《概率论与数理统计》（缪柏其、张伟平）参考答案

٩(□•̀ω•́□)۶

“我们无法选择命运赐予什么，却能选择追逐怎样的明天！”

Jinghao Chen PB23010359（生僻字实在打不出来 ΩДΩ）

最后更新时间：**2025** 年 **6** 月 **30** 日

序

了解到大家貌似没有这本书的参考答案, 为了提高大家的学习效率, 特意制作了这份参考答案, 希望能对大家有所帮助 („ · ω · „). 我尽量让每道题都保持和原书一致, 不过也有个别字符存在微调, 以便阅读通畅. 因为电子版的图有点糊, 所以有些简单的图表可能我就自己做了, 但是如果有些图实在难做可能就只能从电子版截图放上去了, 希望大家见谅. 以及由于课程只上到第九章, 所以第十章的习题还没有解答, 如果有同学能提供答案的话欢迎补充.

在此感谢___________________ 对这份答案的检查校对.

注 1 如果发现解答有问题或有其他建议, 欢迎发邮件至 chenjinghao@mail.ustc.edu.cn 或 stardust.math26@gmail.com.

注 2 "➢"的后面是我个人对于该题相关的知识点的整理或者理解, 一般除了基本定义以外的性质都会加上这些注释, 方便对课本或者相关知识不熟的同学进行学习. 可能有些地方不够严谨, 欢迎大家批评指正. (不出意外的话我应该还会在担任这门课助教前或者期间，根据本班进度做一份讲义，可能也会解释部分习题中的内容.)

注 3 封面用的是 Wallpaper 中的" 山景", 模板改编自https://www.latexstudio.net/index/details/index/mid/4438.html。如果有侵权，请联系我删除 (╥~~╥).

目录

第 1 章　事件及其概率

1. 写出下列各试验的样本空间及指定事件的样本点.
(1) 连续两次掷骰子, A={第一次掷出的点数比第二次的大}, B={两次点数相等}, C={两次点数之和为 10};
(2) 连续 3 次掷硬币, 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 且大于 1/3}.

解:

(1) 样本空间: $\Omega=\{(i,j)|i,j\in\{1,2,\cdots,6\}\}$;

样本点: $\omega(A)=(i,j)$, $i>j$; $\omega(B)=(i,i)$, $i=1,2,\cdots,6$; $\omega(C)=(i,j)$, $i+j=10$.

(2) 样本空间: $\Omega=\{(a,b,c)|(a,b,c)\in\{0,1\}^3\}$;

样本点: $\omega(A)=(0,b,c)$, $(b,c)\in\{0,1\}^2$; $\omega(B)=(a,b,c)$, $a+b+c=2$; $\omega(C)=(a,a,a)$, $a\in\{0,1\}$.

(3) 样本空间: $\Omega=\{(x,y)|x^2+y^2\le 1\}$;

样本点: $\omega(A)=(x,y)$, $x^2+y^2<\frac{1}{4}$; $\omega(C)=(x,y)$, $\frac{1}{9}<x^2+y^2<\frac{1}{4}$.

2. 某炮弹射击目标 3 次, 记 A_i={第 i 次击中目标}(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.

解:
(1) $A_1\overline{A_2}\,\overline{A_3}\cup\overline{A_1}A_2\overline{A_3}\cup\overline{A_1}\,\overline{A_2}A_3$;
(2) $A_1\cup A_2\cup A_3=\overline{\overline{A_1}\,\overline{A_2}\,\overline{A_3}}$;
(3) $A_1\cap(A_2\cup A_3)$;
(4) $\overline{A_1}\,\overline{A_2}\,\overline{A_3}\cup A_1\overline{A_2}\,\overline{A_3}\cup\overline{A_1}A_2\overline{A_3}\cup\overline{A_1}\,\overline{A_2}A_3$.

3. 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: (1) $A\overline{B}$; (2) $\overline{A}\cup B$; (3) $\overline{AB}$; (4) $\overline{\overline{A}\,\overline{B}}$.

解:
(1) $A\overline{B}=\emptyset$; (2) $\overline{A}\cup B=\Omega=[0,2]$; (3) $\overline{AB}=[0,\frac{1}{2}]\cup(1,2]$; (4) $\overline{\overline{A}\,\overline{B}}=A\cup B=(\frac{1}{4},\frac{3}{2}]$.

4. 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 人喜欢奶糖, 570 人喜欢巧克力糖和夹心糖, 356 人喜欢巧克力糖和奶糖, 348 人喜欢夹心糖和奶糖, 以及 297 人喜欢全部三种糖果. 证明这一消息有误.

解:
记 A, B, C 分别为喜欢巧克力糖, 夹心糖, 奶糖, 则 $|\Omega|=1000$, $|A|=811$, $|B|=752$, $|C|=418$.
由容斥原理:

$$|A\cup B\cup C|=|A|+|B|+|C|-|A\cup B|-|A\cup C|-|B\cup C|+|A\cap B\cap C|$$

$\Longrightarrow |A\cup B\cup C|=1004>1000=|\Omega|=|A|+|B|+|C|$ 矛盾.
➢ 集合势的容斥原理: $|\bigcup_{i=1}^{n}A_i|=\sum_{k=1}^{n}(-1)^{k-1}\sum_{1\le i_1<\cdots<i_k\le n}|A_{i_1}\cap A_{i_2}\cdots\cap A_{i_k}|$.
➢ 事实上还要检查子集关系, 即 $|A\cap B|\le\min\{|A|,|B|\}$.

5. 一小区居民订阅报纸的统计数字如下: 订甲种报纸的占 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) 只订甲种报纸的; (2) 只订一种报纸的; (3) 至少订一种报纸的; (4) 两种报纸都不订的.

解:
记 A, B 分别为订甲种报纸, 订乙种报纸, 则 $\frac{|A|}{|\Omega|}=0.4$, $\frac{|B|}{|\Omega|}=0.25$, $\frac{|AB|}{|\omega|}=0.15$.
(1) $p_1=\frac{|A\setminus B|}{|\Omega|}=\frac{|A|-|AB|}{|\Omega|}=0.25$;
(2) $p_2=\frac{|(A\cup B)\setminus AB|}{\Omega}=0.35$;
(3) $p_3=\frac{|A\cup B|}{|\Omega|}=\frac{|A|+|B|-|AB|}{|\Omega|}=0.5$;
(4) $p_4=1-p_3=0.5$.

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

解:
(1) $P(AB)_{\max}=0.7$, 当且仅当 $A\subset B$ 时取得最大值;
(2) $P(AB)_{\min}=0.5$, 当且仅当 $A\cup B=\Omega$ 时取得最小值.
➢ 第 7 题的特殊情况.

7. 设 $P(A)=a>0, P(B)=b>0$, 试确定 $P(AB)$ 的取值范围.

解:
$P(AB)\in[\max\{0,a+b-1\},\min\{a,b\}]$.
➢ 省去了讨论的过程, 可以画 venn 图理解.

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

解:
$ABC\subset AC \Longrightarrow P(ABC)=0 \xRightarrow{\text{容斥原理}} P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)=\frac{3}{4}$.

9. 一个班有 30 个同学, 求 12 个月中有 6 个月恰好包含 2 个人生日, 另外 6 个月恰好包含 3 个人生日的概率.

解:
$P=\dfrac{\binom{12}{6}\cdot\frac{30!}{(2!)^6(3!)^6}}{12^{30}}\approx 3.46\times10^{-4}$.
➢ 将不同的同学安置在不同的月份内，选定六个月后, 同一个月内的 n 个人前后排列等价, 故除以 $n!$, 即“多重排列”问题.

10. 一停车场的 A 区域一排有 12 个停车位, 某人发现其中有 8 个停车位停了车, 而 4 个空的停车位是连着的, 这种现象是随机的吗?

解:
$P=\dfrac{12-4+1}{A_{12}^4}\approx 1.82\%\in(0.01,0.05]\Longrightarrow$ 从宽标准看, 该现象是小概率事件; 从严标准看, 该现象是随机的.

➢ 有关小概率事件及宽、严标准, 见课本 P14.

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

解:

记 $A_i, 1\le i\le 4$ 分别为缺少梅花、方块、红心、黑桃的牌, 则由容斥原理: $P=1-\frac{\bigcup\limits_{i=1}^{4}A_i}{\binom{52}{10}}=1-\frac{4\binom{39}{10}-6\binom{26}{10}+4\binom{13}{10}}{\binom{52}{10}}\approx 0.9988.$

12. 甲、乙两选手进行乒乓球单打比赛, 已知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?

解:
三局两胜制下, 甲胜的概率为: $p_1=p^2+2(1-p)p^2$
五局三胜制下, 甲胜的概率为: $p_2=p^3+3(1-p)p^3+\binom{4}{2}p^3(1-p)^2$
观察到 $p_2-p_1=3p^2(p-1)^2(2p-1)>0$, 故 $p>\frac{1}{2}$ 时, 五局三胜制对甲更有利.
➢ 这个现象的直观理解是甲的平均水平高于乙, 因此比赛场次越多则偶然性越小, 本质涉及到后面的大数定律以及中心极限定理.

13. * 甲、乙二人约定了这样一个赌博规则: 有无穷多个盒子, 编号为 n 的盒子中有 n 个红球、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, 若抽出白球算甲胜, 否则乙胜. 你认为这规则对谁更有利?

解:

记随机变量 N: 甲掷硬币直到第一次出现正面所需的次数, 事件 A: 甲从盒子中抽到白球 (即甲胜), 则

$$P(N=n)=(\frac{1}{2})^{n-1}\cdot\frac{1}{2}=(\frac{1}{2})^n,\qquad P(A|N=n)=\frac{1}{n+1}$$

由全概率公式:

$$P(A)=\sum_{n=1}^{\infty}P(N=n)\cdot P(A|N=n)=\sum_{n=1}^{\infty}\frac{1}{2^n}\cdot\frac{1}{n+1}$$

① 放缩估计:

$P(A|N=n)=\frac{1}{n+1}\le\frac{1}{2}$ 且当 $n\ge 2$ 时, 不等号严格成立, 故 $P(A)<\frac{1}{2}\sum\limits_{n=1}^{\infty}\frac{1}{2^n}=\frac{1}{2}$. 因此, 该规则对甲不公平.

② 直接计算:

由泰勒展开式: $\sum\limits_{k=1}^{\infty}\frac{x^k}{k}=-\ln(1-x)$, $|x|<1$, 令 $x=\frac{1}{2}$, 即得 $P(A)=2\sum\limits_{m=2}^{\infty}\frac{(\frac{1}{2})^m}{m}=2\ln 2-1\approx 0.3862<0.5$. 因此, 该规则对甲不公平.

➢ 全概率公式; 概率估计/泰勒展开.

14. * 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) 每个人恰有一个邻座;

(3) 关于中央对称的两个座位至少有一个空着.

解:

总坐法数: $A_N^n=\frac{N!}{(N-n)!}$.

(1) 任何人都没有邻座, 即在 $N-n$ 个空位中插入 n 个人, 则坐法数为: $A_{N-n+1}^n \Longrightarrow$ 若 $N\ge 2n-1$, 则 $p_1=\frac{A_{N-n+1}^n}{A_N^n}=\frac{(N-n+1)!(N-n)!}{(N-2n+1)!N!}$, 否则 $p_1=0$.

(2) 每个人恰有一个邻座 $\Longleftrightarrow$ 所有人成对相邻, 且每一对都没有邻座.

故若 $2\nmid n$, 则 $p_2=0$; 若 $2\mid n$ 且 $N\ge\frac{3}{2}n-1$, 则由 (1) 及每对内部的排列得, 坐法数为: $A_{N-n+1}^{\frac{n}{2}}\cdot 2^{\frac{n}{2}}$ $\Longrightarrow p_2=2^{\frac{n}{2}}\frac{(N-n+1)!(N-n)!}{(N-\frac{3}{2}n+1)!N!}$, 否则 $p_2=0$.

(3) 若 $n>\lceil\frac{N}{2}\rceil$, 则 $p_3=0$; 若 $n\le\lceil\frac{N}{2}\rceil$, 则从 $\lceil\frac{N}{2}\rceil$ 对中选 n 对排列, 再内部排列得, 坐法数为: $A_{\lfloor\frac{N}{2}\rfloor}^n\cdot(2!)^n+(\lceil\frac{N}{2}\rceil-\lfloor\frac{N}{2}\rfloor)\cdot A_{\lfloor\frac{N}{2}\rfloor}^{n-1}\cdot(2!)^{n-1}\Longrightarrow p_3=\frac{A_{\lfloor\frac{N}{2}\rfloor}^n\cdot(2!)^n+(\lceil\frac{N}{2}\rceil-\lfloor\frac{N}{2}\rfloor)\cdot A_{\lfloor\frac{N}{2}\rfloor}^{n-1}\cdot(2!)^{n-1}}{A_N^n}$.

➢ 第三题为了叙述与答案的简洁采用了上下取整, 不习惯的同学可以根据奇偶分类讨论.

15. * (伯川德问题) 在半径为 1 的圆内任取一条弦, 在下列条件下求弦长大于 $\sqrt{3}$ 的概率:
(1) 弦的端点等可能地落在圆周上;
(2) 弦的中点等可能地落在圆内;
(3) 弦的中点等可能地落在与弦垂直的直径上.
(由本问题可知, 在不同的样本空间中考虑等可能性会得出不同的结论, 所以在计算相关的概率时, 一定要清楚样本空间是什么.)

解:
(1) 对该圆上的任意一点 A, 另一端点 B 在圆上均匀分布, 则弦 AB 对应的圆心角 θ 满足在 $[0,\pi]$ 上均匀分布, 则 $|AB|>\sqrt{3}\iff\theta\in[\frac{2}{3}\pi,\pi]\implies p_1=\frac{1}{3}$.
(2) 记弦的中点到圆心的距离为 d, 则弦长 $L=2\sqrt{1-d^2}$, 故 $L>\sqrt{3}\iff d<\frac{1}{2}\implies p_2=\frac{(\frac{1}{2})^2\pi}{1^2\pi}=\frac{1}{4}$.
(3) $d<\frac{1}{2}\implies p_3=\frac{\frac{1}{2}}{1}=\frac{1}{2}$.

16. * 从一个给定的 $2n+1$ 边的正多边形的顶点中, 随机 (即等可能地) 选择 3 个不同的顶点, 求该正多边形的中心恰好位于这三个顶点所确定的三角形内部的概率.

