为

$$f_Y(y)=\begin{cases}\frac{1}{2}e^{-y/2}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

(1) 求 (X,Y) 的联合密度函数; (2) 求二次方程 $a^2+2Xa+Y=0$ 有实根的概率.

57. 设随机向量 (X,Y,Z) 的密度函数为

$$f(x,y,z)=\begin{cases}\frac{1}{8\pi^3}(1-\sin x\sin y\sin z), & 0\leqslant x,y,z\leqslant 2\pi,\\ 0, & \text{其他}.\end{cases}$$

证明 X,Y,Z 两两独立但不相互独立.

58. 设随机向量 (X,Y) 的密度函数为

$$f(x,y)=\begin{cases}\frac{1+xy}{4}, & |x|<1,|y|<1,\\ 0, & \text{其他}.\end{cases}$$

证明 X,Y 不独立但是 X^2,Y^2 是相互独立的.

第四章 数字特征

1. **期望：**

$$E(X)=\sum_{i=1}^{\infty}x_ip_i(\text{离散型}),E(X)=\int_{-\infty}^{\infty}xf(x)dx(\text{连续型})$$

注 4.0.1. $E(X)=\infty$ 则不存在，如 *Cauchy* 分布 $\int_{-\infty}^{\infty}|x|p(x)dx=\frac{2}{\pi}\int_0^{\infty}\frac{x}{1+x^2}dx=\infty$

性质 4.0.1. 期望的性质

- $E\left(c_1X_1+c_2X_2+\cdots+c_nX_n\right)=c_1EX_1+c_2EX_2+\cdots+c_nEX_n$
- $E\left(X_1X_2\cdots X_n\right)=EX_1EX_2\cdots EX_n$(独立的随机变量)
- $Eg(X)=\begin{cases}\sum_i g\left(a_i\right)p_i, & \sum_i\left|g\left(a_i\right)\right|p_i<\infty;\\ \int_{-\infty}^{+\infty}g(x)f(x)dx, & \int_{-\infty}^{+\infty}|g(x)|f(x)dx<\infty.\end{cases}$ (随机变量函数的期望)

条件期望：

$$E(Y\mid X=x)=\begin{cases}\int_{-\infty}^{+\infty}yf(y\mid x)dy, & (X,Y)\text{ 为连续型};\\ \sum_i a_ip_i, & (X,Y)\text{ 为离散型}.\end{cases}$$

全期望公式：

$$EX=E\{E[X\mid Y]\}$$

2. **方差：**

$$\operatorname{Var}(X)=\sigma^2=E(X-EX)^2=EX^2-(EX)^2$$

注 4.0.2. 正态分布的标准化随机变量：$X^*=\dfrac{X-EX}{\sqrt{\operatorname{Var}(X)}}$

性质 4.0.2. 方差的性质

- $0\leqslant\operatorname{Var}(X)=EX^2-(EX)^2$
- $\operatorname{Var}(X)\leqslant E(X-C)^2,C=EX$ 等号成立
- $\operatorname{Var}(aX+bY)=a^2\operatorname{Var}(X)+b^2\operatorname{Var}(Y)$
- 乘积的方差大于方差的乘积：

$$Var(XY)=VarX\cdot VarY+VarX(EY)^2+VarY(EX)^2\geqslant VarX\cdot VarY.$$

- 与协方差：

$$\begin{aligned}&Var(X\pm Y)=VarX+VarY\pm2\operatorname{Cov}(X,Y),\\&Var\left(\sum_{i=1}^n a_iX_i\right)=\sum_{i=1}^n a_i^2VarX_i+2\sum_{1\leqslant i<j\leqslant n}a_ia_j\operatorname{Cov}\left(X_i,X_j\right).\end{aligned}$$

3. **矩：设 X 为随机变量，c 为常数，r 为正整数，则 $E\left[(X-c)^r\right]$ 称为 X 关于 c 点的 r 阶矩.**

- **平均绝对差** $E|X-EX|$
- **矩母函数** $g(t)=Ee^{tX}$，**其中** $t\in\mathbb{R}$.
- **特征函数** $\phi(t)=Ee^{itX}$，**其中** $t\in\mathbb{R},i$ **为虚数**

第 1 节 不同分布的数字特征

1 离散型分布

I. 0-1 分布

X	1	0
p_k	p	$1-p$

$$E(X)=1\times p+0\times(1-p)=p, E(X^2)=p^2 \tag{4.1}$$

$$Var(X)=E\left(X^2\right)-[E(X)]^2=p(1-p)=pq \tag{4.2}$$

II. 二项分布 $X\sim\mathbf{B}(n,p), P\{X=k\}=\binom{n}{k}p^kq^{n-k},\quad k=0,1,2,\cdots,n,\quad q=1-p$

$$\begin{aligned}E(X)&=\sum_{k=0}^{n}kp_k=\sum_{k=0}^{n}k\binom{n}{k}p^kq^{n-k}=\sum_{k=0}^{n}k\frac{n!}{k!(n-k)!}p^kq^{n-k}\\&=\sum_{k=1}^{n}\frac{n!}{(k-1)!(n-k)!}p^kq^{n-k}=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!(n-k)!}p^{k-1}q^{n-k}\\&=np\sum_{k=1}^{n-1}\binom{n-1}{k-1}p^{k-1}q^{(n-1)-(k-1)}=np(p+q)^{n-1}=np\end{aligned} \tag{4.3}$$

$$\begin{aligned}E\left(X^2\right)=E[X(X-1)+X)]&=E[X(X-1)]+E(X)\\&=\sum_{k=0}^{n}k(k-1)\binom{n}{k}p^kq^{n-k}+np\\&=\sum_{k=0}^{n}k(k-1)\frac{n!}{k!(n-k)!}p^kq^{n-k}+np\\&=n(n-1)p^2\sum_{k=2}^{n}\frac{(n-2)!}{(k-2)!(n-k)!}p^{k-2}q^{n-k}+np\\&=n(n-1)p^2(p+q)^{n-2}+np=n(n-1)p^2+np\end{aligned} \tag{4.4}$$

$$Var(X)=E\left(X^2\right)-[E(X)]^2=npq \tag{4.5}$$

III. 负二项分布 $X\sim\mathbf{NB}(p), P\{X=k\}=\binom{k-1}{r-1}p^r(1-p)^{k-r},\quad k=r,r+1,r+2,\cdots,+\infty$

$$\begin{aligned}E(X)&=\sum_{k=r}^{\infty}kP\{X=k\}=\sum_{x=r}^{\infty}k\binom{k-1}{r-1}p^r(1-p)^{x-r}\\&=\frac{r}{p}\sum_{k=r}^{\infty}\binom{k}{r}p^{r+1}(1-p)^{x-r}=\frac{r}{p}\sum_{k=r}^{\infty}P\{X=k+1\}=\frac{r}{p}\end{aligned} \tag{4.6}$$

$$\begin{aligned}E(X^2)&=\sum_{k=r}^{\infty}k^2P\{X=k\}=\sum_{x=r}^{\infty}k^2\binom{k-1}{r-1}p^r(1-p)^{x-r}\\&=\sum_{x=r}^{\infty}k^2\binom{k-1}{r-1}p^r(1-p)^{x-r}=\frac{r}{p}\sum_{k=r}^{\infty}k\binom{k}{r}p^{r+1}(1-p)^{x-r}\\&=\frac{r}{p}\sum_{k=r}^{\infty}(k+1-1)\binom{k}{r}p^{r+1}(1-p)^{x-r}=\frac{r}{p}\left(\frac{r+1}{p}-1\right)=\frac{r(r-p+1)}{p^2}\end{aligned} \tag{4.7}$$

$$Var(X) = E\left(X^2\right) - [E(X)]^2 = \frac{r(1-p)}{p^2} \tag{4.8}$$

IV. Poisson 分布 $X \sim \mathbf{Poi}(\lambda), P\{X=k\} = \frac{\lambda^k}{k!}e^{-\lambda}, \quad k=0,1,2,3,\ldots,\lambda>0$

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!}e^{-\lambda} \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{(j)!} = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda \end{aligned} \tag{4.9}$$

$$\begin{aligned} E\left(X^2\right) &= E[X(X-1)+X] = E[X(X-1)] + E[X] \\ &= \sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} + \lambda = \lambda^2 e^{-\lambda} \cdot e^{\lambda} + \lambda = \lambda^2 + \lambda \end{aligned} \tag{4.10}$$

$$Var(X) = E\left(X^2\right) - [E(X)]^2 = \lambda \tag{4.11}$$

V. 几何分布 $X \sim \mathbf{G}(p), P\{X=k\} = p(1-p)^{k-1}, \quad k=1,2,3,\ldots,+\infty$

$$\begin{aligned} E(X) &= \sum_{j=1}^{\infty} j p_j = \sum_{j=1}^{\infty} j q^{j-1} p = \frac{p}{q} \sum_{j=1}^{\infty} j q^j = \frac{p}{q} \lim_{n\to\infty} \sum_{j=1}^{n} j q^j \\ &= \frac{p}{q} \lim_{n\to\infty} \frac{q\left(1-q^n-nq^n+nq^{n+1}\right)}{(1-q)^2} = \frac{1}{p} \end{aligned} \tag{4.12}$$

$$\begin{aligned} E(X^2) &= \sum_{j=1}^{\infty} j^2 p_j = \sum_{j=1}^{\infty} j^2 q^{j-1} p = \frac{p}{q} \sum_{j=1}^{\infty} j^2 q^j = \frac{p}{q} \lim_{n\to\infty} \sum_{j=1}^{n} j^2 q^j \\ &= \frac{p}{q}\left[\frac{q}{(1-q)^2}\right]' = \frac{p}{q}\frac{1+q}{(1-q)^3} = \frac{2-p}{p^2} \end{aligned} \tag{4.13}$$

$$Var(X) = E\left(X^2\right) - [E(X)]^2 = \frac{1-p}{p} \tag{4.14}$$

以上推导中有两个和式使用了错位相减法：

(a) $S = \sum_{j=1}^{\infty} j q^j$

$$\begin{aligned} S &= 1 + 2q + 3q^2 + \ldots + nq^{n-1} \\ qS &= q + 2q^2 + 3q^3 + \ldots (n-1)q^{n-1} + nq^n \\ \Rightarrow (1-q)S &= 1 + q + q^2 + q^3 + \ldots + q^{n-1} - nq^n = \frac{1-q^n}{1-q} - nq^n \\ \therefore S &= \frac{1-q^n}{(1-q)^2} - \frac{nq^n}{1-q} \end{aligned}$$

(b) $T = \sum_{j=1}^{\infty} j^2 q^j$

$$\begin{aligned} T &= 1 + 2^2 q + 3^2 q^2 + \ldots + n^2 q^{n-1} = \left(q + 2q^2 + 3q^3 + \ldots + nq^n\right)' = S' \\ &= \left(\frac{q}{(1-q)^2}\right)' = \frac{1-q^2}{(1-q)^4} = \frac{1+q}{(1-q)^3} \end{aligned}$$

VI. **超几何分布** $X\sim\mathbf{H}(N,M,k),P\{X=k\}=\dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},M,N,n\in\mathbb{N}$

$$\begin{aligned}E(X)&=\sum_{k=0}^{n}k\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}=\frac{n}{N}\sum_{k=0}^{n}k\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N-1}{n-1}}=\frac{nM}{N}\sum_{k=0}^{n}\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N-1}{n-1}},\because\binom{N}{n}=\frac{N}{n}\binom{N-1}{n-1}\\&=\frac{nM}{N}\sum_{k=0}^{n}\frac{\binom{M}{k}\binom{N-M}{n-1-(k-1)}}{\binom{N-1}{n-1}}=\frac{nM}{N}\sum_{k=0}^{n}P\{X=k-1\}=\frac{nM}{N}\end{aligned}\tag{4.15}$$

$$\begin{aligned}\text{由于}E(X(X-1))&=\sum_{k=2}^{n}k(k-1)\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}=\frac{n(n-1)}{N(N-1)}\sum_{k=2}^{n}k(k-1)\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N-2}{n-2}}\\&=\frac{n(n-1)M(M-1)}{N(N-1)}\sum_{k=2}^{n}\frac{\binom{M-2}{k-2}\binom{N-M}{n-k}}{\binom{N-2}{n-2}}=\frac{n(n-1)M(M-1)}{N(N-1)}\sum_{k=2}^{n}P\{X=k-2\}\\&=\frac{n(n-1)M(M-1)}{N(N-1)}\Rightarrow E(X^2)=\frac{nM(nM-M-n-N+2)}{N(N-1)}\end{aligned}\tag{4.16}$$

$$Var(X)=E\left(X^2\right)-[E(X)]^2=\frac{nM}{N}\frac{N-M}{N}\frac{N-n}{N-1}\tag{4.17}$$

2 连续型分布

I. **均匀分布** $X\sim\mathbf{U}[a,b],f(x)=\dfrac{1}{b-a}I_{(a\leqslant x\leqslant b)}$

$$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_a^b x\cdot\frac{1}{b-a}dx=\frac{a+b}{2}\tag{4.18}$$

$$E\left(X^2\right)=\int_{-\infty}^{+\infty}x^2f(x)dx=\int_a^b x^2\cdot\frac{1}{b-a}dx=\frac{a^2+ab+b^2}{3}\tag{4.19}$$

$$Var(X)=E\left(X^2\right)-[E(X)]^2=\frac{1}{3}\left(a^2+ab+b^2\right)-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}\tag{4.20}$$

II. **指数分布** $X\sim\mathbf{Exp}(\lambda),f(x)=\lambda e^{-\lambda x}I_{(x\geqslant 0)}$

$$E(X)=\int_{-\infty}^{\infty}|x|f(x)dx=\int_0^{\infty}xf(x)dx=\frac{1}{\lambda}\int_0^{\infty}\lambda xe^{-\lambda x}d\lambda x=\frac{1}{\lambda}\int_0^{\infty}ue^{-u}du=\frac{1}{\lambda}\tag{4.21}$$

$$E\left(X^2\right)=\int_0^{\infty}x^2f(x)dx=\int_0^{\infty}x^2\cdot\lambda e^{-\lambda x}dx=\frac{1}{\lambda^2}\int_0^{\infty}\lambda x\lambda xe^{-\lambda x}d\lambda x=\frac{\Gamma(3)}{\lambda^2}=\frac{2}{\lambda^2}\tag{4.22}$$

$$Var(X)=E\left(X^2\right)-[E(X)]^2=\frac{1}{\lambda^2}\tag{4.23}$$

III. **正态分布** $X\sim\mathbf{N}(\mu,\sigma),f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}I_{(-\infty\leqslant x\leqslant+\infty)}$

$$E(X)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(\sigma t+\mu)e^{-\frac{t^2}{2}}dt=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}xe^{-\frac{t^2}{2}}dt=\mu\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\mu\tag{4.24}$$

$$\begin{aligned} Var(X) = E[X-E(X)]^2 &= \int_{-\infty}^{+\infty}(x-\mu)^2\cdot\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x \\ &= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2\mathrm{e}^{\frac{t^2}{2}}\,\mathrm{d}t \quad \left(t=\frac{x-\mu}{\sigma}\right) \\ &= \frac{\sigma^2}{\sqrt{2\pi}}\left[t\left(-\mathrm{e}^{-\frac{t^2}{2}}\right)\Big|_{-\infty}^{+\infty}+\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t\right] \\ &= \frac{\sigma^2}{\sqrt{2\pi}}\cdot\sqrt{2\pi}=\sigma^2 \end{aligned} \tag{4.25}$$

$$E(X^2)=\mathrm{Var}(X)+E^2(X)=\mu^2+\sigma^2 \tag{4.26}$$

标准正态分布 k 阶原点矩公式:

$$\begin{aligned} E\left(X^k\right) &= \int_{-\infty}^{+\infty}x^k\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx \\ &= \frac{1}{k+1}x^{k+1}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty}+\frac{1}{k+1}\int_{-\infty}^{+\infty}x^{k+2}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx \\ &= \frac{1}{k+1}E\left(X^{k+2}\right)=\begin{cases}(k-1)!! & ,k=2i,i=1,2,3\cdots \\ 0 & ,k=2i-1\end{cases} \end{aligned} \tag{4.27}$$

IV. Gamma 分布 $X\sim\Gamma(\alpha,\beta)(\alpha,\beta>0)$,$f(x)=\dfrac{x^{\alpha-1}\beta^\alpha e^{-\beta x}}{\Gamma(\alpha)}I_{(0,+\infty)}$

$$M(t)=\int_0^{+\infty}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}e^{tx}dx=\int_0^{+\infty}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{(t-\beta)x}dx=\beta^\alpha\int_0^{+\infty}\frac{1}{\Gamma(\alpha)}x^{\alpha-1}e^{(t-\beta)x}dx$$

$$t<\beta\text{ 时,令}v=(\beta-t)x\text{,原式}=\frac{\beta^\alpha}{\beta-t}\int_0^{+\infty}\frac{1}{\Gamma(\alpha)}\left(\frac{v}{\beta-t}\right)^{\alpha-1}e^{-v}dv$$

$$M(t)=\left(\frac{\beta}{\beta-t}\right)^\alpha\frac{1}{\Gamma(\alpha)}\int_0^{+\infty}v^{\alpha-1}e^{-v}dv=\left(\frac{\beta}{\beta-t}\right)^\alpha\frac{1}{\Gamma(\alpha)}\Gamma(\alpha)=\left(\frac{\beta}{\beta-t}\right)^\alpha \tag{4.28}$$

$$M'(t)=\beta^\alpha(\beta-t)^{-\alpha-1}\alpha\Rightarrow EX=M'(0)=\frac{\alpha}{\beta} \tag{4.29}$$

$$M''(t)=\beta^\alpha(\beta-t)^{-\alpha-2}\alpha(\alpha+1)\Rightarrow EX^2=M''(0)=\frac{\alpha(\alpha+1)}{\beta^2} \tag{4.30}$$

$$Var(X)=EX^2-(EX)^2=\frac{\alpha}{\beta^2} \tag{4.31}$$

V. Beta 分布 $X\sim\mathrm{Be}(\alpha,\beta)(\alpha,\beta>0)$, $f(x)=\dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1u^{\alpha-1}(1-u)^{\beta-1}du}I_{(0,1)}$

$$f(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}I_{(0,1)}=\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}I_{(0,1)}$$

$$\begin{aligned}M(t)&=\int_0^1\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}e^{tx}dx\\&=\int_0^1\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}\sum_{k=0}^{\infty}\frac{(tx)^k}{k!}dx=\sum_{k=0}^{\infty}\int_0^1\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}\frac{(tx)^k}{k!}dx\\&=\sum_{k=0}^{\infty}\int_0^1\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha+k-1}(1-x)^{\beta-1}\frac{t^k}{k!}dx=\sum_{k=0}^{\infty}\int_0^1\frac{x^{\alpha+k-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}\frac{t^k}{k!}dx\\&=\sum_{k=0}^{\infty}\frac{\int_0^1x^{\alpha+k-1}(1-x)^{\beta-1}dx}{\mathrm{B}(\alpha,\beta)}\frac{t^k}{k!}=\sum_{k=0}^{\infty}\frac{\mathrm{B}(\alpha+k,\beta)}{\mathrm{B}(\alpha,\beta)}\frac{t^k}{k!}\\&=\sum_{k=0}^{\infty}\frac{\frac{\Gamma(\alpha+k)\Gamma(\beta)}{\Gamma(\alpha+\beta+k)}}{\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}}\frac{t^k}{k!}=1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}\end{aligned}\tag{4.32}$$

$$\begin{aligned}E(X)&=\int_0^1\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}xdx=\int_0^1\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha}(1-x)^{\beta-1}dx\\&=\frac{\int_0^1x^{\alpha}(1-x)^{\beta-1}dx}{\mathrm{B}(\alpha,\beta)}=\frac{\mathrm{B}(\alpha+1,\beta)}{\mathrm{B}(\alpha,\beta)}\\&=\frac{\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}}{\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}}=\frac{\alpha}{\alpha+\beta}\end{aligned}\tag{4.33}$$

$$\begin{aligned}E\left(X^2\right)&=\int_0^1\frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}x^2dx=\frac{\int_0^1x^{\alpha+1}(1-x)^{\beta-1}dx}{\mathrm{B}(\alpha,\beta)}\\&=\frac{\mathrm{B}(\alpha+2,\beta)}{\mathrm{B}(\alpha,\beta)}=\frac{\frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)}}{\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}}=\frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)}\end{aligned}\tag{4.34}$$

$$Var(X)=E\left(X^2\right)-(EX)^2=\frac{\alpha(\alpha+1)}{(\alpha+\beta+1)(\alpha+\beta)}-\frac{\alpha^2}{(\alpha+\beta)^2}=\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}\tag{4.35}$$

VI. Student-t 分布

性质6.2.4 $X\sim\chi_n^2\Rightarrow E(X)=n,\quad Var(X)=\Sigma_{i=1}^n(E(X_i^4-(EX_i^2)^2)=2n.$

VII. F 分布

性质6.2.6

$$\mathrm{E}(t(n))=0,\quad n\geqslant 2$$
$$\mathrm{Var}(t(n))=\frac{n}{n-2},\quad n\geqslant 3$$

VIII. Rayleigh 分布

$$R=\sqrt{Y}=\sqrt{X_1^2+X_2^2},f_Y(y)=\frac{1}{2\sigma^2}\mathrm{e}^{-\frac{Y}{2\sigma^2}}\Rightarrow f_R(r)=\frac{r}{\sigma^2}\mathrm{e}^{-\frac{r^2}{2\sigma^2}}\quad r\geqslant 0\tag{4.36}$$

$$E\left[R^k\right]=\left(2\sigma^2\right)^{\frac{k}{2}}\Gamma\left(1+\frac{k}{2}\right)\tag{4.37}$$

$$E[R]=\sqrt{2}\sigma\frac{\Gamma(\dfrac{n+1}{2})}{\Gamma(\dfrac{n}{2})} \tag{4.38}$$

$$\begin{aligned}E(X)&=\int_0^{+\infty}x\times\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}dx=\int_0^{+\infty}\frac{x^2}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}dx=-\int_0^{+\infty}xd\left(e^{-\frac{x^2}{2\sigma^2}}\right)\\&=-xe^{-\frac{x^2}{2\sigma^2}}\Big|_0^{+\infty}+\int_0^{+\infty}e^{-\frac{x^2}{2\sigma^2}}dx=0+\sqrt{2\pi}\sigma\int_0^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx\\&=\sqrt{2\pi}\sigma\times\frac{1}{2}=\sigma\sqrt{\frac{\pi}{2}}\end{aligned} \tag{4.39}$$

$$\begin{aligned}\mathrm{Var}(X)&=\int_0^{+\infty}x^2\times\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}dx-E^2(X)=\int_0^{+\infty}\frac{x^3}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}dx-E^2(X)\\&=-\int_0^{+\infty}x^2d\left(e^{-\frac{x^2}{2\sigma^2}}\right)-E^2(X)=-x^2e^{-\frac{x^2}{2\sigma^2}}\Big|_0^{+\infty}+2\int_0^{+\infty}xe^{-\frac{x^2}{2\sigma^2}}dx-E^2(X)\\&=0-2\sigma^2e^{-\frac{x^2}{2\sigma^2}}\Big|_0^{+\infty}-E^2(X)=2\sigma^2-E^2(X)\\&=2\sigma^2-\left(\sigma\sqrt{\frac{\pi}{2}}\right)^2=\frac{4-\pi}{2}\sigma^2\end{aligned} \tag{4.40}$$

3 不同分布数字特征表

	分布名称	参数	概率密度	期望	方差	特征函数
离散型	0-1 分布	p	$\begin{pmatrix}0 & 1\\ 1-p & p\end{pmatrix}$	p	$p(1-p)$	e^{ict}
	二项分布	n,p	$\binom{n}{k}p^k(1-p)^{n-k}$ $k=0,1,2,\cdots,n$	np	$np(1-p)$	$1-p+pe^{it}$
	负二项分布	r,p	$\binom{k-1}{r-1}p^r(1-p)^{k-r}$, $k=r,r+1,r+2,\cdots,+\infty$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{pe^{it}}{1-(1-p)e^{it}}\right)^r$
	Poisson 分布	λ	$\frac{\lambda^k}{k!}e^{-\lambda}$, $k=0,1,2,3,\ldots,+\infty,\lambda>0$	λ	λ	$e^{\lambda(e^{it}-1)}$
	几何分布	p	$p(1-p)^{1-k},k\in\mathbb{N}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^{it}}{1-(1-p)e^{it}}$
	超几何分布	M,n,N	$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},M,N,n\in\mathbb{N}$	$\frac{nM}{N}$	$\frac{nM}{N}\frac{N-M}{N}\frac{N-n}{N-1}$	
连续型	均匀分布	a,b	$\frac{1}{b-a}I_{(a\leqslant x\leqslant b)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{it}(e^b-e^a)}{it(b-a)}$
	指数分布	λ	$\lambda e^{-\lambda x}I_{(x\geqslant 0)}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-it}$
	正态分布	μ,σ^2	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}I_{(-\infty\leqslant x\leqslant+\infty)}$	μ	σ	$e^{i\mu t-\frac{1}{2}\sigma^2 t}$
	χ^2 分布	n	$\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}I_{(x>0)}$	n	$2n$	$\frac{\lambda}{\lambda-it}$

4 常见分布可加性

$$\begin{aligned}
&\text{若} X \sim B(n, p), Y \sim B(m, p), \text{则} X+Y \sim B(n+m, p)\ (\text{注意} p \text{ 相同});\\
&\text{若} X \sim Poi\left(\lambda_1\right), Y \sim Poi\left(\lambda_2\right), \text{则} X+Y \sim Poi\left(\lambda_1+\lambda_2\right);\\
&\text{若} X \sim N\left(\mu_1, \sigma_1^2\right), Y \sim N\left(\mu_2, \sigma_2^2\right), \text{则} X+Y \sim N\left(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2\right);\\
&\text{若} X \sim \chi^2(n), Y \sim \chi^2(m), \text{则} X+Y \sim \chi^2(n+m).
\end{aligned}$$

第 2 节 协方差和相关系数

1. 协方差:

$$\operatorname{Cov}(X, Y)=E(X-EX)(Y-EY)=\mathrm{E}(X, Y)-(\mathrm{E}(X) \mathrm{E}(Y))$$

性质 4.2.1. 协方差的性质

- $\operatorname{Cov}(X, Y)=\operatorname{Cov}(Y, X), \operatorname{Cov}(X, X)=\operatorname{Var}(X)$
- $\operatorname{Cov}(X, Y)=EXY-EXEY$, 显然若 X, Y 相互独立, 则 $\operatorname{Cov}(X, Y) \quad=0$
- $\operatorname{Cov}\left(a_1 X_1+a_2 X_2, b_1 Y_1+b_2 Y_2\right)=\sum_{i=1}^2 \sum_{j=1}^2 a_i b_j \operatorname{Cov}\left(X_i, Y_j\right)$

协方差矩阵

$$\begin{aligned}
\Sigma &=\left(b_{ij}\right)=\left(\operatorname{cov}\left(\xi_i, \xi_j\right)\right) \geqslant 0 \\
&=\begin{pmatrix}
\operatorname{Var}\left(\xi_1\right) & \operatorname{cov}\left(\xi_1, \xi_2\right) & \ldots & \operatorname{cov}\left(\xi_1, \xi_n\right) \\
\operatorname{cov}\left(\xi_2, \xi_1\right) & \operatorname{Var}\left(\xi_2\right) & \ldots & \operatorname{cov}\left(\xi_2, \xi_n\right) \\
\vdots & \vdots & \ddots & \vdots \\
\operatorname{cov}\left(\xi_n, \xi_1\right) & \operatorname{cov}\left(\xi_n, \xi_2\right) & \ldots & \operatorname{Var}\left(\xi_n\right)
\end{pmatrix}
\end{aligned}$$

2. 相关系数

$$\rho_{X, Y}=\frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var} X} \cdot \sqrt{\operatorname{Var} Y}}=\operatorname{Cov}\left(X^*, Y^*\right)=\frac{\mathrm{E}(X, Y)-(\mathrm{E}(X) \mathrm{E}(Y))}{\sqrt{\operatorname{Var} X} \cdot \sqrt{\operatorname{Var} Y}}$$

性质 4.2.2. 相关系数的性质

- 不相关不一定独立 ($\rho_{X, Y}=0$ 推不出独立)
- 独立一定不相关 (独立推出 $\rho_{X, Y}=0$)
- $\left|\rho_{X, Y}\right| \leqslant 1$, 等号成立当且仅当 X, Y 之间存在严格的线性关系:
  $\rho_{X, Y}=1$, 则存在 $a>0, b \in \mathbb{R}$ 使得 $X=aY+b$ (正相关)
  $\rho_{X, Y}=-1$, 则存在 $a<0, b \in \mathbb{R}$ 使得 $X=aY+b$ (负相关)

$\left|\rho_{XY}\right|=1 \Leftrightarrow$ 存在常数 a, b, 使 $P\{Y=aX+b\}=1$

$$,\text{其中} a=\begin{cases}\sqrt{\dfrac{VarY}{VarX}}, & \rho_{XY}=1, \\ -\sqrt{\dfrac{VarY}{VarX}}, & \rho_{XY}=-1,\end{cases} \quad b=EY-aEX\,.$$

第 3 节 思考题

1. 试求下列随机变量 X 的期望 $\mathbb{E}X$ 和方差 $\mathrm{Var}(X)$. (1) 设 X 服从参数为 p 的几何分布, 其中 $0<p<1$, 即

$$\mathbb{P}(X=k)=p(1-p)^{k-1},\quad k=1,2,3,\cdots$$

(2) 设 X 服从参数为 r 和 p 的帕斯卡 (Pascal) 分布 (或负二项分布), 其中 $r\geqslant 1$ 为正整数, $0<p<1$, 即

$$\mathbb{P}(X=k)=\begin{pmatrix}k-1\\r-1\end{pmatrix}p^r(1-p)^{k-r},\quad k=r,r+1,r+2,\cdots$$

解: 根据公式4.12、4.14、4.6、4.8, 得到

$$(1)\ EX=\frac{1}{p};\mathrm{Var}(X)=\frac{1-p}{p^2}.\ (2)\ EX=\frac{r}{p};\mathrm{Var}(X)=\frac{r(1-p)}{p^2}.$$

2. 美国男子职业篮球联赛 (NBA) 季后赛的总决赛采用七战四胜制, 即哪支球队先获得四场比赛的胜利即可获得该年度的总冠军. 假设 A, B 两队势均力敌, 即每场各队获胜的概率都为 $p=0.5$, 以 X 表示一届总决赛的比赛场次, 试求 $\mathbb{E}X$. 若 A 队每场获胜的概率均为 $p=0.6$ 呢?

