统计学在线资源

概率统计习题集

中国科学技术大学统计与金融系编

高等教育出版社
2018年1月

目录

第一部分

概率论

第一章　概率论基础

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$A=\{1/2<x\le 1\}$, $B=\{1/4<x\le 3/2\}$, 写出下列各事件下列事件
   (1) $A\overline{B}$,　(2) $\overline{A}\cup B$,　(3) $\overline{AB}$,　(4) $\overline{\overline{A}\ \overline{B}}$.

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?

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$(1 \le k \le a+b)$.

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, $m = 0, 1, 2, 3, 4$.

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.

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点间到达某地闯关过河, 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, 二年级有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. 甲乙两选手进行乒乓球单打比赛, 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. 计算玩家赢的概率.

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$1\sim9$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$1\sim9$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$m(1\leq m\leq9)$ 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点的概率是多少?

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$0\sim9$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. 投掷两枚均匀的骰子, 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. 已知抛掷硬币的过程中有3次正面向上, 计算$N=4$的概率.

47. 有两箱同种类型的零件. 第一箱装 50 只, 其中10只一等品; 第二箱装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$\mathbb{P}(B|A) = \mathbb{P}(B|\overline{A})$, 则事件$A$与$B$ 独立.

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

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

$$\mathbb{P}(AB|C) = \mathbb{P}(A|C)\mathbb{P}(B|C)$$

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 $\mathbb{P}(C) = 0.5$, $\mathbb{P}(A|C) = \mathbb{P}(B|C) = 0.9$, $\mathbb{P}(A|\bar{C}) = 0.2$, $\mathbb{P}(B|\bar{C}) = 0.1$, 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=0.008 $n = 25$, 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.

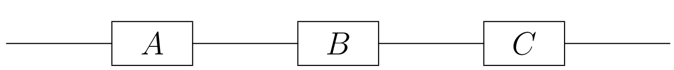
60. 某炮弹单发命中率为p, 连续向某个目标发射m次, 试求
(1) 击中目标的概率是多少?
(2) 恰好击中10次的概率是多少?
(3) 至少击中3次的概率是多少?

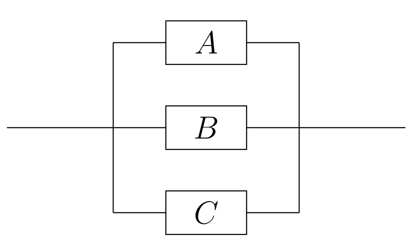
61. 某人从家到学校要经过4个红绿灯口, 1、3红绿灯口出现红灯的概率均为0.4, 2、4路口出现红灯的可能性为0.6, 且相互独立. 求某人从家到学校至少碰到两个红灯的概率是多少?

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.

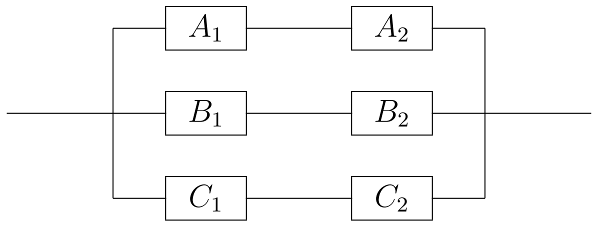
63. 连续投掷一均匀硬币三次, 事件$A = \{$至多有一次正面$\}$, 事件$B = \{$三次中有正面有反面$\}$, 求证两事件A 与B相互独立.

64. 投掷两枚均匀的骰子, 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.

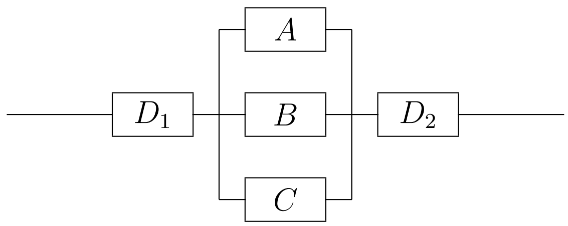
65. 求下列各系统能正常工作的概率, 其中框图中的字母代表元件, 字母相同但下标不同的都是同一种元件, 只是装配在不同的位置上, A, B, C, D 类元件能正常工作的概率分别为 p_A, p_B, p_C, p_D.
(1)


(2)


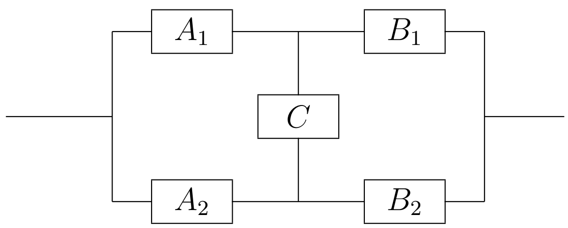
(3)



(4)



(5)*



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. 假设罐子中有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. 计算当罐子中有$m(m \geq 2)$个球时, 黑球个数为$k(1 \leq k \leq m$的概率.

72. 假设A罐子中有10个红球和6个白球, B罐子中是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, 从达拉斯到芝加哥有3个航班, 从北京直飞芝加哥有2个航班. 这些航班的票务之间完全独立, 买到票的概率都是$p(0 < p < 1)$. 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. 罐子中有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个罐子, 其中第k个罐子里有$k-1$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

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 $i(0 \leq i \leq 6)$ 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 元买一注后的下述事件
   $$A=\{\text{获奖}\}, \quad B=\{\text{获得一等奖或二等奖}\},$$
   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. 试求 X 的分布律.

6. 同时掷三枚均匀的骰子, 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箱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亏损两万元. 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 $\mathbb{P}(X \le 0)$;$\mathbb{P}\left(\frac{1}{2} < X \le \frac{3}{2}\right)$;$\mathbb{P}(1 \le X \le 2)$ 和 $\mathbb{P}(1 < X \le 2)$.