解:
对圆上三点, 要使圆心落在三点构成的三角形内当且仅当这三点不位于同一个半圆内. 做正多边形的外接圆, 则所有顶点均位于圆上且正多边形的中心即为圆心. 故正多边形的顶点也需满足这一充要条件.
在 $2n+1$ 个顶点中任取一个顶点作为起点. 在该起点固定的条件下, 剩下两个顶点可以任取, 从而形成三角形. 记三角形三条边之间间隔的边数分别为 x, y, z, 即求三元组 (x,y,z), s.t. $x+y+z=2n+1$, 且 $1\le x,y,z\le n$.
这里采取补集法, 先计算所有满足 $x+y+z=2n+1$ 的 (x,y,z) 数量: (相当于将 $2n+1$ 分配到三个正整数中, 可以采用隔板法, 在 $2n$ 个缝隙中选 2 个隔板), 则有 $\binom{2n}{2}$ 种分配方法.
记 A, B, C 分别为使得 $x\ge n+1$, $y\ge n+1$, $z\ge n+1$ 的事件, 计算 $|A\cup B\cup C|$. 先计算单个事件的势: $x+y+z=2n+1$ 且 $x\ge n+1\implies$ 令 $x'=x-n$, 则 $x'+y+z=n+1$ 且 $x'\ge 1$, 同理由隔板法可得: $|A|=\binom{n}{2}$. 由对称性, $|A|=|B|=|C|=\binom{n}{2}$. 显然 $A\cap B, A\cap C, B\cap C, A\cap B\cap C=\varnothing$. 故由容斥原理:
$|A\cup B\cup C|=|A|+|B|+|C|-|A\cap B|-|A\cap C|-|B\cap C|+|A\cap B\cap C|=3\binom{n}{3}$
取补集即得满足条件的样本点数为: $\binom{2n}{2}-\binom{n}{3}=\frac{n(n+1)}{2}$. 由于先前我们的起点是固定的, 故总的样本点数为: $(2n+1)\frac{n(n+1)}{2}$. 由对称性, 每个顶点作为起点的概率相同, 即每个三角形会被取到三次, 故最终的样本点数为: $\frac{n(n+1)(2n+1)}{6}\implies p=\frac{\frac{n(n+1)(2n+1)}{6}}{\binom{2n+1}{3}}=\frac{n+1}{2(2n-1)}$.

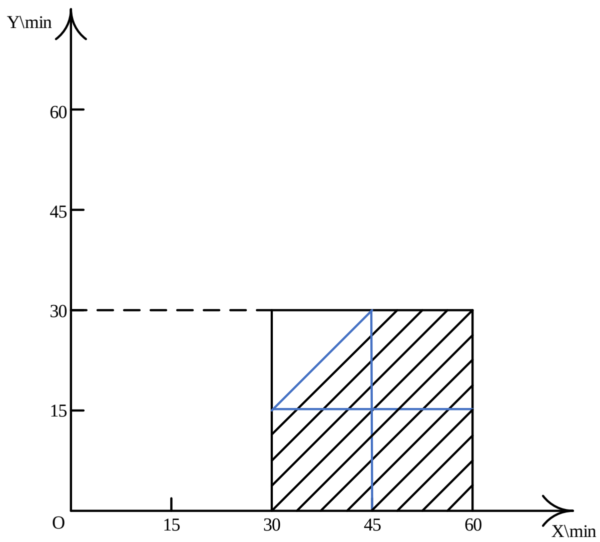
➢ 容斥原理; 隔板法 (这里采取的隔板法中每个元素都为正整数, 即将 n 个元素分给 k 个元素, 样本点数为: $\binom{n-1}{k}$; 若元素可以为 0, 则样本点数为: $\binom{n+k-1}{k-1}$. 两者本质是一样的, 将每个元素减一即可, 这道题也可以先将每条边标准化为可取到 0 的非负整数, 再进行计算, 本质不变).
➢ 若边为偶数同样可以以这种方式计算, 规定好" 中心落在三角形的边上" 是否算作" 在内部" 即可, 这里就不进行拓展了.

17. 甲、乙两人约定在下午 3:00 和 4:00 之间到某公交始发站乘公交车, 该公交始发站每隔 15min 发出一辆公交车. 假定甲、乙两人在这期间到达为等可能. 现约定见车就乘, 求甲、乙同乘一辆车的概率.

解:
以甲、乙到达时间为横、纵坐标作图即可, 若第一辆车在 $3{:}x, x\in[0,15)$ 发出, 则 $p=\frac{x^2+3\times15^2+(15-x)^2}{60^2}$.
➢ 几何概型.

18. 在一次游戏过程中有两队需要合作, 甲队将于 12:30 到 1:00 间到达某地闯关过河, 乙队将于 12:00 到 12:30 到达此地为甲队准备船只, 准备船只需要 15min, 请问甲队到达即能过河的概率是多少?

解:
记甲、乙到达时间分别为 $12{:}X, 12{:}Y$, 则 $X\in[30,60]$, $Y\in[0,30]$. 甲队到达即能过河的条件是 $Y+15\le X$, 故 $p=\frac{1}{30^2}\iint\limits_{Y+15\le X} dXdY=\frac{1}{900}(\int_0^{15}\int_{30}^{60}dXdY+\int_{15}^{30}\int_{Y+15}^{60}dXdY)=\frac{7}{8}$.



➢ 几何概型.

19. * (博雷尔-坎泰利 (Borel-Cantelli) 引理) 设 $\{A_n, n \geq 1\}$ 为事件列. 事件 $B_n = \bigcup\limits_{k=n}^{\infty} A_k$ 表示事件 $A_k, A_{k+1}, \cdots$ 中至少有一个发生, 而事件 $C = \bigcap\limits_{n=1}^{\infty} B_n$ 表示事件 $B_1, B_2, \cdots$ 同时发生. 所以事件 C 表示事件 $A_n, n \geq 1$ 中有无穷个事件发生, 我们把事件 C 记为 $\{A_n,\ \text{i.o.}\}$(i.o. 是无限频繁 (infinitely often) 的缩写). 试证明如下的博雷尔-坎泰利引理. 对于事件列 $A_n, n \geq 1$,

(1) 如果 $\sum\limits_{k=1}^{\infty} P(A_k) < \infty$, 则 $P(A_n,\ \text{i.o.}) = 0$;

(2) 如果 $\{A_k, k \geq 1\}$ 相互独立, $\sum\limits_{k=1}^{\infty} P(A_k) = \infty$, 那么 $P(A_n,\ \text{i.o.}) = 1$.

提示:

(1) 若 $i < j$, 则事件 $B_i \supset B_j$, 所以

$$P(A_n,\ \text{i.o.}) = \lim_{n\to\infty} P(B_n),$$

再用不等式

$$P(\bigcup_{k=n}^{\infty} A_k) \leq \sum_{k=n}^{\infty} P(A_k);$$

(2) 对 $0 < x < 1$, 利用不等式 $\ln(1-x) > -x$.

解:

先说明一下上下限事件的定义:

$\{A_n\}$ 的上限事件:

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m,$$

它表示 A_n 发生无穷多次, 因为 $\omega \in RHS \Longleftrightarrow \omega$ 属于无穷多个 A_n, 记

$$\{A_n, i.o.\} = \limsup_{n\to\infty} A_n$$

$\{A_n\}$ 的下限事件:

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m,$$

它表示 A_n 至多只有有限个不发生, $\omega \in RHS \Longleftrightarrow \exists n \in \mathbb{N}$, s.t. $\omega \in \bigcap\limits_{m=n}^{\infty} A_m$.

下面是 Borel-Cantelli 引理的证明:
设 $A=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m$. 对第一部分 (1)

$$\mathbb{P}(A)\leq\mathbb{P}\left(\bigcup_{m=n}^{\infty}A_m\right)\leq\sum_{m=-n}^{\infty}\mathbb{P}(A_m)\to 0.$$

对第二部分 (2), 由 (1) 知 $\Leftarrow$. 另一方面, 只需 $\mathbb{P}(A^c)=0$. 而

$$A^c=\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty}A_m^c,$$

所以

$$\mathbb{P}\left(\bigcap_{m=n}^{\infty}A_m^c\right)=\lim_{r\to\infty}\mathbb{P}\left(\bigcap_{m=n}^{r}A_m^c\right)\xlongequal{\text{独立性}}\lim_{r\to\infty}\prod_{m=n}^{r}(1-\mathbb{P}(A_m))\leq e^{-\sum_{m=n}^{\infty}\mathbb{P}(A_m)}=0.$$

因此 $\mathbb{P}(A^c)=0$.

➢ 这部分直接照搬刘党政老师的《简明概率论》讲义.

➢ (1) 的本质是实分析中的 Borel-Cantelli 引理:

设 (X,Γ,μ) 是测度空间, $A_n\in\Gamma(\forall n)$. 则

$$\sum_{n=1}^{\infty}\mu(A_n)<\infty\Longrightarrow\mu(\overline{\lim_{n\to\infty}})=0.$$

20. * 如果把 $P(A|B)>P(A)$ 理解为“B 对 A 有促进作用”, 那么直观上似乎能有如下的结论: 由 $P(A|B)>P(A)$ 及 $P(B|C)>P(B)$ 推出 $P(A|C)>P(A)$(意思是 B 促进了 A, C 促进了 B, 故 C 促进了 A). 举一简例说明上述直观看法不对.

解:

记 A: 骰子点数为 4 或 6, B: 骰子点数为 2 或 4, C: 骰子点数为 1, 2 或 4. 则有: $P(A)=\frac{1}{3}$, $P(B)=\frac{1}{2}$, $P(C)=\frac{1}{2}$, $P(A|B)=\frac{1}{2}$, $P(B|C)=\frac{2}{3}$, $P(A|C)=\frac{1}{3}$, 满足 $P(A|B)>P(A)$, $P(B|C)>P(B)$ 但 $P(A|C)=P(A)$.

➢ 本质是条件概率不存在传递性.

21. 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.

解:

根据判别式 $\Delta=B^2-4C,(B,C)\in\{1,2,3,4,5,6\}^2$, 穷举即可, 结果为 $p_{\text{实}}=\frac{19}{36}$, $p_{\text{重}}=\frac{1}{18}$.

22. 某路公交汽车共有 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. 试求下列各事件发生的概率:
(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.

解:
(1) $p_1 = \frac{A_{10}^8}{10^8} \approx 0.018$.
(2) $p_2 = \frac{10}{10^8} = 1 \times 10^{-7}$.
(3) $p_3 = \frac{\binom{8}{3} \cdot 9^5}{10^8} \approx 0.033$

23. * 设某地有 $n+1$ 个微信群, 某个造谣者向第二个微信群转发了谣言, 而第二个微信群中有人向第三个微信群转发该谣言, 如此下去. 在每一步中, 谣言的接收群都是随机从其余 n 个微信群中选取的.
(1) 求谣言传播了 r 次后还没有回到第一个造谣者所在的群的概率;
(2) 求没有一个微信群两次收到谣言的概率;
(3) 若每次随机地向 m 个微信群传播谣言, 回答上面两个问题.

解:
(这类题似乎倾向于将造谣者转发的过程定义为造谣, 后面的才算传播, 但我个人还是将造谣者转发的过程也定义为传播, 总之不影响对知识点的掌握.)
(1) $p_1 = (\frac{n-1}{n})^{r-1}$.
(2) $p_2 = \frac{A_n^r}{n^r} = \frac{n!}{(n-r)!n^r}$.
(3.1) $p_1' = (\frac{\binom{n-1}{m}}{\binom{n}{m}})^{r-1} = (\frac{n-m}{n})^{r-1}$.
(3.2) 先在 n 个群中选取 rm 个不同的群, 有 $\binom{n}{rm}$ 种选择, 再将这些群分配到 r 步, 分配方式有 $\frac{(rm)!}{(m!)^r}$ 种. 故 $p_2' = \frac{\binom{n}{rm}\frac{(rm)!}{(m!)^r}}{[\binom{n}{m}]^r} = \frac{[(n-m)!]^r}{(n!)^r(n-rm)!}$.

24. 有两箱同种类型的零件. 第一箱装 50 只, 其中 10 只为一等品; 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) 在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.

解:
(1) 记挑到第一箱为事件 A, 第 i 次 $(i=1,2)$ 取到的零件是一等品为事件 B_i, 则由全概率公式:

$$P(B_1)=P(B_1|A)P(A)+P(B_1|\overline{A})P(\overline{A})=0.4.$$

(2)
① $P(B_2|B_1)=\frac{P(B_1B_2)}{P(B_1)}$, 同样由全概率公式:

$$P(B_1B_2)=P(B_1B_2|A)P(A)+P(B_1B_2|\overline{A})P(\overline{A}),$$

其中 $P(B_1B_2|A)=\frac{10}{50}\times\frac{9}{49}$, $P(B_1B_2|\overline{A})=\frac{18}{30}\times\frac{17}{29}\Longrightarrow P(B_2|B_1)=\frac{690}{1421}\approx 0.4856$.
② 由贝叶斯公式:

$$P(A|B_1)=\frac{P(B_1|A)P(A)}{P(B_1)}=\frac{1}{4}\qquad P(\overline{A}|B_1)=1-P(A|B_1)=\frac{3}{4},$$

再由 $P(B_2|B_1)=P(B_2|AB_1)P(A|B_1)+P(B_2|\overline{A}B_1)P(\overline{A}|B_1)$, 其中 $P(B_2|AB_1)=\frac{9}{49}$, $P(B_2|\overline{A}B_1)=\frac{17}{29}\Longrightarrow$ $P(B_2|B_1)=\frac{690}{1421}\approx 0.4856$.
➢ 全概率公式; 贝叶斯公式.

25. 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, 并再加入 a 支同色的铅笔. 求第一次、第二次取到红色铅笔且第三次、第四次取到黑色铅笔的概率.

解:
记第 i 次 $(i\geq 1)$ 取到红色铅笔为事件 A_i, 则:

$$P(A_1A_2\overline{A_3}\,\overline{A_4})=P(A_1)P(A_2|A_1)P(\overline{A_3}|A_1A_2)P(\overline{A_4}|A_1A_2A_3)=\frac{b(a+b)r(r+a)}{(r+b)(r+a+b)(r+2a+b)(r+3a+b)}.$$

26. 某工厂的一、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%, 2%. 现从该厂随机抽取一件产品.
(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.

解:
记该产品是次品为事件 A, 该产品是由一、二、三号车间生产的分别为事件 $B_i, 1 \le i \le 3$.
(1) 由全概率公式:

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) = \frac{7}{600}.$$

(2) 由贝叶斯公式:

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A)} = \frac{3}{7}.$$

➢ 全概率公式; 贝叶斯公式.

27. 设男性色盲的概率为 0.05, 女性色盲的概率为 0.0025, 现发现从人群中任选的一人为色盲患者, 求此人为男性的概率.

解:
设该人是男性为事件 A, 该人是色盲患者为事件 B. 由贝叶斯公式:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})} = \frac{20}{21} \approx 0.9524.$$

➢ 贝叶斯公式.

28. 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 份、15 份和 25 份, 其中女生报名表分别为 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.

解:

设选择第 i 个地区为事件 $A_i, 1 \le i \le 3$, 第 j 次抽到女生为事件 $B_j, 1 \le i \le 2$.

(1) 由全概率公式:

$$P(B_1) = P(B_1|A_1)P(A_1) + P(B_1|A_2)P(A_2) + P(B_1|A_3)P(A_3) = \frac{29}{90} \approx 0.322.$$

(2) 同样由全概率公式:

$$P(B_1|\overline{B_2}) = \frac{P(B_1\overline{B_2})}{P(\overline{B_2})} = \frac{\sum_{i=1}^{3} P(B_1\overline{B_2}|A_i)P(A_i)}{\sum_{i=1}^{3} P(\overline{B_2}|A_i)P(A_i)},$$

其中对于不放回随机抽样, 第二次抽到男生的概率与第一次相等, 故 $P(\overline{B_2}|A_1) = \frac{7}{10}$, $P(\overline{B_2}|A_2) = \frac{8}{15}$, $P(\overline{B_2}|A_3) = \frac{4}{5} \Longrightarrow P(B_1|\overline{B_2}) = \frac{20}{61} \approx 0.328$.