3. 设随机变量 X 的期望存在, 试证明: (1) 若 X 为非负整值随机变量, 则

$$\mathbb{E}X=\sum_{n=1}^{\infty}\mathbb{P}(X\geqslant n)=\sum_{n=0}^{\infty}\mathbb{P}(X>n)$$

(2) 若 X 为非负连续型随机变量, 且分布函数为 $F(x)$, 则

$$\mathbb{E}X=\int_0^{\infty}(1-F(x))\mathrm{d}x$$

(3) 若 X 为非负随机变量, 则 (2) 中的结论依然成立.

4. 平均要取多少个 $(0,1)$ 中的随机数才能让它们的和超过 1?

5. 设随机变量 $X\sim B(n,p)$, 且 $\mathbb{E}X=2.4,\mathrm{Var}(X)=1.68$, 则二项分布中的参数 n 和 p 各是多少?

6. (2017 年全国考研试题) 设随机变量 X 的分布函数 $F(x)=0.5\Phi(x)+0.5\Phi\left(\frac{x-4}{2}\right)$, 其中 $\Phi(x)$ 为标准正态分布函数, 则 $\mathbb{E}X=$

7. 设随机变量 X 的密度函数为

$$f(x)=c\mathrm{e}^{-x^2+x},\quad -\infty<x<\infty$$

试求常数 c, 及 $\mathbb{E}X$ 和 $\mathrm{Var}(X)$.

8. 设随机变量 X 的密度函数为

$$f(x)=\frac{x^n\mathrm{e}^{-x}}{n!},\quad x>0$$

试求 $\mathbb{E}X$ 和 $\mathrm{Var}(X)$.

9. 设 X 为一个连续型随机变量, 试对下列各种情形, 分别求 $\mathbb{E}X$ 和 $\mathrm{Var}(X)$. (1) 若 X 的密度函数为

$$f(x)=\frac{x}{\sigma^2}\exp\left\{-\frac{x^2}{2\sigma^2}\right\},\quad x>0$$

其中 $\sigma>0$ 为常数, 则称 X 服从瑞利 (Rayleigh) 分布. (2) 若 X 的密度函数为

$$f(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1},\quad 0<x<1$$

其中 $\alpha,\beta>0$ 为常数, $\Gamma(x)$ 为伽马 (Gamma) 函数, 则称 X 服从贝塔 (Beta) 分布. (3) 若 X 的密度函数为

$$f(x)=\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}\exp\left\{-\left(\frac{x}{\lambda}\right)^{k}\right\},\quad x>0$$

其中 $k,\lambda>0$ 为常数, 则称 X 服从威布尔 (Weibull) 分布.

10. 设随机变量 X 的密度函数为

$$f(x)=ax^2+bx+c,\quad 0<x<1$$

且已知 $\mathbb{E}X=0.5,\mathrm{Var}(X)=0.15$. 试求常数 a,b,c.

11. 设随机变量 X 服从参数为 λ 的指数分布, 试求 $\mathbb{P}(X>\sqrt{\mathrm{Var}(X)})$.

12. 有 $n(n\geqslant 2)$ 个人参加聚会, 每个人都将自己的帽子放入同一个箱子中, 经充分混合后每人随机取一顶, 试求取到自己原来帽子的人数的期望.

解: E(取到自己帽子人数)= $n\cdot\frac{1}{n}=1$

取帽子问题

设 B_j 表示第 j 个人取到自己帽子的概率。则至少有一人戴对这个事件可以描述为

$$\mathrm{B}_1 \cup \mathrm{B}_2 \cup \cdots \cup \mathrm{B}_\mathrm{n}$$

我们的任务是计算这个集合的概率。由于

$$\mathbb{P}\left(\mathrm{B}_\mathrm{k}\right) = \frac{(n-1)!}{n!} = \frac{1}{n}, \quad \mathbb{P}\left(\mathrm{B}_\mathrm{k} \cap \mathrm{B}_\mathrm{j}\right) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}, \quad \cdots$$

由容斥公式

$$\begin{aligned}\mathbb{P}\left(\mathrm{B}_1 \cup \mathrm{B}_2 \cup \cdots \cup \mathrm{B}_\mathrm{n}\right) =& \sum_{k=1}^{n} \mathbb{P}\left(\mathrm{B}_\mathrm{k}\right) - \sum_{j<k} \mathbb{P}\left(\mathrm{B}_\mathrm{j} \cap \mathrm{B}_\mathrm{k}\right) + \sum_{i<j<k} \mathbb{P}\left(\mathrm{B}_\mathrm{i} \cap \mathrm{B}_\mathrm{j} \cap \mathrm{B}_\mathrm{k}\right) \\ & - \cdots + (-1)^{n+1} \mathbb{P}\left(\mathrm{B}_1 \cap \mathrm{B}_1 \cap \cdots \cap \mathrm{B}_\mathrm{n}\right) \\ =& n \times \frac{1}{n} - \binom{n}{2} \frac{(n-2)!}{n!} + \binom{n}{3} \frac{(n-3)!}{n!} - \cdots + (-1)^{n+1} \frac{1}{n!} \\ =& 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n+1} \frac{1}{n!}\end{aligned}$$

即有

$$\mathbb{P}\left(\mathrm{B}_1 \cup \mathrm{B}_2 \cup \cdots \cup \mathrm{B}_\mathrm{n}\right) = \sum_{k=1}^{n} (-1)^{k+1} \frac{1}{k!}$$

使用 $\exp(x)$ 的幂级数展开, 得到

$$\lim_{n\to\infty} \mathbb{P}\left(\mathrm{B}_1 \cup \mathrm{B}_2 \mid \cup \cdots \cup \mathrm{B}_\mathrm{n}\right) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k!} = 1 - \exp(-1) \approx 0.63$$

同时得到没有一个帽子戴对的概率满足

$$\lim_{n\to\infty} \mathbb{P}\left(\overline{\mathrm{B}_1} \cap \overline{\mathrm{B}_2} \cap \cdots \cap \overline{\mathrm{B}_\mathrm{n}}\right) = \sum_{k=0}^{n} (-1)^{k} \frac{1}{k!} \to \exp(-1)$$

恰有 k 人戴对的概率

$$\begin{aligned}P_k &= \sum_{1 \leqslant n_1, \cdots, n_n \leqslant n} \mathbb{P}\left(\mathrm{B}_{n_1} \mathrm{B}_{n_2} \cdots \mathrm{B}_{n_k}\right) \left(1 - \mathbb{P}\left(\mathrm{B}_{n_{k+1}} \cup \cdots \mathrm{B}_{n_n}\right)\right) \\ &= \binom{n}{k} \frac{1}{n(n-1)\cdots(n-k+1)} \left(1 - \sum_{m=1}^{n-k} (-1)^{m+1} \frac{1}{m!}\right) \\ &= \frac{1}{k!} \sum_{m=0}^{n-k} (-1)^m \frac{1}{m!}\end{aligned}$$

当 $n \to \infty$ 时, 有

$$P_k \to \frac{\exp(-1)}{k!}$$

13. 设 $\mathbb{P}(X=-1) = \mathbb{P}(X=1) = 1/2$, 试求 X 的中位数. 0

14. 设 X 为标准正态随机变量, 试求 X 的 α 分位数 $(0<\alpha<1)$.　　$Z_\alpha=\Phi^{-1}(1-\alpha)$

15. 设 X 为随机变量, 则使得 $\mathbb{E}(X-a)^2$ 达最小的常数 $a=\mathbb{E}X$

    解：$E(X-a)^2=E(X-EX+EX-a)^2=E(X-EX)^2+(EX-a)^2\geqslant E(X-EX)^2$

16. 设 X 为随机变量, 则使得 $\mathbb{E}|X-a|$ 达最小的常数 $a=\mathbb{E}X$

    解：采用距离和最小说明。

17. 假设在 20000 件产品中有 1000 件次品, 若从中不放回地任取出 150 件进行检验, 试求查得次品数的期望.

18. 将 n 个球依次放入 n 个盒子中, 假设每个球放入每个盒子中是等可能的, 试求放完后空盒子个数的期望, 以及当 $n\to\infty$ 时空盒子的平均比例.

19. 某零食厂商设计了一种营销策略, 即在产品中放入一套有趣的卡片. 假设这套卡片由 $n=12$ 张不同的卡通人物头像组成, 且在每袋零食中随机放入其中一张. 某人想聚齐这套卡片, 设他一共需要买 X_n 袋该零食. 试求: (1) $\mathbb{E}[X_n]$ (2) $\lim_{n\to\infty}\frac{\mathbb{E}[X_n]}{n\ln n}$.

彩券问题

某商场发行 n 种购物券，每一次在该商场购物, 即可获得一张种类随机选取的购物券。该商场规定，如果能够收集齐所有的购物券, 那么就可以得到该商场的大奖。试问需要平均购物多少次才能获得大奖?
解答: 设收集所有购物券所需的购物次数为 X, 则可以写为

$$X=C_1+C_2+\cdots+C_n$$

其中 $C_k,k>1$ 为在已经收集到 $k-1$ 种购物券的条件下, 收集到第 k 种购物券所需的购物次数, $C_1=1$ 。由于手中已有 $k-1$ 种, 收集到新的购物券的可能性有 $n-k+1$ 种, 所有 C_k 是服从几何分布, 每次要么收集到新的 (概率为 $(n-k+1)/n$), 要不还是原来的 $k-1$ 种, 等待收集新的所需试验次数是几何分布, 即

$$C_k\sim \text{Geo}\left(\frac{n-k+1}{n}\right)$$

$$EC_k=\frac{n}{n-k+1}$$

因此

$$\begin{aligned}EX&=E\left[C_1+C_2+\cdots+C_n\right]\\&=1+EC_2+\cdots+EC_n\\&=1+\frac{n}{n-1}+\cdots+\frac{n}{1}\\&=n\left(\frac{1}{n}+\frac{1}{n-1}+\cdots+1\right)\\&=n(\ln n+\gamma+o(1))\end{aligned}$$

20. 现有 $n(n \geqslant 1)$ 个袋子, 各装有 a 只白球和 b 只黑球. 先从第一个袋子中随机摸出一球, 然后把它放入第二个袋子中, 混合后再从第二个袋子随机摸出一球放入第三个袋子中, 照此做法依次下去, 最后从第 n 个袋子中随机摸出一球. 将这 n 次摸球中所摸出的白球总个数记为 W_n, 试求 $\mathbb{E}[W_n]$.

21. (2010 年全国考研试题) 设随机变量 X 的概率分布为 $\mathbb{P}(X=k)=\frac{C}{k!}, k=0,1,2,\cdots$, 则 $\mathbb{E}[X^2]=$

22. 设随机变量 X 服从参数为 $\lambda>0$ 的 Poisson 分布, 已知 $\mathbb{E}[(X-1)(X-2)]=5$, 试求 λ 的值.

23. 设随机变量 X 只能取有限个正值 $x_1, x_2, \cdots, x_k(k \geqslant 2)$, 证明:

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}[X^{n+1}]}{\mathbb{E}[X^n]}=\max_{1 \leqslant i \leqslant k} x_i$$

24. 设 X 为一随机变量, 若对任意实数 t, 期望 $\mathbb{E}\left[\mathrm{e}^{tX}\right]$ 存在, 则称它为 X 的矩母函数, 且记为 $\psi(t)$. 试对下列常见的分布求其矩母函数: (1) 二项分布 $B(n, p)$; (2) 参数为 λ 的 Poisson 分布; (3) 参数为 λ 的指数分布; (4) 正态分布 $N\left(\mu, \sigma^2\right)$.

解: 矩母函数 $M(t)=\mathbb{E}\left[\mathrm{e}^{tX}\right]$，特征函数 $\varphi(t)=\mathbb{E}\left[\mathrm{e}^{itX}\right]$ 见4.1.3不同分布数字特征表

(1) 若 X 服从伯努利分布 $B(1, p)$, 则 X 满足 $P(x=1)=p, P(x=0)=1-p=q$

$$M(t)=pe^t+1-p$$

若 X 服从二项分布 $B(n, p)$, 因为服从二项分布的变量可以看作 n 个独立相同的服从伯努利分布的变量之和, 因此 $M(t)=\left(pe^t+1-p\right)^n, \varphi(t)=\left(pe^{it}+1-p\right)^n$

(2)

$$M(t)=\sum_{k=0}^{+\infty} \frac{e^{-\lambda} \lambda^k}{k!} e^{tk}=e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\left(\lambda e^t\right)^k}{k!}=e^{\lambda\left(e^t-1\right)}$$

(3)

$$M(t)=\int_0^{+\infty} \lambda e^{-\lambda x} e^{tx} dx=\frac{\lambda}{\lambda-t}$$

(4)

$$\begin{aligned}
M(t) & =\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2 \pi} \sigma} e^{-\frac{(x-\mu)^2}{2 \sigma^2}} e^{tx} dx, \omega=\frac{x-\mu}{\sigma} \\
& =\frac{1}{\sqrt{2 \pi}} \int_{-\infty}^{+\infty} e^{-\frac{\omega^2}{2}+t(\omega \sigma+\mu)} d\omega \\
& =e^{\mu t+\frac{t^2 \sigma^2}{2}} \frac{1}{\sqrt{2 \pi}} \int_{-\infty}^{+\infty} e^{-\frac{(\omega-t \sigma)^2}{2}} d\omega=\exp \left(\mu t+\frac{t^2 \sigma^2}{2}\right)
\end{aligned}$$

25. 设随机变量 X 服从参数为 λ 的指数分布. (1) 对任意常数 $c>0$, 证明 cX 服从参数为 λ / c 的指数分布. (2) 对任意正整数 $n \geqslant 1$, 计算 $\mathbb{E}[X^n]$.

解: 指数分布 $X \sim \mathbf{Exp}(\lambda), f(x)=\lambda e^{-\lambda x} I_{(x \geqslant 0)}, f(x)=\lambda e^{-\lambda Y / c} I_{(x \geqslant 0)}, Y \sim \mathbf{Exp}(\lambda / c)$

$$E\left(X^n\right)=\int_0^{\infty} x^n f(x) dx=\int_0^{\infty} x^n \cdot \lambda e^{-\lambda x} dx=\frac{1}{\lambda^n} \int_0^{\infty}(\lambda x)^n e^{-\lambda x} d\lambda x=\frac{\Gamma(n+1)}{\lambda^n}=\frac{n!}{\lambda^n}$$

26. 设随机变量 X 的密度函数为 $f(x)=2(x-1), 1<x<2$, 试求随机变量 $Y=\mathrm{e}^X$ 和 $Z=\frac{1}{X}$ 的数学期望.

解:

$$\begin{aligned}
P\{Y\leqslant y\} &= P\left\{e^x\leqslant y\right\}=P\{x\leqslant \ln y\} \\
&= \int_1^{\ln y} f(x)dx=\left(x^2-2x\right)\Big|_1^2 \\
&= (\ln y)^2-2\ln y+1 \Rightarrow f_Y(y)=F_Y'(y)=\frac{2\ln y}{y}-\frac{2}{y}
\end{aligned}$$

$$E(Y)=\int_e^{e^2} y\left(\frac{2\ln y}{y}-\frac{2}{y}\right)dy=2y(\ln y-2)\Big|_e^{e^2}=2e$$

$$\begin{aligned}
P\{Z\leqslant z\} &= P\left\{\frac{1}{X}\leqslant z\right\}=P\left\{X\geqslant\frac{1}{2}\right\} \\
&= \int_{\frac{1}{2}}^{2} f(x)dx=-\frac{1}{z^2}+\frac{2}{z}+2\Rightarrow f_z(z)=\frac{2}{z^3}-\frac{2}{z^2}
\end{aligned}$$

$$E(Z)=\int_{\frac{1}{2}}^{1} z\left(\frac{2}{z^3}-\frac{2}{z^2}\right)dz=\left(-2\ln z-\frac{2}{z}\right)\Big|_{\frac{1}{2}}^{1}=2-2\ln 2$$

27. 设随机变量 X 服从参数为 μ 和 σ^2 的对数正态分布, 即 $\ln X\sim N\left(\mu,\sigma^2\right)$, 试求 X 的密度函数 $p(x)$, 期望 $\mathbb{E}X$ 和方差 $\operatorname{Var}(X)$.

解:

$$\begin{aligned}
&Y=\ln X\sim N\left(\mu,\sigma^2\right), f(y)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mu)^2}{2\sigma^2}} \\
&P\{X\leqslant x\}=P\{Y\leqslant \ln x\}=\int_{-\infty}^{\ln x}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mu)^2}{2\sigma^2}}dy=\Phi(\ln x)
\end{aligned}$$

$$\Rightarrow p(x)=\Phi'(\ln x)=\frac{1}{\sqrt{2\pi}\sigma x}\exp\left\{-\frac{(\ln x-\mu)^2}{2\sigma^2}\right\},\quad x>0$$

$$E(X)=\int_0^{+\infty}xf(x)dx=\int_0^{+\infty}\frac{x}{x\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right\}dx$$
$$(\text{令}\frac{\ln x-\mu}{\sigma}=t,\text{则}x=e^{\sigma t+\mu},t\in(-\infty,+\infty),\text{则}dx=\sigma e^{\sigma t+\mu}dt\,)$$
$$=\frac{e^{\mu}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}+\sigma t}dt,\ (\text{将}-\frac{t^2}{2}+\sigma t\text{ 配方})$$
$$=\frac{e^{\mu}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{1}{2}(t-\sigma)^2+\frac{\sigma^2}{2}}dt=\frac{e^{\mu+\frac{\sigma^2}{2}}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{1}{2}(t-\sigma)^2}dt,\left(\text{令}\frac{t-\sigma}{\sqrt{2}}=m,\text{ 则}dt=\sqrt{2}dm\right)$$
$$=\frac{e^{\mu+\frac{\sigma^2}{2}}}{\sqrt{2\pi}}\sqrt{2}\int_{-\infty}^{+\infty}e^{-m^2}dm,\left(\int_{-\infty}^{+\infty}e^{-m^2}dm=\sqrt{\pi}\right)=e^{\mu+\frac{\sigma^2}{2}}$$
$$E\left(X^2\right)=\int_0^{+\infty}x^2f(x)dx=\int_0^{+\infty}\frac{x}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right\}dx$$
$$(\text{同理, 令}\frac{\ln x-\mu}{\sigma}=t\,)$$
$$=\frac{e^{2\mu}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}+2\sigma t}dt,\quad\left(\text{将}-\frac{t^2}{2}+2\sigma t\text{ 配方}\right)$$
$$=\frac{e^{2\mu}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{1}{2}(t-2\sigma)^2+2\sigma^2}dt$$
$$=\frac{e^{2\mu+2\sigma^2}}{\sqrt{2\pi}}\sqrt{2}\int_{-\infty}^{+\infty}e^{-\left(\frac{t-2\sigma}{\sqrt{2}}\right)^2}d\left(\frac{t-2\sigma}{\sqrt{2}}\right)=e^{2\mu+2\sigma^2}$$
$$\text{所以, }Var(X)=E\left(X^2\right)-(EX)^2=e^{2\mu+2\sigma^2}-\left(e^{\mu+\frac{\sigma^2}{2}}\right)^2=e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$$

28. 设随机变量 X 服从区间 $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ 上的均匀分布. 试求期望 $\mathbb{E}[\sin X],\mathbb{E}[\cos X]$ 及 $\mathbb{E}[X\cos X]$.

解:

$$f(x)=\frac{1}{\pi}I_{\left(-\frac{\pi}{2}\leqslant x\leqslant\frac{\pi}{2}\right)}$$
$$E(\sin X)=\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin x dx=0$$
$$E(\cos X)=\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos x dx=\frac{2}{\pi}$$
$$E(X\cos X)=\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}x\cos x dx=\frac{1}{\pi}(x\sin x+\cos x)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}$$

29. 某鲜花店为了迎接情人节的销售高峰期, 决定购进一批玫瑰花. 已知在情人节期间, 每出售一束玫瑰花会获利 a 元, 若未售出, 最终则会灭损 b 元, 该店预测玫瑰花情人节期间的销售量 (单位: 束) 会服从 $[m,n]$ 上的均匀分布. 试问为了平均获利最大, 该店应购进多少束玫瑰花?

解: 设 y 为每日购进玫瑰花，x 为每日能卖出的玫瑰花

$$L(y)=\begin{cases}ay, y\leqslant x\\ ax-b(y-x)\quad y>x\end{cases},f(x)=\frac{1}{n-m}I_{(m\leqslant x\leqslant n)}$$
$$E(L(y))=\int_m^n L(y)f(x)dx=\int_m^y\frac{(a+b)x-by}{n-m}dx+\int_y^n\frac{ay}{n-m}dx$$
$$\frac{d}{dy}EL(y)=\frac{mb+an}{n-m}-\frac{b}{n-m}y+\frac{-a}{n-m}y=0\Rightarrow y=\frac{an+bm}{a+b}\text{且 y 为最大值}$$

30. 设 X 为一随机变量, 它的符号函数定义为

$$\operatorname{sgn}(X)=\begin{cases}1, & X>0\\ 0, & X=0\\ -1, & X<0\end{cases}$$

(1) 若 $X\sim U(-2,1)$, 试求 $\operatorname{Var}(\operatorname{sgn}(X))$. (2) 若 X 服从标准正态分布, 试求 $\mathbb{E}[\operatorname{sgn}(X)\cdot X]$.

31. 设随机变量 X 的密度函数为

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

试求 $\mathbb{E}[\min\{|X|,1\}]$.

解:

$$\begin{aligned}E(\min\{|x|,1\})&=\int_{-\infty}^{+\infty}\min\{|x|,1\}\frac{1}{\pi\left(1+x^2\right)}dx\\&=2\int_{-\infty}^{-1}\frac{dx}{\pi\left(1+x^2\right)}+\int_{-1}^{1}\frac{xdx}{\pi\left(1+x^2\right)}\\&=\frac{2}{\pi}\arctan x\Big|_{-\infty}^{-1}+\ln\left(1+x^2\right)\Big|_{-1}^{1}=\frac{1}{2}\end{aligned}$$

32. (2014 年全国考研试题) 设随机变量 X 的分布为 $\mathbb{P}(X=1)=\mathbb{P}(X=2)=\frac{1}{2}$, 在给定 $X=i$ 的条件下, 随机变量 Y 服从均匀分布 $U(0,i)(i=1,2)$. (1) 求 Y 的分布函数; (2) 求期望 $\mathbb{E}Y$.

33. 设二维随机变量 (X,Y) 服从二维正态分布 $N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$, 其中 $\mu_1=\mu_2=1,\sigma_1^2=\sigma_2^2=0.5,\rho=0.5$, 记

$$Z=|X-Y|,\quad U=\max(X,Y),\quad V=\min(X,Y)$$

(1) 求 Z 的密度函数与期望 $\mathbb{E}[Z]$; (2) 分别求数学期望 $\mathbb{E}[U]$ 和 $\mathbb{E}[V]$.

34. 假设有 $n(n\geqslant 3)$ 个不同的盒子与 m 个相同的小球, 每个小球独立地以概率 p_k 落入第 k 个盒子 $(k=1,2,\ldots,n)$. 分别以 $X_1,X_2,\ldots,X_n$ 表示落入各个盒子的球数. 试求 (1) $\mathbb{E}\left[X_2\mid X_1=k\right]$ 和 $\operatorname{Var}\left(X_2\mid X_1=k\right)$. (2) $\mathbb{E}\left[X_1+X_2\right]$ 和 $\operatorname{Var}\left(X_1+\cdots+X_k\right),k=1,\ldots,n$.

35. 设某投资者希望投资两个金融产品, 设两个金融产品在一年后的价值 (X,Y) 服从二元正态分布 $N\left(\mu,2\mu,\sigma^2,3\sigma^2,0.5\right)$, 其中负值表示损失, 正值表示收益. 设其希望找到最优的投资组合, 即找 $\omega\in[0,1]$ 使得 $\omega X+(1-\omega)Y$ 的夏普比率 Sharpe $\text{Ratio}_{(\omega X+(1-\omega)Y)}=\frac{\mathbb{E}[\omega X+(1-\omega)Y]}{\sqrt{\operatorname{Var}(\omega X+(1-\omega)Y)}}$ 达到最大.

二维正态分布的性质

$$(X,Y)\sim N\left(a,b,\sigma_1^2,\sigma_2^2,\rho\right)$$

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\exp\left[-\frac{1}{2\left(1-\rho^2\right)}\left(\frac{(x-a)^2}{\sigma_1^2}-\frac{2\rho\,(x-a)\,(y-b)}{\sigma_1\sigma_2}+\frac{(y-b)^2}{\sigma_2^2}\right)\right]$$

(1) $X\sim N\left(a,\sigma_1^2\right),\quad Y\sim N\left(b,\sigma_2^2\right)$

(2) $X\mid_{Y=y}\sim N\left(a+\rho\frac{\sigma_1}{\sigma_2}(y-b),\sigma_1^2\left(1-\rho^2\right)\right)$

$$p(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{1}{\sqrt{2\pi\left(1-\rho^2\right)}\sigma_1}\exp\left[-\frac{\left(x-a-\rho\frac{\sigma_1}{\sigma_2}(y-b)\right)^2}{2\sigma_1^2\left(1-\rho^2\right)}\right]$$

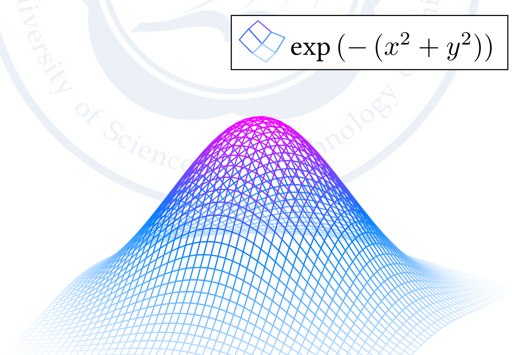
(3) $\forall c,d\in\mathbb{R},\quad cX+dY\sim\left(ca+db,c^2\sigma_1^2+d^2\sigma_2^2+2cd\rho\sigma_1\sigma_2\right)$

(4) $\rho=0$，**两变量是不相关的，也是独立的**

(5) **协方差的性质：**

$$\begin{aligned}\operatorname{Cov}(X,Y)&=E(X-EX)(Y-EY)=\rho\sigma_1\sigma_2\\&=EXY-EX\cdot EY=EXY-ab\\\Rightarrow EXY&=ab+\rho\sigma_1\sigma_2\end{aligned}$$

二维正态分布概率密度函数



解：

$$E(\omega x+(1-\omega)Y)=\omega EX+(1-\omega)EY$$

$$\operatorname{Var}(\omega X+(1-\omega)Y)=\omega^2\operatorname{Var}(X)+(1-\omega)^2Var(Y)+2\omega(1-\omega)\operatorname{Cov}(X,Y)$$

$$\begin{aligned}\frac{\mathbb{E}[\omega X+(1-\omega)Y]}{\sqrt{\operatorname{Var}(\omega X+(1-\omega)Y)}}&=\frac{\omega\mu+2(1-\omega)\mu}{\sqrt{\omega^2\sigma^2+(1-\omega)^2 3\sigma^2+\sqrt{3}\omega(1-\omega)\sigma^2}}\\&=\frac{(2-\omega)\mu}{\sigma\sqrt{(4-\sqrt{3})\omega^2+(\sqrt{3}-6)\omega+3}},\text{设}f(\omega)=\frac{(2-\omega)^2}{(4-\sqrt{3})\omega^2+(\sqrt{3}-6)\omega+3}\end{aligned}$$

$$f'(\omega)=\frac{(-2+2\omega)\left((4-\sqrt{3})\omega^2+(\sqrt{3}-6)\omega+3\right)-[(8-2\sqrt{3})\omega+(\sqrt{3}-6)](2-\omega)^2}{\left((4-\sqrt{3})\omega^2+(\sqrt{3}-6)\omega+3\right)^2}=0$$

$$\Rightarrow\omega=\frac{-8+2\sqrt{3}+\sqrt{34-15\sqrt{3}}}{\sqrt{3}-4}\approx 0.75137$$

36. 设某两个风险 (X, Y) 服从二元正态分布 $N\left(\mu, 2\mu, \sigma^2, 2\sigma^2, \sqrt{2}/4\right)$, 某投资者购买了一个基于这两只风险和的金融衍生品 (欧式看涨期权), 即到期收益为

$$(X+Y-3\mu)_+ = \max\{X+Y-3\mu, 0\}$$

(1) 求到期收益的均值 $\mathbb{E}\left[(X+Y-3\mu)_+\right]$; (2) 求到期收益的方差 $\operatorname{Var}\left((X+Y-3\mu)_+\right)$.