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

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

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 $\mathbb{P}(X \ge 1) = 0.51$, 试求 $\mathbb{P}(Y \ge 1)$.

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 $\mathbb{P}(X \ge 1)$ 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赞成票的概率是 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, 然后掷三个骰子, 若赌徒所押的数字出现 i 次, $i=1,2,3$, 则赌徒赢 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未能孵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况的概率大约是多少? 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) 若一个人买了 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

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

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

33. 设随机变量 X 的分布函数为

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

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

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) $(0,1)$

35. (2010年全国考研试题) 设随机变量 X 的分布函数为

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

则 $\mathbb{P}(X=1)$ 等于(　).

(A) 0　　(B) 1　　(C) $\frac{1}{2}-\mathrm{e}^{-1}$　　(D) $1-\mathrm{e}^{-1}$

36. 设随机变量 X 的分布函数为

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

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

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 $0\leq a<b\leq 1$, $F(b)-F(a)$ 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

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

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

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

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

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

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

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

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

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) $f_1(x)F_2(x)+f_2(x)F_1(x)$

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 $\mathbb{P}(Y\le a+1|Y>a)=$ _____.

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

$$\mathbb{P}(X\le t+x|X>t)=\mathbb{P}(X\le x).$$

(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

$$\mathbb{P}(X \le m+n | X > n) = \mathbb{P}(X \le m).$$

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 $\mathbb{P}(|X| \le 2)$;

(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 $\mathbb{P}(0 \le X \le 4), \mathbb{P}(X > 2.4)$ 和 $\mathbb{P}(|X| > 2)$;

(2) 试求常数 c, 使得 $\mathbb{P}(X > c) = 2\mathbb{P}(X \le c)$.

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 200V, $200 \sim 240$V 和超过 240V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 $f(x)=\begin{cases} af_1(x), & x\le 0,\\ bf_2(x), & x>0,\end{cases}$ $(a>0,b>0)$ 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$X_1\sim N(0,1),X_2\sim N(0,2^2),X_3\sim N(5,3^2)$, $P_i=\mathbb{P}(-2\le X_i\le 2)$ $(i=1,2,3)$, 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 $p=\mathbb{P}(X\le \mu+\sigma^2)$, 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 的增加而减少

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 $Z=|X-\frac{\pi}{2}|$ 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 $X \sim U(-\frac{\pi}{2}, \frac{\pi}{2})$, 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

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

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

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

(1) 求随机变量 Y 的分布函数.

(2) 求概率 $\mathbb{P}(X \leq Y)$.

第三章　多维随机变量及其分布

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 $\mathbb{P}(X=1|Z=0)$;

(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

Y \ 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) $f_1(x)f_2(y)(1 + (2F_1(x) - 1)(2F_2(y) - 1))$

(C) $0.5(f_1(x) + f_2(y))$　　(D) $0.5(f_1(x)f_1(y) + f_2(x)f_2(y))$

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

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 $\mathbb{P}(X \le 2, Y \le -2) = 0.25$, 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) 不是二元正态的.

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

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

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

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 $f_{Y|X}(y|x)$.

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 $\mathbb{P}(X\le 0.25|Y=0.5)$.

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, $X\sim N(0,\sigma_1^2)$, $Y\sim N(0,\sigma_2^2)$, 其中 $\sigma_1,\sigma_2>0$ 为常数. 引入随机变量

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

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 $f_{Y|X}(y|x)$.

(2) 求条件概率 $\mathbb{P}(X\le 1|Y\le 1)$.

27. 设二维随机变量 (X,Y) 的密度函数为

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

(1) 求 c 的值;

(2) 求 (X,Y) 落在圆 $\{(x,y):x^2+y^2\le r^2\}$ $(r<R)$ 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 $\mathbb{E}(X|X+Y=1)$.

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

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

(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 $f_X(x) = \begin{cases} xe^{-x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$, 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 $\{(x, y) : 0 \leq x \leq 2, 0 \leq y \leq 2\}$ 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 $f_{X|Y}(x|y)$.

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 $f_{X|Y}(x|y)$;

(4) $\mathbb{E}(X|Y = y)$.

37. 设随机向量 (X, Y) 服从 $\{(x, y) : |x+y| \leq 1, |x-y| \leq 1\}$ 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 $\{(x, y) : 0 \leq x \leq 2, 0 \leq y \leq 1\}$ 内的均匀分布, 记

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

求 (U, V) 的联合概率分布.

39. 假设有 $n\ (n \geq 3)$ 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 $(X_2, \cdots, X_n)$ 的条件分布.

40. 设 (X, Y, Z) 服从单位球 $\{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$ 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 $(0, 1)$ 上的均匀分布，则 $\mathbb{P}(X^2 + Y^2 \leq 1)$=(　).
(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 $\mathbb{P}(X=k|X+Y=n)=$ _____, 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 是否相互独立.

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 $f_Y(y)=\begin{cases} 1, & 0\le y\le 1 \\ 0, & \text{其他} \end{cases}$, 记 $Z=X+Y$.
(1) 求 $\mathbb{P}(Z\le 1/2|X=0)$;
(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 $\{(x,y): x^2+y^2\le 1\}$ 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 $\mathbb{P}(X+Y\le 1)$.

55. 设 (X,Y) 的联合分布函数为 $F(x,y)=\begin{cases} \frac{(1-(x+1)e^{-x})y}{1+y}, & x>0,\ y>0, \\ 0, & \text{其他}. \end{cases}$

(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

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

(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

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

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. 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 $r\geq 1$ 为正整数, $0<p<1$, 即

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

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

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

   (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 $F(x) = 0.5\Phi(x) + 0.5\Phi(\frac{x-4}{2})$, 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

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

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

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

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 $n(n \geq 2)$ 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.