➢ 全概率公式, 不放回随机抽样的性质.

29. 罐子中有 25 个球, 其中 20 个红球、5 个黑球, 从中不放回取出 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, 求扔掉的球里至少有两个红球的概率.

解:

设扔掉的球中红球的数量为随机变量 X, 记从剩下的 20 个球中摸到红球为事件 A. 则由贝叶斯公式:

$$P(X \ge 2|A) = \frac{\sum_{i=2}^{5} P(A|X=i)P(X=i)}{\sum_{i=0}^{5} P(A|X=i)P(X=i)},$$

其中 $P(X=i) = \frac{\binom{20}{i}\binom{5}{5-k}}{\binom{25}{5}} \Rightarrow P(X \ge 2|A) = \frac{1767}{1771} \approx 0.9977$.

➢ 贝叶斯公式; 这里的 X 服从的是参数为 (N=25, M=20, n=5) 的超几何分布 (Hypergeometric Distribution), 即 $X \sim H(25, 20, 5)$.

30. 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)* 上一问中的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?

解:

记该人确为带菌者为事件 A, 该人第 i 次 $(1 \le i \le 2)$ 检测呈阳性为事件 B_i.

(1) 由贝叶斯公式:

$$P(A|B_1) = \frac{P(B_1|A)P(A)}{P(B_1|A)P(A) + P(B_1|\overline{A})P(\overline{A})} = \frac{95}{104} \approx 0.91$$

(2) 由两次检测的独立性:

$$P(A|B_1B_2) = \frac{P(B_1B_2|A)P(A)}{P(B_1B_2|A)P(A) + P(B_1B_2|\overline{A})P(\overline{A})} = \frac{9025}{9034} \approx 0.999.$$

➢ 贝叶斯公式.

31. 桌上有 3 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 3 个笔筒装有 3 支红芯圆珠笔、3 支蓝芯圆珠笔. 笔筒外表看起来一模一样, 先随机取一个笔筒, 任取一支笔出来.

(1) 试求取得红芯圆珠笔的概率;

(2) 在已知取得红芯圆珠笔的条件下, 问笔从哪个笔筒中取出的概率最大?

解:

记选取第 i 个笔筒为事件 $A_i, 1 \le i \le 3$, 取到红芯圆珠笔为事件 B.

(1) 由全概率公式:

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) = \frac{1}{2}.$$

(2) 由贝叶斯公式:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)}.$$

$\Longrightarrow P(A_1|B) = \frac{2}{9}, P(A_2|B) = \frac{4}{9}, P(A_3|B) = \frac{1}{3} \Longrightarrow$ 笔从第二个笔筒取出的可能性最大.

➢ 全概率公式; 贝叶斯公式.

32. 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, 则真实信号是 0 的概率是多少?

解:

记真实信号是 0 为事件 A, 接收到的信号是 0 为事件 B. 由贝叶斯公式:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\overline{A})P(\overline{A})} = \frac{196}{197} \approx 0.995.$$

➢ 贝叶斯公式.

33. 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 只白球和 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. (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, 求从甲袋中取的两球中有白球的概率.

解:

记从甲袋放入乙袋的球中白球的数量为随机变量 X, 取出从乙袋中取出的球是白球为事件 A.

(1) 由全概率公式:

$$P(A) = \sum_{i=0}^{2} P(A|X=i)P(X=i) = \frac{38}{77} \approx 0.49.$$

(2) 由贝叶斯公式:

$$P(X \geq 1|A) = \frac{\sum_{i=1}^{2} P(A|X=i)P(X=i)}{\sum_{i=0}^{2} P(A|X=i)P(X=i)} = \frac{55}{57} \approx 0.96.$$

➢ 全概率公式; 贝叶斯公式.

34. 证明: 若 $P(B|A) = P(B|\overline{A})$, 则事件 A 与 B 独立.

解:

$$P(B|A) = P(B|\overline{A}) \Longleftrightarrow \frac{P(AB)}{P(A)} = \frac{P(\overline{A}B)}{P(\overline{A})} = \frac{P(B) - P(AB)}{1 - P(A)} \Longleftrightarrow P(AB) = P(A)P(B) \Longleftrightarrow A, B\text{独立}.$$

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

解:

$$P(B|A) > P(B) \Longleftrightarrow P(AB) > P(A)P(B) \Longleftrightarrow P(\overline{A}B) < P(\overline{A})P(B) \Longleftrightarrow P(\overline{A}\,\overline{B}) > P(\overline{A})P(\overline{B}).$$

➢ 与证明过程一致, 由于 $\overline{\overline{A}} = A$, 这道题的逆命题也是成立的.

36. 事件 A 与事件 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 与事件 B 相互独立吗?

解:
$P(A)+P(B) = P(A\cup B)+P(A\cap B) = 0.22$, 反证法: 设 A,B 相互独立, 则 $P(A)P(B) = P(A\cap B) = 0.1$
$\Longrightarrow (P(A)-P(B))^2 = (P(A)+P(B))^2 - 4P(A)P(B) = -0.3516$ 矛盾. 故 A,B 不相互独立.
➢ 本质是韦达定理.

37. 对于三个事件 A,B,C, 若

$$P(AB|C) = P(A|C)P(B|C)$$

成立, 则称 A 与 B 关于 C 条件独立. 若已知 A 与 B 关于 C 与 $\overline{C}$ 条件独立, 且 $P(C) = 0.5$, $P(A|C) = P(B|C) = 0.9$, $P(A|\overline{C}) = 0.2$, $P(B|\overline{C}) = 0.1$, 试求 $P(A)$, $P(B)$, $P(AB)$, 并证明 A 与 B 不相互独立.

解:
由全概率公式:

$$P(A) = P(A|C)P(C)+P(A|\overline{C})P(\overline{C}) = 0.55, \qquad P(B) = P(B|C)P(C)+P(B|\overline{C})P(\overline{C}) = 0.5,$$

A 与 B 关于 C 与 $\overline{C}$ 条件独立 $\Longrightarrow$

$$P(AB) = P(AB|C)P(C)+P(AB|\overline{C})P(\overline{C}) = P(A|C)P(B|C)P(C)+P(A|\overline{C})P(B|\overline{C})P(\overline{C}) = 0.42.$$

38. 对于同一目标进行三次独立射击第一、二、三次射中的命中率分别为 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.

解:

记第 i 次 $(i=1,2,3)$ 射中为事件 A_i, 则:

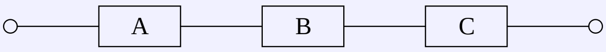
(1) 由独立性:

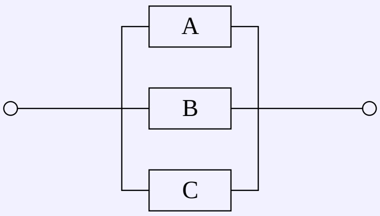
$$p_1 = P(A_1\overline{A_2}\,\overline{A_3}) + P(\overline{A_1}A_2\overline{A_3}) + P(\overline{A_1}\,\overline{A_2}A_3) = 0.26.$$

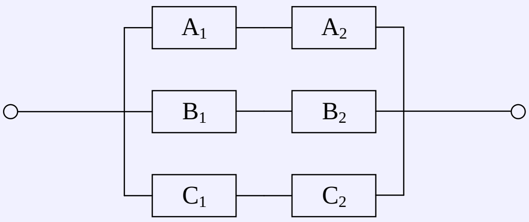
(2)

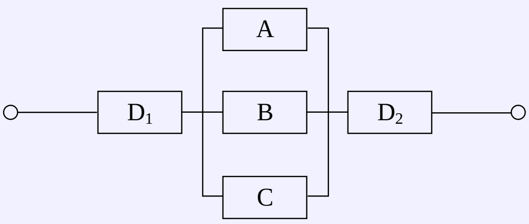
$$p_2 = 1 - P(\overline{A_1}\,\overline{A_2}\,\overline{A_3}) = 0.96.$$

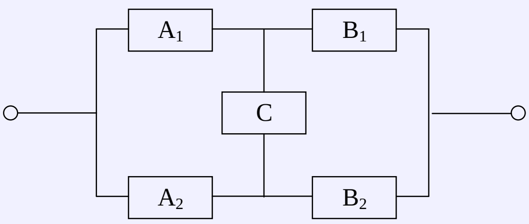
39. 求下列图中各系统能正常工作的概率, 其中框图中的字母代表元件, 字母相同但下标不同的都是同一种元件, 只是装配在不同的位置上, 元件 A, B, C, D 能正常工作的概率分别为 p_A, p_B, p_C, p_D.

(1) 

(2) 

(3) 

(4) 

(5)* 

解:

(1) $p_1 = p_A p_B p_C$.

(2) $p_2 = 1-(1-p_A)(1-p_B)(1-p_C)$.

(3) $p_3 = 1-(1-p_A^2)(1-p_B^2)(1-p_C^2)$.

(4) $p_4 = p_D^2[1-(1-p_A)(1-p_B)(1-p_C)]$.

(5) 该图共有四条通路: $A_1 \longrightarrow B_1$; $A_2 \longrightarrow B_2$; $A_1 \longrightarrow C \longrightarrow B_2$; $A_2 \longrightarrow C \longrightarrow B_1$.

① 记该四条通路连通分别为事件 $M_i, 1 \le i \le 4$. 则: $P(M_1) = P(M_2) = p_A p_B$, $P(M_3) = P(M_4) = p_A p_B p_C$, $P(M_1M_2) = p_A^2 p_B^2$, $P(M_1M_3) = P(M_2M_4) = p_A p_B^2 P_C$, $P(M_1M_4) = P(M_2M_3) = p_A^2 p_B p_C$, $P(M_3M_4) = p_A^2 p_B^2 p_C$, $P(M_iM_jM_k) = p_A^2 p_B^2 p_C, \forall 1 \le i,j,k \le 4 = P(M_1M_2M_3M_4)$. 由容斥原理:

$$p_5 = P(\bigcup_{i=1}^{4} M_i) = \sum_{k=1}^{4}(-1)^{k-1}\sum_{1\le i_1<\cdots<i_k\le 4} P(M_{i_1}\cap M_{i_2}\cdots\cap M_{i_k}) = 2p_Ap_B[1+p_C(1-p_A)(1-p_B)] - p_A^2p_B^2.$$

② 记 K_i 为 K_i 能正常工作 $(K \in \{A,B,C\}, i \in \{\emptyset,1,2\})$.

C 能正常工作时: 系统能正常工作 $\Longleftrightarrow$ A_1, A_2 至少有一个能正常工作且 B_1, B_2 至少有一个能正常工作, 即:

$$p_{5|C} = P(A_1\cup A_2)P(B_1\cup B_2) = [1-(1-p_A)^2][1-(1-p_B)^2].$$

C 不能正常工作时: 系统能正常工作 $\Longleftrightarrow$ A_1, B_1 能正常工作或 A_2, B_2 能正常工作, 即:

$$p_{5|\overline{C}} = 1-(1-p_Ap_B)^2.$$

综上, 由全概率公式:

$$p_5 = p_{5|C}P(C) + p_{5|\overline{C}}P(\overline{C}) = 2p_Ap_B[1+p_C(1-p_A)(1-p_B)] - p_A^2p_B^2.$$

➢ 容斥原理; 全概率公式.

40. 一个电路共有三个继电器, 当第一个继电器断开, 或者第二个、第三个同时断开时, 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.

解:

$$p = 0.3 + 0.4\times0.6 - 0.3\times0.4\times0.6 = 0.468.$$

➢ 容斥原理.

第 2 章 随机变量及其分布

1. 双色球是目前彩票中最受欢迎的玩法之一. 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 33 个号码组成, 蓝色球号码区由 1 到 16 共 16 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. 若某单位投注中分别有 $i(0 \le i \le 6)$ 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中奖等级情况, 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:

$$A=\{\text{中奖}\}, \quad B=\{\text{中一等奖或二等奖}\}$$

并求出它们发生的概率.

解:

记匹配 i 个红色球号码为事件 A_i, 匹配蓝色球号码为事件 B, 随机变量 X 为:

等级	未获奖	一等奖	二等奖	三等奖	四等奖	五等奖	六等奖
X	0	1	2	3	4	5	6

则 $P(A_i)=\dfrac{\binom{6}{i}\binom{27}{6-i}}{\binom{33}{6}}$, $P(B)=\dfrac{1}{16}$.

(1) 显然各个不同的 $A_i \cap B$(或 $A_i \cap B^c$) 交集为空且 A_i, B(或 B^c) 相互独立, 故 $P((A_mB_1)\cup(A_nB_2)) = P(A_mB_1)+P(A_nB_2)$, $B_1, B_2 = B$或B^c 且 $P(A_iB)=P(A_i)P(B)$(或 $P(A_iB^c)=P(A_i)P(B^c)$). 计算可得 X 的分布律如下:

X	0	1	2	3	4	5	6
$P(X=x)$	$\frac{16532100}{17721088}$	$\frac{1}{17721088}$	$\frac{15}{17721088}$	$\frac{162}{17721088}$	$\frac{7695}{17721088}$	$\frac{137475}{17721088}$	$\frac{1043640}{17721088}$

(2) $A=\{X\neq 0\}$, $B=\{X=1\text{或}2\} \Longrightarrow P(A)=1-P(X=0)=\frac{1189988}{17721088}\approx 0.0671$, $P(B)=P(X=1)+P(X=2)=\frac{16}{17721088}\approx 9.03\times 10^{-7}$.

2. 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 5 次投进了 4 个球, 则第 6 次他的投篮命中率为 4/5). 求他最终在这 100 次投篮中投进次数的分布律.

解:
记 S_n 为 n 次投篮中投进的次数, 由数学归纳法:
已知 $n=3$ 时, $P(S_3=1)=\frac{1}{2}=P(S_3=2)=\frac{1}{2}$. 设 $P(S_{n-1}=k)=\frac{1}{n-2},\forall 1\le k\le n-2$, 则:

$$P(S_n=k)=\frac{k-1}{n-1}P(S_{n-1}=k-1)+\left(1-\frac{k}{n-1}\right)P(S_{n-1}=k)=\frac{1}{n-1}. \qquad \square$$

综上所述, S_{100} 服从在 $\{1,2,\cdots,99\}$ 的等可能分布, 即 $P(S_{100}=k)=\frac{1}{99},\forall 1\le k\le 99$.
➢ 数学归纳法.
➢ 这道题实际上是一个随机强化模型, 依据过去经验不断更新决策概率, 这里服从等可能分布的本质原因是初始条件的对偶性. 与之等价的数学模型: 将 "命中" 等效为 "从箱中抽出红球","未命中" 等效为 "从箱中抽出蓝球", 命中则投入一个红球, 未命中则投入一个蓝球.