解: 本题求解的关键在于合理的换元积分，用 (u, v) 替换 (x, y), 利用二维正态分布概率密度函数公式:

$$u = x-\mu, v = y-2\mu, \quad \to \operatorname{Cov}(x, y) = \operatorname{Cov}(u, v) \Rightarrow \rho_{xy} = \rho_{uy},$$

$$\therefore (u, v) \sim \left(0, 0, \sigma^2, 2\sigma^2, \frac{\sqrt{2}}{4}\right), z = u+v.$$

$$f(x, y) = \frac{1}{2\pi\sqrt{2}\sigma^2\sqrt{\frac{7}{8}}} \exp\left\{-\frac{4}{7}\left[\left(\frac{x-\mu}{\sigma}\right)^2 - \frac{\sqrt{2}}{2}\frac{x-\mu}{\sigma}\frac{y-2\mu}{\sqrt{2}\sigma} + \frac{(y-2\mu)^2}{2\sigma^2}\right]\right\}$$

$$f(u, v) = \frac{1}{2\pi\sqrt{2}\sigma^2\sqrt{\frac{7}{8}}} \exp\left[-\frac{4}{7}\cdot\left(\frac{u^2}{\sigma^2} - \frac{\sqrt{2}}{2\sigma}\frac{uv}{\sqrt{2}\sigma} + \frac{v^2}{2\sigma^2}\right)\right]$$

$$f_z(z) = \int_{-\infty}^{+\infty} f(u, z-u)du = \frac{1}{2\sigma\sqrt{2\pi}} \exp\left(-\frac{z^2}{8\sigma^2}\right)$$

$$E[(X+Y-3\mu)_+] = \int_0^{+\infty} \frac{z}{2\sigma\sqrt{2\pi}} \exp\left(-\frac{z^2}{8\sigma^2}\right) dz$$

$$= -\frac{2\sigma}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{8\sigma^2}\right)\bigg|_0^1 = \frac{\sqrt{2}\sigma}{\sqrt{\pi}}$$

$$\operatorname{Var}[(x+y-3\mu)+] = Ez^2 - (Ez)^2$$

$$Ez^2 = \int_0^{+\infty} \frac{z^2}{2\sigma\sqrt{2\pi}} \exp\left(-\frac{z^2}{8\sigma^2}\right) dz$$

$$= \frac{8\sigma^2}{\sqrt{\pi}} \int_0^{+\infty} \left(\frac{z}{2\sqrt{2}\sigma}\right)^2 \exp\left(-\left(\frac{z}{2\sqrt{2}\sigma}\right)^2\right) d\left(\frac{z}{2\sqrt{2}\sigma}\right)$$

$$= \frac{8\sigma^2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{4} = 2\sigma^2,$$

$$\therefore \operatorname{Var}[(X+Y-3\mu)_+] = 2\sigma^2 - \frac{2\sigma^2}{\pi} = 2\sigma^2\left(1-\frac{1}{\pi}\right)$$

37. 假设随机变量 X 有分布律 $\mathbb{P}(X=0) = \mathbb{P}(X=1) = \mathbb{P}(X=2) = 1/3$, 随机变量 Y 满足在 $X=k$ 的条件下服从均值为为 k, 方差为 1 的正态分布, 即 $[Y \mid X=k] \sim N(k, 1)$ (1) 求随机变量 Y 的概率密度函数和期望; (2) 求随机变量 $X+Y$ 的分布函数; (3) 求随机变量 X 和 Y 协方差.

38. 设有 n 个人, 每个人都将自己的帽子放入同一个箱子中, 经充分混合后, 每人再随机从中选取一顶. 求选中自己帽子的人数的期望.

39. 设一日光灯的寿命服从均值为 $1/\lambda$ 的指数分布, 某间办公室有 n 个日光灯, 不同的日光灯的寿命相互独立. 求该办公室有日光灯亮的时间的期望与方差.

40. 设一日光灯的寿命服从均值为 5 的指数分布, 某间办公室有 2 个日光灯与 2 个备用日光灯, 设不同的日光灯的寿命相互独立. 如果有灯坏了则换上备用的日光灯, 求该办公室有日光灯正常工作的时间的期望.

41. 设 X_1, X_2 是相互独立的随机变量, 服从参数为 2 的指数分布, 求 $\mathbb{E}(\min\{X_1, X_2\})$ 和 $\mathbb{E}(\max\{X_1, X_2\})$.

42. 设 X_1, X_2 是相互独立的随机变量, 服从参数为 2 的 Poisson 分布, 则 $\mathbb{E}(\min\{X_1, X_2\}) =$

43. 设 $X_1, X_2, \ldots, X_n$ 为一列随机变量, 下列说法错误的是 () (A) 如果 $X_1, \ldots, X_n$ 相互独立, $\mathbb{E}[X_1 \cdots X_n] = \mathbb{E}[X_1] \cdots \mathbb{E}[X_n]$ (B) $\mathbb{E}[X_1 + \cdots + X_n] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n]$ (C) 如果 X_1 与 X_2 独立, $\mathbb{E}[|X_1 X_2| + X_3] = \mathbb{E}[|X_1|]\mathbb{E}[|X_2|] + \mathbb{E}[X_3]$ (D) 如果 X_1 与 X_2 独立, $X_2 > 0, \mathbb{E}[X_1/X_2] = \mathbb{E}[X_1]/\mathbb{E}[X_2]$.

44. 设 X_1, X_2, X_3 服从球面 $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ 上面的均匀分布, 求 $X_1 + X_2 + X_3$ 的数学期望 $\mathbb{E}[X_1 + X_2 + X_3]$ 和方差 $\operatorname{Var}(X_1 + X_2 + X_3)$.

45. 设某人连续独立地投郑一枚均匀的骰子（即投出 $1, 2, 3, 4, 5, 6$ 的概率均为 $1/6$), 直到点数大于等于 10 时停止. (a) 求投郑次数的期望; (b) 求停止时点数和的期望.

46. 设某金融公司有 20 个基金经理相互独立地做投资理财. 设每个基金经理每个月的收益情况服从均值为 50 万，方差为 10 (万 ^{2}) 的正态分布. 求该金融公司的每个月收益的期望和方差.

47. 设某投资者分别购买了两只不同板块的股票, 分别花了 2 万元. 按照经验, 两只股票一周的收益率 (收益率为 r 指的是投资 1 单位的资金后, 到期日收到 $1 + r$ 单位的收益.) 均服从 $[-0.1, 0.2]$ 上的均匀分布, 且相关系数为 0.2. 求其一周后投资的股票价值总和的均值和方差.

48. 设一个离散随机变量 X 的概率分布函数为 $\mathbb{P}(X = i) = p_i \geqslant 0, i = 1, \ldots, n, \sum_{i=1}^n p_i = 1$. 其熵 $H(X)$ 定义为

$$H(X) = \sum_{i=1}^{n} -p_i \log p_i$$

其中 $0 \log 0 = 0$. 设 U 为取值 $\{1, \ldots, n\}$ 上的均匀分布, 即 $\mathbb{P}(U = i) = 1/n, i = 1, \ldots, n$ (1) 求 U 的熵 $H(U)$. (2) 证明 $H(X) \leqslant H(U)$. 提示: 利用 Jensen 不等式: 设 X 为一个随机变量, ϕ 为一个凸函数（二阶导数非负）, 则

$$\phi(\mathbb{E}[X]) \leqslant \mathbb{E}[\phi(X)]$$

49. 设随机变量 X 与 Y 相互独立, 且 $X \sim N(3, 2), Y \sim U(1, 2), Z = 2X$, 令 $W = X - Y + Z - 1$, 则 $\operatorname{Var}(W) =$

50. 设随机变量 X 服从 Poisson 分布, 参数为 λ, 若 $\mathbb{E}(X-2)(X-3) = 2$, 则 $\lambda =$

51. 设 X_1, X_2, X_3 是相互独立的随机变量, 服从参数为 2 的 Poisson 分布, 令 $Y = (X_1 + X_2 + X_3)/3$, 则 $\operatorname{Var}(Y) =$

52. 设 X_1 和 X_2 是独立的指数随机变量, 均值分别为 1 和 2. 定义

$$Y = \min\{X_1, X_2\} \text{ 和} Z = \max\{X_1, X_2\}.$$

求 (1)$\mathbb{E}[Y]$ 和 $\mathbb{E}[Z]$; (2) $\operatorname{Var}(Y)$ 和 $\operatorname{Var}(Z)$.

53. 一袋中有张卡片, 分别记为 $1,2,\ldots,n$, 从中有放回地抽取出 k 张, $k \geqslant 1$, 以 X 表示所得号码之和, 则 $\mathbb{E}[X]=$; $\operatorname{Var}(X)=$

54. 设 $X \sim B(3,p)$ (参数为 $(3,p)$ 的二项分布) 和 $Y \sim U(0,2)$ (区间 $(0,2)$ 上的均匀分布), 已知 X 与 Y 相互独立, 则 $\operatorname{Var}(2X+3Y)=$

55. 在长为 1 的线段上任取两点 A 和 B, 试求线段 AB 长度的数学期望和方差.

56. 设 X 与 Y 为两个随机变量. 若两者的相关系数 $\rho_{X,Y}=1$, 则有 ( ) A. 存在 $a \neq 0$ 使得 $Y=aX+b$ B. 存在 $a>0$ 使得 $Y=aX+b$ C. 存在 $a<0$ 使得 $Y=aX+b$ D. 以上均不对.

57. 考虑甲乙两个小朋友随机分割一长度为 a 的树枝, 设 X 与 Y 为分别代表甲乙两个小朋友所得部分树枝的长度. 则两者的相关系数 $\rho_{X,Y}$ 为多少?

58. 设 X 与 Y 为两个随机变量. 已知 $\operatorname{Var}(X)=9, \operatorname{Var}(Y)=16, \operatorname{Cov}(X,Y)=-3$, 则 X 与 Y 的相关系数 $\rho_{X,Y}=$

59. 设 X 与 Y 是两个方差存在有限的随机变量, 如果 $Y=aX+b,(a\neq 0)$, 那么 X 与 Y 的相关系数 $\rho_{X,Y}$ 为 ( ) A. 1 B. -1 C. $\frac{a}{|a|}$ D. 以上均不对.

60. 设二元随机变量 (X,Y) 服从 $\{(x,y):|x|+|y|\leqslant 1\}$ 上的均匀分布, 求 $\operatorname{Cov}(X,Y)$.

61. 已知二元随机变量 (X,Y) 有概率分布如下:

X	-1	0	1
-1	0.1	0.2	0.2
0	0.05	0.1	0.15
1	0.05	0.05	0.1

求 $\operatorname{Cov}(X,Y)$ 和 $\operatorname{Cov}\left(X^2,Y^2\right)$.

62. 郑两颗均匀骰子, 以 X 表示第一颗骰子郑出的点数, Y 表示两颗骰子所郑出的点数中的最大值. (1) 求 X,Y 的数学期望与方差. (2) 求 $\operatorname{Cov}(X,Y)$.

63. 设随机变量 X,Y 相互独立, 具有共同分布 $N\left(\mu,\sigma^2\right)$. 设 α,β 为两个常数. (1) 求 $\operatorname{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)$. (2) 当 α,β 取何值时, $\alpha X+\beta Y$ 与 $\alpha X-\beta Y$ 相互独立.

64. 设随机变量 $(X,Y)\sim N\left(\mu,\mu,\sigma^2,\sigma^2,\rho\right)$, 其中 $\rho>0$. 问是否存在两个常数 α,β 使得 $\operatorname{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)=0$? 如存在请求出, 否则请说明原因.

65. 若随机变量 X 和 Y 的协方差 $\operatorname{Cov}(X,Y)=0$, 则有 (A. $\operatorname{Var}(X+Y)=\operatorname{Var}(X)\operatorname{Var}(Y)$ B. $\operatorname{Var}(X+Y)=\operatorname{Var}(X)+\operatorname{Var}(Y)$ C. X 和 Y 独立 D. X 和 Y 不独立

66. 设随机变量 (X,Y) 服从区域 $\mathbb{S}=\{(x,y):|x|+|y|\leqslant 1\}$ 中的均匀分布. (1) 求 $\operatorname{Cov}(X,Y)$. (2) X 与 Y 是否独立?

67. 设某人购买了保险, 其一年内发生的汽车事故的次数是一个随机变量 N, 其中 N 以概率 $1/3,1/2,1/6$ 取值 $0,1,2$. 每次事故的索赔额服从均值为 2000 的指数分布, 但其保险合同中规定了免赔额为 700 , 即只赔付超过 700 的部分. 求保险公司赔付给此人的事故金额的均值和标准差.

68. 设随机变量 X 服从 $[-1,1]$ 上的均匀分布, 令 $Y=\cos X$, 证明 X 和 Y 的协方差为 0, 但是不独立.

69. 投资组合是将总资本按一定比例分配于各种投资, 以分散和降低风险, 所谓风险通常以方差来度量. 现假设某两种投资的回报率 X,Y 都是随机变量, 投资的风险 (即方差) 为 $\operatorname{Var}(X)=\operatorname{Var}(Y)=\sigma^2$. 假设 $\rho_{XY}=-0.5$, 即两种投资呈负相关. 记投资组合中两种投资的比例分别为 π 和 $1-\pi$, 则投资组合的回报率为 $Z=\pi X+(1-\pi)Y$. (1) 试证明该投资组合 Z 的风险小于将所有资本投资于其中一个的风险. (2) 求使得投资组合风险最小的分配比例 π.

70. (1) 证明

$$\operatorname{Cov}(X_1,X_2)=\operatorname{Cov}(X_1,\mathbb{E}[X_2\mid X_1])$$

(2) 假设存在常数 $c,\mathbb{E}[X_2\mid X_1]=1+cX_1$, 证明

$$c=\operatorname{Cov}(X_1,X_2)/\operatorname{Var}(X_1)$$

71. 若 $\mathbb{E}[X_2\mid X_1]=1$, 证明

$$\operatorname{Var}(X_1X_2)\geqslant\operatorname{Var}(X_1)$$

72. 设 $X_1,\ldots,X_n,\ldots$ 是独立同分布的随机变量, 其分布为几何分布 $\operatorname{Geo}(p)$; 假设它们与另一个随机变量 N 独立, 且 N 服从二项分布 $B(n,q)$, $\quad p,q\in(0,1)$. 令 $S=\sum_{i=1}^N X_i$. 则 $\mathbb{E}[S\mid N=n]=\underline{\operatorname{Var}[S\mid N=n]}=\operatorname{Var}[S]=$

73. 设随机变量 X 服从正态分布 $N(\mu_1,\sigma_1^2)$, 令 $Y=|X|$, 则 $\mathbb{E}(X\mid Y=x)=$ 和 $\operatorname{Var}(X\mid Y=x)=$

74. 设 $N(t)$ 是一个依赖于变量 t 的随机变量, 对 $t>0$, $N(t)$ 服从分布为

$$\mathbb{P}(N(t)=k)=e^{-\lambda t}\frac{(\lambda t)^k}{k!},k=0,1,2,\ldots$$

设 T 是一个均值为 a, 方差为 $b>0$ 的非负随机变量. 求 (1) $\operatorname{Cov}(T,N(T))$; $\quad$ (2) $\operatorname{Var}(N(T))$.

75. 设一个随机变量 X 有概率质量函数: $p_i=\mathbb{P}(X=i),i=0,1,2,\ldots$, 设均值 $a=\mathbb{E}[X]$ 和 $\operatorname{Var}(X)=\sigma^2$ 已知, 求 $\mathbb{E}[X\mid X>0]$ 和 $\operatorname{Var}(X\mid X>0)$.

76. 经统计, 每周移民到某地区的家庭数 N 服从均值为 3 的泊松分布, 即

$$\mathbb{P}(N=k)=e^{-3}\frac{3^k}{k!},\quad k=0,1,2,\ldots$$

如果每个家庭的人数是独立的, 分别以概率 1/8, 1/4, 1/2, 1/8 取值 1, 2, 3 和 4. 求未来四周内移民到该地区人数的期望值.

77. 设二元随机变量 (X,Y) 服从二元正态分布 $N(1,2,4,9,3)$, 则 $\mathbb{E}[X\mid Y=2]=1,\mathbb{E}(XY^2+Y\mid Y=1)$

解: 参考352, $\mathbb{E}(XY^2+Y\mid Y=1)=E(X|Y=1)+1=1+0.3\cdot\frac{2}{3}(1-2)+1=1.8$

78. 设某公交车站于每小时的 15 分, 30 分, 45 分发车, 设某顾客不知发车时间, 在每小时内的任一时刻随机到达车站, 求乘客候车时间的数学期望.

79. 设 X 服从参数为 $(2,p)$ 的二项分布, Y 在给定 $X=i$ 条件下服从参数为 i 的 Poisson 分布, $i=0,1,2$. (参数为 0 的 Poisson 分布理解为 0 点的退化分布) 则 $\mathbb{E}[X]=$

80. 设某人在一周内有两天会购买彩票, 每次买一注彩票, 每注两块钱, 每注彩票以 $1/10$ 的概率中五元钱, 其余获得 0 元, 设此人每天的首次购买彩票如果中了五元则再买一注, 否则就回家, 且第二次购买无论结果如何均回家. 求此人每周的收益的数学期望.

81. (1) 设随机变量 X 与 Y 独立, 均服从泊松分布, 参数分别为 λ 与 μ. 对任何给定的非负整数 $k\leqslant m$, 求 $\mathbb{P}(X=k\mid X+Y=m)$ 及 $\mathrm{E}(X\mid X+Y=m)$. (2) 设随机变量 X 与 Y 独立, 均服从二项分布 $B(n,p)$, 对任何给定的非负整数 $k\leqslant m$, 求 $\mathbb{P}(X=k\mid X+Y=m)$ 及 $\mathbb{E}(X\mid X+Y=m)$.

82. 设随机变量 X 与 Y 独立, 且服从相同的分布. 求 $\mathbb{E}(X\mid X+Y=z)$.

83. 设随机变量 X 与 Y 分别是均值为 $1/\lambda$ 和 $1/\mu$ 的独立的指数随机变量. (1) 证明在条件 $X>Y$ 下, 随机变量 $\min\{X,Y\}$ 和 $X-Y$ 是相互独立的. (2) 证明对任意正数 $c>0$,

$$\mathbb{E}[\min\{X,Y\}\mid X>Y+c]=\mathbb{E}[\min\{X,Y\}\mid X>Y]=\mathbb{E}[\min\{X,Y\}]=\frac{1}{\lambda+\mu}$$

第五章 极限理论

第 1 节 大数定律

1 弱大数律

反映了平均结果的稳定性

$$\lim_{n\to+\infty} P(|\frac{S_n}{n}-\mu|\geqslant\varepsilon)=0,\forall\varepsilon>0$$

$$\frac{S_n}{n}\stackrel{n\to\infty}{\longrightarrow} EX,\quad \forall\varepsilon>0,\quad EX^2<+\infty$$

2 Марков -Чебышёвым 不等式

$$P(|X|\geqslant\varepsilon)\leqslant\frac{EX^2}{\varepsilon^2},\quad P(|X-\mu|\geqslant\varepsilon)\leqslant\frac{\mathrm{Var}(X)}{\varepsilon^2}$$

证明:

$$\mathbb{P}(A)=EI(A)=E(1\cdot I_{(|x|\geqslant\varepsilon)})\leqslant E(\frac{x^2}{\varepsilon^2}I_{(|x|\geqslant\varepsilon)})=\frac{EX^2}{\varepsilon^2}$$

第 2 节 中心极限定理 Central Limit Theorem

$$\frac{\sum_{i=1}^n X_j-E\left(\sum_{j=1}^n X_j\right)}{\sqrt{\mathrm{Var}\left(\sum_{i=1}^n X_j\right)}}\text{当 n 很大近似服从正态分布}\to N(0,1)$$

Ляпунóв 的分析:

$$\begin{aligned}Ee^{itX} &=Ee^{it\frac{\sum(x_j-\mu)}{\sqrt{n\sigma^2}}}=\prod_{j=1}^{j}Ee^{it\frac{x_j-\mu}{\sqrt{n}\sigma}}\\ &=\left(Ee^{it\frac{x_j-\mu}{\sqrt{n}\sigma}}\right)^n=E\left(1+\frac{it\left(x_j-\mu\right)}{\sqrt{n}\sigma}+\frac{i^2t^2}{2n\sigma^2}\left(x_j-\mu\right)^2+o\left(\frac{1}{n}\right)\right)^n\\ &=\left(1-\frac{1}{2}\frac{t^2}{n}+o(\frac{1}{n})\right)=e^{-\frac{1}{2}t^2}(n\to+\infty)\Rightarrow X\sim N(0,1)\end{aligned}$$

(a) Lindeberg–Lévy CLT

随机变量序列 $X_1,X_2,\ldots X_n,\ldots$ 独立同分布，且 $EX_i=\mu,DX_i=\sigma^2(i=1,2,3\ldots)$，则对任意实数 x，有

$$\lim_{n\to\infty}P\left\{\frac{\sum_{i=1}^n X_i-n\mu}{\sqrt{n}\sigma}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt=\Phi\left(\frac{x-n\mu}{\sqrt{n}\sigma}\right)$$

(b) De Moivre–Laplace CLT

设 $X\sim B(n,p)(0<p<1)(n=1,2,\ldots)$，则对任意实数 x，有

$$\lim_{n\to\infty}P\left\{\frac{X_n-np}{\sqrt{np(1-p)}}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt=\Phi\left(\frac{x-np}{\sqrt{np(1-p)}}\right)$$

第 3 节 思考题

1. 某设餐厅每年平均接待 365×400 名顾客, 设每位顾客的消费额 (元) 服从 $(100, 800)$ 上的均匀分布, 且顾客的消费额是相互独立的. 则该餐厅每天的平均营业额为

2. 设 X, Y 为两个非负随机变量, 有有限的期望 $\mathbb{E}[X]=2, \mathbb{E}[Y]=3$ 和协方差 $\operatorname{Cov}(X, Y)=-5$. 则类似切比雪夫不等式可得到概率 $\mathbb{P}(XY>\varepsilon), \quad \varepsilon>0$, 的上界为

3. 设 $X_1, \ldots, X_n, \ldots$ 是一列独立同分布的随机变量, 均值为 2 , 方差为 2 , 记
$$Y_n=\frac{X_1 X_2+X_3 X_4+\cdots+X_{2n-1} X_{2n}}{n}, n \in \mathbb{N}$$
则 Y_n 依概率收敛于

4. 设 $(X_1, Y_1) \ldots, (X_n, Y_n), \ldots$ 是一列独立同分布的二维随机变量, 均值均为 2 , 方差均为 2 , 设对任意的 $n \in \mathbb{N}, \operatorname{Cov}(X_n, Y_n)=1$, 记
$$Z_n=\frac{X_1 Y_1+\cdots+X_n Y_n}{n}, n \in \mathbb{N}$$
则 Z_n 依概率收敛于

5. 设 $X_1, \ldots, X_n, \ldots$ 为一列独立同分布的随机变量, 均值为 a, 方差为 b, 则 $\frac{1}{n} \sum_{i=1}^n X_i^2$ 依概率收敛于

6. 设在每次试验中, 某事件 A 发生的概率为 0.2, 分别利用切比雪夫不等式及中心极限定理估计, 在 500 次独立重复试验中, 事件 A 发生的次数在 80 到 120 之间的概率.

7. 设 $X_1, \ldots, X_n, \ldots$ 为一列独立同分布的随机变量, 满足 $\mathbb{E}\left[X_i^k\right]=\alpha_k, k=1,2,3,4$, 则利用中心极限定理说明 $\sum_{i=1}^n X_i^2$ 渐近分布是什么.

8. 假设某地区的房屋入住率是 20%, 以 X 表示随机抽查 100 个房屋中有人居住的户数. 求有人居住的户数不少于 15 户且不多于 25 户的概率的近似值.

9. 设某高校共有 1000 人, 选课系统中设定一门课必须有 50 人选中方可开课, 假设每人选择某一门课程的概率是 4%, 且每个学生选择与否相互独立, 求该课程可开课的概率.

10. 设各零件的重量都是随机变量, 它们相互独立, 且服从相同的分布, 其数学期望为 0.5 kg, 标准差为 0.1 kg, 问 5000 只零件的总重量超过 2510 kg 的概率是多少?

11. 设某沿海城市实行人口移民限制, 每个月移民至该城市的家庭数为 8 , 如果每个家庭的人数是独立的, 分别以概率 $1/8, 1/4, 1/2, 1/8$ 取值 $1, 2, 3$ 和 4 . 求未来三年内不超过 760 人移民到该城市的近似概率.

12. 设某次考试共 100 道选择题, 每道题为 4 选 1 . 设某考生一点也不懂, 完全靠蒙, 即随机选, 分别求选对 20 题以上和 40 题以上的概率.

13. (1) 一个复杂的系统由 100 个相互独立起作用的部件所组成. 在整个运行期间每个部件损坏的概率为 0.10. 为了使整个系统起作用, 至少必须有 85 个部件正常工作, 求整个系统起作用的概率. (2) 一个复杂的系统由 n 个相互独立起作用的部件所组成, 且必须至少有 80% 的部件工作才能使整个系统正常工作. 每个部件的可靠性为 0.90. 问 n 至少为多大才能使系统的可靠性不低于 0.95 ?

14. 设某自动取款机每天有 200 次取款, 设每次的取款额 (百元) 服从

$$\{1,2,3,4,5,6,7,8,9,10\}$$

上的均匀分布, 且每次取款额是相互独立的. 试求该存款机要至少存多少钱才能保证以 95% 的概率不会出现余额不足.

15. 某种计算机在进行加法时, 要对每个加数进行取整. 设每次取整的误差相互独立且服从 $(-0.5,0.5)$ 上的均匀分布. (1) 若现要进行 1500 次加法运算, 求误差总和的绝对值超过 15 的概率. (2) 若要保证误差总和的绝对值不超过 10 的概率不小于 0.90, 至多只能进行多少次加法运算?

16. 设某生产线上组装每件产品的时间服从指数分布, 平均需要 10 分钟, 且各产品的组装时间是相互独立的. (1) 试求组装 100 件产品需要 15 小时至 20 小时的概率. (2) 保证有 95% 的可能性, 问 16 小时内最多可以组装多少件产品.

17. 设某保险公司每年平均承保车险的车辆数为 2400 , 每个参保车辆所交保费为 5000 . 设每年内每个参保车辆的事故数 (即索赔次数) 服从参数 (速率) 为 2 的泊松分布, 即

$$\mathbb{P}(X=k)=e^{-2}\frac{2^k}{k!},\quad k=0,1,2\dots$$

且每次事故的索赔额度服从 $[1000,5000]$ 上的均匀分布. 求平均每年保险公司盈利 2000000 的概率.

第六章 数理统计基本概念

第 1 节 基本概念

第 2 节 三大分布

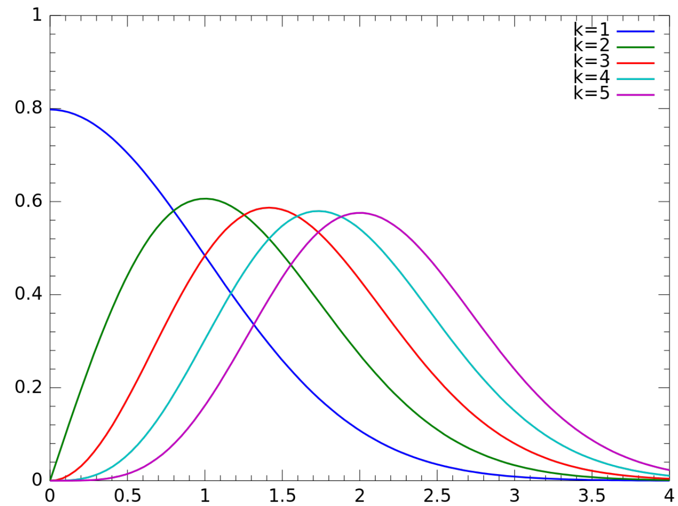
1 χ-distribution

设 $X_1, X_2, \ldots, X_n$, i.i.d. $\sim N(0,1)$，则

$$X = \sum_{i=1}^{n} X_i^2$$

的分布为具有自由度 n 的 χ^2 分布，
记作 $X \sim \chi^2(n)$ 或 $X \sim \chi_n^2$

$$p(x;n) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \quad x > 0$$



χ-distribution 概率密度函数

性质 6.2.1. $X_j^2 \sim \Gamma(\frac{1}{2}, \frac{1}{2})$, $\quad X = \sum_{j=1}^{n} X_j^2 \sim \Gamma(\frac{n}{2}, \frac{1}{2})$. $\Rightarrow P(Y = X_j^2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} y^{-\frac{1}{2}}$

性质 6.2.2. $n \geqslant 3$, 单峰，右偏

性质 6.2.3. $X \sim \chi_m^2, Y \sim \chi_n^2$ 且独立，则 $X + Y \sim \chi_{m+n}^2$

性质 6.2.4. $X \sim \chi_n^2 \Rightarrow E(X) = n, Var(X) = \Sigma_{i=1}^{n}(E(X_i^4 - (EX_i^2)^2) = 2n \because EX^{2n} = (2n-1)!!$

★ **定理 6.2.1.** 设随机样本 $X_1, \cdots, X_n \sim N(\mu, \sigma^2)$，样本均值 $\bar{X} = \dfrac{\sum_{j=1}^{n} X_j}{n}$ 且样本方差 $S^2 = \sum_{j=1}^{n} \dfrac{(X_j - \bar{X})^2}{(n-1)}$，则 *(1) $\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$;(2) $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$;(3) $\bar{X}$ 和 S^2 相互独立.*

证明. (1) $\bar{X} = \dfrac{1}{n}\sum_{j=1}^{n} X_j \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$ (服从正态分布随机变量的线性组合仍为正态分布)

(2)
$$\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} = \sum_{i=1}^{n} \frac{\left[(X_i - \bar{X}) + (\bar{X} - \mu)\right]^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} + \frac{n(\bar{X} - \mu)^2}{\sigma^2}$$

因为 $\displaystyle\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_n^2, \frac{n(\bar{X} - \mu)^2}{\sigma^2} \sim \chi_1^2 \Rightarrow \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} \sim \chi_{n-1}^2$

(3) $\displaystyle\because \bar{X} = \sum_{l=1}^{n} \frac{1}{n} X_l; \qquad X_i - \bar{X} = \sum_{l=1}^{n}\left(\delta_{il} - \frac{1}{n}\right) X_l,$ 其中 $\delta_{il} = I_{(i=l)}$

$$\Rightarrow \operatorname{Cov}\left(X_i - \bar{X}, \bar{X}\right) = \sigma^2 \sum_{l=1}^{n}\left(\delta_{il} - \frac{1}{n}\right)\frac{1}{n} = \sigma^2\left(\frac{1}{n} - n\frac{1}{n^2}\right) = 0$$ 故 $\bar{X}$ 和 S^2 相互独立

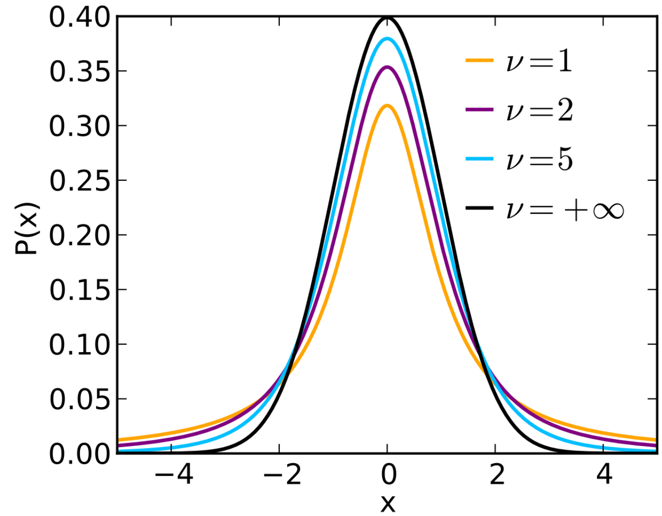
□

2 t-distribution

设 $X \sim N(0,1), K \sim \chi^2(n)$ 且二者独立，则

$$T = \frac{X}{\sqrt{K/n}}$$

的分布称为自由度为 n 的 t 分布，
记作 $T \sim t(n)$ 或 $T \sim t_n$.