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

14. 设 X 为标准正态随机变量, 试求 X 的 α 分位数 $(0 < \alpha < 1)$.

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

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

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) $\lim\limits_{n\to\infty} \frac{\mathbb{E}[X_n]}{n \ln n}$.

20. 现有 $n(n \geq 1)$ 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 $x_1, x_2, \cdots, x_k (k \geq 2)$, 证明:

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

24. 设 X 为一随机变量, 若对任意实数 t, 期望 $\mathbb{E}[e^{tX}]$ 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 $N(\mu,\sigma^2)$.

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 $n\geq 1$, 计算 $\mathbb{E}[X^n]$.

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 $Y=e^X$ 和 $Z=\frac{1}{X}$ 的数学期望.

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

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

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, 最终则会亏损 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?

30. 设 X 为一随机变量, 它的符号函数定义为

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

    (1) 若 $X\sim U(-2,1)$, 试求 $\mathrm{Var}(\mathrm{sgn}(X))$.

    (2) 若 X 服从标准正态分布, 试求 $\mathbb{E}[\mathrm{sgn}(X)\cdot X]$.

31. 设随机变量 X 的密度函数为

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

    试求 $\mathbb{E}[\min\{|X|,1\}]$.

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 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 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 $(n\geq 3)$ 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) $\mathbb{E}[X_2|X_1=k]$ 和 $\mathrm{Var}(X_2|X_1=k)$.
(2) $\mathbb{E}[X_1+X_2]$ 和 $\mathrm{Var}(X_1+\cdots+X_k)$, $k=1,\ldots,n$.

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 $N(\mu,2\mu,\sigma^2,3\sigma^2,0.5)$, 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

$$\text{Sharpe Ratio}_{(\omega X+(1-\omega)Y)}=\frac{\mathbb{E}[\omega X+(1-\omega)Y]}{\sqrt{\mathrm{Var}(\omega X+(1-\omega)Y)}}$$

达到最大.

36. 设某两个风险 (X,Y) 服从二元正态分布 $N(\mu,2\mu,\sigma^2,2\sigma^2,\sqrt{2}/4)$, 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 $\mathbb{E}[(X+Y-3\mu)_+]$;
(2) 求到期收益的方差 $\mathrm{Var}((X+Y-3\mu)_+)$.

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 $[Y|X=k]\sim N(k,1)$.

(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 $\mathrm{Var}(X_1 + X_2 + X_3)$.

45. 设某人连续独立地投掷一枚均匀的骰子（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) 求投掷次数的期望；

    (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 $\mathbb{P}(X=i)=p_i\geq 0$, $i=1,\ldots,n$, $\sum_{i=1}^n p_i=1$. 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 $H(X)\leq H(U)$.

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

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

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 $\mathrm{Var}(W)=$ ________.

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 $\mathrm{Var}(Y)=$ ________.

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) $\mathrm{Var}(Y)$ 和 $\mathrm{Var}(Z)$.

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, $k\geq 1$, 以 X 表示所得号码之和，则 $\mathbb{E}[X]=$ ________; $\mathrm{Var}(X)=$ ________.

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 $\mathrm{Var}(2X+3Y)=$ ________.

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 $\mathrm{Var}(X)=9, \mathrm{Var}(Y)=16$，$\mathrm{Cov}(X,Y)=-3$，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 $\{(x,y):|x|+|y|\leq 1\}$ 上的均匀分布，求 $\mathrm{Cov}(X,Y)$.

61. 已知二元随机变量 (X,Y) 有概率分布如下：

X \ Y	-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 $\mathrm{Cov}(X,Y)$ 和 $\mathrm{Cov}(X^2,Y^2)$.

62. 掷两颗均匀骰子, 以 X 表示第一颗骰子掷出的点数, Y 表示两颗骰子所掷出的点数中的最大值.
(1) 求 X, Y 的数学期望与方差.　(2) 求 $\mathrm{Cov}(X,Y)$.

63. 设随机变量 X,Y 相互独立, 具有共同分布 $N(\mu,\sigma^2)$. 设 α,β 为两个常数.
(1) 求 $\mathrm{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)$.
(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 $(X,Y)\sim N(\mu,\mu,\sigma^2,\sigma^2,\rho)$, 其中 $\rho>0$. 问是否存在两个常数 α,β 使得 $\mathrm{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)=0$? 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 $\mathrm{Cov}(X,Y)=0$，则有（　）
A. $\mathrm{Var}(X+Y)=\mathrm{Var}(X)\mathrm{Var}(Y)$　　B．$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$
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 $\mathbb{S}=\{(x,y):|x|+|y|\leq 1\}$ 中的均匀分布.
(1) 求 $\mathrm{Cov}(X,Y)$.　(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 $\mathrm{Var}(X)=\mathrm{Var}(Y)=\sigma^2$. 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
$$\mathrm{Cov}(X_1,X_2)=\mathrm{Cov}(X_1,\mathbb{E}[X_2|X_1]).$$
(2) 假设存在常数 c, $\mathbb{E}[X_2|X_1]=1+cX_1$, 证明
$$c=\mathrm{Cov}(X_1,X_2)/\mathrm{Var}(X_1).$$

71. 若 $\mathbb{E}[X_2|X_1]=1$, 证明
$$\mathrm{Var}(X_1X_2)\geq \mathrm{Var}(X_1).$$

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

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

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
$$\mathbb{P}(N(t)=k)=e^{-\lambda t}\frac{(\lambda t)^k}{k!},\ \ k=0,1,2,\dots,$$
设 T 是一个均值为 a, 方差为 $b>0$ 的非负随机变量.
求 (1) $\mathrm{Cov}(T,N(T))$; (2) $\mathrm{Var}(N(T))$.