3. 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, 不按期到达但无损坏的占 20%, 按期到达但货物有损坏的占 10%, 货物不按期到达又有损坏的占 10%. 以 X 记该物流公司用车将一箱货物运到目的地后的毛利润. 试求 X 的分布律.

解:
显然, X 的分布律为:

X	100	80	50	-60
$P(X=x)$	0.6	0.2	0.1	0.1

4. 设某游乐场的一部设备在一天内发生故障的概率为 0.2, 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 2 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.

解:

记五个工作日内发生故障的次数为随机变量 X, 则 X 服从参数为 $n=5$, $p=0.2$ 的二项分布, 即 $X \sim B(5, 0.2)$, 则 $P(X=k) = \binom{5}{k} \cdot (0.2)^k \cdot (0.8)^{5-k}$, $0 \le k \le 5$. 故毛利润的分布律为:

毛利润 (万元)	10	5	0	-2
X	0	1	2	3, 4, 5
P	0.32768	0.4096	0.2048	0.05792

➢ 二项分布.

5. 设随机变量 X 的分布律如下:

X	-1	1	2
P	0.25	0.5	0.25

(1) 试求 X 的分布函数 $F(x)$;

(2) 试求概率 $P(X \le 0)$, $P(0.5 < X \le 1.5)$, $P(1 \le X \le 2)$ 和 $P(1 < X \le 2)$.

解:

(1)

$$F(x) = \begin{cases} 0, & x < -1, \\ 0.25, & -1 \le x < 1, \\ 0.75, & 1 \le x < 2, \\ 1, & x \ge 2. \end{cases}$$

(2) $P(X \le 0) = F(0) = 0.25$, $P(0.5 < X \le 1.5) = F(1.5) - F(0.5) = 0.5$, $P(1 \le X \le 2) = F(2) - F(1) + P(X = 1) = 0.75$, $P(1 < X \le 2) = F(2) - F(1) == 0.25$.

6. 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.

解:

显然 X 的分布律为:

X	1	2	3
P	0.8	$\frac{8}{45}\approx 0.1778$	$\frac{1}{45}\approx 0.0222$

分布函数为:

$$F(x)=\begin{cases}0, & x<1,\\ \frac{4}{5}, & 1\le x<2,\\ \frac{44}{45}, & 2\le x<3,\\ 1, & x\ge 3.\end{cases}$$

7. 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 0.4. 若以 X 表示 A 发生时的累计试验次数, 试求概率 $P(X\text{为偶数})$ 和 $P(X>2)$.

解:

X 服从参数为 p 的几何分布, 即 $X\sim Ge(p)$; $P(X=k)=(1-p)^{k-1}p$, $k=1,2,\cdots$, 其中 $p=0.4$. 则:

$$P(X\text{为偶数})=\sum_{k=1}^{\infty}P(X=2k)=\sum_{k=1}^{\infty}(1-p)^{2k-1}p=\frac{1-p}{2-p};$$

$$P(X>2)=1-P(X=1)-P(X=2)=(1-p)^2.$$

代入 $p=0.4$, 得:

$$P(X\text{为偶数})=0.375,\qquad P(X>2)=0.36.$$

➢ 几何分布.

8. 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 $P(X\ge 1)$ 及 X 最有可能的取值.

解:

X 服从参数为 $n=20$, $p=0.2$ 的二项分布, 即 $X\sim B(20,0.2)$; $P(X=k)=\binom{20}{k}\cdot(0.2)^k\cdot(0.8)^{20-k}$, $k=0,1,\cdots,20$. 则:

$$P(X \geq 1) = 1 - P(X=0) \approx 0.9885.$$

X 最有可能的取值即二项分布的众数, 由对应公式(见注释)可知:

$$X_{\text{众}} = \lfloor (n+1)p \rfloor = 4.$$

➢ 二项分布的众数 (高中人教版附录里似乎有提到接下来的公式, 这里再证一遍): 设 $X \sim B(n,p)$, 则:

$$k_{\text{众}} = \lfloor (n+1)p \rfloor$$

证: $\dfrac{P(X=k+1)}{P(X=k)} = \dfrac{\binom{n}{k+1}p^{k+1}(1-p)^{n-k-1}}{\binom{n}{k}p^k(1-p)^{n-k}} = \dfrac{(n-k)p}{(k+1)(1-p)}$ 关于 k 单调递减, 故令 $\dfrac{(n-k)p}{(k+1)(1-p)} \leq 1 \iff k \geq (n+1)p - 1$. 从而 $k_{\text{众}} = \lfloor (n+1)p \rfloor$ 是最小的满足 $\dfrac{P(X=k+1)}{P(X=k)} \leq 1$ 的整数 (注意当 $(n+1)p \in \mathbb{Z}$ 时, $\frac{P((n+1)p)}{P((n+1)p-1)} = 1$, 即有两个众数: $(n+1)p$, $(n+1)p-1$). (!!! 考试时建议不要直接套公式, 而是按这种方式给出完整计算或证明 (‘ · ω · ′).)

9. 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. 若 A 不出现, 则 B 也不出现; 若 A 只出现一次, 则 B 出现的概率是 0.6; 若 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.

解:

记结果 A 出现的次数为随机变量 X, 则 $X \sim B(4, 0.3)$, 即 $P(X=k) = \binom{4}{k} \cdot (0.3)^k \cdot (0.7)^{4-k}$, $0 \leq k \leq 4$

(1) 由全概率公式:

$$P(B) = \sum_{k=0}^{4} P(B|X=k)P(X=k) = 0.59526.$$

(2) 由贝叶斯公式:

$$P(X=1|B) = \frac{P(B|X=1)P(X=1)}{P(B)} \approx 0.4149.$$

➢ 全概率公式; 贝叶斯公式; 二项分布.

10. 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?

解:

记甲队胜四场比赛所需的场次为随机变量 X, 则 X 服从参数为 $r=4$, $p=0.6$ 的负二项分布 (帕斯卡分布), 即 $X \sim NB(4,0.6)$, 则:

$$P(X=i)=\binom{i-1}{4-1}\cdot(0.4)^4\cdot(0.6)^{i-4}, i=4,5,6,7$$

$\Longrightarrow$

i	4	5	6	7
P	0.1296	0.20736	0.20736	0.165888

$p_{乙,4}=1-\sum_{i=4}^{7}P(X=i)=0.289792<p_{乙,2}=0.352 \Longrightarrow$"三局两胜制" 对乙队更有利 (原理同第一章第 12 题).

11. 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).

解:

记 Y 为三个骰子中押的数字出现的次数, 则 $Y \sim B(3,\frac{1}{6})$, 即 $P(Y=k)=\binom{3}{k}\cdot(\frac{1}{6})^k\cdot(\frac{5}{6})^{3-k}$, $k=0,1,2,3$.
则: $P(x)=P(Y), X=1,2,3$; $P(X=-1)=P(Y=0) \Longrightarrow X$ 的分布律为:

X	-1	1	2	3
P	$\frac{125}{216}$	$\frac{25}{72}$	$\frac{5}{72}$	$\frac{1}{216}$

➢ 二项分布.

12. 设某种昆虫单只每次产卵的数量服从参数为 λ 的泊松分布, 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的个数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?

解:

设一只昆虫一次产卵的数量为随机变量 N, 则 $N \sim P(\lambda)$, 即 $P(N=k)=\frac{\lambda^k}{k!}e^{-\lambda}$, $k=0,1,2,\cdots$. 此外, $Y|N=n \sim B(n,p)$, $Z|N=n \sim B(n,1-p) \Longrightarrow$ 由全概率公式:

$$\begin{aligned}P(Y=k)=\sum_{n=k}^{\infty}P(N=n)P(Y=k|N=n)&=\sum_{n=k}^{\infty}\frac{\lambda^n}{n!}e^{-\lambda}\binom{n}{k}p^k(1-p)^{n-k}\\&=e^{-\lambda}\frac{p^k}{k!}\sum_{n=k}^{\infty}\frac{\lambda^n(1-p)^{n-k}}{(n-k)!}\\&\xlongequal{m=n-k}e^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{m=0}^{\infty}\frac{\lambda^m(1-p)^m}{m!}=e^{-\lambda p}\frac{(\lambda p)^k}{k!}.\end{aligned}$$

故 $Y\sim P(\lambda p)$. 同理, $Z\sim P(\lambda(1-p))$. 再计算:

$$\begin{aligned}P(Y=y,Z=z)=P(N=y+z)P(Y=y,Z=z|N=y+z)&=e^{-\lambda}\frac{\lambda^{y+z}}{(y+z)!}\binom{y+z}{y}p^y(1-p)^z\\&=e^{-\lambda p}\frac{(\lambda p)^y}{y!}\cdot e^{-\lambda(1-p)}\frac{(\lambda(1-p))^z}{z!}\\&=P(Y=y)P(Z=z)\end{aligned}$$

$\Longrightarrow$ Y,Z 相互独立.

➢ 泊松分布; 二项分布.

13. 一个系统包含了 1000 个零件, 各个零件是否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. 试利用泊松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.

解:

记 n 为零件数, p 为出故障的概率, X 为一个月内出故障的零件个数, 则 $X\sim B(n=1000,p=0.001)$. 由于 $n=1000\geq 30$(或 $n\geq 100$) 且 $np=1\leq 5$(或 $np\leq 10$), 故可以使用 “泊松逼近定理” 进行估算: 令 $\lambda=np=1$, 则 $P(X=0)\approx\frac{\lambda^0}{0!}e^{-\lambda}=e^{-1}\approx 0.3679$.

➢ 泊松逼近定理及其判断 (使用) 条件, 见课本 P57 定理 2.2 及下方的正文.

14. 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 0.02. 利用泊松分布, 试求在 400 份保单中最终至少赔付两份保单的概率. 结果精确到小数点后三位.

解:

记 n 为保单数, p 为保险期内死亡概率, X 为需要赔付的保单数, 则 $X\sim B(n=400,p=0.02)$. 由于 $n=400\geq 100$ 且 $np=8\leq 10$, 故可以使用 “泊松逼近定理” 进行估算: 令 $\lambda=np=10$, 则 $P(X\geq 2)=1-P(X=0)-P(X=1)\approx 1-9e^{-8}\approx 0.997$.

➢ 泊松逼近定理及其判断 (使用) 条件, 见课本 P57 定理 2.2 及下方的正文.

15. 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 10S 内至少出现一个误码的概率. 在 100s 内呢? 结果精确到小数点后三位.

解:

记 $n_i, i=1,2$ 为指定秒内出现的字符数, p 为出现误码的概率, $X_i, i=1,2$ 为指定秒内出现的误码数, 则 $X_1\sim B(n_1=5.12\times10^6, p=10^{-7})$. 由于 $n_1=5.12\times10^6\geq30$(或 $n_1\geq100$) 且 $n_1p=0.512\leq5$(或 $n_1p\leq10$), 故可以使用 "泊松逼近定理" 进行估算: 令 $\lambda_1=n_1p=1$, 则 $P(X_1\geq1)=1-P(X_2=0)\approx1-e^{-\lambda_1}\approx0.401$. 同理可得: $P(X_2)\approx0.994$.

➢ 泊松逼近定理及其判断 (使用) 条件, 见课本 P57 定理 2.2 及下方的正文.

16. 某航空公司知道预订航班的乘客有 0.05 的概率最终不会来搭乘, 为了盈利更多, 他们的政策是接受比实际座位更多的预订. 若一个恰有 50 个座位的航班一共被预订了 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.

解:

记 n 为预订的票数, p 为乘客不来搭乘的概率, X 为最终不来搭乘的乘客数量, 则 $X\sim B(n=52, p=0.05)$. 由于 $n=52\geq30$ 且 $np=2.6\leq5$, 故可以使用 "泊松逼近定理" 进行估算: 令 $\lambda=np=2.6$, 则 $P(X\leq1)\approx3.6e^{-2.6}\approx0.27$.

➢ 泊松逼近定理及其判断 (使用) 条件, 见课本 P57 定理 2.2 及下方的正文.

17. 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 0.95 的概率中奖?

解:

(1) 记总彩票数为 N, 有奖彩票数为 M, 购买的彩票数为 n, 中奖次数为 X, 则实际上 X 服从超几何分布, 即 $X\sim H(N,M,n)$. 但因为 $1\times10^6\gg100$, 所以每次买完一张彩票后, 下一张彩票中奖的概率 p 几乎不变, 即可以将该问题近似为二项分布: 令 $p=\frac{M}{N}=0.0001\Longrightarrow X\sim B(n=100, p=0.0001)$. 又因为 $n=100\geq30$ 且 $np=0.01\ll5$, 由泊松逼近定理, $P(X\geq1)=1-P(X=0)\approx1-e^{-0.01}=0.00995$.

(2) 要使 $1-e^{-n'p} \geq 0.95 \iff n' \geq 29957$. 经检验, 此时 $\frac{n'}{N} = 0.29957$ 仍处于较小水平且 $n'p = 2.9957 \leq 5$, 仍满足二项分布近似与泊松逼近的条件.

➢ 泊松逼近定理及其判断 (使用) 条件, 见课本 P57 定理 2.2 及下方的正文.

➢ 第二题也可以直接用二项分布的概率计算, 在处理不等式时对多项式取对数, 由于 p 较小, 再使用 $\ln(1-p) = \approx -p$ 进行近似, 结果与泊松逼近一致, 因为这就是它的本质.

➢ 关于超几何分布和二项分布何时能近似, 貌似书上并没有提到, 但这里仅从百分比来判断依旧是合理的. 实际上利用超几何分布进行精确计算得到的结果是 29512 注, 但由于做题不太可能要求短时间内计算那么庞大的组合数, 所以不需要用超几何分布.

18. 设随机变量 X 的分布函数为:

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

试求: (1) $P(X=k)$, $k = 1,2,3$;　　(2) $P(1/2 < x < 3/2)$.

解: (注: $f(x^{\mp}) = \lim\limits_{y \to x^{\mp}} f(y)$ 指 f 在 x 处的左右极限.)

(1) $P(X=k) = F(k) - F(k^-) \Longrightarrow P(X=1) = \frac{1}{4}$; $P(X=2) = \frac{1}{12}$; $P(X=3) = \frac{1}{6}$.

(2) $P(\frac{1}{2} < x < \frac{3}{2}) = F(\frac{3}{2}^-) - F(\frac{1}{2}) = \frac{1}{2}$.

19. 设随机变量 X 的分布函数为:

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

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

解: (注: $f(x^{\mp})=\lim\limits_{y\to x^{\mp}} f(y)$ 指 f 在 x 处的左右极限.)

$F(1^-)=a+b=F(1)-P(X=1)=\frac{3}{4}$, $F(-1^+)=-a+b=F(-1)=\frac{1}{8}\Longrightarrow a=\frac{5}{16}$, $b=\frac{7}{16}$.

20. 设随机变量 X 的密度函数为:

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

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

解:

$P(1<x<2)=\int_1^2 ax\mathrm{d}x=\int_2^3 b\mathrm{d}x=P(2<X<3)\Longleftrightarrow \frac{ax^2}{2}\Big|_1^2=\frac{3}{2}a=b$. 又由于 $\int_{-\infty}^{\infty} f(x)\mathrm{d}x=\frac{3}{2}a+b=1$. 故解得: $a=\frac{1}{3}$, $b=\frac{1}{2}$.