Student-t-distribution 概率密度函数

性质 6.2.5.

$$p(t;n) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1+\frac{t^2}{n}\right)^{-\frac{n}{2}}$$

性质 6.2.6.

$$\mathrm{E}(t(n)) = 0, \quad n \geqslant 2$$

$$\mathrm{Var}(t(n)) = \frac{n}{n-2}, \quad n \geqslant 3$$

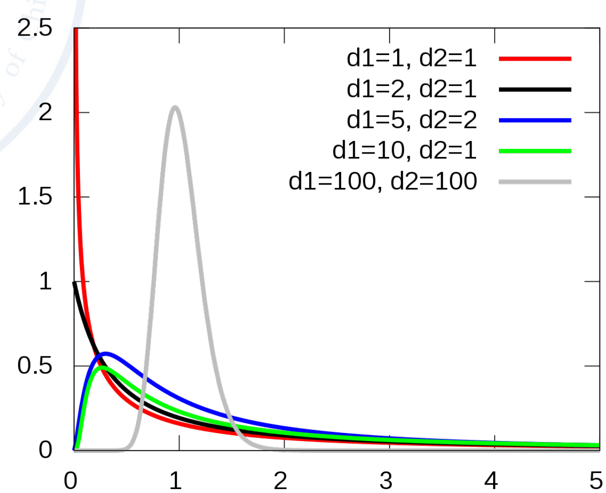
3 F-distribution

设 $X \sim \chi^2(m), Y \sim \chi^2(n)$ 且二者独立，则

$$F = \frac{X/m}{Y/n}$$

的分布称为具有自由度 (m,n)
(或第一自由度为 m ，第二自由度为 n) 的 F 分布，
记为 $F \sim F(m,n)$ 或 $F \sim F_{m,n}$

$$p(x;n,m) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} m^{\frac{m}{2}} n^{\frac{n}{2}} \frac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{m+n}{2}}}, \quad x > 0$$



F-distribution 概率密度函数

性质 6.2.7.

$$X \sim F_{m,n} \Leftrightarrow \frac{1}{X} \sim F_{n,m}$$

性质 6.2.8. 记其上 α 分位数为 $F_\alpha(m,n)$ ，则：

$$F_{1-\alpha}(m,n) = [F_\alpha(n,m)]^{-1}$$

第 3 节 思考题

1. 如何理解“样本既可以视为是随机变量, 也可以是具体的数值”?

解: 试验之前, 样本观测值是未知的, 所以可以看成是随机变量;
试验完成后, 样本又是一组确定的值, 故可以看成是一组确定的值.
视样本为随机变量时候方便研究其性质.

2. 设某人进行射击练习, 他独立射击 5 次, 结果分别为 $8,9,7,10$ 和 6 环. 则总体是什么? 样本是什么?

解: 总体：该人射击的所有可能环数，记为 $\Omega=\{1,2,3,\cdots,10\}$
样本：射击5 次的环数为$X_1,\ldots,X_5$, 样本值分别为$8,9,7,10$ 和6.

3. 从全班同学中随机选择 5 名同学, 则总体和样本分别指什么?

解: 总体为全班同学, 样本为选择的 5 名同学. 抽样结束后, 则样本值为特定的 5 名同学.

4. 盒中有 4 个黑球, 2 个白球. 现从中随机摸出 2 个球, 记录其中的白球数目. 试分别写出不放回和有放回两种摸球方式下的总体分布.

5. 调查 50 个人对某件事情是 (1) 否 (0) 支持, 假设每个人对该事情支持的可能性为 p, 各人之间相互独立. 则总体分布是什么? 若其中 10 个人的调查结果为 $x_1,\ldots,x_{10}$ (其中 x_i 只取 0 或 1), 则抽样分布是什么?

6. 测量一个物体的长度, 试写出总体, 并解释该总体的合理性.

7. 考虑某工厂生产的灯管寿命, 则总体又是什么? 解释你做法的合理性.

解: 总体为指数分布/威布尔分布. 由于寿命为非负值, 且失效率往往会随时间增加 (为简单记, 也常假设为常数), 因此指数分布或威布尔分布是合适的.

8. 一个总体有 N 个元素, 其指标分别为 $a_1>a_2>\ldots>a_N$, 指定自然数 $M<N, n<N$, 在 $(a_1,\ldots,a_M)$ 中不放回的随机抽出 m 个, 在 $(a_{M+1},\ldots,a_N)$ 中不放回地随机抽出 $n-m$ 个. 写出所得样本的分布.

解：样本分布

$$P\{X_1=x_1,\ldots,X_n=x_n\}=\frac{1}{A_M^m A_{N-M}^{n-m}}$$

其中 $(x_1,\ldots,x_m)\subset\{a_1,\ldots,a_M\},\quad (x_{m+},\ldots,x_n)\subset\{a_{M+1},\ldots,a_N\}$

9. 假设总体 X 服从两点分布 $B(1,p)$, 其中 p 为末知参数, $X=(X_1,\ldots,X_5)$ 为从此总体中抽取的简单样本, 试
(1) 写出样本空间和抽样分布.
(2) 指出 $X_1+X_2, \min_{1\leqslant i\leqslant 5} X_i, X_5+2p, X_5-\mathbb{E}X_1, (X_5-X_1)^2/\operatorname{Var}(X_1)$ 哪些是统计量, 哪些不是, 为什么?
(3) 若样本观察值 $(X_1,\ldots,X_n)$ 中有 m 个 $1, n-m$ 个 0, 求经验分布函数.

解: (1) 样本空间为 $\Omega = \{(x_1, \ldots, x_5) : x_i = 0, 1, 1 \leqslant i \leqslant 5\}$,
抽样分布为 $\mathbb{P}(X_1 = x_1, \ldots, X_5 = x_5) = p^{\sum_{i=1}^5 x_i}(1-p)^{5-\sum_{i=1}^5 x_i}$.
(2) $X_1 + X_2$ 和 $\min_{1\leqslant i\leqslant 5} X_i$ 为统计量, 其余因为依赖于末知参数 p, 因此不是统计量.
(3)

$$F_n(x) = \begin{cases} 0, & x < x_1, \\ \dfrac{i}{n}, & x_k \leqslant x < x_{k+1}, (x_k\text{为第 i 个取 1 的随机变量}, 1 \leqslant i \leqslant m) \\ 1, & x \geqslant x_n \end{cases}$$

10. 设样本量为 10 的一个样本值为

$$0.4, 0.3, -0.3, -0.1, 1.7, 0.6, -0.1, 0.9, 2.6, 0.5$$

试计算经验分布函数.

解:

$$F_n(x) = \begin{cases} 0, & x < -0.3 \\ 0.1, & -0.3 \leqslant x < -0.1 \\ 0.3, & -0.1 \leqslant x < 0.3 \\ 0.4, & 0.3 \leqslant x < 0.4 \\ 0.5, & 0.4 \leqslant x < 0.5 \\ 0.6, & 0.5 \leqslant x < 0.6 \\ 0.7, & 0.6 \leqslant x < 0.9 \\ 0.8, & 0.9 \leqslant x < 1.7 \\ 0.9, & 1.7 \leqslant x < 2.6 \\ 1, & 2.6 \leqslant x \end{cases}$$

1 经验 (样本) 分布函数

(对于 $X_1, \cdots, X_n$):

$$\begin{aligned} F_n(x) &= \frac{x_1, x_2, \cdots, x_n \text{ 中小于等于} x \text{ 的样本值个数}}{n} \\ &= \begin{cases} 0, & x < x_1, \\ \dfrac{k}{n}, & x_k \leqslant x < x_{k+1}, (k = 1, 2, \cdots, n-1) \\ 1, & x \geqslant x_n \end{cases} \end{aligned}$$

11. 随机地取 7 只活塞环, 测得它们的直径为 (以 mm 计)

$$74.001, \quad 74.005, \quad 74.003, \quad 74.000, \quad 73.908, \quad 74.006, \quad 74.002$$

试求样本均值和样本标准差.

解：　利用 CASIO®fx-991N X 计算器的统计功能, 选择单变量统计，输入这 7 个数据，之后按 OPTN 键，选择单变量计算，得到一系列结果为，

$\bar{x}$ ——平均值
Σx ——数据总和
Σx^2 ——数据平方和
$\sigma^2 x$ ——总体方差
σx ——总体标准差
$s^2 x$ ——样本方差
sx ——样本标准差
n ——数据个数
$\min(x)$ ——最小值
Q_1 ——下四分位数 (第一四分位数)
Med ——中位数
Q_3 ——上四分位数（第三四分位数）
$\max(x)$ ——最大值

得到样本均值 73.989, 样本标准差 0.0351

12. (2003 年全国硕士研究生入学考试试题) 设随机变量 $X \sim t_n(n>1), Y=\frac{1}{X^2}$, 则 (A) $Y \sim \chi_n^2$ (B) $Y \sim \chi_{n-1}^2$ (C) $Y \sim F_{n,1}$ (D) $Y \sim F_{1,n}$

解：　(C) $\because X=\dfrac{N}{\sqrt{Kn}}, K \sim \chi_n^2, N \sim N(0,1) \Rightarrow Y=\dfrac{K}{nN^2} \sim F_{n,1}$

13. (2005 年全国硕士研究生入学考试试题) 设 $X_1, \ldots, X_n$ 为来自总体 $N(0,1)$ 的简单随机样本, $\bar{X}$ 为样本均值, S^2 为样本方差, 则 (A) $n\bar{X} \sim N(0,1)$ (B) $nS^2 \sim \chi_n^2$ (C) $\frac{(n-1)\bar{X}}{S} \sim t_{n-1}$ (D) $\frac{(n-1)X_1^2}{\sum_{i=2}^n X_i^2} \sim F_{1,n-1}$, 选 D

(A) $\because \bar{X} \sim N(0,\dfrac{1}{n}) \Rightarrow n\bar{X} \sim N(0,n)$ 故 A 错误, 重要关系：$Var(\bar{X})=\dfrac{1}{n}Var(X)$

(B) $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$ 故 B 错误

(C) $\dfrac{n\bar{X}\sigma}{S} \sim t_{n-1}$ 故 C 错误

(D) $\dfrac{(n-1)X_1^2}{\sum_{i=2}^n X_i^2}=\dfrac{X_1^2}{\sum_{i=2}^n X_i^2/(n-1)} \sim F_{1,n-1}$

14. 设随机变量 X_1, X_2 相互独立同分布于标准正态分布, 求 $Y=\frac{(X_1-X_2)^2}{(X_1+X_2)^2}$ 的分布.

15. 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, 令 $T=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$. 试求 a, b 使统计量 T 服从 χ^2 分布.

16. 设 $X_1, X_2, \ldots, X_9$ 为独立同分布的正态随机变量, 记

$$Y_1=\frac{1}{6}(X_1+\ldots+X_6), \quad Y_2=\frac{1}{3}(X_7+X_8+X_9), \quad S^2=\frac{1}{2}\sum_{i=7}^9 (X_i-Y_2)^2$$

试求 $Z=\frac{\sqrt{2}(Y_1-Y_2)}{S}$ 的分布.

17. 设 $X_1, X_2, \ldots, X_{15}$ 是独立同分布的随机变量, 服从正态分布 $N(0, 2^2)$. 试求

$$Y = \frac{X_1^2 + \ldots + X_{10}^2}{2\left(X_{11}^2 + \ldots + X_{15}^2\right)}$$

的概率分布.

18. 设 $X_1, \ldots, X_n$ 为从下列总体中抽取的简单样本, 试求样本均值 $\bar{X}$ 的分布: (1) 正态总体 $N(a, \sigma^2)$; (2) 参数为 λ 的 Poisson 总体; (3) 参数为 λ 的指数分布.

解: (1) $\bar{X} \sim N\left(a, \frac{\sigma^2}{n}\right)$

(2) $n\bar{X} \sim Poi(n\lambda)$

(3) $2n\lambda\bar{X} \sim \chi_{2n}^2$

19. 设 $X_1, \ldots, X_n$ 是从两点分布 $B(1, p)$ 中抽取的简单样本, $0 < p < 1$, 记 $\bar{X}$ 为样本均值, 求 $S_n^2 = \frac{1}{n}\sum_{i=1}^n \left(X_i - \bar{X}\right)^2$ 的期望.

20. 设 $X_1, \ldots, X_n$ 为来自正态总体 $N(a, \sigma^2)$ 的一个简单随机样本, $\bar{X}$ 和 S_n^2 分别表示样本均值和样本方差, 又设 $X_{n+1} \sim N(a, \sigma^2)$ 且与 $X_1, \ldots, X_n$ 独立, 试求统计量 $\frac{X_{n+1} - \bar{X}}{S_n}\sqrt{\frac{n}{n+1}}$ 的分布.

21. 设 $X_1, \ldots, X_m$ 为来自正态总体 $N(\mu_1, \sigma^2)$ 的一个简单随机样本, $Y_1, \ldots, Y_n$ 为来自正态总体 $N(\mu_2, \sigma^2)$ 的一个简单随机样本, 且 $X_1, \ldots, X_n$ 和 $Y_1, \ldots, Y_n$ 相互独立, $\bar{X}$ 和 $\bar{Y}$ 分别表示它们的样本均值, S_{1m}^2 和 S_{2n}^2 分别表示它们的样本方差, α 和 β 是两个给定的实数, 试求

$$T = \frac{\alpha\left(\bar{X} - \mu_1\right) + \beta\left(\bar{Y} - \mu_2\right)}{\sqrt{\frac{(m-1)S_{1m}^2 + (n-1)S_{2n}^2}{n+m-2} \cdot \left(\frac{\alpha^2}{m} + \frac{\beta^2}{n}\right)}}$$

的分布.

22. 设 $X_{(1)}, \ldots, X_{(n)}$ 为从均匀分布 $U(0, 1)$ 中抽取的次序统计量, (1) 样本量 n 为多大时, 才能使 $\mathbb{P}\left(X_{(n)} \geqslant 0.99\right) \geqslant 0.95$? (2) 求极差 $R_n = X_{(n)} - X_{(1)}$ 的期望;

第七章　参数估计

第 1 节　点估计

1　矩估计

第 1 步 给出总体矩形式

$$\begin{cases} EX = \mu \\ Var(X) = \sigma^2 \end{cases}$$

第 2 步 用样本矩替代估计 (矩估计不唯一，能用低阶矩不用高阶)

$$\begin{cases} \bar{X} = \hat{\mu} \\ \dfrac{1}{n}\Sigma(X_j - \bar{X})^2 = \hat{\sigma^2} \end{cases}$$

2　极大似然估计

第 1 步 写出似然函数

$$L(\theta) = L(x_1, x_2, \cdots, x_n; \theta_1, \theta_2, \cdots, \theta_k) = \prod_{i=1}^{n} p(x_i; \theta)$$

其中,θ 为描述总体的参数，通过对样本概率密度的累乘，计算最大可能性的 L，则为极大似然估计

e.g. 二项分布 $B(n, p)$

$$L(p) = \prod_{i=1}^{n} p^{x_j}(1-p)^{1-x_j} = p^{\Sigma x_j}(1-p)^{n-\Sigma x_j}$$

第 2 步 研究此函数的性质 (单调则在边界取) 不单调则常取对数研究极值 (或用 Hessen 矩阵)

$$\frac{\partial L}{\partial \theta} = 0 / \frac{\partial \ln L}{\partial \theta} = 0$$

几个常见的 MLE(Maximum likelihood estimation):

(a) 正态分布　$X_1, \cdots, X_n \sim N(\mu, \sigma^2)$.

$$\begin{aligned} L\left(\mu, \sigma^2\right) &= f\left(x_1, \cdots, x_n\right) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ &= (2\pi)^{-\frac{n}{2}} \left(\sigma^2\right)^{-\frac{n}{2}} e^{-\frac{\sum_{j=1}^{n}(x_j-\mu)^2}{2\sigma}} \\ \Rightarrow l\left(\mu, \sigma^2\right) &= \ln L\left(\mu, \sigma^2\right) \\ &= -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \sum_{j=1}^{n}\frac{(x-\mu)^2}{2\sigma^2} \end{aligned}$$

$$\begin{cases} \frac{\partial l}{\partial \mu} = \frac{\pm 2\sum_{j=1}^{n}(x_j-\mu)}{2\sigma^2} = 0 \\ \frac{\partial l}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{\sum_{j=1}^{n}(x_j-\mu)}{2}\frac{1}{(\sigma^2)^2} = 0 \end{cases} \Rightarrow \begin{cases} \mu = \bar{x} \\ \sigma^2 = \frac{1}{n}\sum_{j=1}^{n}\left(x_j-\mu\right)^2 \end{cases}$$

(b) $X_1,\cdots,X_n \sim f(x)=\frac{1}{\theta}e^{-\frac{1}{\theta}(x-\mu)}I_{(x,\mu,\theta>0)}$

$$L(\mu,\theta)=\prod_{j=1}^{n}\frac{1}{\theta}e^{-\frac{1}{\theta}(x_j-\mu)}I_{(x_j>\mu)}=\frac{1}{\theta^n}e^{-\frac{1}{\theta}\sum_{j=1}^{n}(x_j-\mu)}\prod_{j=1}^{n}I_{\left(x_{(1)}>\mu\right)}$$

$\mu<x_{(1)}, L(\mu,\theta)\leqslant L\left(x_{(1)},\theta\right), x_{(1)}$为样本取值最小值, $x_{(n)}$为样本取值最大值,

$$\Rightarrow \hat{\mu}=x_{(1)}=\min\left\{x_1,\cdots,x_n\right\}$$

$$L(\hat{\mu},\theta)=\ln L(\hat{\mu},\theta)=-n\ln\theta-\frac{1}{\theta}\sum_{j=1}^{n}\left(x_j-x_{(1)}\right)$$

$$\frac{\partial L}{\partial\theta}=-\frac{n}{\theta}+\frac{1}{\theta^2}\sum_{j=0}^{n}\left(x_j-x_{(1)}\right)$$

$$\Rightarrow \hat{\theta}=\frac{1}{n}\sum_{j=1}^{n}\left(x_j-x_{(1)}\right)$$

(c) $X_1,\cdots,X_n \sim U[a,b]$.

$$L(a,b)=f\left(x_1,\cdots,x_n\right)=\prod_{j=1}^{n}\frac{1}{b-a}I_{(a<x_j<b)}=\frac{1}{(b-a)^n}I_{\left(a<x_{(1)}<x_{(n)}<b\right)}$$

$$L(a,b)\leqslant\frac{1}{\left(x_{(n)}-x_{(1)}\right)^n},\quad \hat{\theta}=x_{(n)}=\max\left\{x_1,\cdots,x_n\right\}$$

3 估计优良性准则

	相合性 Consistency	均方误差 Mean Squared Error	无偏性 Unbiasedness	有效性 Effciency
定义	统计量 $\hat{\theta}_n=\theta(X_1,X_2,\cdots,X_n)$ 满足 $P\{\mid\hat{\theta}_n-\theta\mid>\varepsilon\}\overset{n\to\infty}{=}0$	$MSE=E(\hat{\theta}-\theta)^2$ $=Var(\hat{\theta})+(E\hat{\theta}-\theta)^2$	$E\hat{\theta}(X_1,X_2,\cdots,X_n)=\theta$	$\hat{\theta}_1,\hat{\theta}$是$\theta$的无偏估计 $\mathrm{Var}(\hat{\theta}_1)\leqslant\mathrm{Var}\left(\hat{\theta}_2\right)$, 称 $\hat{\theta}_1$ 更有效
解释	$\Leftrightarrow\begin{cases}\lim\limits_{n\to\infty}E\hat{\theta}_n=0\\ \lim\limits_{n\to\infty}\mathrm{Var}\left(\hat{\theta}_n\right)=0\end{cases}$	$=EL(\hat{\theta},\theta)=E(\hat{\theta}-\theta)^2$ $=E(\hat{\theta}-E(\hat{\theta})+E\hat{\theta}-\theta)^2$ $=\mathrm{Var}(\hat{\theta})+E(E\hat{\theta}-\theta)^2$	$\hat{\theta}(X_1,X_2,\cdots,X_n)$是$\theta$的无偏估计	方差越小的无偏估计约有效

注：有时候，利用有效性比较矩估计和 MLE 的参数 θ 下，二者的方差，比较优劣

第 2 节 区间估计

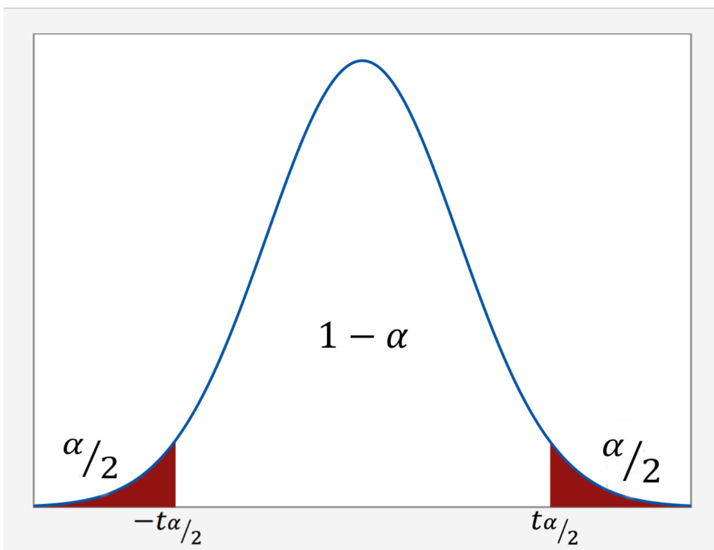
置信区间 Confidential Interval

$$\text{Neyman: }\left[\hat{\theta}_1(x_1,\cdots,x_n),\hat{\theta}_2\left(x_1,\cdots,x_n\right)\right]\text{ 估计}\theta,\left[\hat{\theta}_1,\hat{\theta}_2\right]\text{ 包含 }\theta\text{ 概率}$$

$$P\left\{\hat{\theta}_1\leqslant\theta\leqslant\hat{\theta}_2\right\}\geqslant 1-\alpha$$

$$\left[\hat{\theta}_1,\hat{\theta}_2\right]\text{ 称为 }\theta\text{ 区间估计 (水平为 }1-\alpha)$$

置信区间的面积表示（如 t 分布）



1 枢轴变量法

以服从 $X_1,\cdots,X_n\sim N(\mu,\sigma^2)$ 正态分布为例：

第 1 步 找出 μ 的良好估计：如 $\bar{X}$

第 2 步 写出 $\bar{X}$ 估计量的分布：如正态分布 $\bar{X}\sim N(\mu,\dfrac{\sigma^2}{n})$，枢轴变量 $Z=\dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\sim N(0,1)$

第 3 步 通过定义的概率，找到两侧分位数 (分位数满足 $-U_\alpha=U_{1-\alpha}$)：

$$P\{c<Z<d\}=1-\alpha,\quad c=u_{-\frac{\alpha}{2}},d=u_{\frac{\alpha}{2}}$$

第 4 步 反解出良好估计的参数 θ 范围, 得到 $\bar{X}$ 满足

$$-u_{\frac{\alpha}{2}}\leqslant\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\leqslant u_{\frac{\alpha}{2}}$$

$$\Rightarrow\text{置信区间：}\left[\bar{X}-\frac{u_{\frac{\alpha}{2}}\sigma}{\sqrt{n}},\bar{X}+\frac{u_{\frac{\alpha}{2}}\sigma}{\sqrt{n}}\right]$$

◊ 正态总体

1. μ 的 $1-\alpha$ 置信区间

(a) σ^2 已知

$$P\left\{-u_{\frac{\alpha}{2}} \leqslant \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \leqslant u_{\frac{\alpha}{2}}\right\}=1-\alpha$$
$$\Rightarrow \text{CI:} \quad \left[\bar{X}-\frac{\sigma u_{\frac{\alpha}{2}}}{\sqrt{n}}, \bar{X}+\frac{\sigma u_{\frac{\alpha}{2}}}{\sqrt{n}}\right]$$

(b) σ^2 未知

$$P\left\{-t_{n-1}(\frac{\alpha}{2}) \leqslant \frac{\sqrt{n}(\bar{X}-\mu)}{S} \leqslant t_{n-1}(\frac{\alpha}{2})\right\}=1-\alpha$$
$$\Rightarrow \text{CI:} \quad \left[\bar{X}-\frac{St_{n-1}(\frac{\alpha}{2})}{\sqrt{n}}, \bar{X}+\frac{St_{n-1}(\frac{\alpha}{2})}{\sqrt{n}}\right]$$

2. σ^2 的 $1-\alpha$ 置信区间

(a) μ 已知

$$P\left\{-\chi_n^2(\frac{\alpha}{2}) \leqslant \frac{nS^{\star 2}}{\sigma^2}=\frac{\sum(x_j-\mu)^2}{\sigma^2} \leqslant \chi_n^2(\frac{\alpha}{2})\right\}=1-\alpha$$
$$\Rightarrow \text{CI:} \quad \left[-\frac{nS^{\star 2}}{\chi_n^2(\frac{\alpha}{2})}, \frac{nS^{\star 2}}{\chi_n^2(\frac{\alpha}{2})}\right]$$

(b) μ 未知

$$P\left\{-\chi_{n-1}^2(\frac{\alpha}{2}) \leqslant \frac{(n-1)S^2}{\sigma^2} \leqslant \chi_{n-1}^2(\frac{\alpha}{2})\right\}=1-\alpha$$
$$\Rightarrow \text{CI:} \quad \left[-\frac{(n-1)S^2}{\chi_{n-1}^2(\frac{\alpha}{2})}, \frac{(n-1)S^2}{\chi_{n-1}^2(\frac{\alpha}{2})}\right]$$

◊ 两总体情形

1. $\mu_1-\mu_2$ 的 $1-\alpha$ 置信区间

(a) σ_1^2, σ_2^2 已知

$$\bar{X}-\bar{Y} \sim N\left(\mu_1-\mu_2, \frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}\right)$$
$$Z=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}\sigma_1^2+\frac{1}{m}\sigma_2^2}}$$
$$\Rightarrow \text{CI:} \quad \bar{X}-\bar{Y}+u_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_n^2}{n}+\frac{\sigma_2^2}{m}}$$

(b) σ_1^2, σ_2^2 未知

$$S_w^2=\tfrac{1}{n+m-2}\left(\sum\left(X_i-\bar{X}\right)^2+\sum\left(Y_j-\bar{Y}\right)^2\right)$$
$$\Rightarrow \text{CI:} \quad \bar{X}-\bar{Y} \pm \sqrt{\tfrac{1}{m}+\tfrac{1}{n}}{S_w}^2 t_{n+m-2}\left(\tfrac{\alpha}{2}\right)$$

2. σ_1^2/σ_2^2 的 $1-\alpha$ 置信区间

(a) μ_1, μ_2 已知

$$\frac{S_1^{*2}/\sigma_1^2}{S_2^{*2}/\sigma_2^2} \sim F_{n,m}$$

$$\Rightarrow \text{CI: } \left[\frac{S_1^{*^2}}{S_2^{*^2}}/F_{n,m}\left(\frac{\alpha}{2}\right) \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^{*^2}}{S_2^{*^2}}/F_{n,m}\left(1-\frac{\alpha}{2}\right)\right]$$

(b) μ_1, μ_2 未知

$$\frac{S_1^{*2}/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n-1,m-1}$$

$$\Rightarrow \text{CI: } \left[\frac{S_1^2}{S_2^2}/F_{n-1,m-1}\left(\frac{\alpha}{2}\right) \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2}{S_2^2}/F_{n-1,m-1}\left(1-\frac{\alpha}{2}\right)\right]$$

2 可靠度

真实值要么落在 CI 内，要么不在。

3 置信界

$$P\left(\theta \leqslant \bar{\theta}\left(X_1, \cdots, X_n\right)\right) = 1-\alpha \Leftrightarrow \text{置信上界：} (-\infty, \bar{\theta}]$$

$$P\left(\theta \geqslant \bar{\theta}\left(X_1, \cdots, X_n\right)\right) = 1-\alpha \Leftrightarrow \text{置信下界：} [\underline{\theta}, +\infty)$$

注 7.2.1. 大样本法（视为渐进正态分布）：

$$\text{二项样本：} \frac{\sum X_j - np}{\sqrt{npq}} \xrightarrow{Asymptotic} N(0,1)$$

$$\text{正态样本：} \frac{\sqrt{n}(\sum X_j - \mu)}{S} \xrightarrow{A} N(0,1) \Rightarrow \frac{\sqrt{n}(\sum X_j - \mu)}{S} \sim t_{n-1}$$

第 3 节　思考题

1. 设总体 X 的概率分布如下表,

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 $0<\theta<\frac{1}{2}$ 为末知参数. 现从此总体中抽出一样本量为 100 的简单随机样本, 其中 0 出现了 10 次, 1 出现了 53 次, 2 出现了 16 次, 3 出现了 21 次. 试求 θ 的矩估计.