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 $\mathrm{Var}(X) = \sigma^2$ 已知，求 $\mathbb{E}[X|X > 0]$ 和 $\mathrm{Var}(X|X > 0)$.

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

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

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 $\mathbb{E}[X|Y = 2] =$ ________, $\mathbb{E}(XY^2 + Y|Y = 1) =$ ________.

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 $k \le m$, 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 $k \le m$, 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 $c > 0$,

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

第五章　极限理论

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 $\mathrm{Cov}(X, Y) = -5$. 则类似切比雪夫不等式可得到概率 $\mathbb{P}(XY > \epsilon)$， $\epsilon > 0$, 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
$$Y_n = \frac{X_1X_2 + X_3X_4 + \cdots + X_{2n-1}X_{2n}}{n}, \ \ n \in \mathbb{N}$$
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 $n \in \mathbb{N}$, $\mathrm{Cov}(X_n, Y_n) = 1$, 记
$$Z_n = \frac{X_1Y_1 + \cdots + X_nY_n}{n}, \ \ n \in \mathbb{N}$$
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 $\mathbb{E}[X_i^k] = \alpha_k, k = 1, 2, 3, 4,$ 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 0.5kg, 标准差为 0.1kg, 问 5000 只零件的总重量超过 2510kg 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

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

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.

第二部分

数理统计

第六章　数理统计的基本概念

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”?

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?

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?

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.

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.

9. 假设总体 X 服从两点分布 $B(1, p)$, 其中 p 为未知参数, $X = (X_1, \ldots, X_5)$ 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 $X_1 + X_2, \min\limits_{1 \leq i \leq 5} X_i, X_5 + 2p, X_5 - \mathbb{E}X_1, (X_5 - X_1)^2/\mathrm{Var}(X_1)$ 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.

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.

11. 随机地取 8 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.

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) $Y \sim F_{1,n}$

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

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 $Z=\dfrac{\sqrt{2}(Y_1-Y_2)}{S}$ 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

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

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.

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 $S_n^2=\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2$ 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

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

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 $\mathbb{P}(X_{(n)}\geq 0.99)\geq 0.95$?

(2) 求极差 $R_n=X_{(n)}-X_{(1)}$ 的期望;

第七章　参数估计

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 $0<\theta<\frac{1}{2}$ 为未知参数. 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 $0<p_1,p_2<1$ 为未知参数. 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 $X_1,\ldots,X_n$ 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) $p(x;\theta)=\frac{1}{\theta},\quad x=0,1,2,\ldots,\theta-1$, 其中 θ (正整数)是未知参数.

(2) $p(x;\theta)=\binom{m}{x}\theta^x(1-\theta)^{m-x},\quad x=0,1,\ldots,m.$

(3) $p(x;\theta)=(x-1)\theta^2(1-\theta)^{x-2},\quad x=2,3,\ldots;0<\theta<1.$

(4) $p(x;\theta)=-\frac{1}{\ln(1-\theta)}\frac{\theta^x}{x},\quad x=1,2\ldots.$

(5) $p(x;\theta)=\frac{\theta^x}{x!}e^{-\theta},\quad x=0,1,2,\ldots.$

4. 设 $X_1,\ldots,X_n$ 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

$$f(x)=\begin{cases}\dfrac{4x^2}{\theta^3\sqrt{\pi}}e^{-\frac{x^2}{\theta^2}}\ , & x\geq 0\\ 0 & \text{其他}\ .\end{cases}$$

设 $X_1,X_2,\ldots,X_n$ 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

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

$X_1,X_2,\ldots,X_n$ 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 $X_1,X_2,\ldots,X_n$ 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未引起严重事故的概率 p 的极大似然估计.

r	0	1	2	3	4	5	≥ 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	≥ 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

$$f(x)=\begin{cases}\frac{1}{\lambda}e^{-x/\lambda}, & 0<x<\infty\\ 0, & x\leq 0\end{cases}$$

从这批产品中抽取 n 个作寿命试验, 规定到第 r 个 $(0<r\leq n)$ 电子元件失效时就停止试验. 这样获得前 r 个次序统计量 $X_{(1)}\leq X_{(2)}\leq\ldots\leq X_{(r)}$ 和 n 个电子元件总试验时间 $T=\sum\limits_{i=1}^{r}X_{(i)}+(n-r)X_{(r)}$.

(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 $\theta>0$ 为未知参数. 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$X_1,\ldots,X_n$，试求未知参数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 $-\infty<\mu<+\infty$, $\sigma^2>0$ 为未知参数. 求 $\theta=\mathbb{P}(X\geq 2)$ 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

$$f(x;\theta)=\begin{cases}\theta, & 0<x<1\\ 1-\theta, & 1\le x<2\\ 0, & \text{其它}.\end{cases}$$

其中 $0<\theta<1$ 为未知参数. 现从该总体中简单随机样本 $X_1,\dots,X_n$, 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

$$F(x;\theta)=\begin{cases}1-e^{-x^2/\theta}, & x\ge 0\\ 0, & \text{其他}\end{cases}$$

其中 θ 为未知参数且大于零, $X_1,X_2,\dots,X_n$ 为来自总体 X 的简单随机抽样.

(1) 求 $\mathbb{E}X,\mathbb{E}[X^2]$;

(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 $\lim\limits_{n\to\infty}\mathbb{P}(|\hat{\theta}-a|\ge\epsilon)=0$?

18. 设总体的数学期望为 μ, $X_1,\dots,X_n$ 是来自总体 X 的样本. 假设 $a_1,\dots,a_n$ 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 $T=\max(X_1,X_2,X_3)$.

(1) 求 T 的概率密度;

(2) 确定 a, 使得 aT 为 θ 的无偏估计.