21. 设随机变量 X 的密度函数为:

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

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

解:

(1) $\int_{-\infty}^{\infty} f(x)\mathrm{d}x=a\arctan x\big|_{-\infty}^{\infty}=a\pi=1\Longrightarrow a=\frac{1}{\pi}$.

(2) $F(x)=P(X\le x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=\frac{1}{\pi}\arctan t\big|_{-\infty}^{x}=\frac{1}{2}+\frac{1}{\pi}\arctan x$.

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

22. 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).

解:

$S_{总} = \int_0^2 2x - x^2 \mathrm{d}x = \frac{4}{3}$; $F(x) = \frac{1}{S_{总}} P(X \le x) = \frac{3}{4} \int_0^x 2t - t^2 \mathrm{d}t = \frac{-x^3 + 3x^2}{4}, 0 \le x \le 2 \Longrightarrow f(x) = F'(x) = \frac{-3x^2 + 6x}{4}, 0 \le x \le 2.$

23. 设连续型随机变量 X 的分布函数为:

$$F(x) = \begin{cases} 0, & x < 1, \\ ax\ln x + bx + 1, & 1 \le x < e, \\ 1, & x \ge e, \end{cases}$$

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

解:

(1) X 是连续型随机变量 $\Longrightarrow F(x)$ 连续 $\Longrightarrow F(1) = b + 1 = 0, F(e) = (a + b)e + 1 = 1 \Longrightarrow a = 1, b = -1.$

(2) $f(x) = F'(x) = \ln x, 1 \le x < e.$

24. 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.

解:

方程 $x^2 + Xx + 1 = 0$ 有实根 $\Longleftrightarrow \Delta = X^2 - 4 \ge 0 \Longleftrightarrow |X| \ge 2 \Longrightarrow p = \frac{3}{5}.$

25. 某城际列车从早上 6:00 开始每 15min 发出一趟列车, 假设某乘客达到车站的时间服从 7:00 到 7:30 的均匀分布, 若忽略买票等其他时间, 试求该乘客等车时间少于 5 min 的概率.

解:

$$p = \frac{(15-10) + (30-25)}{30} = \frac{1}{3}.$$

26. 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.

解:

$$p = 3 \times \left(\frac{4-2}{4-1}\right)^2 \times \left(\frac{2-1}{4-1}\right) + \left(\frac{4-2}{4-1}\right)^3 = \frac{20}{27}.$$

27. 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 $0 \le a < b \le 1$, $F(b)-F(a)$ 的值仅与差 $b-a$ 有关. 试证明 X 服从 $(0,1)$ 上的均匀分布.

解:

设 $F(x)-F(y) = G(x-y)$, 则对 $\forall x \in (0,1)$, 有 $G(x) = F(x)-F(0)$, 从而 $G(x)-G(y) = F(x)-F(y) = G(x-y), \forall x,y \in (0,1)$. 即 G 满足柯西函数方程. 由于对 $\forall x \in (0,1)$, $\exists a,b \in (0,1)$, s.t.$x = a-b$ $\Longrightarrow$ $|G(x)| = |F(a)-F(b)| \le 2$, 即 G 在正测集 $(0,1)$ 上有界, 从而 (证明见注释)$G(x) = G(1)x \Longrightarrow$ $F(x) = G(x)+G(0)$ 关于 x 是线性的 $\Longrightarrow f(x) \equiv c \in (0,1)$, 即 X 服从 $(0,1)$ 上的均匀分布.

➢ 柯西方程及其解的描述与证明如下 (基本照搬周民强《实变函数论 (第三版)》P86 例 2):

设有定义在 $\mathbb{R}$ 上的函数 $f(x)$, 满足

$$f(x+y) = f(x)+f(y), \qquad x,y \in \mathbb{R},$$

且在 $E \subset \mathbb{R}\ (m(E)>0)$ 上有界, 则 $f(x) = cx (x \in \mathbb{R})$, 其中 $c = f(1)$.

证:

(i) 首先, 由题设知, 对 $r \in \mathbb{Q}$, 必有 $f(r) = rf(1)$.

(ii) 其次, 由 $m(E)>0$ 可知, 存在区间 $I: I \subset E-E$. 不妨设 $|f(x)| \le M (x \in E)$, 又对任意的 $x \in I$, 有 $x', x'' \in E$, 使得 $x = x'-x''$, 则

$$|f(x)| = |f(x')-f(x'')| \le |f(x')|+|f(x'')| \le 2M.$$

记 $I = [a,b]$, 并考查 $[0,b-a]$. 若 $x \in [0,b-a]$, 则 $x+a \in [a,b]$. 从而由 $f(x) = f(x+a)-f(a)$ 可知, $|f(x)| \le 4M$, $x \in [0,b-a]$. 记 $b-a = c$, 这说明

$$|f(x)| \le 4M, x \in [0,c].$$

易知

$$|f(x)| \le 4M, x \in [-c,c].$$

已知对任意的 $x \in \mathbb{R}$ 以及自然数 n, 均存在有理数 r, 使得 $|x-r| < c/n$, 因此我们得到

$$|f(x)-xf(1)| = |f(x-r)+rf(1)-xf(1)| = |f(x-r)+(r-x)f(1)| \le \frac{4M+c|f(1)|}{n}.$$

根据 n 的任意性 (r 的任意性), 即得 $f(x) = xf(1)$. (其实 (ii) 也能由 Steinhaus 定理直接得到)

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

解:
记 T 为检修时间, 则 $T\sim Exp(\lambda=1)\Longrightarrow f(t)=\lambda e^{-\lambda t}=e^{-t}$, $F(t)=1-e^{-\lambda t}=1-e^{-t}$, $t\geq 0$.
(1) $P(T>2)=e^{-2}\approx 0.1353$.
(2) 由指数分布的无记忆性, $P(T>4|T>2)=P(T>2)=e^{-2}\approx 0.1353$.
➢ 指数分布的无记忆性, 见课本 P68

29. 设顾客在某银行的窗口等待服务的时间 X 服从参数为 $\lambda=\dfrac{1}{5}$ 的指数分布 (单位: min). 假设某顾客一旦等待时间超过 10 min 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.

解:
$X\sim Exp(\lambda=\frac{1}{5})\Longrightarrow f(t)=\lambda e^{-\lambda t}=\frac{1}{5}e^{-\frac{1}{5}t}$, $F(t)=1-e^{-\lambda t}=1-e^{-\frac{1}{5}t}$, $t\geq 0$. 故 $P(X\leq 10)=F(10)=1-e^{-2}$
$\Longrightarrow p=1-P(X\leq 10)^5\approx 0.5167$.

30. (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要条件是对任意的常数力 $t,x>0$, 均有:

$$P(X\leq t+x|X>t)=P(X\leq x);$$

(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要条件是对任意的正整数 m, n, 均有:

$$P(X\leq m+n|X>n)=P(X\leq m).$$

解:
(1)

$$\text{证}\quad P(X\leq t+x|X>t)=P(X\leq x)\Longleftrightarrow\text{证}\quad P(X>t+x|X>t)=P(X>x)$$

必要性显然, 充分性:

$$P(X>t+x|X>t)=\frac{1-F(t+x)}{1-F(t)}=1-F(x)=P(X>x)\Longleftrightarrow 1-F(t+x)=[1-F(t)][1-F(x)]$$

记 $G(x)=1-F(x)$, 则 $G(x+y)=G(x)G(y), \forall x,y>0$ 满足柯西乘法方程 (或 $\Longrightarrow \ln G(x+y)=\ln G(x)+\ln G(y)$ 满足柯西加法方程). 又由 $G(0)=1-F(0)>1-F(1)=G(1)$, 从而 $\exists\lambda>0$, s.t.$G(x)=e^{-\lambda x}$, 即 $F(x)=1-e^{-\lambda x} \Longrightarrow X\sim Exp(\lambda)$.

(2) 必要性显然, 充分性见课本 P54.

➢ 关于柯西方程, 见 27 题的注释.

31. 设随机变量 $X\sim N(1,4)$,

(1) 试求概率 $P(0\le X\le 4)$, $P(X>2.4)$ 和 $P(|X|>2)$;

(2) 试求常数 c, 使得 $P(X>c)=2P(X\le c)$.

解:

将 X 标准化变换为 $\dfrac{X-\mu}{\sigma}$, 则

(1) 查表 (课本附表 1) 得: $P(0\le X\le 4)=\Phi(1.5)-\Phi(-0.5)=\approx 0.6247$; $P(X>2.4)=1-\Phi(0.7)\approx 0.2420$; $P(|X|>2)=\Phi(0.5)+\Phi(-1.5)=0.3753$.

(2) $P(X>c)=2P(X\le c) \Longrightarrow \Phi(\frac{c-1}{2})=\frac{1}{3} \stackrel{\text{查表}}{\Longrightarrow} \frac{c-1}{2}\approx 0.43 \Longrightarrow c\approx 0.14$.

➢ 正态函数标准化, 正态函数的对称性, 见课本 P70-71.

32. 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 96Ω 和 104Ω 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 服从区间 $(95,105)$ 上的均匀分布;

(2) 若 R 服从正态分布 $N(100,4)$.

解:

(1) $p_1=\frac{104-96}{105-95}=0.8$.

(2) $p_2=P(96\le R\le 104)=\Phi(2)-\Phi(-2)\approx 0.9544$.

➢ 正态函数标准化, 正态函数的对称性, 见课本 P70-71.

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

解:

(1) $P(X_1 \le 50) = \Phi(2) < \Phi(2.5) = P(X_2 \le 50) \Longrightarrow$ 选择环城公路.

(2) $P(X_1 \le 45) = \Phi(1.5) > \Phi(1.25) = P(X_2 \le 45) \Longrightarrow$ 选择市区路线.

➢ 正态函数标准化, 见课本 P70-71.

34. 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.

解:

易知 X 的分布律为:

X	2	3	4	5	6	7	8	9	10	11	12
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

35. 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.

解:

记三个骰子的点数分别为 X_1, X_2, X_3, 则 $P(X \le k) = P(X_1 \le k, X_2 \le k, X_3 \le k) = P(X_1 \le k)P(X_2 \le k)P(X_3 \le k) = \left(\frac{k}{6}\right)^3, 1 \le k \le 6 \Longrightarrow P(X = k) = P(X \le k) - P(X \le k-1) = \frac{k^3-(k-1)^3}{216}$. X 的分布律为:

X	1	2	3	4	5	6
P	$\frac{1}{216}$	$\frac{7}{216}$	$\frac{19}{216}$	$\frac{37}{216}$	$\frac{61}{216}$	$\frac{91}{216}$

➢ 这里涉及的是最大值, 对应的一般情况是次序统计量, 习题课会提到.

36. 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) $Y_3 = (X-1)^2$.

解:

(1) Y_1 的分布律为:

Y_1	-3	-1	1	3
P	0.4	0.1	0.3	0.2

(2) Y_2 的分布律为:

Y_2	0	1	2
P	0.3	0.3	0.4

(3) Y_3 的分布律为:

Y_3	0	1	4
P	0.1	0.7	0.2

37. 设连续型随机变量 X 的分布函数为:

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

(1) 试求常数 a, b 的值;
(2) 试求随机变量 $Y = 3 - \sqrt[3]{X}$ 的密度函数 $p(y)$;
(3) 试证明 X 与 $1/X$ 具有相同的分布.

解:

(1)

$$\begin{cases} F(-\infty) = a - \dfrac{b\pi}{2} = 0, \\ F(\infty) = a + \dfrac{b\pi}{2} = 1. \end{cases} \Longrightarrow \begin{cases} a = \dfrac{1}{2}, \\ b = \dfrac{1}{\pi}. \end{cases}$$

(2) $f(x) = F'(x) = \dfrac{1}{\pi(1+x^2)}$. $Y = 3 - \sqrt[3]{X}$ 关于 X 严格单调, 且 $X = (3-Y)^3$ 可导, 故由密度函数变换公式:

$$p(y) = f((3-Y)^3)|[(3-Y)^3]'| = \frac{3(3-y)^2}{\pi[1+(3-y)^6]}, -\infty < y < \infty.$$

(3) 由密度函数变换公式:

$$f_{\frac{1}{X}}(x)=f_X(\frac{1}{x})\left|\left(\frac{1}{x}\right)'\right|=\frac{1}{\pi(1+\frac{1}{x^2})}\cdot\frac{1}{x^2}=\frac{1}{\pi(1+x^2)}=f_X(x),$$

即 X 与 $1/X$ 具有相同的分布.

➢ 密度函数变换公式, 见课本 P73.

38. 设粒子运动速度服从正态分布, 求该粒子动能的分布.

解:

记速度为随机变量 V, 动能为随机变量 E, 质量为常数 M, 则 $E=\frac{1}{2}MV^2$, 其中 $V\sim N(\mu,\sigma)$.

$V(E)=\pm\sqrt{\dfrac{2E}{M}}\Longrightarrow|V'(E)|=\dfrac{1}{\sqrt{2EM}}$. 由推广的密度变换公式:

$$\begin{aligned}f_E(e)&=\frac{1}{\sqrt{2eM}}\left[f_V\left(\sqrt{\frac{2e}{M}}\right)+f_V\left(-\sqrt{\frac{2e}{M}}\right)\right]\\&=\frac{1}{\sqrt{2\pi eM}\sigma}\left[e^{-\frac{(\frac{2e}{M}-\mu)^2}{2\sigma^2}}+e^{-\frac{(-\frac{2e}{M}-\mu)^2}{2\sigma^2}}\right]\end{aligned}$$

关于具体的分布类型, 当 $\mu\neq 0$ 时, E 服从缩放的非中心卡方分布; 当 $\mu=0$ 时, E 服从缩放的卡方分布/伽马分布, 即 $E\sim\frac{1}{2}M\sigma^2\chi^2(1)$ / $E\sim Ga(\alpha=\frac{1}{2},\beta=\frac{1}{\frac{1}{2}m\sigma^2})$.

➢ 推广的密度函数变换公式, 见课本 P74 的注.

➢ 卡方分布; 伽马分布, 见课本 P193. 因为我感觉书上并不是很全, 所以这里简单补充一下相关的知识:

自由度为 n 的卡方分布:

$$f_{\chi^2(n)}(x)=\frac{1}{\Gamma(n/2)2^{n/2}}e^{-x/2}x^{n/2-1}I_{(0,\infty)}(x)$$

伽马分布:

$$f_{Ga}(x;\alpha,\beta)=\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$$

自由度为 n 的卡方分布和伽马分布的关系: 令 $\alpha=\frac{n}{2},\beta=\frac{1}{2}$, 则 $\chi^2(n)\Longleftrightarrow Ga(\frac{n}{2},\frac{1}{2})$.

若 X 服从标准正态分布, 则 X^2 服从自由度为 1 的卡方分布, 即 $X^2\sim\chi^2(1)$.

(这些分布的证明均可由分布函数的定义或运用密度函数变换公式推得, 这里不详细展开.)

39. 设元件寿命 X 服从指数分布 $Exp(\lambda)$, 求 $Y=XI_{(t,\infty)}(X)$ 的分布.