2. 设总体 X 的概率分布如右表, 其中 $0<p_1,p_2<1$ 为末知参数. 现从此总体中抽出一样本量为 n 的简单随机样本, 其中 1 出现了 n_1 次, 2 出现了 n_2 次, 3 出现了 n_3 次. 试求 p 的矩估计.

X	1	2	3
P	p_1	p_2	$1-p_1-p_2$

3. 设 $X_1,\dots,X_n$ 是总体 X 的一个简单随机样本, 试求在 X 具有下列概率分布时参数 θ 的矩估计. (1) $p(x;\theta)=\frac{1}{\theta},\quad x=0,1,2,\dots,\theta-1$, 其中 θ (正整数) 是末知参数. (2) $p(x;\theta)=\begin{pmatrix} m \\ x \end{pmatrix}\theta^x(1-\theta)^{m-x},\quad x=0,1,\dots,m$ (3) $p(x;\theta)=(x-1)\theta^2(1-\theta)^{x-2},\quad x=2,3,\dots;0<\theta<1$. (4) $p(x;\theta)=-\frac{1}{\ln(1-\theta)}\frac{\theta^x}{x},\quad x=1,2\dots$ (5) $p(x;\theta)=\frac{\theta^x}{x!}e^{-\theta},\quad x=0,1,2,\dots$

4. 设 $X_1,\dots,X_n$ 是总体 X 的一个简单随机样本, 试求在 X 具有下列概率密度时参数 θ 的矩估计. (1) $f(x;\theta)=\begin{cases}\frac{2}{\theta^2}(\theta-x), & 0<x<\theta\\ 0 & \text{其他.}\end{cases}$ (2) $f(x;\theta)=\begin{cases}(\theta+1)x^\theta, & 0<x<1,\theta>0\\ 0 & \text{其他.}\end{cases}$ (3) $f(x;\theta)=\begin{cases}\sqrt{\theta}x^{\sqrt{\theta}-1}, & 0<x<1,\theta>0\\ 0 & \text{其他.}\end{cases}$ (4) $f(x;\theta)=\begin{cases}\theta c^\theta x^{-(\theta+1)}, & x>c(c>0\text{ 已知}),\theta>1\\ 0 & \text{其他.}\end{cases}$ (5) $p(x;\theta)=\begin{cases}\frac{6x}{\theta^3}(\theta-x), & 0<x<\theta\\ 0 & \text{其他}\end{cases}$ (6) $p(x;\theta)=\begin{cases}\frac{\theta^2}{x^3}e^{-\theta/x}, & x>0\\ 0 & \text{其他}\end{cases},\theta>0$

5. 总体 X 的概率密度函数为

$$f(x)=\begin{cases}\frac{4x^2}{\theta^3\sqrt{\pi}}e^{-\frac{x^2}{\theta^2}}, & x\geqslant 0\\ 0 & \text{其他.}\end{cases}$$

设 $X_1,X_2,\dots,X_n$ 是取自总体 X 的简单随机样本. (1) 求 θ 的矩估计量 $\hat{\theta}$. (2) 求 $\hat{\theta}$ 的方差.

6. (2007 年全国硕士研究生入学考试试题) 总体 X 的概率密度函数为

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

$X_1,X_2,\dots,X_n$ 是取自总体 X 的简单随机样本. $\bar{X}$ 为样本平均值, (1) 求 θ 的矩估计量 $\hat{\theta}$. (2) 判断 $4\bar{X}^2$ 是否为 θ^2 的无偏估计量, 并说明理由.

7. (1) 设 $X_1,X_2,\dots,X_n$ 是来自总体 X 的一个样本, 且 X 服从参数为 λ 的泊松分布. 求 $\mathbb{P}(X=0)$ 的极大似然估计. (2) 下表统计了某铁路局 122 个扳道员五年内由于操作失误引起的严重事故

情况, 其中 r 表示一扳道员某五年内引起严重事故的次数, s 表示扳道员人数. 假设扳道员由于操作失误在五年内所引起的严重事故的次数服从泊松分布. 求一个扳道员在五年内末引起严重事故的概率 p 的极大似然估计.

r	0	1	2	3	4	5	$\geqslant 6$
s	44	42	21	9	4	2	0

8. 设电话总机在某一段时间内接到呼叫的次数服从泊松分布. 观察一分钟内接到的呼叫次数, 设共观察了 40 次, 得如下数据:

接到的呼叫次数	0	1	2	3	4	5	$\geqslant 6$
观察到的次数	5	10	12	8	3	2	0

试求泊松分布参数 λ 的极大似然估计.

9. 试求第1题参数的极大似然估计.

10. 试求第3题各情形下参数的极大似然估计.

11. 试求第4题各情形下参数的极大似然估计,

12. * 某电子元件寿命服从指数分布, 其密度函数为

$$f(x)=\begin{cases}\frac{1}{\lambda}e^{-x/\lambda}, & 0<x<\infty\\ 0, & x\leqslant 0\end{cases}$$

从这批产品中抽取 n 个作寿命试验, 规定到第 r 个 $(0<r\leqslant n)$ 电子元件失效时就停止试验. 这样获得前 r 个次序统计量 $X_{(1)}\leqslant X_{(2)}\leqslant\ldots\leqslant X_{(r)}$ 和 n 个电子元件总试验时间 $T=\sum_{i=1}^{r}X_{(i)}+(n-r)X_{(r)}$. (1) 证明: $2T/\lambda$ 服从自由度为 $2r$ 的 χ^2 分布, 即 $2T/\lambda\sim\chi^2_{2r}$. (2) 求 λ 的矩估计.

13. 人体中某个基因的形态有三种, 分别是 AA, Aa, aa, 每个人的基因型只可能为这三种形态之一. 设总体中该基因的基因型概率分布律如下表, 其中 $\theta>0$ 为末知参数. 现从总体中随机抽取 n 个人, 其中 n_1 个人具有基因型 AA, n_2 个人为 Aa, n_3 个人为 aa. 试求 θ 的极大似然估计.

基因型	AA	Aa	aa
概率	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

14. 设从均匀总体 $U(\theta_1,\theta_2)$ 中抽取一组简单样本 $X_1,\ldots,X_n$, 试求末知参数 θ_1,θ_2 的极大似然估计.

15. 设 $X_1,\ldots,X_n$ 是抽自正态总体 $N(\mu,\sigma^2)$ 的简单随机样本, 其中 $-\infty<\mu<+\infty, \sigma^2>0$ 为末知参数. 求 $\theta=\mathbb{P}(X\geqslant 2)$ 的极大似然估计.

16. (2006 年全国硕士研究生入学考试试题) 设总体 X 的概率密度函数为

$$f(x;\theta)=\begin{cases}\theta, & 0<x<1\\ 1-\theta, & 1\leqslant x<2\\ 0, & \text{其它.}\end{cases}$$

其中 $0<\theta<1$ 为末知参数. 现从该总体中简单随机样本 $X_1,\ldots,X_n$, 求 θ 的极大似然估计.

17. (2014 年研究生入学考试试题) 设总体 X 的分布函数为

$$F(x;\theta)=\begin{cases}1-e^{-x^2/\theta}, & x\geqslant 0\\ 0, & \text{其他}\end{cases}$$

其中 θ 为未知参数且大于零, $X_1,X_2,\ldots,X_n$ 为来自总体 X 的简单随机抽样. (1) 求 $\mathbb{E}X,\mathbb{E}\left[X^2\right]$; (2) 求 θ 的极大似然估计量 $\hat{\theta}$; (3) 是否存在实数 a, 使得对任何 $\epsilon>0$, 都有 $\lim_{n\to\infty}\mathbb{P}(|\hat{\theta}-a|\geqslant\epsilon)=0$?

18. 设总体的数学期望为 $\mu, X_1,\ldots,X_n$ 是来自总体 X 的样本. 假设 $a_1,\ldots,a_n$ 是任意常数, 且 $\sum_{i=1}^n a_i\neq 0$. 验证 $\sum_{i=1}^n a_iX_i/\sum_{i=1}^n a_i$ 是 μ 的无偏估计量.

19. (2016 年研究生入学考试试题) 设总体 X 的概率密度为

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

其中 $\theta\in(0,+\infty)$ 为未知参数, X_1,X_2,X_3 为总体 X 的简单随机抽样, 令 $T=\max\left(X_1,X_2,X_3\right)$ (1) 求 T 的概率密度; (2) 确定 a, 使得 aT 为 θ 的无偏估计.

20. 设 $X_1,\ldots,X_n$ 是来自总体 X 的一个简单随机样本, $\mathbb{E}X=\mu,\operatorname{Var}(X)=\sigma^2$. (1) 确定常数 c 使得 $c\sum_{i=1}^{n-1}\left(X_{i+1}-X_i\right)^2$ 为 σ^2 的无偏估计. (2) 记 $\bar{X},S^2$ 分别是样本均值和样本方差. 确定常数 c 使 $\bar{X}^2-cS^2$ 是 μ^2 的无偏估计.

21. 设从均值为 μ, 方差为 σ^2 的总体中, 分别抽取容量为 n_1,n_2 的两个独立样本. 设 $\bar{X}_1,\bar{X}_2$ 分别是两样本的均值. 试证明对于任意常数 $a, Y=a\bar{X}_1+(1-a)\bar{X}_2$ 是 μ 的无偏估计, 并确定常数 a 使 Y 的方差达到最小.

22. 设有 k 台仪器, 第 i 台仪器测量的标准差为 $\sigma_i, i=1,\ldots,k$. 用这些仪器独立地对某一物理量 θ 各测一次, 分别得到 $X_1,X_2,\ldots,X_k$. 设仪器都没有系统误差, 即 $\mathbb{E}X_i=\theta, i=1,\ldots,k$. 问 $a_1,\ldots,a_k$ 应取何值方能使 $\hat{\theta}=\sum_{i=1}^k a_iX_i$ 估计 θ 时, $\hat{\theta}$ 是无偏的, 并且 $\operatorname{Var}(\hat{\theta})$ 最小?

23. 设 $X_1,\ldots,X_n$ 是抽自均匀分布 $U(\theta,c\theta)$ 的简单随机样本, 其中 $c>1$ 为常数, $\theta>0$ 为未知参数. (1) 试求 θ 的极大似然估计. (2) 试求 θ 的矩估计, 并验证其是否具有无偏性.

24. 设 $X_1,\ldots,X_n$ 为从下述几何分布中抽出的简单随机样本,

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

分别求出 p^{-1} 和 p^{-2} 的无偏估计.

25. 假设如第题, 并假定 $p_2=2p_1=2p$. 记 p 的矩估计为 $\hat{p}$, 现定义

$$\hat{p}_1=\frac{n_1}{n},\quad \hat{p}_2=\frac{n_2}{2n},\quad \hat{p}_3=\frac{1}{3}\left(1-\frac{n_3}{n}\right)$$

试验证它们的无偏性并确定何者的方差最小.

26. (2010 年全国硕士研究生入学考试试题) 设总体 X 的概率分布为

X	1	2	3
$\mathbb{P}$	$1-\theta$	$\theta-\theta^2$	θ^2

其中 $\theta \in (0,1)$ 末知, 以 N_i 来表示来自总体 X 的简单随机样本 (样本量为 n) 中等于 i 的个数 $(i=1,2,3)$, 试求常数 a_1, a_2, a_3, 使得 $T=\sum_{i=1}^{3} a_i N_i$ 为 θ 的无偏估计量, 并求 T 的方差.

27. 一袋中有 N 个均匀硬币, 其中 θ 个是普通的硬币, 其余 $N-\theta$ 个两面都是正面. 现从袋中随机摸出一个把它连郑两次, 记下结果, 但是不看它属于哪种硬币, 又把它放回袋中, 如此重复 n 次. 如果掷出 $0,1,2$ 次正面的次数分别是 n_0, n_1, n_2 次 $(n_0+n_1+n_2=n)$, 试分别用矩估计法和极大似然法这两种方法估计袋中普通硬币数 θ.

28. * 设 $X_1, \ldots, X_n$ 是从总体 X 中抽出的简单随机样本, 已知 X 服从概率密度函数

$$f(x)=\begin{cases}\frac{1}{\sigma} e^{-\frac{x-\theta}{\sigma}}, & x>\theta \\ 0, & \text{其它}\end{cases}$$

其中 $\sigma>0$ 为一已知常数, 而 θ 是末知参数 (1) 试求 θ 的矩估计 $\hat{\theta}_1$ 和极大似然估计 $\hat{\theta}_2$. (2) 验证 $\hat{\theta}_1, \hat{\theta}_2$ 的无偏性. 如果不是无偏的话, 你是否可以将其修正得到 θ 的无偏估计 $\tilde{\theta}_1, \tilde{\theta}_2$? (3) 比较 $\tilde{\theta}_1$ 与 $\tilde{\theta}_2$ 何者为优 (即方差较小).

29. 设样本 $X_1, \ldots, X_n$ 抽自正态分布 $N\left(0, \sigma^2\right)$, 求 σ 的矩估计: (1) 利用 $\mathbb{E}\left|X_1\right|=\sqrt{\frac{2}{\pi}} \sigma$; (2) 利用 $\sigma=\sqrt{\operatorname{Var}\left(X_1\right)}$.

30. 设 $X_1, \ldots, X_n$ 是来自正态总体 $X \sim N\left(a, \sigma^2\right)$ 的简单随机样本, 求 $\mathbb{P}(X>1)$ 的矩估计量.

31. 若 $Y=\mathrm{e}^X$, 而 $X \sim N\left(a, \sigma^2\right)$, 则随机变量 Y 的分布称为对数正态分布. 设 $Y_1, \ldots, Y_n$ 是从总体 Y 中抽取的简单随机样本, 求 a 和 σ^2 的矩估计和极大似然估计.

32. 设总体 X 的密度为 $\frac{1}{2\sigma} \exp\{-|x-a|/\sigma\}$, 其中 $\sigma>0$ 和 a 为末知参数. 设 $X_1, \ldots, X_n$ 为抽自此总体的简单随机样本, 求 a 和 σ 的矩估计和极大似然估计.

33. 设总体 X 服从 Weibull 分布, 密度函数为

$$f(x, \lambda)=\begin{cases}\lambda \alpha \cdot x^{\alpha-1} e^{-\lambda x^\alpha} & x>0 \\ 0 & \text{其它}\end{cases} \quad \lambda>0, \alpha>0 .$$

设 $X_1, \ldots, X_n$ 为此总体中抽取的简单样本. 若 α 已知, 求 λ 的矩估计和极大似然估计.

34. 设 $X_1, \ldots, X_n$ 为从总体 X 中抽取的随机样本, X 服从均匀分布 $U\left(\theta-\frac{1}{2}, \theta+\frac{1}{2}\right)$, 求 θ 的极大似然估计.

35. 设 $X_1, \ldots, X_n$ 是来自均匀分布 $U(\theta, 2\theta)$ 的简单随机样本, 其中 $0<\theta<+\infty$, 求 θ 的极大似然估计, 它是 θ 的无偏估计吗? 如果不是, 试对它略作修改, 得到 θ 的一个无偏估计.

36. 设 $X_1, \ldots, X_m$ 和 $Y_1, \ldots, Y_n$ 分别来自总体 $N\left(\mu_1, \sigma^2\right)$ 和 $N\left(\mu_2, \sigma^2\right)$ 的两组独立样本, 求 μ_1、μ_2 和 σ^2 的极大似然估计.

37. 设 $X_1, \ldots, X_m$ 和 $Y_1, \ldots, Y_n$ 分别来自总体 $N\left(\mu, \sigma^2\right)$ 和 $N\left(\mu, 2\sigma^2\right)$ 的两组独立样本, 求 μ 和 σ^2 的极大似然估计.

38. 为了估计湖中有多少条鱼, 从中拱出 1000 条, 标上记号后放回湖中, 然后再捻出 150 条鱼, 发现其中有 10 条鱼有记号. 问湖中有多少条鱼, 才能使 150 条鱼中出现 10 条带记号的鱼的概率最大?

39. 一个罐子中装有黑白两种球, 今有放回地抽取一个大小为 n 的样本, 其中有 k 个白球, 则罐中白黑球比例的极大似然估计为

40. 设 $X_1,\ldots,X_n$ 为抽自指数分布

$$f(x,\theta)=\mathrm{e}^{-(x-\mu)}, x\geqslant\mu, -\infty<\mu<+\infty$$

的简单样本. (1) 试求 μ 的极大似然估计 $\hat{\mu}^*, \hat{\mu}^*$ 是 μ 的无偏估计吗? 如果不是, 试对它作修改, 以得到 μ 的无偏估计 $\hat{\mu}^{**}$. (2) 试求 μ 的矩估计 $\hat{\mu}$, 并证明它是 μ 的无偏估计. (3) 试问 $\hat{\mu}^{**}$ 和 $\hat{\mu}$ 哪一个有效?

41. 设 $X_1,\ldots,X_n$ 为来自正态总体 $N\left(a,\sigma^2\right)$ 的一个简单随机样本, $S^2=\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\bar{X}\right)^2$ 是样本方差, 证明 S^2 是 σ^2 的相合估计.

42. 设 $X_1,\ldots,X_n$ 为来自均匀分布总体 $U(0,\theta)$ 的一个简单随机样本, 求 θ 的极大似然估计 $\hat{\theta}_n$, 证明 $\hat{\theta}_n$ 是 θ 的相合估计.

43. 设 $X_1,\ldots,X_n$ 为来自指数分布总体 $\operatorname{Exp}(\lambda)$ 的一个简单随机样本, 已知 $\bar{X}$ 为 $1/\lambda$ 的无偏估计. 则 $1/\bar{X}$ 是否为 λ 的无偏估计?

44. 设 $X_1,\ldots,X_n$ 是来自均匀分布总体 $U\left(\theta_1,\theta_2\right)$ 的简单随机样本. 试求 θ_1 和 θ_2 的极大似然估计 $\hat{\theta}_1$ 和 $\hat{\theta}_2$;

45. 设 $X_1,\ldots,X_n$ 自下列总体中抽取的简单样本,

$$f(x,\theta)=\begin{cases}1, & \text{当}\theta-\frac{1}{2}\leqslant x\leqslant\theta+\frac{1}{2}\\ 0, & \text{其他}\end{cases}, -\infty<\theta<+\infty.$$

证明: 样本均值 $\bar{X}$ 及 $\max_{1\leqslant i\leqslant n}X_i+\frac{n+3}{2(n+1)}$ 都是 θ 的无偏估计, 问何者更有效?

46. 设 X_1,X_2,X_3 i.i.d. 服从均匀分布 $U(0,\theta)$, 试证 $\frac{4}{3}\max_{1\leqslant i\leqslant 3}X_i$ 及 $4\min_{1\leqslant i\leqslant 3}X_i$ 都是 θ 的无偏估计量, 哪个更有效?

47. 设从均值为 μ, 方差为 σ^2 的总体中, 分别抽取样本量为 n_1,n_2,n_3 的三组独立简单随机样本. $\bar{X}_1,\bar{X}_2$ 和 $\bar{X}_3$ 分别是三组样本的均值. 1) 证明: 对任意的常数 a,b,c 满足 $a+b+c=1$, 则 $a\bar{X}_1+b\bar{X}_2+c\bar{X}_3$ 是 μ 的无偏估计. 2) 求 a,b,c 使得 $\operatorname{Var}\left(a\bar{X}_1+b\bar{X}_2+c\bar{X}_3\right)$ 达到最小.

48. 设 $X_1,\ldots,X_n$ 为抽自均匀分布 $U(0,\theta)$ 的简单随机样本. 1) 选取适当的参数 a_n,b_n,c_n 使得, $\hat{\theta}_1=a_n\bar{X}, \hat{\theta}_2=b_n\min\left(X_1,\ldots,X_n\right)$ 和 $\hat{\theta}_3=c_n\max\left(X_1,\ldots,X_n\right)$ 都是 θ 的无偏估计. 2) 比较 $\hat{\theta}_i, i=1,2,3$ 哪个更有效.

49. 指数分布总体 $X\sim\operatorname{Exp}\left(\theta^{-1}\right), X_1,\ldots,X_n$ 为简单随机样本. 1) 选取适当的参数 a_n,b_n 使得, $\hat{\theta}_1=a_n\bar{X}$ 和 $\hat{\theta}_2=b_n\min\left(X_1,\ldots,X_n\right)$ 都是 θ 的无偏估计. 2) 比较 $\hat{\theta}_i, i=1,2$ 哪个更有效.

50. 设 $X_1,\ldots,X_n$ 为抽自密度函数为 $f(x,\theta)$ 的简单随机样本

$$f(x,\theta)=\begin{cases} e^{-(x-\theta)}, & x\geqslant\theta \\ 0, & x<\theta \end{cases}$$

1) 选取适当的参数 a_n,b_n 使得, $\hat\theta_1=\bar X+a_n$ 和 $\hat\theta_2=\min(X_1,\ldots,X_n)+b_n$ 都是 θ 的无偏估计. 2) 比较 $\hat\theta_i,i=1,2$ 哪个更有效.

51. 设从总体

X	0	1	2	3
P	$\theta/2$	θ	$3\theta/2$	$1-3\theta$

抽取的一个简单样本 $X_1,\cdots,X_{10}$ 的观察值为 $(0,3,1,1,0,2,0,0,3,0)$, 1) 求 θ 的矩估计量 $\hat\theta_M$ 和极大似然估计量 $\hat\theta_L$, 并求出估计值. 2) 上述估计量是否为无偏的? 若不是, 请作修正. 3) 比较修正后的两个估计量, 指出那个更有效.

52. 设 $X_1,\ldots,X_n$ 为抽自均匀分布 $U(0,\theta)$ 的简单随机样本. 证明下述断言: $\hat\theta=\max(X_1,\ldots,X_n)$ 为 θ 的相合估计但不是无偏估计.

53. 设 $\hat\theta_n$ 是参数 θ 的无偏估计, 其中 n 为样本量. 方差 $\operatorname{Var}\left(\hat\theta_n\right)>0$ 且 $\lim_{n\to\infty}\operatorname{Var}\left(\hat\theta_n\right)=0$. 证明: $\hat\theta_n^2$ 是 θ^2 的相合估计但不是无偏估计.

54. 设 $X_1,\ldots,X_n$ 为抽自均匀分布 $U(\theta-1,\theta+1)$ 的简单随机样本. θ 的矩估计为 $\hat\theta=\bar X$. 证明: 当 $n\to\infty$ 时, $\sqrt{3n}(\hat\theta-\theta)\to_d N(0,1)$.

55. 设 $X_1,\ldots,X_n$ 为抽自均匀分布 $U(0,\theta)$ 的简单随机样本. 对任给的 $\alpha\in(0,1)$, 求常数 c_n 使得 $[\max(X_1,\ldots,X_n),c_n\max(X_1,\ldots,X_n)]$ 为 θ 的 $1-\alpha$ 置信区间.

56. 设 $X_1,\ldots,X_n$ 为抽自指数分布 $\operatorname{Exp}(\lambda)$ 的简单随机样本. 对任给的 $\alpha\in(0,1)$, 1) 证明 $\sum_{i=1}^n 2\lambda X_i$ 服从 χ^2_{2n} 分布. 2) 求 λ 的 $1-\alpha$ 置信区间.

解: (1)

$$\sum_{j=1}^n X_j\sim\text{Erlang}(n,\lambda)\Rightarrow 2\lambda\sum_{j=1}^n X_j\sim\text{Erlang}\left(n,\frac{\lambda}{2\lambda}\right)=\Gamma(\frac{2n}{2},\frac{1}{2})$$

$$\therefore\sum_{i=1}^n 2\lambda X_i\sim\chi^2_{2n}$$

(2) 置信区间为 $\left[\dfrac{\chi^2_{2n}(1-\frac{\alpha}{2})}{\sum_{i=1}^n 2X_i},\dfrac{\chi^2_{2n}(\frac{\alpha}{2})}{\sum_{i=1}^n 2X_i}\right]$

57. 设 $X_1,\ldots,X_n$, 为抽自正态总体 $N(3,5^2)$ 的简单随机样本. 如果要求其样本均值位于区间 $(1,5)$ 的概率不小于 0.95, 问样本容量 n 至少应取多大?($U_0.025=1.96$)

58. 设一个测量仪器的误差 X 服从正态分布 $N(0,\sigma^2)$, 其标准差 $\sigma=10$, 若测量仪器无系统误差, 问至少重复测量几次, 才能以 99% 的把握, 保证平均误差的绝对值小于 2?

59. 设总体 $X\sim N(\mu,\sigma^2)$, σ^2 已知, 样本量为 n, 置信度为 $1-\alpha$, 样本均值为 $\bar X$. 若样本均值变大, 样本量和置信度不变, 则总体均值 μ 的置信区间长度 () (A) 变长 (B) 变短 (C) 保持不变 (D) 不能确定

60. 设总体 $X \sim N(\mu, \sigma^2)$, σ^2 未知, 样本量为 n, 置信度为 $1-\alpha$, 样本均值为 $\bar{X}$, 样本方差为 S^2. 若样本方差变大, 样本量和置信度不变, 则总体均值 μ 的置信区间长度 () (A) 变长 (B) 变短 (C) 保持不变 (D) 不能确定

61. 令 $X_1, \ldots, X_{10}$ 是从 $N(\mu, 4)$ 中抽取的随机样本, 假设样本均值 $\bar{X} = 48$, 求 μ 的 95% 置信区间.

62. 2016 年在某一商学院毕业的硕士生中随机抽取了 40 位, 他们的平均起薪是 8000 人民币, 样本标准差是 900 人民币, 求这一届毕业生平均起薪的 95% 置信区间.

63. 随机从一批钉子中抽取 9 枚, 测得其长度 (cm) 为:

$$2.15, 2.13, 2.10, 2.14, 2.15, 2.16, 2.12, 2.11, 2.13$$

假设钉子长度服从正态分布, 分别在下面两种情况下, 求出总体均值的 90% 置信区间: (1) $\sigma = 0.01$ (2) σ 未知.

64. 从某一小学的一年级学生中随机选了 9 名男生和 9 名女生, 测量他们的身高 (cm) , 假设身高服从正态分布.

男孩	126	131	120	125	116	126	117	130	117
女孩	122	123	124	125	125	118	120	120	114

(1) 求这所小学一年级学生平均身高的 95% 置信区间; (2) 求这所小学一年级男孩平均身高的 95% 置信区间; (3) 求这所小学一年级女孩平均身高的 95% 置信区间.

65. 假设 0.4, 2.5, 1.8, 0.7 是来自总体 X 的简单随机样本. 已知 $Y = \ln(X)$ 服从正态 $N(\mu, 1)$. (1) 求 X 的数学期望 $a = \mathbb{E}X$; (2) 求 μ 的 95% 置信区间; (3) 求 a 的 95% 置信区间.

66. 一个无线通讯公司, 考虑改变按分钟收费为包月不限时间. 公司预计新的策略会增加顾客每个月的通话时间. 为了验证这个结论, 公司随机抽取了 900 个包月客户，他们一个月平均使用时间是 220 分钟，样本标准差是 90 分钟. 同时也随机抽取了 800 个按流量收费的客户, 其一个月平均使用时间和标准差分别为 160 和 80 , 假设使用时间服从正态分布. (1) 求包月客户平均使用时间的 95% 置信区间; (2) 求按流量收费的客户平均使用时间的 95% 置信区间.

67. 一家企业更换了领导, 采取了新的经营策略. 随机选取公司 11 种商品, 更换经营策略前后一个季度的销量 (万元) 如表, 假设销量服从正态分布.

前	69.3	38.0	131.4	123.1	127.3	57.7	95.7	89.4
后	72.5	33.5	132.1	129.8	121.2	54.0	104.6	92.6

(1) 更换经营策略前平均销量的 95% 置信区间; (2) 更换经营策略后平均销量的 95% 置信区间; (3) 更换经营策略前后平均销量差异的 95% 置信区间.

68. 令 $1.7, 4, 2.3, 3.2$ 是从 $N(2.5, \sigma^2)$ 中抽取的随机样本, 求 σ^2 的 95% 置信区间.