20. 设 $X_1,\dots,X_n$ 是来自总体 X 的一个简单随机样本, $\mathbb{E}X=\mu,\mathrm{Var}(X)=\sigma^2$.
(1) 确定常数 c 使得 $c\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计.
(2) 记 $\overline{X}$, S^2 分别是样本均值和样本方差. 确定常数 c 使 $\overline{X}^2-cS^2$ 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 $\overline{X}_1$, $\overline{X}_2$ 分别是两样本的均值. 试证明对于任意常数 a, $Y = a\overline{X}_1 + (1-a)\overline{X}_2$ 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 σ_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 $\hat{\theta} = \sum_{i=1}^k a_i X_i$ 估计 θ 时, $\hat{\theta}$是 无偏的, 并且 $\mathrm{Var}(\hat{\theta})$ 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 2 题, 并假定 $p_2 = 2p_1 = 2p$. 记 p 的矩估计为 $\hat{p}$, 现定义
$$\hat{p}_1 = \frac{n_1}{n}, \quad \hat{p}_2 = \frac{n_2}{2n}, \quad \hat{p}_3 = \frac{1}{3}(1 - \frac{n_3}{n}).$$
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 $\theta \in (0, 1)$ 未知, 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掷两次, 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 $X_1,...,X_n$ 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 是未知参数

(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 $\widetilde{\theta}_1$, $\widetilde{\theta}_2$?

(3) 比较 $\widetilde{\theta}_1$ 与 $\widetilde{\theta}_2$ 何者为优 (即方差较小).

29. 设样本 $X_1,...,X_n$ 抽自正态分布 $N(0,\sigma^2)$, 求 σ 的矩估计:
(1) 利用 $\mathbb{E}|X_1|=\sqrt{\frac{2}{\pi}}\sigma$; (2) 利用 $\sigma=\sqrt{\mathrm{Var}(X_1)}$.

30. 设 $X_1,...,X_n$ 是来自正态总体 $X\sim N(a,\sigma^2)$ 的简单随机样本, 求 $\mathbb{P}(X>1)$ 的矩估计量.

31. 若 $Y=\mathrm{e}^X$, 而 $X\sim N(a,\sigma^2)$, 则随机变量 Y 的分布称为对数正态分布. 设 $Y_1,...,Y_n$ 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 为未知参数. 设 $X_1,...,X_n$ 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 $X_1,...,X_n$ 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 $X_1,...,X_n$ 为从总体 X 中抽取的随机样本, X 服从均匀分布 $U(\theta-\frac{1}{2},\theta+\frac{1}{2})$, 求 θ 的极大似然估计.

35. 设 $X_1,...,X_n$ 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 $N(\mu_1,\sigma^2)$ 和 $N(\mu_2,\sigma^2)$ 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 $N(\mu,\sigma^2)$ 和 $N(\mu,2\sigma^2)$ 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, 从中捞出 1000 条, 标上记号后放回湖中, 然后再捞出 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

$$f(x,\theta)=\mathrm{e}^{-(x-\mu)}, x\geq\mu, -\infty<\mu<+\infty$$

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 $N(a,\sigma^2)$ 的一个简单随机样本, $S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\bar{X})^2$ 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 $Exp(\lambda)$ 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 $U(\theta_1,\theta_2)$ 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,

$$f(x,\ \theta)=\begin{cases}1, & \text{当}\ \theta-\frac{1}{2}\leq x\leq\theta+\frac{1}{2}\\ 0, & \text{其他}\end{cases},\quad -\infty<\theta<+\infty.$$

证明: 样本均值 $\bar{X}$ 及 $\max\limits_{1\leq i\leq n}X_i+\frac{n+3}{2(n+1)}$ 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 $\frac{4}{3}\max\limits_{1\le i\le 3} X_i$ 及 $4\min\limits_{1\le i\le 3} X_i$ 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 $\mathrm{Var}(a\bar{X}_1+b\bar{X}_2+c\bar{X}_3)$ 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, $\hat{\theta}_1 = a_n\bar{X}$, $\hat{\theta}_2 = b_n\min(X_1,\ldots,X_n)$ 和 $\hat{\theta}_3 = c_n\max(X_1,\ldots,X_n)$ 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 $X \sim Exp(\theta^{-1})$, $X_1, \ldots, X_n$ 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 $\hat{\theta}_2 = b_n\min(X_1,\ldots,X_n)$ 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

$$f(x,\theta) = \begin{cases} e^{-(x-\theta)}, & x \ge \theta \\ 0, & x < \theta \end{cases}$$

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 $\mathrm{Var}(\hat{\theta}_n) > 0$ 且 $\lim\limits_{n\to\infty} \mathrm{Var}(\hat{\theta}_n) = 0$. 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 $Exp(\lambda)$ 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.

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?

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	93.8	102.0	73.3
后	72.5	33.5	132.1	129.8	121.2	54.0	104.6	92.6	119.4	84.7	85.1

(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 未知.

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 未知.

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未知, 求r/N 的 $1-\alpha$ 置信区间;

(2) 若只有N未知, 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　假设检验

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_α 为标准正态的上 α 分位数, 现有4个拒绝域： $V_1 = \{4|\bar{X}| \geq u_{0.05}\}$, $V_2 = \{4|\bar{X}| \leq u_{0.45}\}$, $V_3 = \{4\bar{X} \geq u_{0.10}\}$, $V_4 = \{4\bar{X} \leq -u_{0.10}\}$.