解:

$t \le 0$ 时, $Y = X \sim Exp(\lambda)$

$t > 0$ 时, $P(Y=0) = P(x \le t) = \int_0^t \lambda e^{-\lambda x} \mathrm{d}x = 1 - e^{-\lambda t}$, $f_Y(y) = f_X(y) = \lambda e^{-\lambda y}, \forall y \in [t, \infty)$.

$$\Longrightarrow F_Y(y) \begin{cases} 0, & y < 0 \\ 1 - e^{-\lambda t}, & 0 \le y \le t \\ 1 - e^{-\lambda y}, & y > t \end{cases}$$

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

解:

Y_1, Y_2, Y_3 均关于 X 单调, 由密度函数变换公式:

(1)

$$X = \ln Y_1 \in (0,1) \Longrightarrow f_{Y_1}(y) = f_X(\ln y) \cdot |(\ln y)'| = \frac{1}{y},\ y \in (1, e).$$

(2)

$$X = Y_2^{-1} \in (0,1) \Longrightarrow f_{Y_2}(y) = f_X(1/y) \cdot |(1/y)'| = \frac{1}{y^2},\ y \in (1, \infty).$$

(3)

$$X = e^{-\lambda Y_3} \in (0,1) \Longrightarrow f_{Y_3}(y) = f_X(e^{-\lambda y})|(e^{-\lambda y})'| = \lambda e^{-\lambda y},\ y \in (0, \infty).$$

➢ 密度函数变换公式, 见课本 P73.

41. 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.

解:

(1) 由密度函数变换公式:

$$X=\arctan Y_1\text{单调},\in(-\frac{\pi}{2},\frac{\pi}{2})\Longrightarrow f_{Y_1}(y)=f_X(\arctan y)|(\arctan y)'|=\frac{1}{\pi(1+y^2)},\ y\in\mathbb{R}.$$

(2) 由推广的密度变换公式:

$$X=\pm\arccos Y_2\text{分段单调},\in(-\frac{\pi}{2},\frac{\pi}{2})\Longrightarrow f_{Y_2}(y)=\sum f_X(\pm\arctan y)|(\pm\arctan y)'|=\frac{2}{\pi\sqrt{1-y^2}},\ y\in(0,1).$$

➢ 密度函数变换公式, 见课本 P73; 推广的密度函数变换公式, 见课本 P74 的注.

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

解:

$$P(Y\le y)=P(F(X)\le y)\xlongequal{\text{严格单调递增}}P(X\le F^{-1}(y))=F(F^{-1}(y))=y\Longrightarrow f_Y(y)=1,\ y\in(0,1).$$

43. 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 $Y_2=1-e^{-X}$ 的密度函数.

解:

$$X\sim Exp(1)\Longrightarrow X>0\Longrightarrow Y_1=X^2,\ Y_2=1-e^{-X}\text{在}(0,\infty).$$

由密度函数变换公式:

(1)

$$X=\sqrt{Y_1}\Longrightarrow f_{Y_1}(y)=f_X(\sqrt{y})|(\sqrt{y})'|=\frac{e^{-\sqrt{y}}}{2\sqrt{y}},\ y\in(0,\infty).$$

(2)

$$X=-\ln(1-Y_2)\Longrightarrow f_{Y_2}(y)=f_X(-\ln(1-y))|(-\ln(1-y))'|=e^{\ln(1-y)}\frac{1}{1-y}=1,\ y\in(0,1).$$

➢ 密度函数变换公式, 见课本 P73.

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

解:

$g(x)$ 在 $(0,1)$ 上单调增 $\Longrightarrow$ 由密度函数变换公式:

$$f_g(x) = f_X\left(g^{-1}(y)\right)\left|\left(g^{-1}\right)'(y)\right| = 2\left(1-g^{-1}(y)\right)\left(g^{-1}\right)'(y) = e^{-y} \Longleftrightarrow \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{e^{-y}}{2(1-x(y))},$$

解微分方程即得 $g(x) = y = -2\ln(1-x)$, $x \in (0,1)$.

➢ 密度函数变换公式, 见课本 P73.

45. 设随机变量 X 服从参数为 λ 的指数分布, 以 $Y=[X]$ 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 $Z = X-[X]$. 试求随机变量 Y 和 Z 各自的分布, 且它们是否相互独立?

解:

$$P(Y=k) = P(k \le x < k+1) = \int_k^{k+1} \lambda e^{-\lambda x}\mathrm{d}x = e^{-\lambda k} - e^{-\lambda(k+1)} = (1-e^{-\lambda})e^{-\lambda k}.$$

故 Y 服从参数为 $p = 1-e^{-\lambda}$ 的几何分布, 即 $Y \sim Ge(p=1-e^{-\lambda})$.

$$\begin{aligned} P(k \le X < k+z) = \int_k^{k+z} \lambda e^{-\lambda x}\mathrm{d}x = e^{-\lambda k}(1-e^{-\lambda z}) \Longrightarrow F_Z(z) = P(Z \le z) \quad &= P(X-[X] \le z) \\ &= \sum_{k=0}^{\infty} P(k \le X < k+z) \\ &= (1-e^{-\lambda z})\sum_{k=0}^{\infty} e^{-\lambda k} = \frac{1-e^{-\lambda z}}{1-e^{-\lambda}}. \end{aligned}$$

$\Longrightarrow f_Z(z) = F_Z'(z) = \dfrac{\lambda e^{-\lambda z}}{1-e^{-\lambda}}$, $z \in [0,1)$. 故 Z 服从 $[0,1)$ 上的截断指数分布 (Truncated Exponential Distribution), 即 $Z \sim TruncExp(\lambda; 0, 1)$.

考虑 Y, Z 的联合分布:

$$f_{Y,Z}(k,z) = f_X(k+z) = \lambda e^{-\lambda(k+z)} = P(Y=k) f_Z(z) \Longrightarrow Y, Z\text{独立}.$$

➢ 截断指数分布 (Truncated Exponential Distribution), 若我们希望变量只在区间 $[a,b]$ 上有分布, 而原始指数分布是 $Exp(\lambda)$, 故而为保持概率空间的规范性, 其截断版本的密度函数为:

$$f_{X|a\le X\le b}(X) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda a} - e^{\lambda b}},$$

记为 $X \sim Exp(\lambda)|a \le X \le b$ 或 $X \sim TruncExp(\lambda; a, b)$.

46. 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:

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

试求 Y 的密度函数 $p(y)$.

解:

Y 在 $X \in (0,1)$ 与 $[1,\infty)$ 分段单调且为双射 (即各段值域没有交集). 故由密度函数变换公式:

$$p(y) = \begin{cases} f_X(y) = \lambda e^{-\lambda y}, & y \in [1,\infty) \\ f_X(\sqrt{-y})|\sqrt{-y}'| = \dfrac{\lambda e^{-\lambda\sqrt{-y}}}{2\sqrt{-y}}, & y \in (-1,0) \\ 0, & otherwise \end{cases}$$

➢ 密度函数变换公式, 见课本 P73.

47. 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) $Y_1 = e^X$; (2) $Y_2 = |X|$; (3) $Y_3 = 2X^2 + 1$.

解:

(1) $Y_1(X)$ 在 $\mathbb{R}$ 上严格单调递增 $\Longrightarrow$ 由密度函数变换公式:

$$f_{Y_1}(y) = f_X(\ln y)|(\ln y)'| = \frac{e^{-(\ln y)^2/2}}{\sqrt{2\pi}y}, \ y > 0.$$

(2) $Y_2(X)$ 在 $(-\infty,0)$ 与 $(0,\infty)$ 上分段单调 $\Longrightarrow$ 由推广的密度变换公式:

$$f_{Y_2}(y) = f_X(y) + f_X(-y) = 2f_X(y) = \frac{2}{\sqrt{\pi}}e^{-y^2/2}, \ y \geq 0.$$

(3) $Y_3(X)$ 在 $(-\infty,0)$ 与 $(0,\infty)$ 上分段单调 $\Longrightarrow$ 由推广的密度变换公式:

$$f_{Y_3}(y) = \sum f_X(\sqrt{\pm\frac{y-1}{2}})\left|\pm\sqrt{\frac{y-1}{2}}'\right| = \frac{e^{-(y-1)/4}}{2\sqrt{\pi(y-1)}}, \ y > 1.$$

➢ 密度函数变换公式, 见课本 P73; 推广的密度函数变换公式, 见课本 P74 的注.

48. 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:

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

(1) 求随机变量 Y 的分布函数;

(2) 求概率 $P(X\le Y)$.

解:

(1) 先确定 a 的值:

$$\int_0^3 f(x)\mathrm{d}x=\frac{9}{a}=1\Longrightarrow a=9.$$

$P(Y\le y<1)=0$, $P(Y=1)=\int_2^3 f(x)\mathrm{d}x=\frac{19}{27}$, $P(Y=2)=\int_0^1 f(x)\mathrm{d}x=\frac{1}{27}$, $F_Y(y)=P(X=1)+\int_1^y f(x)\mathrm{d}x=\frac{y^3}{27}+\frac{2}{3}$, $1<y<2\Longrightarrow$

$$F_Y(y)=\begin{cases}0, & y<1\\ \frac{y^3}{27}+\frac{2}{3}, & 1\le y<2\\ 1, & y\ge 2\end{cases}$$

(2) $P(X\le Y)=P(X<2)=1-P(Y=1)=\dfrac{8}{27}$.

49. * 设随机变量 $X\sim U(0,1)$, 求下列随机变量的分布函数或密度函数:

(1) $Y=\frac{X}{1-X}$; (2) $Z=XI_{(a,1]}(X)$, 其中 $0<a<1$; (3) $W=X^2+XI_{[0,b]}(X)$, 其中 $0<b<1$.

解:

(1) $Y(X)$ 在 $(0,1)$ 上严格单调递增 $\Longrightarrow$ 由密度函数变换公式:

$$f_Y(y)=f_X(\frac{y}{1+y})\left|\left(\frac{y}{1+y}\right)'\right|=\frac{1}{(1+y)^2}.$$

(2) 显然有:

$$F_Z(z)=\begin{cases}0, & z<0\\ a, & 0\le y\le a\\ z, & a<z\le 1\\ 1, & z>1\end{cases}$$

(3) 当 $X\in[0,b]$ 时, $W=X^2+X\in[0,b^2+b]$; 当 $X\in(b,1]$ 时, $W=X^2\in(b^2,1]$. 两值域存在重合 $(b^2,\min\{b^2+b,1\}]$. 故:

$$F_W(w)=\begin{cases}\dfrac{-1+\sqrt{1+4w}}{2}, & 0\le w<b^2\\ \dfrac{-1+\sqrt{1+4w}}{2}+\sqrt{w}-b, & b^2\le w\le\min\{b^2+b,1\}\\ \dfrac{-1+\sqrt{1+4w}}{2}+1-b, & 1<w\le b^2+b\\ \sqrt{w}, & b^2+b<w\le 1\end{cases}$$

50. * 设随机变量 $X\sim N(0,1)$, 证明对任意 $x>0$, 有:

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}-\frac{1}{x^3}\right)e^{-\frac{x^2}{2}}\le P(X>x)\le\frac{1}{\sqrt{2\pi}x}e^{-\frac{x^2}{2}}.$$

解:

两次分部积分即可:

$$\begin{aligned}P(X>x)=\frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}\mathrm{d}t &=\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}e^{-x^2/2}-\int_x^\infty\frac{1}{t^2}e^{-t^2/2}\mathrm{d}t\right)\le\frac{1}{\sqrt{2\pi}x}e^{-\frac{x^2}{2}}\\ &=\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}e^{-x^2/2}-\frac{1}{x^3}e^{-x^2/2}+3\int_x^\infty\frac{1}{t^4}e^{-t^2/2}\right)\ge\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}-\frac{1}{x^3}\right)e^{-\frac{x^2}{2}}\end{aligned}$$

51. * (两端带吸收壁的随机游动) 设赌徒甲有本金 a 元, 赌徒乙有本金 b 元, 每局甲赢的概率为 p, 乙赢的概率为 $q=1-p$, 没有平局. 设每局输赢都是一元. 求赌徒甲输光的概率. 如果赌徒乙是赌场老板, 赌徒甲的资本 a 相对于赌场老板的赌本 b 而言是 $o(b)$, 设 $p=q=1/2$, 你对该结论有什么看法?
提示: 把上述模型视为一个质点在直线上左右游动, 设质点初始位置为 a. 质点向左游动一步的概率为 q, 表示乙赢甲输; 质点向右游动一步的概率为 p, 表示甲赢乙输. 以 p_n 表示质点位置在 n 而质点随机游动最终被 0 点吸收的概率 (即甲输光的概率), q_n 表示质点位置在 n 而质点随机游动最终被 $a+b$ 点吸收的概率 (即乙输光的概率). 由题意:

$$p_0=1, q_0=0,$$

$$p_{a+b}=0, q_{a+b}=1.$$

若在某时刻, 质点位于 $x=n$, 则它被 $x=0$ 吸收有两种方式来实现: 一种是接下来向右移动而最终被 $x=0$ 吸收; 另一种是向左移动而最终被 $x=0$ 吸收. 运用全概率公式, 得:

$$p_n=p\cdot p_{n+1}+q\cdot p_{n-1},\ n=1,2,\cdots,a+b-1,$$

这样我们得到一个二阶差分方程, 对应的特征方程为:

$$px^2-x+q=0.$$

设特征方程解为 x_1, x_2, 则通解为 $p_n=C_1x_1^n+C_2x_2^n=C_1+C_2\left(\frac{q}{p}\right)^n$, 利用边界条件即可得出甲最终输光的概率为:

$$p_a=\frac{\left(\frac{q}{p}\right)^a-\left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}}.$$

当 $p=q=\frac{1}{2}$ 时, 考虑当 $\frac{q}{p}\longrightarrow 1$ 时, p_a 的极限, 也可以直接从差分方程出发求出 p_a.

解:
按照提示所说计算即可, 或者直接求出递推式 $p(p_{n+1}-p_n)=q(p_n-p_{n-1})$, 最后得到 $p_a=\frac{\left(\frac{q}{p}\right)^a-\left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}}$.
令 $\frac{q}{p}\to 1$, 求 p_a 的极限, 或直接从头计算, 得到 $p_a=\frac{b}{a+b}$. 显然, 若 $a=o(b)$, 即 $\lim\limits_{b\to\infty}\frac{a}{b}=0$, 则 $\lim\limits_{b\to\infty}p_a=0$, 即甲几乎必定输光.

看法: 单局游戏公平不代表全局游戏公平, 当对手资本远大于自身资本时, 资本小的一方几乎必然破产. 根本原因在于随机游走的波动性: 虽然期望的收益为 0, 但方差的存在导致资本较少的一方更易触及吸收壁. (提示: 珍爱生命, 远离赌博.)(。í _ ì。)

第 3 章　多维随机变量及其分布

1. 箱中装有 6 个球, 其中红、白、黑球的个数分别为 1, 2, 3. 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) 的联合分布函数.