69. 令 $X_1, \ldots, X_9$ 是从 $N(\mu, \sigma^2)$ 中抽取的随机样本, 假设样本均值 $\bar{X} = 48$, 样本方差为 $\sigma^2 = 64$. (1) 求 μ 的 95% 置信区间; (2) 求 σ^2 的 95% 置信区间.

70. 铅的密度服从正态分布, 如果观测 25 次, 算得样本均值为 3.2, 样本标准差为 0.031. (1) 求铅密度期望的 95% 置信区间; (2) 求铅密度标准差的 95% 置信区间.

71. 一批零件的长度 $X \sim N(\mu, \sigma^2)$, 从这批零件中随机地抽取 10 件, 测得长度值分别为 (单位: mm): $49.5, 50.4, 49.7, 51.1, 49.4, 49.7, 50.8, 49.9, 50.3, 50.0$. 在下列条件下求这批零件长度总体方差 σ^2 的 95% 置信区间. (1) $\mu = 50$ mm. (2) μ 末知.

72. 假设用机器包装精盐的重量服从正态分布. 现从生产线上随机地抽取 10 袋, 测得其重量为 (单位: 克) : $501.5, 500.7, 492.0, 504.7, 483.0, 512.8, 504.0, 490.3, 486.0, 520.0$. 试在下列条件下分别求总体方差的 95% 和 90% 置信区间. (1) $\mu = 500$ g. (2) μ 末知.

73. 随机抽取 16 发子弹做试验, 测得子弹速度的样本标准方差为 $S = 12$, 假设子弹速度服从正态分布 $N(\mu, \sigma^2)$, 分别求 σ 和 σ^2 的 95% 置信区间.

74. 试求第 66 题中, 1) 包月客户使用时间方差的 95% 置信区间, 2) 按流量收费的客户使用时间方差的 95% 置信区间.

75. 试求第 67 题中, (1) 更换经营策略前销量方差的 95% 置信区间; (2) 更换经营策略后销量方差的 95% 置信区间; (3) 更换经营策略前后销量差异方差的 95% 置信区间.

76. 假设到一商场的顾客有 p 的概率购买商品, 现随机抽取了 500 个顾客, 其中 15 个购买了商品. 求 p 的 95% 置信区间.

77. 假设某一生产线上商品的次品率为 p, 现随机抽取了 1000 个商品, 其中 5 个为次品. 求 p 的 95% 置信区间.

78. 假设湖中有 N 条鱼（N 很大）, 现钓出 r 条鱼, 做上标记后放回湖中. 一段时间后, 再钓出 s 条鱼 (设 s 远大于 r), 结果其中有 t 条标有记号 (s, t 已知）. (1) 若 r, N 末知, 求 r/N 的 $1-\alpha$ 置信区间; (2) 若只有 N 末知, 求 N 的 $1-\alpha$ 置信区间.

79. 某一地区高校中, 副教授以上职称有 1000 名. 高校的管理部门想了解具有高级职称教师中, 有基础研究课题的教师所占的比例. 于是随机抽取了 200 人组成一随机样本, 经调查有 80 人有基础研究课题. 求具有高级职称的教师中, 有基础研究课题的教师人数的 95% 置信区间.

80. 设 $X_1, \ldots, X_n$, 为抽自均匀分布 $U(\theta, 0)$ 的简单样本. 对任给的 $\alpha \in (0,1)$, 采用 θ 的极大似然估计, 构造 θ 的一个置信系数为 $1-\alpha$ 的置信下限和置信上限.

81. 设一农作的单位面积产量服从正态分布 $N(80, \sigma^2)$, 其标准差 $\sigma = 5$, 问至少需要几块试验田, 才能有 99% 的把握保证这些试验田的单位面积平均产量大于 75 ?

82. 试求 62 题中, 这一届毕业生平均起薪的 95% 置信下限.

83. 试分别在下面两种情况下, 求 17 题中, 这批钉子总体标准差的 95% 置信上限: (1) $\mu = 2.12$ (2) μ 未知.

84. 求 64 题中, (1) 一年级学生平均身高的 95% 置信下限; (2) 一年级男孩平均身高的 95% 置信下限; (3) 一年级女孩平均身高的 95% 置信下限.

85. 为研究某种轮胎的磨损情况, 随机的选取了 9 个轮胎, 每个轮胎行驶到磨坏为止. 记录所行使的路程 (km) 如下: $42350, 40297, 43176, 41010, 42657, 44210, 41879, 39678, 43520$. 假设这些数据服从正态分布, 求这个这种轮胎平均行使的路程的 95% 置信下限.

86. 为了了解一批灯泡的使用寿命, 共测试了 16 个灯泡的寿命, 得到寿命的平均为 $1600h$, 样本标准差为 $15h$. 假设寿命服从正态分布 $N(\mu, \sigma^2)$. (1) 求 μ 的 95% 置信下限; (2) 求 σ^2 的 95% 置信上限.

第八章 假设与检验

第 1 节 统计术语

假设检验研究内容：根据抽样后获得的样本来检查抽样前假设合理与否。

1 基本概念

I. 原（零/解消）假设 (Null Hypothesis) 和备择假设 (Alternative Hypothesis)：

$$H_0: \mu = 500 \leftrightarrow H_1: \mu = 490$$

II. 接受域/拒绝域：

III. 检验统计量：如 $Z = \dfrac{\sqrt{n}\left(\bar{X} - \mu_0\right)}{\sigma}$(相当于 H_0 代入枢轴变量)

IV. 假设种类：

(1) 简单假设：$H_0: \theta = \theta_0 \leftrightarrow H_1: \theta = \theta_1$; (2) 双侧假设：$H_0: \theta = \theta_0 \leftrightarrow H_1: \theta \neq \theta_0$;

(3) 右侧假设$H_0: \theta = \theta_0 \leftrightarrow H_1: \theta > \theta_0$; (4) 左侧假设$H_0: \theta = \theta_0 \leftrightarrow H_1: \theta < \theta_0$

V. 假设提法：

原则 1： 保护对象；

原则 2： 维持现状；

原则 3： 验证备择。

2 两类错误

第 I 型错误 实际上H_0 成立但是它被拒绝（弃真）

第 II 型错误 实际上H_0 不成立但是它没有被拒绝（取伪）

决策 \ 事实	H_0 成立	H_1 成立
不拒绝 H_0	不犯错	第 II 型错误
拒绝 H_0	第 I 型错误	不犯错

第 2 节 检验

1 显著性检验

在控制 I 类错误的基础上, 尽量少犯第 II 类错误

$$P_{H_0}(T < \tau) \leqslant \alpha(\text{显著性水平, 犯第一类错误概率})$$

$$\alpha = P\{\text{拒绝域的条件}|H_0\}$$

2 一般步骤

第 1 步 根据问题, 提出假设检验问题

$$H_0 : \theta \in \Theta_0 \leftrightarrow H_1 : \theta \in \Theta_1.$$

其中 H_0 为零假设或原假设, 而 H_1 为对立假设或备择假设.

第 2 步 根据参数的估计方法构造一个适当的检验统计量 $T = T(X_1, \cdots, X_n)$, 其中 $X_1, \cdots, X_n$ 为从总体中抽得的一个样本.

第 3 步 根据对立假设的形状构造一个检验的拒绝域 $W = \{T(X_1, \cdots, X_n) \in A\}$, 其中 A 为一个集合, 通常是一个区间. 比如拒绝域可取为 $\{T(X_1, \cdots, X_n) > \tau\}$, 则称 τ 为临界值.

第 4 步 对任意的 $\theta \in \Theta_0$, 犯第 I 类错误的概率 $P_\theta(T(X_1, \cdots, X_n) \in \mathrm{A})$ 小于或等于某个指定正的常数 α), 则称 α 为显著性水平.

第 5 步 结合 T 在 H_0 下的分布, 定出 A.

功效函数 Power

$$\beta_\Psi(\theta) = P(\text{在检验}\Psi\text{下拒绝}H_0)$$

其中检验 Ψ: 当 $\bar{X} \geqslant c$ 时接受 H_0 ，不然否定 H_0

$$P(\text{犯第 I 型错误}) = \begin{cases} \beta_\Psi(\theta), & \theta \in \Theta_0 \\ 0, & \theta \in \Theta_1 \end{cases}, P(\text{犯第 II 型错误}) = \begin{cases} 0, & \theta \in \Theta_0 \\ 1 - \beta_\Psi(\theta), & \theta \in \Theta_1 \end{cases}$$

3 假设检验步骤

第 1 步 求出未知参数 θ 的一个较优的点估计 $\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n)$, 如极大似然估计.

第 2 步 以 $\hat{\theta}$ 为基础, 寻找一个检验统计量

$$T = t(X_1, \cdots, X_n)$$

且使得当 $\theta = \theta_0$ 时, T 的分布已知 (如 $N(0,1), t_n, F_{m,n}$), 从而容易通过查表或计算得到这个分布的分位数, 用以作为检验的临界值.

第 3 步 以检验统计量 T 为基础, 根据对立假设 H_1 的实际意义, 寻找适当形状的拒绝域 (它是关于 T 的一个或两个不等式, 其中包含一个或两个临界值).

第 4 步 当零假设成立时, 犯第 I 类错误的概率小于或等于给定的显著性水平 α, 这给出一个关于临界值的方程, 解出临界值, 它 (们) 等于 T 的分位数, 这样即确定了检验的拒绝域.

第 5 步 如果给出样本观测值, 则可算出检验统计量的样本观测值, 如落在拒绝域中则可拒绝零假设, 否则不能.

4 常见假设检验的拒绝域

一样本正态总体的假设检验

检验对象	检验统计量	分布	拒绝域
$\mu\,(\sigma^2$ 已知)	$Z=\dfrac{\sqrt{n}\,(\bar{X}-\mu_0)}{\sigma}$	N(0,1)	$\begin{cases}\lvert Z\rvert>u_{\frac{\alpha}{2}}\\ Z>u_\alpha\\ Z<-u_\alpha\end{cases}$
$\mu(\sigma^2$ 未知)	$T=\dfrac{\sqrt{n}\,(\bar{X}-\mu_0)}{S}$	t_{n-1}	$\begin{cases}\lvert T\rvert>t_{n-1}(\frac{\alpha}{2})\\ T>t_{n-1}(\alpha)\\ T<-t_{n-1}(\alpha)\end{cases}$
$\sigma^2\,(\mu$已知)	$\chi^2=\dfrac{\sum_{i=1}^n (X_i-\mu_0)^2}{\sigma_0^2}$	χ_n^2	$\begin{cases}\chi^2>\chi_n^2(\frac{\alpha}{2})\text{ 或者}\chi^2<\chi_n^2(1-\frac{\alpha}{2})\\ \chi^2>\chi_n^2(\alpha)\\ \chi^2<\chi_n^2(1-\alpha)\end{cases}$
$\sigma^2\,(\mu$未知)	$\chi^2=\dfrac{\sum_{i=1}^n (X_i-\bar{X})^2}{\sigma_0^2}$	χ_{n-1}^2	$\begin{cases}\chi^2>\chi_{n-1}^2(\frac{\alpha}{2})\text{ 或者}\chi^2<\chi_{n-1}^2(1-\frac{\alpha}{2})\\ \chi^2>\chi_{n-1}^2(\alpha)\\ \chi^2<\chi_{n-1}^2(1-\alpha)\end{cases}$

注：有关均值的检验: 对立假设分别为 $\mu\neq\mu_0,\mu>\mu_0$ 和 $\mu<\mu_0$；

有关方差的检验: 对立假设分别为 $\sigma^2\neq\sigma_0^2,\sigma^2>\sigma_0^2$ 和 $\sigma^2<\sigma_0^2$. 注意使用: $S^2=\dfrac{1}{n-1}\left(\sum_{i=1}^n X_i^2-n\bar{X}^2\right)$

计算检验统计量分母: $\sum_{i=1}^n (x_i-\mu)^2=(n-1)S^2+n\mu^2$

两样本正态总体的假设检验

检验对象	检验统计量	分布	拒绝域
均值 (方差已知)	$Z=\dfrac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}$	N(0,1)	$\begin{cases}\lvert Z\rvert>u_{\frac{\alpha}{2}}\\ Z>u_\alpha\\ Z<-u_\alpha\end{cases}$
均值 (方差未知)	$T=\dfrac{\bar{X}-\bar{Y}}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}$	t_{m+n-2}	$\begin{cases}\lvert T\rvert>t_{m+n-2}(\frac{\alpha}{2})\\ T>t_{m+n-2}(\alpha)\\ T<-t_{m+n-2}(\alpha)\end{cases}$
方差 (均值已知)	$F=\dfrac{\frac{\sum_{i=1}^m (X_i-\mu_1)^2}{m}}{\frac{\sum_{j=1}^n (Y_j-\mu_2)^2}{n}}$	$F_{m,n}$	$\begin{cases}F>F_{m,n}(\frac{\alpha}{2})\text{ 或}F<\dfrac{1}{F_{n,m}(\frac{\alpha}{2})}\\ F>F_{m,n}(\alpha)\\ F<\dfrac{1}{F_{n,m}(\alpha)}\end{cases}$
方差 (均值未知)	$F=\dfrac{S_1^2}{S_2^2}$	$F_{m-1,n-1}$	$\begin{cases}F>F_{m-1,n-1}(\frac{\alpha}{2})\text{ 或}F<\dfrac{1}{F_{n-1,m-1}(\frac{\alpha}{2})}\\ F>F_{m-1,n-1}(\alpha)\\ F<\dfrac{1}{F_{n-1,m-1}(\alpha)}\end{cases}$

注：有关均值的检验: 对立假设分别为 $\mu_1\neq\mu_2,\mu_1>\mu_2$ 和 $\mu_1<\mu_2$.

有关方差的检验: 对立假设分别为 $\sigma_1^2\neq\sigma_2^2,\sigma_1^2>\sigma_2^2$ 和 $\sigma_1^2<\sigma_2^2$, 且假定方差相等

区分两样本/成对数据：$Y=X_1-X_2$（常是前后数据差）, 此处的观测不独立仍可视为一样本

方差齐性检验

$$
\begin{aligned}
& H_0=\sigma_1^2=\sigma_2^2 \longleftrightarrow H_1^2 \quad \sigma_1^2 \neq \sigma_2^2 \\
& \hat{\sigma_1}^2=S_1^2, \frac{(n-1) S_1^2}{\sigma_1^2} \sim \chi_{m-1}^2 \\
& \hat{\sigma_2}^2=S_2^2, \frac{(m-1) S_2^2}{\sigma_2^2} \sim \chi_{n-1}^2 \\
& \Rightarrow F_{n-1, m-1}\left(1-\frac{\alpha}{2}\right)<\frac{S_1^2}{S_2^2}<F_{n-1, m-1}\left(\frac{\alpha}{2}\right)
\end{aligned}
$$

两样本 t-检验 (相当于上面框图里的方差相同但未知)

$$
T=\frac{\bar{X}-\bar{Y}}{S_w \sqrt{\frac{1}{m}+\frac{1}{n}}} \sim t_{m+n-2}
$$

$$
\text{其中，} S_w^2=\frac{1}{n+m-2}\left(\sum_{i=1}^m\left(X_i-\bar{X}\right)^2+\sum_{j=1}^n\left(Y_j-\bar{Y}\right)^2\right)
$$

第 3 节 拟合优度检验

1 拟合优度检验方法/ Pearson-χ^2 检验方法:

基于样本得到 F 的估计 $\hat{F}_n$, 计算某种偏差 $D\left(\hat{F}_n, F\right)$, 例如 $\sup_{x \in R}\left|\hat{F}_n(x)-F(x)\right|$. 当 H_0 正确时, 由于 $\hat{F}_n$ 是 F 的相合估计, 偏差 $D\left(\hat{F}_n, F\right)$ 应该很小.

1. 离散总体，不含未知参数

类别	a_1	a_2	$\cdots$	a_k
理论频数	np_1	np_2	$\cdots$	np_k
观测频数	n_1	n_2	$\cdots$	n_k

$$
\text{检验统计量} T=\sum_{i=1}^k \frac{\left(n_i-n p_i\right)^2}{n p_i}=\sum \frac{(O-E)^2}{E} \tag{8.1}
$$

$$
\text{拒绝域: } T>\chi_\alpha^2(k-1) \tag{8.2}
$$

2. 离散总体，含未知参数 (通过极大似然估计 MLE 进行参数估计，减小 r 个自由度)

$$
\text{检验统计量} \chi^2=\sum_{i=1}^k \frac{\left(n_i-n \hat{p}_i\right)^2}{n \hat{p}_i} \xrightarrow{A} \chi_{k-r-1}^2 \tag{8.3}
$$

p 值

$$p\text{ - value } = P_{H_0}\{\text{假设零假设为真时观测到至少与实际观测样本相同极端的样本的概率}\}$$

$$\text{p 值的大小}\begin{cases}\text{大：样本的接受程度高}\\ \text{小：倾向于拒绝}H_0,\text{一般取 }0.05\end{cases}$$

2 列联表检验 Contingency Table

Group	Lung Cancer	No Lung Cancer	Total	Incidence of Lung Cancer
Smoker	84	2916	3000	$\frac{84}{3000}=0.028$
Non-Smoker	87	4913	5000	$\frac{87}{5000}=0.0174$
Total	171	7829	8000	

	1	$\cdots$	a	$Total$
1	n_{11}	$\cdots$	n_{a1}	$n_{\cdot 1}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$n_{\cdot j}$
b	n_{1b}	$\cdots$	n_{ab}	$n_{\cdot b}$
$Total$	$n_{1\cdot}$	$n_{i\cdot}$	$n_{a\cdot}$	n

$$\text{MLE：}L\left(p_{1\cdot},\cdots,p_{a\cdot},p_{\cdot 1},\cdots,p_{\cdot b}\right)=\prod_{j=1}^{b}\prod_{i=1}^{a}p_{ij}^{n_{ij}}=\prod_{j=1}^{b}\prod_{i=1}^{a}p_{i\cdot}^{n_{i\cdot}}\prod_{j=1}^{b}p_{\cdot j}^{n_{\cdot j}}$$

$$\text{Lagrange 乘数法：}Q=L\left(p_{i\cdot},p_{\cdot j};i=\overline{1,a},j=\overline{1,b}\right)+\lambda\left(\sum_{i=1}^{a}p_{i\cdot}-1\right)+\mu\left(\sum_{j=1}^{b}p_{\cdot j}-1\right)$$

$$\therefore\begin{cases}\frac{\partial Q}{\partial p_{i\cdot}}=0\\ \frac{\partial Q}{\partial p_{\cdot j}}=0\end{cases}\Rightarrow\hat{p_{i\cdot}}=\frac{n_{i\cdot}}{n},\hat{p_{\cdot j}}=\frac{n_{\cdot j}}{n}$$

$$\chi^2=\frac{\sum(O-E)^2}{E}=\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{\left(n_{ij}-n\frac{n_{i\cdot}}{n}\frac{n_{\cdot j}}{n}\right)^2}{n\cdot\left(\frac{n_{i\cdot}}{n}\frac{n_{\cdot j}}{n}\right)}\sim\chi^2_{ab-(a-1)-(b-1)}$$

$$\chi^2=\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{\left(nn_{ij}-n_{i}.n_{\cdot j}\right)^2}{n\cdot n_{i\cdot}\cdot n_{\cdot j}}\sim\chi^2_{(a-1)(b-1)}\tag{8.4}$$

卡方列联表检验计算器：https://www.socscistatistics.com/tests/chisquare2/default2.aspx

第 4 节　思考题

1. 假设 $X_1,\ldots,X_{16}$ 服从正态分布 $N(\mu,0.16)$. 检验问题
$$H_0:\mu=0.5\leftrightarrow H_1:\mu>0.5,$$
显著水平为 0.05. (1) 检验的拒绝域是什么? (2) $\mu=0.65$ 时犯第二类错误的概率是多少?

2. 产品检验时, 原假设 H_0 : 产品合格. 为了减少次品混入正品的可能性, 在 n 固定的条件下, 显著性水平 α 应取大些还是小些, 为什么?

3. 假设 $X_1,\ldots,X_{16}$ 服从正态分布 $N(\mu,1)$. 检验问题 $H_0:\mu=0\leftrightarrow H_1:\mu=1$, 样本均值为 $\bar{X}, u_\alpha$ 为标准正态的上 α 分位数, 现有 4 个拒绝域: $V_1=\{4|\bar{X}|\geqslant u_{0.05}\}, V_2=\{4|\bar{X}|\leqslant u_{0.45}\}, V_3=\{4\bar{X}\geqslant u_{0.10}\}, V_4=\{4\bar{X}\leqslant -u_{0.10}\}$. (1) 这 4 个拒绝域中的第一类错误概率分别是多少? (2) 比较哪个拒绝域第二类错误最小.

4. 假设随机变量 X 的概率密度函数为 $f(x)=(1+\theta)x^\theta, 0<x<1$, 现考虑假设检验问题:
$$H_0:\theta=5\leftrightarrow H_1:\theta=3.$$
拒绝域为 $\{X>1/2\}$. 试求该检验问题的 I 类和 II 类错误, 以及 $\theta=2$ 时的功效函数值?

5. 设样本 $X_1,\cdots,X_n$ 抽自参数为 λ 的泊松分布总体, 对检验问题
$$H_0:\quad \lambda=\frac{1}{2}\quad\longleftrightarrow\quad H_1:\quad \lambda\neq\frac{1}{2}$$
取检验的拒绝域为 $\left\{(X_1,\cdots,X_n):\sum_{i=1}^{10}X_i\leqslant 1 \text{ 或 } \geqslant 12\right\}$ (1) 求此检验在 $\lambda=0.25,0.5,1$ 处的功效函数值, 并求出该检验的水平. (2) 求犯第一类错误的概率及在 $\lambda=0.25,0.75$ 处犯第二类错误的概率.

6. 设总体为均匀分布 $U(0,\theta), X_1,\cdots,X_n$ 是一组样本. 考虑检验问题
$$H_0:\quad \theta\geqslant 3\quad\longleftrightarrow\quad H_1:\quad \theta<3,$$
拒绝域取为 $W=\left\{X_{(n)}=\max\left\{X_1,\cdots,X_n\right\}\leqslant 2.5\right\}$, (1) 求此检验的功效函数和显著水平. (2) 为使显著水平达到 0.05, 样本量 n 至少应取多大?

7. 设 $0.644, 0.672, 0.070, 0.176, 0.314, 0.295, 0.331, 0.001, 0.490$, 为抽自均匀分布 $U(0,\theta)$ 的简单随机样本. 在 0.05 显著水平下检验 $H_0:\theta=1\leftrightarrow H_1:\theta<1$.

8. 设 $0.455, 0.840, 0.653, 0.443, 0.026, 0.523, 0.284, 0.270, 0.720$, 为抽自均匀分布 $U(\theta,1)$ 的简单随机样本. 在 0.05 显著水平下检验 $H_0:\theta=0\leftrightarrow H_1:\theta>0$.

9. 设 $0.213, 0.626, 1.091, 0.591, 1.083, 0.475, 0.315, 0.328, 4.151$, 为抽自指数分布 $\mathrm{Exp}(\lambda)$ 的简单随机样本. 在 0.05 显著水平下检验 $H_0:\lambda=1\leftrightarrow H_1:\lambda>1$.

10. 设 $X_1,\ldots,X_n$ 为抽自正态总体 $N\left(\mu,5^2\right)$ 的样本. 如果检验问题为 $H_0:\mu=3\leftrightarrow H_1:\mu\neq 3$, 其 0.05 显著水平下拒绝域为 $\{|\bar{X}-3|>2\}$, 问样本容量 n 至少应取多大?

11. 某种产品以往的废品率为 6%, 采用某种技术革新后, 随机抽取 200 个产品进行检查, 其中 8 个废品, 即样本废品率为 4% 相比以往有所降低, 取显著水平为 $\alpha = 0.05$. (1) 此问题的原假设和备择假设分别是什么? (2) 犯第一类的错误的概率是多少? (3) 样本的结果是否能支持备择假设.

12. 假设某一生产线上商品的次品率为 p, 现随机抽取了 1000 个商品, 其中 5 个为次品. 次品率低于 0.52% 即为合格, 在 0.05 显著水平下判断这批产品是否合格.

13. 令 $X_1, \ldots, X_{10}$ 是从 $N(\mu, 16)$ 中抽取的随机样本, 假设样本均值 $\bar{X} = 4.8$. 在 5% 显著水平下, 检验: (1) $H_0 : \mu = 7 \leftrightarrow H_1 : \mu \neq 7$; (2) $H_0 : \mu \geqslant 7 \leftrightarrow H_1 : \mu < 7$. (2) $H_0 : \mu \leqslant 2 \leftrightarrow H_1 : \mu > 2$;

14. 试检验第七章 62 题这一届毕业生平均起薪大于 7700 人民币, 给出 p 值.

15. 在 5% 显著水平下, 试检验第七章 64 题, (1) 这所小学一年级学生平均身高大于 120 cm. (2) 这所小学一年级男孩平均身高大于 120 cm. (3) 这所小学一年级女孩平均身高大于 120 cm.

16. 在 5% 显著水平下, 试检验第七章 66 题, (1) 包月客户 — 个月平均通话时间大于 190 分钟. (2) 按流量收费的客户一个月平均通话时间小于 190 分钟.

17. 在 5% 显著水平下, 试检验第七章 67 更换经营策略前后平均销量是否有显著性的增加.

18. 假设到一商场的顾客有 p 的概率购买商品, 先随机抽取了 500 个顾客, 其中 15 个购买了商品. 5% 显著水平下, 检验 $H_0 : p = 0.02 \leftrightarrow H_1 : p \neq 0.02$.

19. 随机抽取某班 25 名学生的数学考试成绩, 得平均分数为 82 分, 样本标准差为 8 , 已知全年级的数学成绩服从正态分布且平均分数为 87 分, 试问在显著性水平 $\alpha = 0.05$ 下, 能否认为该班的数学平均成绩为 87 分?

20. 2015 年全国人口调查中男女性别比例为 $\mu = 105.02$ (女 =100), 为检验这一比例, 随机抽取了 8 个省份, 其男女性别比例如下表, 假设各省的性别比服从正态分布,

北京	内蒙古	辽宁	安徽	河南	海南	重庆	宁夏
109.45	104.32	100.45	104.90	103.99	110.47	100.60	106.16

在 5% 显著水平下, 检验 $H_0 : \mu = 105.02 \leftrightarrow H_1 : \mu \neq 105.02$.

21. 假设会考成绩服从正态分别, 在某市的一次高中数学会考中, 全市的平均分为 80 分, 标准差为 9 分. 某一高中 180 名学生平均分为 82 分, 该校历年来会考成绩高于全市成绩, 问此次会考该校是否仍然显著高于全市平均成绩 $(\alpha = 0.05)$?

22. 2015 年（年末）全国总人口 137462 万人, 出生率 12.07%; 辽宁省人口 4382 万人, 出生率 7%; 吉林省人口 2753 万人, 出生率 6.17%; 黑龙江人口 3812 万人, 出生率 6.00%. 东三省人口出生率是否显著小于全国人口出生率? 给出 p 值.

23. 为了检验一批灯泡是否合格, 即灯泡寿命的均值 $\mu > 1550h$, 共测试了 16 个灯泡的寿命, 得到寿命的平均为 1600 h, 样本标准差为 $S = 15h$. 假设寿命服从正态分布, 在 5% 显著水平下, 检验这批灯泡是否合格.

24. 用传统工艺加工的某种水果罐头中每瓶维生素 C 的含量平均为 19 毫克, 现采用一种新的加工工艺, 试图减少在加工过程中对维生素 C 的破坏, 抽查了 16 瓶罐头, 测得维生素 C 的含量 (单位: 亭克) 为: $23, 20.5, 21, 20, 22.5, 19, 20, 23, 20.5, 18.8, 20, 19.5, 22, 18, 23, 22$. 已知水果罐头中维生素 C 的含量服从正态分布. 在方差末知的情况下, 问新工艺下维生素的含量是否比旧工艺有所提高 $(\alpha = 0.01)$?

25. 某机器制造出来的肥皀厚度为 5 cm. 今欲了解机器性能是否良好, 随机抽取 10 块肥皀为样本, 测得平均厚度为 5.3 cm, 标准差为 0.3 cm, 试分别在 $0.05, 0.01$ 的显著性水平下检验机器是否工作良好.

26. 令 $1.7, 4, 2.3, 3.2$ 是从 $N(2.5, \sigma^2)$ 中抽取的随机样本, 5% 显著水平下, 检验 $H_0 : \sigma^2 = 1 \leftrightarrow H_1 : \sigma^2 \neq 1$.

27. 令 $X_1, \ldots, X_9$ 是从 $N(\mu, \sigma^2)$ 中抽取的随机样本, 假设样本均值 $\bar{X} = 48$, 样本方差为 $\sigma^2 = 64$, 5% 显著水平下, 检验 (1) $H_0 : \mu \leqslant 40 \leftrightarrow H_1 : \mu > 40$; (2) $H_0 : \sigma^2 \geqslant 70 \leftrightarrow H_1 : \sigma^2 < 70$.

28. 铅的密度服从正态分布, 如果观测 25 次, 算得样本均值为 3.2, 样本标准差为 0.031, 5% 显著水平下, 检验 $H_0 : \sigma = 0.02 \leftrightarrow H_1 : \sigma \neq 0.02$.

29. 某车间生产铜丝, 生产一向比较稳定. 今从产品中随机抽取 10 根检查其折断力, 得数据如下 (单位: kg): 288.8, 294.7, 300.2, 286.6, 290.3, 280.1, 296.4, 295.4, 290.2, 289.2. 假设铜丝的折断力服从正态分布, 问是否可以相信该车间生产的铜丝的折断力的方差是 $16 (\alpha = 0.05)$?