(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 $\{(X_1, \cdots, X_n) : \ \sum_{i=1}^{10} X_i \leq 1 \text{ 或 } \geq 12\}$

(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

$$H_0: \quad \theta \geq 3 \quad \longleftrightarrow \quad H_1: \quad \theta < 3,$$

拒绝域取为 $W=\{X_{(n)}=\max\{X_1,\cdots,X_n\}\leq 2.5\}$,
(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 $Exp(\lambda)$ 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 $N(\mu,5^2)$ 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) $H_0:\mu\geq 7 \leftrightarrow H_1:\mu<7$.
(2) $H_0:\mu\leq 2 \leftrightarrow H_1:\mu>2$;

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120cm.
    (2) 这所小学一年级男孩平均身高大于120cm.
    (3) 这所小学一年级女孩平均身高大于120cm.

16. 在 5% 显著水平下, 试检验第七章66题,
    (1) 包月客户一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1600h, 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: 毫克) 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. 在方差未知的情况下, 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皂厚度为 5cm. 今欲了解机器性能是否良好, 随机抽取 10 块肥皂为样本, 测得平均厚度为 5.3cm, 标准差为 0.3cm, 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) $H_0: \mu \leq 40 \leftrightarrow H_1: \mu > 40$;
(2) $H_0: \sigma^2 \geq 70 \leftrightarrow H_1: \sigma^2 < 70$.

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 $H_0: \sigma \leq 0.01 \leftrightarrow H_1: \sigma > 0.01$ ：
(1) $\mu = 2.12$　　(2) μ 未知.

31. 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

32. 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

33. 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

34. 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(α=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

35. 一个以减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?

36. 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?

37. 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(α=0.05)?

38. 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	$n_2 = 25$	$s_2^2 = 12$

40. 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(α=0.05)?

41. 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买力越高. 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		个人	购买力得分	
	之后	之前		之后	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

42. 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

$$H_0 : \sigma_1^2 \geq \sigma_2^2 \longleftrightarrow H_1 : \sigma_1^2 < \sigma_2^2.$$

43. 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 ($\alpha = 0.1$)?

44. 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(α=0.05)?

甲矿	24.3	20.8	23.7	21.3	17.4
乙矿	18.2	16.9	20.2	16.7	

45. 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(α=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.

47. 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

48. 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 名研究志愿者组成一个样本, 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?

49. 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?

50. 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		零售点	价格/美元	
	豪华型	普通型		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?

51. 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(α=0.05)?

52. 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($\alpha = 0.05$)?

53. 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(α=0.05)?

54. 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(α=0.05)?

55. 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?

56. 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(α=0.01)?

57. 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?

58. 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性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?

59. 袋中装有 8 个球, 其中红球数未知. 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.

60. 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.

61. 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

62. 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?($\alpha = 0.05$)

63. 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?($\alpha = 0.05$)

64. 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	≥ 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.

65. 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 $N(60, 15^2)$.

66. 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 $\alpha = 0.05$).

67. 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 $\alpha = 0.05$)

68. 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

工厂 \ 等级	1	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.

69. 某公司是一家较大的高尔夫球制造商. 管理人员认为：引进某种耐磨损、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?

70. 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					

计算机辅助方法										
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.

71. 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 $\alpha = 0.1$).

73. 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	$20 \le x \le 30$	$30 < x \le 40$	$40 < x \le 50$	$50 < x \le 60$
学生数	5	15	30	51
分数 x	$60 < x \le 70$	$70 < x \le 80$	$80 < x \le 90$	$90 < x \le 100$
学生数	60	23	10	6

74. 袋中装有 8 个球, 其中红球数未知. 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

$$P\{X = k\} = \binom{5}{k}\binom{3}{3-k} \Big/ \binom{8}{3},\quad k = 0, 1, 2, 3.$$

即检验假设 H_0 : 红球的个数为 5.

75. 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、鲈鱼、竹夹鱼和鲇鱼的鱼苗, 现在在鱼塘里获得一个样本如下:

序号	1	2	3	4
种类	鲑鱼	鲈鱼	竹夹鱼	鲇鱼
数量 (条)	132	100	200	168

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.

第九章　回归分析

1. 下表数据是退火温度 x ($^\circ$ C)对黄铜延性 Y (%) 效应的试验结果, Y 是以延长度计算的.

x	300	400	500	600	700	800
y	40	50	55	60	67	70

画出散点图并求 Y 对 x 的线性回归方程.

2. 蟋蟀用一个翅膀在另一个翅膀上快速地滑动, 从而发出吱吱喳喳的叫声. 生物学家知道叫声的频率 x 与气温 Y 具有线性关系. 下表列出了 15 对频率与气温间的对应关系的观察结果:

频率 x_i (叫声数/秒)	20.0	16.0	19.8	18.4	17.1	15.5	14.7	17.1
气温 y_i ($^\circ$ C)	31.4	22.0	34.1	29.1	27.0	24.0	20.9	27.8
频率 x_i (叫声数/秒)	15.4	16.2	15.0	17.2	16.0	17.0	14.4	
气温 y_i ($^\circ$ C)	20.8	28.5	26.4	28.1	27.0	28.6	24.6	

试求 Y 关于 x 的线性回归方程.

3. 以 x 与 Y 分别表示人的脚长 (英寸) 与手长 (英寸), 下面列出了 15 名女子的脚的长度 x 与手的长度 Y 的样本值:

x	9.00	8.50	9.25	9.75	9.00	10.00	9.50	9.00
y	6.50	6.25	7.25	7.00	6.75	7.00	6.50	7.00
x	9.25	9.50	9.25	10.00	10.00	9.75	9.50	
y	7.00	7.00	7.00	7.50	7.25	7.25	7.25	

试求:

(1) Y 关于 x 的线性回归方程 $\hat{y} = \hat{a} + \hat{b}x$.

(2) 求 b 的置信水平为 0.95 的置信区间.

4. 某公司进行一项工资情况的调查,并且把调查资料的摘要登载在其网站上. 根据 2012 年 10 月 1 日的工资数据, 该公司报告, 销售副经理的平均年薪为 142 111 美元, 包括年平均奖金 15 432 美元在内. 假设由 10 名销售副经理组成一个样本, 他们年工资和奖金的统计数据如下.