解:
易得, (X,Y) 的联合概率质量函数/联合分布律为:

$P(X,Y)$ \ Y \ X	0	1	2
0	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{1}{15}$
1	$\frac{1}{5}$	$\frac{2}{15}$	0

从而, 联合分布函数为:

$$F_{X,Y}(x,y)=\begin{cases}0, & if \quad x<0 \quad or \quad y<0,\\ \frac{1}{5}, & if \quad 0\le x<1 \quad and \quad 0\le y<1,\\ \frac{3}{5}, & if \quad 0\le x<1 \quad and \quad 1\le y<2,\\ \frac{2}{3}, & if \quad 0\le x<1 \quad and \quad y\ge 2,\\ \frac{2}{5}, & if \quad x\ge 1 \quad and \quad 0\le y<1,\\ \frac{14}{15}, & if \quad x\ge 1 \quad and \quad 1\le y<2,\\ 1, & if \quad x\ge 1 \quad and \quad y\ge 2.\end{cases}$$

2. 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 $P(X=1|Z=0)$;
(2) 求二维随机变量 (X,Y) 的联合分布函数.

解:
(1) 易得: $P(X=1|Z=0)=\frac{P(X=1,Z=0)}{P(Z=0)}=\frac{4}{9}$.

(2) 易得, (X,Y) 的联合概率质量函数/联合分布律为:

$P(X,Y)$ \ Y / X	0	1	2
0	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{9}$
1	$\frac{1}{6}$	$\frac{1}{9}$	0
2	$\frac{1}{36}$	0	0

从而, 联合分布函数为:

$$F_{X,Y}(x,y)=\begin{cases} 0, & if \quad x<0 \quad or \quad y<0, \\ \frac{1}{4}, & if \quad 0\le x<1 \quad and \quad 0\le y<1, \\ \frac{7}{12}, & if \quad 0\le x<1 \quad and \quad 1\le y<2, \\ \frac{25}{36}, & if \quad 0\le x<1 \quad and \quad y\ge 2, \\ \frac{5}{12}, & if \quad 1\le x<2 \quad and \quad 0\le y<1, \\ \frac{31}{36}, & if \quad 1\le x<2 \quad and \quad 1\le y<2, \\ \frac{35}{36}, & if \quad 1\le x<2 \quad and \quad y\ge 2, \\ \frac{4}{9}, & if \quad x\ge 2 \quad and \quad 0\le y<1, \\ \frac{8}{9}, & if \quad x\ge 2 \quad and \quad 1\le y<2, \\ 1, & if \quad x\ge 2 \quad and \quad y\ge 2. \end{cases}$$

3. 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 的联合分布律.

解:
易得, (X,Y) 的联合分布律为:

$P(X,Y)$ \ X / Y	0	1	2	3
1	0	$\frac{3}{8}$	$\frac{3}{8}$	0
3	$\frac{1}{8}$	0	0	$\frac{1}{8}$

4. 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

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

求 (X_1, X_2) 的联合分布律.

解:
易得, (X_1, X_2) 的联合分布律为:

$P(X,Y)$ X_1 / X_2	0	1
0	0.1	0.8
1	0.1	0

5. 设二维随机向量的联合分布律为

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.

解:

$$\left.\begin{aligned} P(X=-1, X+Y=0) = P(X=-1)P(X+Y=0) \Longleftrightarrow a = (0.2+a)(a+b) \\ a+b+0.2+0.3 = 1 \end{aligned}\right\} \Longrightarrow a = 0.2,\ b = 0.3$$

6. 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所进行的射击次数.
(1) 求二维随机变量 (X, Y) 的联合分布律;
(2) 求 X 和 Y 的边缘分布.

解:

(1) 由几何分布及其无记忆性, (X,Y) 的联合分布律为:

$$P(X=x,Y=y)=P(X=x)P(Y=y|X=x)=(1-p)^{x-1}p\cdot(1-p)^{y-x-1}p=p^2(1-p)^{y-2},\ 1\le x<y.$$

(2) 显然, X 服从参数为 p 的几何分布, Y 服从参数为 $(r=2,p)$ 的负二项分布, 即 $X\sim Ge(p)$, $Y\sim NB(r=2,p)$. 边缘分布律为: $P(X=x)=p(1-p)^{x-1}$, $P(Y=y)=(y-1)p^2(1-p)^{y-2}$.

➢ 几何分布及其无记忆性, 见课本 P54; 负二项分布, 见课本 P53.

7. 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, 求事件 $\{Y=2\}$ 发生的概率.

解:

由全概率公式:

$$P(Y=2)=\sum_{x=1}^{4}P(Y=2|X=x)\cdot P(X=x)=\frac{13}{48}.$$

➢ 全概率公式.

8. 设二维随机变量 (X,Y) 服从正态分布 $N(1,0,1,1,0)$, 求 $P(XY-Y<0)$.

解:

$(X,Y)\sim N(\mu_X=1,\mu_Y=0,\sigma_X^2=1,\sigma_Y^2=1,\rho=0)\Longrightarrow X\sim N(1,1)$, $Y\sim N(0,1)$, 且 $\mathrm{Cov}(X,Y)=\rho\sigma_X\sigma_Y=0\Longrightarrow X$, Y 相互独立 $\Longrightarrow$

$$P(XY-Y<0)=P((X-1)Y<0)=P(X<1,Y>0)+P(X>1,Y<0)=0.5.$$

➢ 二元正态分布的定义, 见课本 P93. 课本上关于二元和多元正态分布的解释比较少, 我会在习题课及讲义里进行详细的补充和推导, 不然这部分内容不好理解. 其中 ρ 代表 X 与 Y 的相关系数, 定义见课本 P146; Cov 代表协方差, 定义见课本 P144. 两者的关系是 $Cox(X,Y)=\rho_{X,Y}\sigma_X\sigma_Y$. 而独立可以推出协方差等于 0, 但反之未必成立. 这里协方差等于 0 能推出独立是由正态分布的定义决定的 (当然这道题不需要知道这些结论, 直接从二元正态分布的定义出发令 $\rho=0$ 即可).

9. 设二维随机变量 (X,Y) 的联合分布函数为

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

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

解:
(1)

$$\begin{cases} \lim\limits_{x\to\infty,y\to\infty} F(x,y) = a(b+\frac{\pi}{2})(c+\frac{\pi}{2}) = 1 \\ \lim\limits_{x\to\infty} F(x,y) = a(b-\frac{\pi}{2})(c+\arctan y) = 0 \\ \lim\limits_{y\to\infty} F(x,y) a(b+\arctan x)(c-\frac{\pi}{2}) = 0 \end{cases} \Longrightarrow a = \frac{1}{\pi^2},\ b = \frac{\pi}{2},\ c = \frac{\pi}{2}.$$

(2) 由二维随机变量在矩形区域内的概率与分布函数的关系:

$$P(X>0,Y>0) = F(+\infty,+\infty) - F(0,+\infty) - F(+\infty,0) + F(0,0) = \frac{1}{4}.$$

(3)
① 先求联合密度函数:

$$f_{X,Y}(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{1}{\pi^2(1+x^2)(1+y^2)}.$$

再积分求得边缘密度函数:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}y = \frac{1}{\pi(1+x^2)}, \qquad f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}x = \frac{1}{\pi(1+y^2)}.$$

② 先求边缘分布函数:

$$F_X(x) = F(x,+\infty) = \frac{1}{2} + \frac{\arctan x}{\pi}, \qquad F_Y(y) = F(+\infty,y) = \frac{1}{2} + \frac{\arctan y}{\pi}.$$

再求导得到边缘密度函数:

$$f_X(x) = F_X'(x) = \frac{1}{\pi(1+x^2)}, \qquad f_Y(y) = F_Y'(y) = \frac{1}{\pi(1+y^2)}.$$

➢ 二维随机变量在矩形区域内的概率与分布函数的关系，见课本 P89.

10. 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 $P(0<X<\pi/4, \pi/4<Y<\pi/2)$.

解:
(1)

$$F(x,y)=\int_0^x\int_0^y \cos u\cos v\,\mathrm{d}v\mathrm{d}u=\sin x\sin y,\ 0<x<\pi/2, 0<y<\pi/2$$

$$\Longrightarrow F(x,y)=\begin{cases}0, & if\quad x\le 0\quad or\quad y\le 0,\\ \sin x\cdot\sin y, & if\quad 0<x<\frac{\pi}{2}\quad and\quad 0<y<\frac{\pi}{2},\\ \sin x, & if\quad 0<x<\frac{\pi}{2}\quad and\quad y\ge\frac{\pi}{2},\\ \sin y, & if\quad x\ge\frac{\pi}{2}\quad and\quad 0<y<\frac{\pi}{2},\\ 1, & if\quad x\ge\frac{\pi}{2}\quad and\quad y\ge\frac{\pi}{2}.\end{cases}$$

(2)

$$P(0<X<\pi/4, \pi/4<Y<\pi/2)=\int_{\frac{\pi}{4}}^{\frac{\pi}{2}}\int_0^{\frac{\pi}{4}}\cos x\cos y\,\mathrm{d}x\mathrm{d}y=\frac{\sqrt{2}-1}{2}.$$

11. 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 $P(X\le 1|Y\le 1)$.

解:

(1)

$$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\int_0^x e^{-x}\mathrm{d}y=xe^{-x}\Longrightarrow f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{1}{x}.\ 0<y<x$$

(2)

$$\left.\begin{array}{r}P(Y\le 1)=\int_0^1\int_y^{+\infty}e^{-x}\mathrm{d}x\mathrm{d}y=1-e^{-1}\\ P(X\le 1,Y\le 1)=\int_0^1\int_y^1 e^{-x}\mathrm{d}x\mathrm{d}y=1-2e^{-1}\end{array}\right\}\Longrightarrow P(X\le 1|Y\le 1)=\frac{P(X\le 1,Y\le 1)}{P(Y\le 1)}=\frac{e-2}{e-1}.$$

12. 设二维随机变量 (X,Y) 的密度函数为

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

求常数 A 及条件密度函数 $f_{Y|X}(y|x)$.

解:

$$\left\{\begin{array}{l}U=X\\ V=Y-X\end{array}\right.\Longrightarrow\left\{\begin{array}{l}X=U\\ Y=U+V\end{array}\right.\Longrightarrow det(J)=\begin{vmatrix}1&0\\1&1\end{vmatrix}=1$$

$$\Longrightarrow\iint\limits_{\mathbb{R}^2}Ae^{-2x^2+2xy-y^2}\mathrm{d}x\mathrm{d}y=\iint\limits_{\mathbb{R}^2}Ae^{-u^2}\cdot e^{-v^2}\mathrm{d}u\mathrm{d}v=\pi A=1\Longrightarrow A=\frac{1}{\pi}.$$

$$f_X(x)=\frac{e^{-x^2}}{\pi}\int_{-\infty}^{+\infty}e^{-(y-x)^2}\mathrm{d}y=\frac{e^{-x^2}}{\pi}\int_{-\infty}^{+\infty}e^{-v^2}\mathrm{d}v=\frac{e^{-x^2}}{\sqrt{\pi}}\Longrightarrow f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{e^{-(y-x)^2}}{\sqrt{\pi}}.$$

13. 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 $P(X\le 0.25|Y=0.5)$.

解:

$$\iint\limits_{0<|x|<y<1} Ax^2\mathrm{d}x\mathrm{d}y = 2\int_0^1\int_0^y Ax^2\mathrm{d}x\mathrm{d}y = \frac{A}{6} = 1 \Longrightarrow A = 6.$$

$$f_Y(y) = \int_{-\infty}^{+\infty} 6x^2\mathrm{d}x = 2\int_0^y 6x^2\mathrm{d}x = 4y^3 \Longrightarrow f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{3x^2}{2y^3}$$

$$\Longrightarrow P(X\le 0.25|Y=0.5) = \int_{-0.5}^{0.25} f_{X|Y}(x|0.5)\mathrm{d}x = \frac{9}{16}.$$

14. 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 的分布律为 $P(Y=0)=P(Y=1)=\dfrac{1}{2}$, 求随机变量 $Z=XY$ 的分布函数.

解:

$$\left.\begin{aligned} P(Z\le z|Y=0) &= I_{[0,+\infty)} \\ P(Z\le z|Y=1) = P(X\le z) &= \Phi(z) \end{aligned}\right\} \Longrightarrow F_Z(z) = P(Y=0)P(Z\le z|Y=0)+P(Y=1)P(Z\le z|Y=1) = \frac{1}{2}I_{[0,+\infty)}+\frac{1}{2}\Phi(z).$$

➢ 示性函数 I_A, 见课本 P49.

15. 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.

解:

$$X\sim N(0,\sigma_1^2), Y\sim N(0,\sigma_2^2) \Longrightarrow X-Y\sim N(0,\sigma_1^2+\sigma_2^2) \Longrightarrow P(Z=1) = P(X\le Y) = P(X-Y\le 0) = \frac{1}{2}.$$

➢ 服从正态分布的随机变量的线性组合: 若 X, Y 相互独立, $X\sim N(\mu_X,\sigma_X^2)$, $Y\sim N(\mu_Y,\sigma_Y^2)$, 则 $aX+bY\sim N(a\mu_X+b\mu_Y, a^2\sigma_X^2+b^2\sigma_Y^2)$, 证明与课本 P109 例 3.22 同理.

16. 设二维随机变量 (X,Y) 的密度函数为

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

(1) 求 c 的值;

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

解:

(1)

$$\iint\limits_{\mathbb{R}^2} f(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{x^2+y^2<R^2} c(R-\sqrt{x^2+y^2})\mathrm{d}x\mathrm{d}y=\int_0^{2\pi}\int_0^R c(R-r)r\mathrm{d}r\mathrm{d}\theta=\frac{\pi R^3}{3}=1\Longrightarrow c=\frac{3}{\pi R^3}.$$

(2)

$$p=\iint\limits_{x^2+y^2\leq r^2} f(x,y)\mathrm{d}x\mathrm{d}y=\int_0^{2\pi}\int_0^r c(R-\rho)\rho\mathrm{d}\rho\mathrm{d}\theta=\frac{r^2(3R-2r)}{R^3}.$$

17. 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

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

(1) 求 (X,Y) 的联合密度函数 $f(x,y)$;

(2) 求 Y 的边缘密度函数 $f_Y(y)$.

解:

(1)

$$f(x,y) = f_X(x) f_{Y|X}(y|x) = \frac{9y^2}{x}, \quad 0 < y < x < 1.$$

(2)

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x = 9y^2 \int_y^1 \frac{1}{x}\mathrm{d}x = -9y^2 \ln y, \quad 0 < y < 1.$$

18. 设随机变量 X 的密度函数为 $f_X(x) = xe^{-x} I_{(0,\infty)}(x)$, 而随机变量 Y 服从 $(0, X)$ 上的均匀分布, 求
(1) (X, Y) 的联合分布函数;
(2) 随机变量 Y 的分布函数.