30. 随机从一批钉子中抽取 9 枚, 测得其长度 (cm) 为: 2.15, 2.13, 2.10, 2.14, 2.15, 2.16, 2.12, 2.11, 2.13, 假设钉子长度服从正态分布, 分别在下面两种情况, 5% 显著水平下, 检验 $H_0 : \sigma \leqslant 0.01 \leftrightarrow H_1 : \sigma > 0.01$: (1) $\mu = 2.12$ (2) μ 本知.

31. 用传统工艺加工的某种水果罐头中每瓶维生素 C 的含量平均为 19 毫克, 现采用一种新的加工工艺, 试图减少在加工过程中对维生素 C 的破坏, 抽查了 16 瓶镘头, 测得维生素 C 的含量 (单位: 亭克) 为: $23, 20.5, 21, 20, 22.5, 19, 20, 23, 20.5, 18.8, 20, 19.5, 22, 18, 23, 22$. 已知水果罐头中维生素 C 的含量服从正态分布. 在方差末知的情况下, 问新工艺下维生素的含量是否比旧工艺有所提高 $(\alpha = 0.01)$?

32. 某机器制造出来的肥皀厚度为 5 cm. 今欲了解机器性能是否良好, 随机抽取 10 块肥皀为样本, 测得平均厚度为 5.3 cm, 标准差为 0.3 cm, 试分别在 $0.05, 0.01$ 的显著性水平下检验机器是否工作良好.

33. 令 $1.7, 4, 2.3, 3.2$ 是从 $N(2.5, \sigma^2)$ 中抽取的随机样本, 5% 显著水平下, 检验 $H_0 : \sigma^2 = 1 \leftrightarrow H_1 : \sigma^2 \neq 1$.

34. 令 $X_1, \ldots, X_9$ 是从 $N(\mu, \sigma^2)$ 中抽取的随机样本, 假设样本均值 $\bar{X} = 48$, 样本方差为 $\sigma^2 = 64$, 5% 显著水平下, 检验 (1) $H_0 : \mu \leqslant 40 \leftrightarrow H_1 : \mu > 40$; (2) $H_0 : \sigma^2 \geqslant 70 \leftrightarrow H_1 : \sigma^2 < 70$.

35. 铅的密度服从正态分布, 如果观测 25 次, 算得样本均值为 3.2, 样本标准差为 0.031, 5% 显著水平下, 检验 $H_0 : \sigma = 0.02 \leftrightarrow H_1 : \sigma \neq 0.02$.

36. 某车间生产铜丝, 生产一向比较稳定. 今从产品中随机抽取 10 根检查其折断力, 得数据如下 (单位: kg): 288.8, 294.7, 300.2, 286.6, 290.3, 280.1, 296.4, 295.4, 290.2, 289.2. 假设铜丝的折断力服从正态分布, 问是否可以相信该车间生产的铜丝的折断力的方差是 $16(\alpha = 0.05)$?

37. 随机从一批钉子中抽取 9 枚, 测得其长度 (cm) 为: $2.15, 2.13, 2.10, 2.14, 2.15, 2.16, 2.12, 2.11, 2.13$, 假设钉子长度服从正态分布, 分别在下面两种情况, 5% 显著水平下, 检验 $H_0 : \sigma \leqslant 0.01 \leftrightarrow H_1 : \sigma > 0.01$: (1) $\mu = 2.12$ (2) μ 本知.

38. 假设从两个独立的正态总体中各得到容量为 10 的样本, 若两个总体的方差相同, 则使用两样本 t 检验时 t 分布的自由度为 (A) 9 (B) 10 (C) 18 (D) 20

39. 假设总体 X 为取值 $0, 1, 2$ 的离散型随机变量, 且取各值的概率分别为 $\mathbb{P}(X = 0) = p, \mathbb{P}(X = 1) = 2p, \mathbb{P}(X = 3) = 1 - 3p$, 其中 p 为参数. 则当使用拟合优度检验时, 检验统计量的渐近卡方分布的自由度为 (A) 3 (B) 2 (C) 1 (D) 0

40. 设 $X_1, \ldots, X_n$ i.i.d $\sim N(\mu, \sigma^2)$, 考虑假设检验问题 $H_0 : \mu = 0 \leftrightarrow H_1 : \mu = 1$, 问下述哪些情形不可能出现: (A) 当水平 $\alpha = 0.01$ 时接受 H_0, 而当水平 $\alpha = 0.05$ 时拒绝 H_0 (B) 当水平 $\alpha = 0.01$ 时拒绝 H_0, 而当水平 $\alpha = 0.05$ 时接受 H_0 (C) 当水平 $\alpha = 0.01$ 和 $\alpha = 0.05$ 时都接受 H_0 (D) 当水平 $\alpha = 0.01$ 和 $\alpha = 0.05$ 时都拒绝 H_0

41. 为了考察 A, B 两种制鞋材料的耐磨性, 用它们制作了 10 双鞋, 其中每双鞋的两只鞋分别用 A, B 两种材料制作 (左、右两只鞋随机地采用 A 或 B). 10 个男孩试穿这 10 双鞋之后的磨损情况如下表所示 (数字代表磨损程度), 问是否可以认为这两种材料的耐磨性无显著差异 $(\alpha = 0.05)$?

	1	2	3	4	5	6	7	8	9	10
A	13.2	8.2	10.9	14.3	10.7	6.6	9.5	10.8	8.8	13.3
B	14.0	8.8	11.2	14.2	11.8	6.4	9.8	11.3	9.3	13.6

42. 一个以械肥为主要目的的健美俱乐部声称, 参加其训练班至少可以使肥胖者平均减少体重 8 kg 以上. 为检验该宣传是否可信, 调查人员随机调查了 9 名参加者, 得到他们训练前后的体重数据如下 (单位: kg):

训练前	104.5	94.0	104.7	96.4	91.6	90.9	92.0	99.9	109.7
训练后	94.2	86.6	97.5	91.7	82.6	83.8	81.3	92.2	101.0

现假设训练前后人的体重均服从正态分布. 问在 0.05 的显著水平下, 是否可以认为该俱乐部的宣传是可信的?

43. 在第五题中, 如果训练前后的数据是对两组人测的, 并假设训练前后人的体重服从方差相同的正态分布, 问在 0.05 的显著水平下, 是否可以认为该俱乐部的宣传是可信的?

44. 装配一个部件可以采用不同的方法, 现在关心的是哪一种方法的效率更高. 现在从两种不同的装配方法中各抽取 12 种产品, 记录各自的装配时间 (单位: 分钟) 如下:

甲方法	30	34	34	35	34	28	34	26	31	31	38	26
乙方法	26	32	22	26	31	28	30	22	31	26	32	29

假设两总体为正态总体, 且方差相等, 问这两种方法的装配时间有无显著不同 ($\alpha = 0.05$) ?

45. 修完一门大学财务会计课程的 352 名学生的总平均分的标准差为 0.940. 退选这门课程的 73 名学生的总平均分的标准差为 0.797. 这些数据是否表明修完与退选财务会计课程的学生在总平均分的方差上有差异? 取显著性水平为 $\alpha = 0.05$. $(F_{351,72}(0.025) = 1.466)$ 39. 两种新的装配方法经过检验后装配时间的方差报告如下. 取 $\alpha = 0.10$, 检验两个总体方差是否相等.

方法	样本容量	样本方差
A	$n_1 = 31$	$s_1^2 = 25$
B	$n_2 = 25$	$s_2^2 = 12$

46. 为了解甲、乙两企业职工工资水平, 分别从两企业各随机抽取若干名职工调查, 得如下数据 (单位: 元):

甲公司	3750	5300	3750	9100	5700	5250	5000	
乙公司	5000	9500	4500	9000	6000	8500	9750	6000

设两企业职工工资分别服从正态分布, 而总体独立且均值方差均未知. 试根据以上数据判断: 两企业职工工资的方差是否相等? 甲企业职工平均工资是否低于乙企业职工平均工资 $(\alpha = 0.05)$?

47. 某市场研究机构用 8 个人组成的样本来给某特定商品的潜在购买力打分. 样本中每个人都分别在看过该产品新的电视广告之前与之后打分. 潜在购买力的分值为 $0 \sim 10$ 分, 分值越高表示潜在购丰力越高. 零假设为看广告后平均得分小于或等于看广告前的平均得分, 拒绝该假设就表明广告有宣传效果, 提高了平均潜在购买力得分. 给定置信水平 $\alpha = 0.05$, 用下列数据检验该假设, 并对广告给予评价.

个人	购买力得分 之后	购买力得分 之前	个人	购买力得分 之后	购买力得分 之前
1	6	5	5	3	5
2	6	4	6	9	8
3	7	7	7	7	5
4	4	3	8	6	6

48. 现有两台天平, 为比较它们的精度, 将一物体分别在两台天平上各称量 9 次, 得数据如下 (单位: g):

甲天平	19.96	19.97	20.06	19.96	20.06	20.01	20.01	19.98	19.98
乙天平	19.90	19.89	20.18	19.91	20.03	20.00	20.00	20.02	19.91

设两台天平的称量结果分别服从 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$, 在显著性水平 0.05 下检验下列假设

$$H_0 : \sigma_1^2 \geqslant \sigma_2^2 \longleftrightarrow H_1 : \sigma_1^2 < \sigma_2^2$$

49. 设有 A 种药随机地给 8 个人服用, 经过一个固定的时间后, 测量病人身体细胞内药的浓度, 其结果为

1.40　1.42　1.41　1.62　1.55　1.81　1.60　1.52

又有 B 种药给其他 6 个人服用, 在同样固定时间后, 测量病人身体细胞内药的浓度, 得数据

1.76　1.41　1.87　1.49　1.67　1.81

设两种药在病人身体细胞内的浓度都服从正态分布, 试问 A 种药在病人身体细胞内的浓度的方差是否小于与 B 种药在病人身体细胞内的浓度的方差 $(\alpha = 0.1)$?

50. 从甲、乙两处煤矿各随机抽取矿石 5 个和 4 个, 分析其含灰率 (%) 得到如下结果假设各煤矿含灰率都服从正态分布且方差相等, 问甲乙两矿的含灰率有无显著差异 $(\alpha = 0.05)$?

甲矿	24.3	20.8	23.7	21.3	17.4
乙矿	18.2	16.9	20.2	16.7	

51. 设总体 $X \sim N(\mu_1, 0.04), Y \sim N(\mu_2, 0.09)$. 现从 X 中抽取的样本观察值为 $2.10, 2.35, 2.39, 2.41, 2.44, 2.56$. 从 Y 中抽取的样本观察值为 $2.03, 2.28, 2.58$, 2.71. 试检验 μ_1 和 μ_2 是否有显著差异 $(\alpha = 0.05)$? 46. 现有两批电子器件, 从中随机抽取若干进行检验, 测得样本的电阻如下 (单位: Ω

A 批	0.140	0.138	0.143	0.142	0.144	0.137
B 批	0.135	0.140	0.142	0.136	0.138	0.140

假设这两批电子器件的电阻均服从正态分布, 试在显著水平 0.05 下比较一下这两批电子器件的电阻有无差异.

52. 通常航空旅客选择出发机场依赖于飞行成本, 某调研机构搜集了从两个机场出发到 8 个城市的样本数据 (以美元计), 以确定从二者中哪个出发成本更高. 某研究者指出显然从机场一出发比从机场二出发成本更高. 以 $\alpha = 0.05$ 作为显著性水平, 用样本数据来验证它们是否支持学者的观点.

目的地	机场一	机场二
A	319	142
B	192	213
C	503	317
D	256	387
E	339	317
F	379	167
G	268	273
H	288	274

53. 2004 年, 有线电视和收音机超过广播电视、录音音乐和日报, 成为使用量最大的两个娱乐媒体. 研究者用 15 名研究志原者组成一个样本, 搜集他们每周看有线电视的时间和听收音机的时间数据 (单位: 小时).

志愿者	有线电视	收音机	志愿者	有线电视	收音机
1	22	25	9	21	21
2	8	10	10	23	23
3	25	29	11	14	15
4	22	19	12	14	18
5	12	13	13	14	17
6	26	28	14	16	15
7	22	23	15	24	23
8	19	21			

(1) 在显著性水平 0.05 下, 检验有线电视和收音机使用量的总体均值之间是否有差异? 实际 p 值是多少? (2) 每周花在看有线电视上的时间样本均值是多少? 每周花在听收音机上的时间的样本均值是多少? 哪个媒体具有较大的使用量?

54. 某报道提供了 2012 年几家著名公司的每股收益的数据. 在 2012 年之前, 财务分析家就预测了这些公司 2012 年的每股收益. 利用下面数据评论实际的和预测的每股收益的差异.

公司	实际每股收益	预计每股收益
A	1.29	0.38
B	2.01	2.31
C	2.59	3.43
D	1.60	1.78
E	1.84	2.18
F	2.72	2.19
G	1.51	1.71
H	2.28	2.18
I	0.77	1.55
J	1.81	1.74

(1) 若 $\alpha = 0.05$, 检验实际的和预测的每股平均收益之间是否存在差异? p 值是多少? 你的结论是什么? (2) 两均值之差的点估计量是多少? 分析家是低估了还是高估了每股的收益? (3) 对于 95% 的置信度, (2) 中估计的边际误差是多少?

55. 某厂家生产豪华型与普通型两种家用自动磨沙器. 由零售点抽样得到的销售价格如下:

零售点	价格/美元 豪华型	零售点 普通型	零售点	价格/美元 豪华型	普通型
1	39	27	5	40	30
2	39	28	6	39	34
3	45	35	7	35	29
4	38	30			

(1) 厂家建议的两种型号的零售价有 10 美元的差价. 在显著性水平 $\alpha = 0.05$ 下, 检验两种型号的平均差价是否为 10 美元. (2) 两种型号平均差价的 95% 的置信区间是多少?

56. 独立地用两种工艺生产同一种产品, 现随机抽取若干产品测得产品的某种性能指标如下:

工艺 A	0.47	1.02	0.33	0.70	0.94	0.85	0.39	0.52	0.47
工艺 B	0.41	1.00	0.46	0.61	0.84	0.87	0.36	0.52	0.51

假设这两种工艺生产的产品指标均服从正态分布, 由此结果是否能判断这两种工艺对产品该性能指标的影响有无显著差异 $(\alpha = 0.05)$?

57. 为比较新旧两种肥料对小麦产量的影响, 研究者选择了面积相等、土壤等条件相同的 12 块地, 分别在 6 块地上施用新旧两种肥料. 对于旧肥料, 得到的产量数据是: $17, 14, 18, 13, 19$ 和 15 ; 而新肥料的产量数据为: $16, 19, 20, 22, 18$ 和 19 . 假设两种肥料的产量分别服从正态分布, 且总体独立, 均值和方差未知. 试根据以上数据判断: (1) 两种肥料产量的方差是否相等 $(\alpha = 0.05)$? (2) 新肥料获得的平均产量是否显著地高于旧肥料 $(\alpha = 0.05)$?

58. 某种大量生产的袋装食品, 按规定不得少于 250 g. 今从一批该食品中任意抽取 50 袋, 发现有 6 袋低于 250 g. 若规定不符合标准的比例超过 5% 就不得出场, 问该批产品是否可以出场 $(\alpha = 0.05)$?

59. 2000 年的全国人口普查表明某城市的 65 岁以上老年人所占的比例为 13.55%, 现在为了调查人口的变动情况, 随机抽取 400 名居民, 发现其中有 57 人年龄在 65 岁以上. 试问该市现在老

年人所占的比例较 2000 年普查时是否有变化 ($\alpha = 0.05$) ?

60. 为检验吸烟与慢性气管炎有无关系, 随机调查了 339 人, 其中 205 名吸烟者中有 43 人患慢性气管炎, 在 134 名不吸烟者中有 13 人患慢性气管炎. 问在显著水平 0.05 下数据是否支持"吸烟者中患慢性气管炎的比例较高"这个结论?

61. 为确定某种肥料的效果, 取 1000 株植物做试验. 在没有施肥的 100 株植物中, 有 53 株长势良好, 在已施肥的 900 株中, 则有 783 株长势良好. 问施肥的效果是否显著 ($\alpha = 0.01$) ?

62. 某汽车协会的一项研究调查是男性还是女性更有可能停车问路. 研究假设: "如果你和你的配偶正在行驶并且迷路, 你会停车问路吗?" 由该协会的典型样本数据得到在 811 名女性中有 300 人说她们会停车问路, 同时在 750 名男性中有 255 人说他们会停车问路. (1) 研究的假设是女性更有可能说她们会停车问路. 建立这个研究的原假设和备择假设. (2) 表示他们会停车问路的女性的百分数是多少? (3) 表示他们会停车问路的男性的百分数是多少? (4) 在 $\alpha = 0.05$ 下检验假设, p 值是多少? 你期待从这个研究中得到什么结论?

63. 某媒体用抽样调查来确定人们如何利用他们的空闲时间. 男性和女性都选择看电视是他们最普遍的活动. 选择看电视作为作为他们最普遍业余活动的男性比率和女性比率可以从如下样本数据中估计出来.

性别	样本容量	看电视
男性	800	248
女性	600	156

(1) 陈述一个假设, 该假设可用于检验选择看电视作为他们最普遍业余活动男佅比率和女性比率之间的差异. (2) 选择看电视作为他们作为他们最普遍业余活动男性样本比率是多少? 相应的女性样本比率是多少? (3) 进行假设检验并计算 p 值. 在显著性水平 $\alpha = 0.05$ 下, 你的结论如何? (4) 总体比率之差的 95% 置信区间估计是多少?

64. 袋中装有 8 个球, 其中红球数末知. 在其中任取 3 个, 记录红球的个数 X, 然后放回. 再任取 3 个, 记录红球的个数, 然后放回. 如此重复进行 112 次, 得到结果如下:

x	0	1	2	3
次数	1	31	55	25

试在 $\alpha = 0.05$ 水平下检验假设 H_0 : 红球的个数为 5 .

65. 有甲、乙、丙三个工厂生产同一种产品. 产品分 $1, 2, 3$ 三个等级 (分别代表高、中、低). 为考察各工厂产品质量是否一致, 从这三个工厂中分别随机抽出产品若干件, 每件鉴定其质量等级, 结果如下:

等级/工厂	甲	乙	丙
1	58	38	30
2	40	44	35
3	11	18	26

试在显著性水平 0.05 下检验假设: 各工厂产品质量一致. 若不一致, 试问哪个厂产品质量较优? 哪个厂的产品质量较劣? 请说明理由.

66. 为了研究蜗牛的种类是否与其生活的珊瑚礁种类有关, 选取了 3 种珊瑚礁作为检验样本, 记为 I, II, III, 记录下 A 和 B 两种蜗牛分别在 3 种珊瑚礁中生存的数目, 得到如下数据. 试问 A 和 B 两种蜗牛的分布是否在 3 种珊瑚礁中都是一样的 ($\alpha = 0.05$) ?

	I	II	III	合计
A	6	8	14	28
B	7	21	5	33
合计	13	29	19	61

67. 某工厂为了了解白班和夜班的产品合格率是否有差异,进行调查得到如下数据

	合格	不合格
白班	232	19
夜班	54	18

试据此判断,产品合格率是否与班次有关? ($\alpha=0.05$)

68. 为了解男性和女性对三种类型的啤酒: 淡啤酒、普通啤酒和黑啤酒的偏好有没有差异,分别调查了 180 位男士和 120 位女士的喜好,得如下数据

	淡啤酒	普通啤酒	黑啤酒
男性	49	31	100
女性	51	20	49

请问男性和女性对这三种类型的啤酒的偏好有显著差异吗? ($\alpha=0.05$)

69. 检查一本书的 150 页,记录各页中印刷错误的个数,其结果为

错误的个数 f_i	0	1	2	3	4	5	6	$\geqslant 7$
含 f_i 个错误的页数	86	40	19	2	0	2	1	0

试在显著性水平 0.05 下检验假设 H_0 : 每页上的印刷错误个数服从泊松分布.

70. 下表给出了从某大学一年级学生中随机抽取的 200 个学生的某次数学考试成绩:

分数	$[20,30]$	$(30,40]$	$(40,50]$	$(50,60]$	$(60,70]$	$(70,80]$	$(80,90]$	$(90,100]$
学生数	5	15	30	51	60	23	10	6

试在显著性水平 0.05 下检验假设成绩是否服从正态分布 $N\left(60,15^2\right)$.

71. 在质量管理中,产品的优良性和稳定性可以分别用均值和方差来体现. 要比较甲乙两厂所生产的电视机的质量是否有差异,对两厂所生产的各 10 台产品的使用情况进行了追踪调查,得到这 20 台电视机的寿命 (单位: 年) 数据如下:

甲	8	7	9	5	12	10	9	10	8	7
乙	10	8	5	7	8	7	11	4	5	6

设电视机的寿命服从正态分布,利用假设检验的知识判断两厂所生产电视机的寿命是否有显著差异 (显著性水平取 $\alpha=0.05$).

72. 对截止目前为止一共 44 位美国总统 (含现任总统 Trump) 的星座进行分析,发现十二星座中天蝎座和水瓶座各有 5 人,双子座和射手座各有 3 人,处女座和白羊座各有两人,而其余星座均有 4 人. 于是有人宣称有些星座擅长当美国总统,而有些星座则不擅长. 结合你所学的知识,说明该说法是否有统计学上的依据? (显著性水平取 $\alpha=0.05$)

73. 设某两家工厂生产同一种产品,产品分 1、2、3 三个等级 (分别表示高、中、低三个等级),现从两家工厂生产的产品中各随机抽取 110 和 100 个产品,检查产品等级后结果如下

等级			
工厂		2	3
甲	58	40	12
乙	40	40	20

试分别在水平 $\alpha=0.1$ 和 $\alpha=0.05$ 下判断两家工厂的产品质量等级是否相同? 并解释你的结论.

74. 某公司是一家较大的高尔夫球制造商. 管理人员认为: 引进某种耐石损、寿命较长的高尔夫球会使该公司的市场占有率增加. 因此, 为了抗磨损, 延长使用寿命, 公司的研究小组开发了一种涂上一层保护膜的新型高尔夫球. 一位研究者关注保护膜对击球距离的影响. 公司希望

新型耐磨的高尔夫球与目前使用的高尔夫球有相同的击球距离. 为比较两种高尔夫球的击球距离, 各取 40 只球来做距离测试. 为了能将两种型号的高尔夫球平均距离的差异归因于制作方法的不同, 检验是用机械击球装置来完成的. 检验结果如下, 其中距离是按最近的码数来测量的.

之前	之后	之前	之后	之前	之后	之前	之后
264	277	270	272	263	274	281	283
261	269	287	259	264	266	274	250
267	263	289	264	284	262	273	253
272	266	280	280	263	271	263	260
258	262	272	274	260	260	275	270
283	251	275	281	283	281	267	263
258	262	265	276	255	250	279	261
266	289	260	269	272	263	274	255
259	286	278	268	266	278	276	263
270	264	275	262	268	264	262	279

(1) 用公式表示并介绍该公司用于比较目前使用的和新型高尔夫球击球距离的假设检验的基本原理. (2) 分析数据, 得出假设检验的结论. 检验的 p 值是多少? 你对公司有何建议? (3) 对每种型号的数据给出描述性的统计汇总. (4) 每种型号的总体均值的 95% 置信区间是多少? 两总体均值差的 95% 置信区间是多少?

75. 空军电子学引导性教程采用一种个人化教学系统, 每位学生观看讲座录像, 然后给予程式化的教材. 每位学生独立学习直至完成训练并通过考试. 人们关心的问题是不同学生完成训练计划的速度的差别. 有些学生能相当快地完成程式化教材, 而另一些需要花较长时间. 学的较快的学生必须等学得慢的学生完成引导性教程后才能一起进行其他方面的训练. 建议的替代系统是使用计算机辅助教学. 在这种方法中, 所有的学生观看相同的讲座录像, 然后每位学生被指派到一个计算机终端来接受进一步的训练. 在整个教程的自我训练中, 由计算机指导学生独立操作. 为了比较建议的和当前的教学方法, 刚入学的 122 名学生被随机安排到这两种教学系统中. 61 名学生使用当前程式化教材, 而另外 61 名学生使用建议的计算机辅助方法. 记录每个学生的学习时间 (以小时记).

当前程式化教材										
76	76	77	74	76	74	74	77	72	78	73
78	75	80	79	72	69	79	72	70	70	81
76	78	72	82	72	73	71	70	77	78	73
79	82	65	77	79	73	76	81	69	75	75
77	79	76	78	76	76	73	77	84	74	74
69	79	66	70	74	72					

74	75	77	78	74	80	73	73	78	76	76
74	77	69	76	75	72	75	72	76	72	77
73	77	69	77	75	76	74	77	75	78	72
77	78	78	76	75	76	76	75	76	80	77
76	75	73	77	77	77	79	75	75	72	82
76	76	74	72	78	71					

(1) 利用适当的描述统计学方法汇总每种方法训练时间的数据. 根据样本资料你能观察到有何异同? (2) 比较两种方法均值上的差异. 讨论你的结论. (3) 计算每一种训练方法的标准差与方差. 进行两种训练方法总体方差相等的假设检验. 讨论你的结论. (4) 关于两种方法之间的差异, 你能得到什么结论? 你有何建议? 请解释.

76. 下表分别给出两位文学家马克. 吐温 (Mark Twain) 的 8 篇小品文以及斯诺特格拉斯 (Snodgrass) 的 10 篇小品文中由 3 个字母组成的单字的比例:

马克. 吐温	0.225	0.262	0.217	0.240	0.230	0.229	0.235	0.217		
斯诺特格拉斯	0.209	0.205	0.196	0.210	0.202	0.207	0.224	0.223	0.220	0.201

设两组数据分别来自正态总体, 且两总体方差相等, 但参数均未知. 两样本相互独立. 问两位作家所写的小品文中包含有 3 个字母组成的单字的比例是否有显著的差异 (取 $\alpha = 0.05$)? 72. 用一种叫“混乱指标”的尺度去衡量工程师的英语文章的可理解性, 对混乱指标的打分越低表示可理解性越高. 分别随机选取 13 篇刊载在工程杂志上的论文, 以及 10 篇未出版的学术报告, 对他们的打分列于下表:

工程杂志上的论文 (数据 I)	未出版的学术报告 (数据 II)					
1.79	1.75	1.67	1.65	2.39	2.51	2.86
1.87	1.74	1.94		2.56	2.29	2.49
1.62	2.06	1.33		2.36	2.58	
1.96	1.69	1.70		2.62	2.41	

设数据 I, II 分别来自正态总体 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2), \mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ 均未知, 两样本独立. (1) 试检验假设 $H_0 : \sigma_1^2 = \sigma_2^2, H_1 : \sigma_1^2 \neq \sigma_2^2$(取 $\alpha = 0.1$). (2) 若能接受 H_0, 接着检验假设 $H_0' : \mu_1 = \mu_2, H_1' : \mu_1 \neq \mu_2$(取 $\alpha = 0.1$).

77. 下面给出了随机选取的某大学一年级学生 (200 名) 一次数学考试的成绩. (1) 画出数据的直方图. (2) 试取 $\alpha = 0.1$ 检验数据来自正态总体 $N(60, 15^2)$.

分数 x	$20 \leqslant x \leqslant 30$	$30 < x \leqslant 40$	$40 < x \leqslant 50$	$50 < x \leqslant 60$
学生数	5	15	30	51

分数 x	$60 < x \leqslant 70$	$70 < x \leqslant 80$	$80 < x \leqslant 90$	$90 < x \leqslant 100$
学生数	60	23	10	6

78. 袋中装有 8 个球, 其中红球数末知. 在其中任取 3 个, 记录红球的个数 X, 然后放回, 再任取 3 个, 记录红球的个数, 然后放回. 如此重复进行了 112 次, 其结果如下:

x	0	1	2	3
次数	1	31	55	25

试取 $\alpha = 0.05$ 检验假设 $H_0 : X$ 服从超几何分布

$$P\{X = k\} = \begin{pmatrix} 5 \\ k \end{pmatrix} \begin{pmatrix} 3 \\ 3-k \end{pmatrix} \Big/ \begin{pmatrix} 8 \\ 3 \end{pmatrix}, k = 0, 1, 2, 3.$$

即检验假设 H_0 : 红球的个数为 5 .

79. 一农场 10 年前在一鱼塘中按比例 $20 : 15 : 40 : 25$ 投放了四种鱼: 鲑鱼、鲇鱼、竹夹鱼和鲇鱼的鱼苗, 现在在鱼塘里获得一个样本如下: 试取 $\alpha = 0.05$, 检验各类鱼数量的比例较 10 年前是否有显著的改变.