销售副经理	年薪/千美元	奖金/千美元	销售副经理	年薪/千美元	奖金/千美元
1	135	12	6	176	24
2	115	14	7	98	7
3	146	16	8	136	17
4	167	19	9	163	18
5	165	22	10	119	11

(1) 以工资为自变量, 画出这些数据的散点图.

(2) 根据散点图. 在年薪和奖金之间显示出什么关系?

(3) 利用最小二乘法, 求出估计的回归方程.

(4) 对估计的回归方程的斜率作出解释.

(5) 如果某销售副经理的年薪是 120 000 美元, 预测他的奖金.

5. 您是否会想到, 可靠性越高的轿车价格越贵? 某媒体评估了 15 部高档轿车. 可靠性的评估采用 5 分制：差 (1), 一般 (2), 好 (3), 很好 (4) 和优秀 (5). 15 部高档轿车的价格和可靠性的评估结果如下表所示.

品牌和型号	可靠性	价格（美元）	品牌和型号	可靠性	价格（美元）
A	4	33150	I	4	34390
B	3	40570	J 135	5	33845
C	5	35105	K	3	36910
D	5	35174	L	4	34695
E	1	42230	M	1	37995
F	3	38225	N	3	36995
G	2	37605	O	3	33890
H	1	37695			

试

(1) 以可靠性评估分为自变量, 建立这些数据的散点图.

(2) 建立最小二乘估计的回归方程.

(3) 根据你的分析, 你是否认为可靠性越高的轿车, 它的价格越贵? 请做出解释.

(4) 如果一辆高档轿车的可靠性评估分是 3, 估计该车的价格.

6. 某媒体对于超过 100HD 高清晰度电视机提供了广泛的测试和评级. 对于每一种型号的高清晰度电视机, 主要根据画面质量进行测试并给出一个总分. 一般情况下, 较高的总分意味着较好的性能.下面是 10 台 42 英寸的电视机的价格和总分的数据.

品牌	价格	得分	品牌	价格	得分
A	2800	62	F	2000	39
B	2800	53	G	4000	66
C	2700	44	H	3000	55
D	3500	50	I	2500	34
E	3300	54	J	3000	39

试

(1) 以价格为自变量, 利用这些数据, 42 英寸电视机总分的估计的回归方程.

(2) 计算 r^2. 估计的回归方程能否为这些数据提供一个好的拟合?

(3) 若某台 42 英寸等离子电视机的价格是 3 200 美元, 估计该台电视机的总分.

7. 某杂志检测了 10 种不同型号供徒步旅行者使用的登山靴. 下表中的数据是被检测的每一种型号的登山靴的支撑力和价格的统计资料. 支撑力用 $1\sim5$ 的等级分来测量, 等级分 1 表示一般水平的支撑力, 等级分 5 表示最好的支撑力.

厂家	支撑力	价格	厂家	支撑力	价格
A	2	120	F	5	189
B	3	125	G	5	190
C	3	130	H	4	195
D	3	135	I	4	200
E	3	150	J	5	220

(1) 利用表中数据建立估计的回归方程, 使该方程能在支撑力的等级分已知时, 用来估计登山靴的价格.

(2) 在 $\alpha = 0.05$ 的显著性水平下, 确定支撑力的等级分和登山靴的价格是否相关.

(3) 在支撑力的等级分已知时, 利用在 (1) 中建立的估计的回归方程来估计登山靴的价格, 你觉得合适吗?

(4) 对于支撑力的等级分为 4 的登山靴, 估计它的价格.

8. 四季餐厅的广告费支出和收入的数据如下.

广告费	收入	广告费	收入	广告费	收入
1	19	6	40	20	54
2	32	10	52		
4	44	14	53		

(1) 设 x 表示广告费支出, y 表示收入. 利用最小二乘法, 求出一条近似这两个变量之间关系的直线.

(2) 在 $\alpha = 0.05$ 的显著性水平下, 检验收入和广告费支出是否相关.

(3) 为准备绘制 $y - \hat{y}$ 关于 $\hat{y}$ 的残差图, 利用 (a) 中的结果, 先求出 $\hat{y}$ 的值.

(4) 从残差分析中你能得出什么结论? 你是应用这个模型呢? 还是寻找一个更好的模型?

9. 存股证 (X) 是某交易所代理一家外国公司股份的证书, 这家外国公司在其本国的银行里持有一定数量的保证金.下表列出了 10 家可能有新的存股证 (X) 的印度公司的价格收益比 (P/E) 和投资收益率 (ROE) 的数据.

公司名称	ROE	P/E	公司名称	ROE	P/E
A	6.43	36.88	F	46.23	95.59
B	13.49	27.03	G	28.9	54.85
C	14.04	10.83	H	54.01	189.21
D	20.67	5.15	I	28.02	75.86
E	22.74	13.35	J	27.04	13.17

(1) 假设 $x = ROE$, $y = P/E$, 利用计算机软件包求出关于 x 和 y 的估计的回归方程.

(2) 绘出标准化残差关于自变量 x 的残差图.

(3) 根据残差图, 你觉得关于误差项和模型形式的假定合理吗?

10. 下面是 10 个主要啤酒品牌的广告费和销售量的数据.

啤酒品牌	广告费（百万美元）	装运量（百万桶）	啤酒品牌	广告费（百万美元）	装运量（百万桶）
A	120	36.3	F	0.1	7.1
B	68.7	20.7	G	21.5	5.6
C	100.1	15.9	H	1.4	4.4
D	76.6	13.2	I	5.3	4.3
E	8.7	8.1	J	1.7	4.3

(1) 求出这些数据的估计的回归方程.