序号	1	2	3	4
种类	鱼	鲇鱼	竹夹鱼	鲇鱼
数量 (条)	132	100	200	168

第九章 线性回归

附录 A t-分布分位数表

n \ P	0.75	0.80	0.85	0.90	0.95	0.975	0.99	0.995	0.9975	0.999	0.9995
1	1	1.376	1.963	3.078	6.314	12.71	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.92	4.303	6.965	9.925	14.09	22.33	31.6
3	0.765	0.978	1.25	1.638	2.353	3.182	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.19	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.61
5	0.727	0.92	1.156	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.44	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.86	2.306	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.1	1.383	1.833	2.262	2.821	3.25	3.69	4.297	4.781
10	0.7	0.879	1.093	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428	3.93	4.318
13	0.694	0.87	1.079	1.35	1.771	2.16	2.65	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.14
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.69	0.865	1.071	1.337	1.746	2.12	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.74	2.11	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.33	1.734	2.101	2.552	2.878	3.197	3.61	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.687	0.86	1.064	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.85
21	0.686	0.859	1.063	1.323	1.721	2.08	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.06	1.319	1.714	2.069	2.5	2.807	3.104	3.485	3.767
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.06	2.485	2.787	3.078	3.45	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.69
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.31	1.697	2.042	2.457	2.75	3.03	3.385	3.646
40	0.681	0.851	1.05	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2	2.39	2.66	2.915	3.232	3.46
80	0.678	0.846	1.043	1.292	1.664	1.99	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.29	1.66	1.984	2.364	2.626	2.871	3.174	3.39
120	0.677	0.845	1.041	1.289	1.658	1.98	2.358	2.617	2.86	3.16	3.373
∞	0.674	0.842	1.036	1.282	1.645	1.96	2.326	2.576	2.807	3.09	3.291

附录 B χ^2 分布分位数表

	P										
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.690	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588
11	2.603	3.816	14.631	17.275	19.675	21.920	22.618	24.725	26.757	29.354	31.264
12	3.074	4.404	15.812	18.549	21.026	23.337	24.054	26.217	28.300	30.957	32.909
13	3.565	5.009	16.985	19.812	22.362	24.736	25.472	27.688	29.819	32.535	34.528
14	4.075	5.629	18.151	21.064	23.685	26.119	26.873	29.141	31.319	34.091	36.123
15	4.601	6.262	19.311	22.307	24.996	27.488	28.259	30.578	32.801	35.628	37.697
16	5.142	6.908	20.465	23.542	26.296	28.845	29.633	32.000	34.267	37.146	39.252
17	5.697	7.564	21.615	24.769	27.587	30.191	30.995	33.409	35.718	38.648	40.790
18	6.265	8.231	22.760	25.989	28.869	31.526	32.346	34.805	37.156	40.136	42.312
19	6.844	8.907	23.900	27.204	30.144	32.852	33.687	36.191	38.582	41.610	43.820
20	7.434	9.591	25.038	28.412	31.410	34.170	35.020	37.566	39.997	43.072	45.315

	P									
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
21	8.034	10.283	26.171	29.615	32.671	35.479	36.343	38.932	41.401	44.522	46.797
22	8.643	10.982	27.301	30.813	33.924	36.781	37.659	40.289	42.796	45.962	48.268
23	9.260	11.689	28.429	32.007	35.172	38.076	38.968	41.638	44.181	47.391	49.728
24	9.886	12.401	29.553	33.196	36.415	39.364	40.270	42.980	45.559	48.812	51.179
25	10.520	13.120	30.675	34.382	37.652	40.646	41.566	44.314	46.928	50.223	52.620
26	11.160	13.844	31.795	35.563	38.885	41.923	42.856	45.642	48.290	51.627	54.052
27	11.808	14.573	32.912	36.741	40.113	43.195	44.140	46.963	49.645	53.023	55.476
28	12.461	15.308	34.027	37.916	41.337	44.461	45.419	48.278	50.993	54.411	56.892
29	13.121	16.047	35.139	39.087	42.557	45.722	46.693	49.588	52.336	55.792	58.301
30	13.787	16.791	36.250	40.256	43.773	46.979	47.962	50.892	53.672	57.167	59.703
31	14.458	17.539	37.359	41.422	44.985	48.232	49.226	52.191	55.003	58.536	61.098
32	15.134	18.291	38.466	42.585	46.194	49.480	50.487	53.486	56.328	59.899	62.487
33	15.815	19.047	39.572	43.745	47.400	50.725	51.743	54.776	57.648	61.256	63.870
34	16.501	19.806	40.676	44.903	48.602	51.966	52.995	56.061	58.964	62.608	65.247
35	17.192	20.569	41.778	46.059	49.802	53.203	54.244	57.342	60.275	63.955	66.619
36	17.887	21.336	42.879	47.212	50.998	54.437	55.489	58.619	61.581	65.296	67.985
37	18.586	22.106	43.978	48.363	52.192	55.668	56.730	59.893	62.883	66.633	69.346
38	19.289	22.878	45.076	49.513	53.384	56.896	57.969	61.162	64.181	67.966	70.703
39	19.996	23.654	46.173	50.660	54.572	58.120	59.204	62.428	65.476	69.294	72.055
40	20.707	24.433	47.269	51.805	55.758	59.342	60.436	63.691	66.766	70.618	73.402

	P									
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
41	21.421	25.215	48.363	52.949	56.942	60.561	61.665	64.950	68.053	71.938	74.745
42	22.138	25.999	49.456	54.090	58.124	61.777	62.892	66.206	69.336	73.254	76.084
43	22.859	26.785	50.548	55.230	59.304	62.990	64.116	67.459	70.616	74.566	77.419
44	23.584	27.575	51.639	56.369	60.481	64.201	65.337	68.710	71.893	75.874	78.750
45	24.311	28.366	52.729	57.505	61.656	65.410	66.555	69.957	73.166	77.179	80.077
46	25.041	29.160	53.818	58.641	62.830	66.617	67.771	71.201	74.437	78.481	81.400
47	25.775	29.956	54.906	59.774	64.001	67.821	68.985	72.443	75.704	79.780	82.720
48	26.511	30.755	55.993	60.907	65.171	69.023	70.197	73.683	76.969	81.075	84.037
49	27.249	31.555	57.079	62.038	66.339	70.222	71.406	74.919	78.231	82.367	85.351
50	27.991	32.357	58.164	63.167	67.505	71.420	72.613	76.154	79.490	83.657	86.661
51	28.735	33.162	59.248	64.295	68.669	72.616	73.818	77.386	80.747	84.943	87.968
52	29.481	33.968	60.332	65.422	69.832	73.810	75.021	78.616	82.001	86.227	89.272
53	30.230	34.776	61.414	66.548	70.993	75.002	76.223	79.843	83.253	87.507	90.573
54	30.981	35.586	62.496	67.673	72.153	76.192	77.422	81.069	84.502	88.786	91.872
55	31.735	36.398	63.577	68.796	73.311	77.380	78.619	82.292	85.749	90.061	93.168
56	32.490	37.212	64.658	69.919	74.468	78.567	79.815	83.513	86.994	91.335	94.461
57	33.248	38.027	65.737	71.040	75.624	79.752	81.009	84.733	88.236	92.605	95.751
58	34.008	38.844	66.816	72.160	76.778	80.936	82.201	85.950	89.477	93.874	97.039
59	34.770	39.662	67.894	73.279	77.931	82.117	83.391	87.166	90.715	95.140	98.324
60	35.534	40.482	68.972	74.397	79.082	83.298	84.580	88.379	91.952	96.404	99.607

	P									
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
61	36.301	41.303	70.049	75.514	80.232	84.476	85.767	89.591	93.186	97.665	100.888
62	37.068	42.126	71.125	76.630	81.381	85.654	86.953	90.802	94.419	98.925	102.166
63	37.838	42.950	72.201	77.745	82.529	86.830	88.137	92.010	95.649	100.182	103.442
64	38.610	43.776	73.276	78.860	83.675	88.004	89.320	93.217	96.878	101.437	104.716
65	39.383	44.603	74.351	79.973	84.821	89.177	90.501	94.422	98.105	102.691	105.988
66	40.158	45.431	75.424	81.085	85.965	90.349	91.681	95.626	99.330	103.942	107.258
67	40.935	46.261	76.498	82.197	87.108	91.519	92.860	96.828	100.554	105.192	108.526
68	41.713	47.092	77.571	83.308	88.250	92.689	94.037	98.028	101.776	106.440	109.791
69	42.494	47.924	78.643	84.418	89.391	93.856	95.213	99.228	102.996	107.685	111.055
70	43.275	48.758	79.715	85.527	90.531	95.023	96.388	100.425	104.215	108.929	112.317
71	44.058	49.592	80.786	86.635	91.670	96.189	97.561	101.621	105.432	110.172	113.577
72	44.843	50.428	81.857	87.743	92.808	97.353	98.733	102.816	106.648	111.412	114.835
73	45.629	51.265	82.927	88.850	93.945	98.516	99.904	104.010	107.862	112.651	116.092
74	46.417	52.103	83.997	89.956	95.081	99.678	101.074	105.202	109.074	113.889	117.346
75	47.206	52.942	85.066	91.061	96.217	100.839	102.243	106.393	110.286	115.125	118.599
76	47.997	53.782	86.135	92.166	97.351	101.999	103.410	107.583	111.495	116.359	119.850
77	48.788	54.623	87.203	93.270	98.484	103.158	104.576	108.771	112.704	117.591	121.100
78	49.582	55.466	88.271	94.374	99.617	104.316	105.742	109.958	113.911	118.823	122.348
79	50.376	56.309	89.338	95.476	100.749	105.473	106.906	111.144	115.117	120.052	123.594
80	51.172	57.153	90.405	96.578	101.879	106.629	108.069	112.329	116.321	121.280	124.839

	P										
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
81	51.969	57.998	91.472	97.680	103.010	107.783	109.232	113.512	117.524	122.507	126.083
82	52.767	58.845	92.538	98.780	104.139	108.937	110.393	114.695	118.726	123.733	127.324
83	53.567	59.692	93.604	99.880	105.267	110.090	111.553	115.876	119.927	124.957	128.565
84	54.368	60.540	94.669	100.980	106.395	111.242	112.712	117.057	121.126	126.179	129.804
85	55.170	61.389	95.734	102.079	107.522	112.393	113.871	118.236	122.325	127.401	131.041
86	55.973	62.239	96.799	103.177	108.648	113.544	115.028	119.414	123.522	128.621	132.277
87	56.777	63.089	97.863	104.275	109.773	114.693	116.184	120.591	124.718	129.840	133.512
88	57.582	63.941	98.927	105.372	110.898	115.841	117.340	121.767	125.913	131.057	134.745
89	58.389	64.793	99.991	106.469	112.022	116.989	118.495	122.942	127.106	132.273	135.978
90	59.196	65.647	101.054	107.565	113.145	118.136	119.648	124.116	128.299	133.489	137.208
91	60.005	66.501	102.117	108.661	114.268	119.282	120.801	125.289	129.491	134.702	138.438
92	60.815	67.356	103.179	109.756	115.390	120.427	121.954	126.462	130.681	135.915	139.666
93	61.625	68.211	104.241	110.850	116.511	121.571	123.105	127.633	131.871	137.127	140.893
94	62.437	69.068	105.303	111.944	117.632	122.715	124.255	128.803	133.059	138.337	142.119
95	63.250	69.925	106.364	113.038	118.752	123.858	125.405	129.973	134.247	139.546	143.344
96	64.063	70.783	107.425	114.131	119.871	125.000	126.554	131.141	135.433	140.755	144.567
97	64.878	71.642	108.486	115.223	120.990	126.141	127.702	132.309	136.619	141.962	145.789
98	65.694	72.501	109.547	116.315	122.108	127.282	128.849	133.476	137.803	143.168	147.010
99	66.510	73.361	110.607	117.407	123.225	128.422	129.996	134.642	138.987	144.373	148.230
100	67.328	74.222	111.667	118.498	124.342	129.561	131.142	135.807	140.169	145.577	149.449

	P										
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
101	68.146	75.083	112.726	119.589	125.458	130.700	132.287	136.971	141.351	146.780	150.667
102	68.965	75.946	113.786	120.679	126.574	131.838	133.431	138.134	142.532	147.982	151.884
103	69.785	76.809	114.845	121.769	127.689	132.975	134.575	139.297	143.712	149.183	153.099
104	70.606	77.672	115.903	122.858	128.804	134.111	135.718	140.459	144.891	150.383	154.314
105	71.428	78.536	116.962	123.947	129.918	135.247	136.860	141.620	146.070	151.582	155.528
106	72.251	79.401	118.020	125.035	131.031	136.382	138.002	142.780	147.247	152.780	156.740
107	73.075	80.267	119.078	126.123	132.144	137.517	139.143	143.940	148.424	153.977	157.952
108	73.899	81.133	120.135	127.211	133.257	138.651	140.283	145.099	149.599	155.173	159.162
109	74.724	82.000	121.192	128.298	134.369	139.784	141.423	146.257	150.774	156.369	160.372
110	75.550	82.867	122.250	129.385	135.480	140.917	142.562	147.414	151.948	157.563	161.581
111	76.377	83.735	123.306	130.472	136.591	142.049	143.700	148.571	153.122	158.757	162.788
112	77.204	84.604	124.363	131.558	137.701	143.180	144.838	149.727	154.294	159.950	163.995
113	78.033	85.473	125.419	132.643	138.811	144.311	145.975	150.882	155.466	161.141	165.201
114	78.862	86.342	126.475	133.729	139.921	145.441	147.111	152.037	156.637	162.332	166.406
115	79.692	87.213	127.531	134.813	141.030	146.571	148.247	153.191	157.808	163.523	167.610
116	80.522	88.084	128.587	135.898	142.138	147.700	149.383	154.344	158.977	164.712	168.813
117	81.353	88.955	129.642	136.982	143.246	148.829	150.517	155.496	160.146	165.900	170.016
118	82.185	89.827	130.697	138.066	144.354	149.957	151.652	156.648	161.314	167.088	171.217
119	83.018	90.700	131.752	139.149	145.461	151.084	152.785	157.800	162.481	168.275	172.418
120	83.852	91.573	132.806	140.233	146.567	152.211	153.918	158.950	163.648	169.461	173.617

	P									
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
121	84.686	92.446	133.861	141.315	147.674	153.338	155.051	160.100	164.814	170.647	174.816
122	85.520	93.320	134.915	142.398	148.779	154.464	156.183	161.250	165.980	171.831	176.014
123	86.356	94.195	135.969	143.480	149.885	155.589	157.314	162.398	167.144	173.015	177.212
124	87.192	95.070	137.022	144.562	150.989	156.714	158.445	163.546	168.308	174.198	178.408
125	88.029	95.946	138.076	145.643	152.094	157.839	159.575	164.694	169.471	175.380	179.604
126	88.866	96.822	139.129	146.724	153.198	158.962	160.705	165.841	170.634	176.562	180.799
127	89.704	97.698	140.182	147.805	154.302	160.086	161.834	166.987	171.796	177.743	181.993
128	90.543	98.576	141.235	148.885	155.405	161.209	162.963	168.133	172.957	178.923	183.186
129	91.382	99.453	142.288	149.965	156.508	162.331	164.091	169.278	174.118	180.103	184.379
130	92.222	100.331	143.340	151.045	157.610	163.453	165.219	170.423	175.278	181.282	185.571
131	93.063	101.210	144.392	152.125	158.712	164.575	166.346	171.567	176.438	182.460	186.762
132	93.904	102.089	145.444	153.204	159.814	165.696	167.473	172.711	177.597	183.637	187.953
133	94.746	102.968	146.496	154.283	160.915	166.816	168.600	173.854	178.755	184.814	189.142
134	95.588	103.848	147.548	155.361	162.016	167.936	169.725	174.996	179.913	185.990	190.331
135	96.431	104.729	148.599	156.440	163.116	169.056	170.851	176.138	181.070	187.165	191.520
136	97.275	105.609	149.651	157.518	164.216	170.175	171.976	177.280	182.226	188.340	192.707
137	98.119	106.491	150.702	158.595	165.316	171.294	173.100	178.421	183.382	189.514	193.894
138	98.964	107.372	151.753	159.673	166.415	172.412	174.224	179.561	184.538	190.688	195.080
139	99.809	108.254	152.803	160.750	167.514	173.530	175.348	180.701	185.693	191.861	196.266
140	100.655	109.137	153.854	161.827	168.613	174.648	176.471	181.840	186.847	193.033	197.451

	P										
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
141	101.501	110.020	154.904	162.904	169.711	175.765	177.594	182.979	188.001	194.205	198.635
142	102.348	110.903	155.954	163.980	170.809	176.882	178.716	184.118	189.154	195.376	199.819
143	103.196	111.787	157.004	165.056	171.907	177.998	179.838	185.256	190.306	196.546	201.002
144	104.044	112.671	158.054	166.132	173.004	179.114	180.959	186.393	191.458	197.716	202.184
145	104.892	113.556	159.104	167.207	174.101	180.229	182.080	187.530	192.610	198.885	203.366
146	105.741	114.441	160.153	168.283	175.198	181.344	183.200	188.666	193.761	200.054	204.547
147	106.591	115.326	161.202	169.358	176.294	182.459	184.321	189.802	194.912	201.222	205.727
148	107.441	116.212	162.251	170.432	177.390	183.573	185.440	190.938	196.062	202.390	206.907
149	108.291	117.098	163.300	171.507	178.485	184.687	186.560	192.073	197.211	203.557	208.086
150	109.142	117.985	164.349	172.581	179.581	185.800	187.678	193.208	198.360	204.723	209.265
151	109.994	118.871	165.398	173.655	180.676	186.914	188.797	194.342	199.509	205.889	210.443
152	110.846	119.759	166.446	174.729	181.770	188.026	189.915	195.476	200.657	207.054	211.620
153	111.698	120.646	167.495	175.803	182.865	189.139	191.033	196.609	201.804	208.219	212.797
154	112.551	121.534	168.543	176.876	183.959	190.251	192.150	197.742	202.951	209.383	213.973
155	113.405	122.423	169.591	177.949	185.052	191.362	193.267	198.874	204.098	210.547	215.149
156	114.259	123.312	170.639	179.022	186.146	192.474	194.384	200.006	205.244	211.710	216.324
157	115.113	124.201	171.686	180.094	187.239	193.584	195.500	201.138	206.390	212.873	217.499
158	115.968	125.090	172.734	181.167	188.332	194.695	196.616	202.269	207.535	214.035	218.673
159	116.823	125.980	173.781	182.239	189.424	195.805	197.731	203.400	208.680	215.197	219.846
160	117.679	126.870	174.828	183.311	190.516	196.915	198.846	204.530	209.824	216.358	221.019

	P										
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
161	118.536	127.761	175.875	184.382	191.608	198.025	199.961	205.660	210.968	217.518	222.191
162	119.392	128.651	176.922	185.454	192.700	199.134	201.076	206.790	212.111	218.678	223.363
163	120.249	129.543	177.969	186.525	193.791	200.243	202.190	207.919	213.254	219.838	224.535
164	121.107	130.434	179.016	187.596	194.883	201.351	203.303	209.047	214.396	220.997	225.705
165	121.965	131.326	180.062	188.667	195.973	202.459	204.417	210.176	215.539	222.156	226.876
166	122.823	132.218	181.109	189.737	197.064	203.567	205.530	211.304	216.680	223.314	228.045
167	123.682	133.111	182.155	190.808	198.154	204.675	206.642	212.431	217.821	224.472	229.215
168	124.541	134.003	183.201	191.878	199.244	205.782	207.755	213.558	218.962	225.629	230.383
169	125.401	134.897	184.247	192.948	200.334	206.889	208.867	214.685	220.102	226.786	231.552
170	126.261	135.790	185.293	194.017	201.423	207.995	209.978	215.812	221.242	227.942	232.719
171	127.122	136.684	186.338	195.087	202.513	209.102	211.090	216.938	222.382	229.098	233.887
172	127.983	137.578	187.384	196.156	203.602	210.208	212.201	218.063	223.521	230.253	235.053
173	128.844	138.472	188.429	197.225	204.690	211.313	213.311	219.189	224.660	231.408	236.220
174	129.706	139.367	189.475	198.294	205.779	212.419	214.422	220.314	225.798	232.563	237.385
175	130.568	140.262	190.520	199.363	206.867	213.524	215.532	221.438	226.936	233.717	238.551
176	131.430	141.157	191.565	200.432	207.955	214.628	216.641	222.563	228.074	234.870	239.716
177	132.293	142.053	192.610	201.500	209.042	215.733	217.751	223.687	229.211	236.023	240.880
178	133.157	142.949	193.654	202.568	210.130	216.837	218.860	224.810	230.347	237.176	242.044
179	134.020	143.845	194.699	203.636	211.217	217.941	219.969	225.933	231.484	238.328	243.207
180	134.884	144.741	195.743	204.704	212.304	219.044	221.077	227.056	232.620	239.480	244.370

	P									
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
181	135.749	145.638	196.788	205.771	213.391	220.148	222.185	228.179	233.755	240.632	245.533
182	136.614	146.535	197.832	206.839	214.477	221.251	223.293	229.301	234.891	241.783	246.695
183	137.479	147.432	198.876	207.906	215.563	222.353	224.401	230.423	236.026	242.933	247.857
184	138.344	148.330	199.920	208.973	216.649	223.456	225.508	231.544	237.160	244.084	249.018
185	139.210	149.228	200.964	210.040	217.735	224.558	226.615	232.665	238.294	245.234	250.179
186	140.077	150.126	202.008	211.106	218.820	225.660	227.722	233.786	239.428	246.383	251.339
187	140.943	151.024	203.052	212.173	219.906	226.761	228.828	234.907	240.561	247.532	252.499
188	141.810	151.923	204.095	213.239	220.991	227.863	229.935	236.027	241.694	248.681	253.659
189	142.678	152.822	205.139	214.305	222.076	228.964	231.040	237.147	242.827	249.829	254.818
190	143.545	153.721	206.182	215.371	223.160	230.064	232.146	238.266	243.959	250.977	255.976
191	144.413	154.621	207.225	216.437	224.245	231.165	233.251	239.386	245.091	252.124	257.135
192	145.282	155.521	208.268	217.502	225.329	232.265	234.356	240.505	246.223	253.271	258.292
193	146.150	156.421	209.311	218.568	226.413	233.365	235.461	241.623	247.354	254.418	259.450
194	147.020	157.321	210.354	219.633	227.496	234.465	236.566	242.742	248.485	255.564	260.607
195	147.889	158.221	211.397	220.698	228.580	235.564	237.670	243.860	249.616	256.710	261.763
196	148.759	159.122	212.439	221.763	229.663	236.664	238.774	244.977	250.746	257.855	262.920
197	149.629	160.023	213.482	222.828	230.746	237.763	239.877	246.095	251.876	259.001	264.075
198	150.499	160.925	214.524	223.892	231.829	238.861	240.981	247.212	253.006	260.145	265.231
199	151.370	161.826	215.567	224.957	232.912	239.960	242.084	248.329	254.135	261.290	266.386
200	152.241	162.728	216.609	226.021	233.994	241.058	243.187	249.445	255.264	262.434	267.541

	P									
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
201	153.112	163.630	217.651	227.085	235.077	242.156	244.290	250.561	256.393	263.578	268.695
202	153.984	164.532	218.693	228.149	236.159	243.254	245.392	251.677	257.521	264.721	269.849
203	154.856	165.435	219.735	229.213	237.240	244.351	246.494	252.793	258.649	265.864	271.002
204	155.728	166.338	220.777	230.276	238.322	245.448	247.596	253.908	259.777	267.007	272.155
205	156.601	167.241	221.818	231.340	239.403	246.545	248.698	255.023	260.904	268.149	273.308
206	157.474	168.144	222.860	232.403	240.485	247.642	249.799	256.138	262.031	269.291	274.460
207	158.347	169.047	223.901	233.466	241.566	248.739	250.900	257.253	263.158	270.432	275.612
208	159.221	169.951	224.943	234.529	242.647	249.835	252.001	258.367	264.285	271.574	276.764
209	160.095	170.855	225.984	235.592	243.727	250.931	253.102	259.481	265.411	272.715	277.915
210	160.969	171.759	227.025	236.655	244.808	252.027	254.202	260.595	266.537	273.855	279.066
211	161.843	172.664	228.066	237.717	245.888	253.122	255.302	261.708	267.662	274.995	280.217
212	162.718	173.568	229.107	238.780	246.968	254.218	256.402	262.821	268.788	276.135	281.367
213	163.593	174.473	230.148	239.842	248.048	255.313	257.502	263.934	269.912	277.275	282.517
214	164.469	175.378	231.189	240.904	249.128	256.408	258.601	265.047	271.037	278.414	283.666
215	165.344	176.283	232.230	241.966	250.207	257.503	259.701	266.159	272.162	279.553	284.815
216	166.220	177.189	233.270	243.028	251.286	258.597	260.800	267.271	273.286	280.692	285.964
217	167.096	178.095	234.311	244.090	252.365	259.691	261.898	268.383	274.409	281.830	287.112
218	167.973	179.001	235.351	245.151	253.444	260.785	262.997	269.495	275.533	282.968	288.261
219	168.850	179.907	236.391	246.213	254.523	261.879	264.095	270.606	276.656	284.106	289.408
220	169.727	180.813	237.432	247.274	255.602	262.973	265.193	271.717	277.779	285.243	290.556

	P									
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
221	170.604	181.720	238.472	248.335	256.680	264.066	266.291	272.828	278.902	286.380	291.703
222	171.482	182.627	239.512	249.396	257.758	265.159	267.389	273.939	280.024	287.517	292.850
223	172.360	183.534	240.552	250.457	258.837	266.252	268.486	275.049	281.146	288.653	293.996
224	173.238	184.441	241.592	251.517	259.914	267.345	269.584	276.159	282.268	289.789	295.142
225	174.116	185.348	242.631	252.578	260.992	268.438	270.681	277.269	283.390	290.925	296.288
226	174.995	186.256	243.671	253.638	262.070	269.530	271.777	278.379	284.511	292.061	297.433
227	175.874	187.164	244.711	254.699	263.147	270.622	272.874	279.488	285.632	293.196	298.579
228	176.753	188.072	245.750	255.759	264.224	271.714	273.970	280.597	286.753	294.331	299.723
229	177.633	188.980	246.790	256.819	265.301	272.806	275.066	281.706	287.874	295.465	300.868
230	178.512	189.889	247.829	257.879	266.378	273.898	276.162	282.814	288.994	296.600	302.012
231	179.392	190.797	248.868	258.939	267.455	274.989	277.258	283.923	290.114	297.734	303.156
232	180.273	191.706	249.908	259.998	268.531	276.080	278.354	285.031	291.234	298.867	304.299
233	181.153	192.615	250.947	261.058	269.608	277.171	279.449	286.139	292.353	300.001	305.443
234	182.034	193.524	251.986	262.117	270.684	278.262	280.544	287.247	293.472	301.134	306.586
235	182.915	194.434	253.025	263.176	271.760	279.352	281.639	288.354	294.591	302.267	307.728
236	183.796	195.343	254.063	264.235	272.836	280.443	282.734	289.461	295.710	303.400	308.871
237	184.678	196.253	255.102	265.294	273.911	281.533	283.828	290.568	296.828	304.532	310.013
238	185.560	197.163	256.141	266.353	274.987	282.623	284.922	291.675	297.947	305.664	311.154
239	186.442	198.073	257.179	267.412	276.062	283.713	286.016	292.782	299.065	306.796	312.296
240	187.324	198.984	258.218	268.471	277.138	284.802	287.110	293.888	300.182	307.927	313.437

	P									
DF	0.995	0.975	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.002	0.001
241	188.207	199.894	259.256	269.529	278.213	285.892	288.204	294.994	301.300	309.058	314.578
242	189.090	200.805	260.295	270.588	279.288	286.981	289.298	296.100	302.417	310.189	315.718
243	189.973	201.716	261.333	271.646	280.362	288.070	290.391	297.206	303.534	311.320	316.859
244	190.856	202.627	262.371	272.704	281.437	289.159	291.484	298.311	304.651	312.450	317.999
245	191.739	203.539	263.409	273.762	282.511	290.248	292.577	299.417	305.767	313.580	319.138
246	192.623	204.450	264.447	274.820	283.586	291.336	293.670	300.522	306.883	314.710	320.278
247	193.507	205.362	265.485	275.878	284.660	292.425	294.762	301.626	307.999	315.840	321.417
248	194.391	206.274	266.523	276.935	285.734	293.513	295.855	302.731	309.115	316.969	322.556
249	195.276	207.186	267.561	277.993	286.808	294.601	296.947	303.835	310.231	318.098	323.694
250	196.161	208.098	268.599	279.050	287.882	295.689	298.039	304.940	311.346	319.227	324.832
300	240.663	253.912	320.397	331.789	341.395	349.874	352.425	359.906	366.844	375.369	381.425
350	285.608	300.064	372.051	384.306	394.626	403.723	406.457	414.474	421.900	431.017	437.488
400	330.903	346.482	423.590	436.649	447.632	457.305	460.211	468.724	476.606	486.274	493.132
450	376.483	393.118	475.035	488.849	500.456	510.670	513.736	522.717	531.026	541.212	548.432
500	422.303	439.936	526.401	540.930	553.127	563.852	567.070	576.493	585.207	595.882	603.446
550	468.328	486.910	577.701	592.909	605.667	616.878	620.241	630.084	639.183	650.324	658.215
600	514.529	534.019	628.943	644.800	658.094	669.769	673.270	683.516	692.982	704.568	712.771
650	560.885	581.245	680.134	696.614	710.421	722.542	726.176	736.807	746.625	758.639	767.141
700	607.380	628.577	731.280	748.359	762.661	775.211	778.972	789.974	800.131	812.556	821.347
750	653.997	676.003	782.386	800.043	814.822	827.785	831.670	843.029	853.514	866.336	875.404
800	700.725	723.513	833.456	851.671	866.911	880.275	884.279	895.984	906.786	919.991	929.329
850	747.554	771.099	884.492	903.249	918.937	932.689	936.808	948.848	959.957	973.534	983.133
900	794.475	818.756	935.499	954.782	970.904	985.032	989.263	1001.630	1013.036	1026.974	1036.826
950	841.480	866.477	986.478	1006.272	1022.816	1037.311	1041.651	1054.334	1066.031	1080.320	1090.418
1000	888.564	914.257	1037.431	1057.724	1074.679	1089.531	1093.977	1106.969	1118.948	1133.579	1143.917

附录 C F-分布分位数表

参考：http://socr.ucla.edu/Applets.dir/F_Table.html

$\alpha = 0.05$	$\mathbf{df_1}$=1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40
$\mathbf{df_2}$= **1**	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251
$\mathbf{df_2}$= **1**	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251

感谢阅读，本书到此