(2) 应用残差分析查明是否存在任何异常值和（或）有影响的观测值. 简短概括一下你的发现和结论.

11. 某大学的一名市场营销学教授感兴趣的问题是, 在一门课程中学生用在学习上的时间和所取得的总学分之间的关系. 搜集了 10 名学生攻读最后一个学期课程的有关数据如下.

用时	总学分	用时	总学分
45	40	65	50
30	35	90	90
90	75	80	80
60	65	55	45
105	90	75	65

(1) 求出表明学生取得的总学分是如何依赖用在学习上的时间的估计的回归方程.

(2) 在 $\alpha = 0.05$ 的显著性水平下, 检验模型的显著性.

(3) 如果某同学用在学习上的时间是 95 小时, 预测他的总学分.

(4) 求出该同学总学分的一个置信水平为 95% 的预测区间.

12. 某公共服务委员会发表的 2012 ~ 2013 年度国家服务报告报道了雇员对工作满意程度的评估情况. 其中的一个调查问题是要求雇员选择 5 个最重要的工作场所因素(从各种因素的清单中), 即他们对自己的工作满意程度影响最大的 5 个因素. 受访者被要求用 5 个因素来表明满意程度.下面的数据给出了雇员前 5 个因素所占的百分比, 以及用雇员在他们目前工作的工作场所对选择的的前 5 个因素是“非常满意”或“满意”所占的百分比来度量相应的满意度评价.

工作场所因素	前5个因素 (%)	满意度评价 (%)
适当的工作量	30	49
创新的机会	38	64
为社会做贡献	40	67
职责明确	40	69
工作时间灵活	55	86
良好的同事关系	60	85
工作趣味性	48	74
职业发展机会	33	43
提升个人能力	46	66
发挥个人特长	50	70
个人努力的肯定	42	53
薪资	47	62
实现个人价值	42	69

(1) 用 5 个因素 (%) 为横轴, 满意度评价 (%) 为纵轴, 画出这些数据的散点图.
(2) 根据在 (1) 中做出的散点图, 在这两个变量之间显出什么关系?
(3) 求出估计的回归方程, 试这个方程能在 5 个因素 (%) 已知时, 用来对满意度评价 (%) 作出预测.
(4) 在 $\alpha = 0.05$ 的显著性水平下, 检验这两个变量之间关系的显著性.
(5) 估计的回归方恒对观测数据的拟合好吗? 请做出解释.
(6) 样本相关系数是多少?

13. 在钢线碳含量对于电阻的效应的研究中, 得到以下的数据:

碳含量 x (%)	0.10	0.30	0.40	0.55	0.70	0.80	0.95
20° C 时电阻 y ($\mu\Omega$)	15	18	19	21	22.6	23.8	26

(1) 画出散点图.

(2) 求线性回归方程 $\hat{y} = \hat{a} + \hat{b}x$.

(3) 求随机误差 ε 的方差 σ^2 的无偏估计.

(4) 检验假设 $H_0 : b = 0,\ H_1 : b \neq 0$.

(5) 若回归效果显著,求 b 的置信水平为 0.95 的置信区间.

(6) 求 $x = 0.50$ 处 $\mu(x) = a + bx$ 的置信水平为 0.95 的置信区间.

(7) 求 $x = 0.50$ 处观察值 Y 的置信水平为 0.95 的预测区间.

14. 槲寄生是一种寄生在大树上部树枝上的寄生植物. 它喜欢寄生在年轻的大树上. 下面给出在一定条件下完成的试验中采集的数据:

大树的年龄 x (年)	3	4	9	15	40
每株大树上槲寄生的株树 y	28	10	15	6	1
	33	36	22	14	1
	22	24	10	9	

(1) 作出 (x_i, y_i) 的散点图.

(2) 令 $z_i = \ln y_i$,作出 (x_i, z_i) 的散点图.

(3) 以模型 $Y = ae^{bx}\varepsilon, \ln\varepsilon \sim N(0, \sigma^2)$ 拟合数据, 其中 a, b, σ^2 与 x 无关. 试求曲线回归方程 $\hat{y} = \hat{a}\exp(\hat{b}x)$.

15. 一种合金在不同浓度的某种添加剂下, 各做三次试验, 得数据如下:

浓度 x	10.0	15.0	20.0	25.0	30.0
抗压强度 y	25.2	29.8	31.2	31.7	29.4
	27.3	31.1	32.6	30.1	30.8
	28.7	27.8	29.7	32.3	32.8

(1) 作散点图.

(2) 以模型 $Y = b_0 + b_1x + b_2x^2 + \varepsilon, \varepsilon \sim N(0, \sigma^2)$ 拟合数据, 其中 $b_0,\ b_1,\ b_2,\ \sigma^2$ 与 x 无关. 求回归方程 $\hat{y} = \hat{b}_0 + \hat{b}_1x + \hat{b}_2x^2$.

16. 某种化工产品的得率 Y 与反应温度 x_1 、反应时间 x_2 及某反应物浓度 x_3 有关. 今得试验结果如下表所示. 其中 $x_1,\ x_2,\ x_3$ 均为二水平且均以编码形式表达.

x_1	-1	-1	-1	-1	1	1	1	1
x_2	-1	-1	1	1	-1	-1	1	1
x_3	-1	1	-1	1	-1	1	-1	1
得率	7.6	10.3	9.2	10.2	8.4	11.1	9.8	12.6

(1) 设 $\mu(x_1, x_2, x_3) = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$,求 Y 的多元线性回归方程.

(2) 若认为反应时间不影响得率, 即认为

$$\mu(x_1, x_2, x_3) = \beta_0 + \beta_1 x_1 + \beta_3 x_3,$$

求 Y 的多元线性回归方程.