解:

(1)

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x) = xe^{-x} I_{(0,\infty)}(x) \cdot \frac{1}{x} I_{(0,x)}(y) = e^{-x}, \quad 0 < y < x.$$

$$\Longrightarrow F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\mathrm{d}v\mathrm{d}u = \begin{cases} 0, & if \quad x \le 0 \quad or \quad y \le 0 \\ 1 - e^{-x}(x+1), & if \quad 0 < x \le y \\ 1 - e^{-y} - ye^{-x}, & if \quad 0 < y < x \end{cases}$$

(3)

① 通过边缘密度函数积分:

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\mathrm{d}x = \int_y^{+\infty} e^{-x}\mathrm{d}x = e^{-y}, \ y > 0 \Longrightarrow F_Y(y) = \int_{-\infty}^{y} f_Y(v)\mathrm{d}v = (1 - e^{-y}) I_{(0,\infty)}(y).$$

② 求联合分布函数的极限:

$$F_Y(y) = \lim_{x \to +\infty} F_{X,Y}(x,y) = (1 - e^{-y}) I_{(0,\infty)}(y).$$

19. 设在 $X = x$ 的条件下, Y 服从参数为 x 的泊松分布, 已知 X 服从参数为 1 的指数分布, 求 Y 的分布律.

解:

已知 $Y|X \sim P(x)$, $X \sim Exp(1)$, 即 $P(Y=k|X=x)=\frac{x^k}{k!}e^{-x}$, $f_X(x)=e^{-x}I_{(0,\infty)}(x)$, 故:

$$P(Y=k)=\int_0^{+\infty}\frac{x^k}{k!}e^{-2x}\mathrm{d}x=\frac{1}{k!}\int_0^{+\infty}x^k e^{-2x}\mathrm{d}x=\frac{1}{k!}\cdot\frac{\Gamma(k+1)}{2^{k+1}}=\frac{1}{2^{k+1}}.$$

➢ 实际上 Y 服从参数为 $p=\frac{1}{2}$ 的几何分布, 不过这里几何分布的定义和书上的不同, Y 代表的第一次成功前进行的试验次数.

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

解:

$0\le x\le 2, 0\le y\le 2 \Longrightarrow z=|x-y|\in[0,2]$. $f_{X,Y}(x,y)=\frac{1}{4}I_{[0,2]}(x)I_{[0,2]}(y)$.

$$F_Z(z)=P(|X-Y|\le z)=P(X-z\le Y\le X+z)=\frac{4-(2-z)^2}{4},\ z\in[0.2]\Longrightarrow f_Z(z)=F'(z)=1-\frac{z}{2},\ z\in[0,2]$$

21. 设 (X,Y) 在区域 G 上服从均匀分布, G 由 $x-y=0$, $x+y=2$ 与 $y=0$ 围成. 试求
(1) 边缘密度函数 $f_X(x)$;
(2) 条件密度函数 $f_{X|Y}(x|y)$.

解:

(1)

$$f_{X,Y}(x,y)=\frac{1}{S}=1\Longrightarrow f_X(x)=\int_0^{\min\{x,2-x\}}f_{X,Y}(x,y)\mathrm{d}y=\begin{cases} x, & 0\le x\le 1\\ 2-x, & 1<x\le 2\\ 0, & x<0 \quad or \quad x>2\end{cases}$$

(2)

$$f_Y(y)=\int_y^{2-y}f_{X,Y}(x,y)\mathrm{d}x=2-2y,\ 0\le y\le 1\Longrightarrow f_{X|Y}(x|y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{1}{2(1-y)},\ y\le x\le 2-y.$$

22. 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) 条件密度函数 $f_{X|Y}(x|y)$;
(4) $P(X\le 0.5|Y=y)$.

解:
(1)
$$f(x,y)=\frac{1}{S}=2.$$
(2)
$$f_1(x)=\int_x^1 f(x,y)\mathrm{d}y=2(1-x),\quad 0\le x\le 1\qquad f_2(y)=\int_0^y f(x,y)\mathrm{d}x=2y,\quad 0\le y\le 1.$$
(3)
$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_2(y)}=\frac{1}{y},\quad 0\le x\le y.$$
(4)
$$P(X\le 0.5|Y=y)=\int_0^{\min\{0.5,y\}} f_{X|Y}(x|y)\mathrm{d}x=\frac{\min\{0.5,y\}}{y}.$$

23. 设 (X,Y) 服从矩形 $\{(x,y):0\le x\le 2, 0\le y\le 1\}$ 内的均匀分布, 记
$$U=\begin{cases}0, & X\le Y,\\ 1, & X>Y,\end{cases}\quad V=\begin{cases}0, & X\le 2Y,\\ 1, & X>2Y,\end{cases}$$
求 (U,V) 的联合分布函数.

解:
$$\begin{cases}P(U=0,V=0)=P(X\le Y)=\int_0^1\int_0^y \frac{1}{2}\mathrm{d}x\mathrm{d}y=\dfrac{1}{4}\\ P(U=0,V=1)=0\\ P(U=1,V=0)=P(Y<X\le 2Y)=\int_0^1\int_Y^{2Y}\frac{1}{2}\mathrm{d}x\mathrm{d}y=\dfrac{1}{4}\\ P(U=1,V=1)=P(X>2Y)=\int_0^1\int_{2Y}^2\frac{1}{2}\mathrm{d}x\mathrm{d}y=\dfrac{1}{2}\end{cases}$$

24. 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) 的联合概率质量函数;
(2) 已知取出的只有一张黑桃, 求此时 Y 的条件概率分布.

解:
(1)

$$P(X=x,Y=y)=\frac{\binom{13}{x}\binom{13}{y}\binom{26}{13-x-y}}{\binom{52}{13}},\quad 0\le x,y\le x+y\le 13.$$

(2)

$$P(X=1,Y=y)=\frac{\binom{13}{1}\binom{13}{y}\binom{26}{12-y}}{\binom{52}{13}},\ P(X=1)=\frac{\binom{13}{1}\binom{39}{12}}{\binom{52}{13}}\Longrightarrow P(Y=y|X=1)=\frac{P(X=1,Y=y)}{P(X=1)}=\frac{\binom{13}{y}\binom{26}{12-y}}{\binom{39}{12}}.$$

25. 假设有 $n\,(n\ge 3)$ 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) $(X_1,X_2,\cdots,X_n)$ 的联合分布函数;
(2) X_k 的边缘分布函数, 其中 $k=1,2,\cdots,n$;
(3) (X_1,X_2) 的边缘分布函数;
(4) 在 $X_1=m_1$ 的条件下, $(X_2,X_3,\cdots,X_n)$ 的条件分布函数.
(➢这里不知道为什么都是求分布函数, 但是我觉得没有什么意义, 大家求密度函数就好了.)

解:
(1) 显然, $(X_1,X_2,\cdots,X_n)$ 服从参数为 $N=m,\ p_1,\ p_2,\cdots,\ P_n$ 的多项分布, 即 $(X_1,X_2,\cdots,X_n)\sim M(N=m,p_1,p_2,\cdots,p_n)$, 从而 $(X_1,X_2,\cdots,X_n)$ 的联合密度函数为:

$$P(X_1=m_1,X_2=m_2,\cdots,X_n=m_n)=\binom{m}{m_1}\binom{m-m_1}{m_2}\cdots\binom{m-\sum\limits_{i=1}^{n-1}m_i}{m_n}p_1^{m_1}p_2^{m_2}\cdots p_n^{m_n}=\frac{m!}{m_1!m_2!\cdots m_n!}p_1^{m_1}p_2^{m_2}\cdots p_n^{m_n}$$

(2) 因为每个球独立地以概率 p_k 落入第 k 个盒子, 故 X_k 服从参数为 $N=m,\ p=p_k$ 的二项分布, 即 $X_k\sim B(N=m,p=p_k)$, 从而 X_k 的边缘密度函数为:

$$P(X_k=m_k)=\binom{m}{m_k}p_k^{m_k}(1-p_k)^{m-m_k}=\frac{m!}{m_k!(m-m_k)!}p_k^{m_k}(1-p_k)^{m-m_k}.$$

(3) 同理 (X_1, X_2) 的边缘分布函数为:

$$P(X_1 = m_1, X_2 = m_2) = \binom{m}{m_1}\binom{m-m_1}{m_2} p_1^{m_1} p_2^{m_2}(1-p_1-p_2)^{m-m_1-m_2} = \frac{m!}{m_1!m_2!(m-m_1-m_2)!} p_1^{m_1} p_2^{m_2}(1-p_1-p_2)^{m-m_1-m_2}.$$

(4)

①

$$P(X_2 = m_2, \cdots, X_n = m_n | X_1 = m_1) = \frac{P(X_1 = m_1, X_2 = m_2, \cdots, X_n = m_n)}{P(X_1 = m_1)} = \frac{(m-m_1)!}{m_2!\cdots m_n!}\left(\frac{p_2}{1-p_1}\right)^{m_2}\cdots\left(\frac{p_n}{1-p_1}\right)^{m_n}.$$

② 令 $m' = m - m_1$, $n' = n-1$, $p_k' = \frac{p_k}{1-p_1}$ 即与 (1) 同理, 故在 $X_1 = m_1$ 的条件下, $(X_2, X_3, \cdots, X_n)$ 的条件分布函数为:

$$P(X_2 = m_2, \cdots, X_n = m_n | X_1 = m_1) = \frac{(m-m_1)!}{m_2!\cdots m_n!}\left(\frac{p_2}{1-p_1}\right)^{m_2}\cdots\left(\frac{p_n}{1-p_1}\right)^{m_n}.$$

➢ 多项分布, 见第 49 题.

26. 设随机变量 X 与 Y 相互独立, X 的概率分布为

$$P(X = i) = \frac{1}{3}, \qquad i = -1, 0, 1,$$

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 $P(Z \le \frac{1}{2} | X = 0)$;

(2) 求 Z 的密度函数.

解:

(1)

$$P(Z \le \frac{1}{2} | X = 0) = P(Y \le \frac{1}{2}) = \int_0^{\frac{1}{2}} f_Y(y)\mathrm{d}y = \frac{1}{2}.$$

(2) 由卷积公式 (这里是连续 + 离散, 实际上从定义出发即可):

$$f_Z(z) = \sum_{i=-1}^{1} P(X = i) f_{Z|X}(z|x = i) = \frac{1}{3}(I_{[-1,0]} + I_{[0,1]} + I_{[1,2]}).$$

➢ 卷积, 见课本 P106 例 3.19.

27. 设随机变量 X 与 Y 相互独立, X 服从参数为 1 的指数分布, Y 服从标准正态分布, 求 $(X,|Y|)$ 的联合密度函数.

解:

$$X与Y相互独立 \Longrightarrow f_{X,|Y|}(x,z)=f_X(x)f_{|Y|}(z)=f_X(x)(f_Y(-z)+f_Y(z))=\frac{2}{\sqrt{2\pi}}e^{-x-z^2/2},\quad x\ge 0, z\ge 0.$$

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

解:
(1)

$$显然 Y|X=n\sim B(n,p)\Longrightarrow P(Y=m|X=n)=\binom{n}{m}p^m(1-p)^{n-m}.$$

(2)

$$P(X=n,Y=m)=P(X=n)P(Y=m|X=n)=\frac{\lambda^n e^{-\lambda}}{n!}\binom{n}{m}p^m(1-p)^{n-m}.$$

29. 设 $X\sim N(\mu,\sigma^2)$, $Y\sim B(1,p)$, 且随机变量 X, Y 相互独立. 求 XY 的分布.

解: 记 $Z=XY$:

$$\begin{cases} 若 Y=0, 则 Z=XY=0; \\ 若 Y=1, 则 Z=X\sim N(\mu,\sigma^2). \end{cases} \Longrightarrow Z=\begin{cases} 0, & 以概率 1-p \\ X\sim N(\mu,\sigma^2), & 以概率 p \end{cases} \Longrightarrow Z\sim(1-p)\delta+p\cdot N(\mu,\sigma^2)$$

XY 服从如上的混合分布.

➢ δ: 狄拉克函数, 满足

$$\delta(x)=\begin{cases}\infty & \text{当}\, x=0\\ 0 & \text{当}\, x\neq 0\end{cases}\quad \text{且}\quad \int_{-\infty}^{\infty}\delta(x)\,dx=1$$

直观理解为一个无限高、无限窄但面积为 1 的尖峰, 所有的质量都集中在 0 点上.
它在积分中的一个性质:

$$\int_{-\infty}^{\infty}\delta(x-a)f(x)\,dx=f(a)$$

这道题之所以要使用狄拉克函数, 本质在于"离散型随机变量"对应的"概率"与"连续型随机变量"对应的"概率密度"的差异: 连续型随机变量在任意一点的概率都是 0, 因为随机变量的概率是依托于样本集的测度比的, 离散型随机变量的点概率可能大于 0 是因为虽然样本点在整个实数集中的测度是 0, 但样本空间的总测度也是 0, 且样本点与样本空间的测度比是非负的, 此时可以简单理解为每个样本点所占的权重相同且非负; 但连续型随机变量的样本点测度为 0 而样本空间的总测度非 0, 故而每一点的概率均为 0, 而非概率密度. 只有对概率密度通过积分的方式求一个正测集上的概率才可能得到正值. 相反的, 要将离散型随机变量的概率转换成概率密度, 则需要借助狄拉克函数:

$$f_X(x)=P(X=a)\delta(x-a)$$

➢ 混合分布: 给定一组分布 $f_1(x),f_2(x),\ldots,f_k(x)$, 和相应的权重 (或混合概率)$\pi_1,\pi_2,\ldots,\pi_k$, 其中: $\pi_i\geq 0$, $\sum_{i=1}^{k}\pi_i=1$ 那么混合分布的密度函数为:

$$f(x)=\sum_{i=1}^{k}\pi_i f_i(x)$$

它表示一个从多个分布中随机选择一个进行采样的过程. 可以想象成先用一个隐变量 (通常是离散型的) 选择一个分布, 然后从该分布中采样.

30. 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

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

(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.

解:

(1)

$$X,Y\text{相互独立} \Longrightarrow f_{X,Y}(x,y)=f_X(x)f_Y(y)=\frac{1}{2}e^{-y/2}I_{(0,1)}(x)I_{(0,+\infty)}(y).$$

(2)

$$\Delta=4(X^2-Y)\geq 0 \Longrightarrow p=P(Y\leq X^2)=\iint_{Y\leq X^2} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y=\int_0^1\int_0^{x^2}\frac{1}{2}e^{-y/2}\mathrm{d}y\mathrm{d}x=\int_0^1(1-e^{-x^2/2})\mathrm{d}x$$

$$=1-\sqrt{2\pi}\int_0^1\phi(x)\mathrm{d}x=1-\sqrt{2\pi}(\Phi(1)-\Phi(0))\approx 0.1446(\text{查表 (课本附表 1)}).$$

31. 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.

解:

$f_{X,Y,Z}(x,y,z)=\frac{3}{4\pi},\ x^2+y^2+z^2\leq 1 \Longrightarrow$

$$f_X(x)=\iint_{y^2+z^2\leq 1-x^2} f_{X,Y,Z}(x,y,z)\mathrm{d}y\mathrm{d}z=\int_0^{2\pi}\int_0^{\sqrt{1-x^2}}\frac{3}{4\pi}\cdot r\mathrm{d}r\mathrm{d}\theta=\frac{3}{4}(1-x^2),\ x\in[-1,1].$$

32. 设随机变量 X,Y 的分布律分别为

X	0	1
P	$\frac{1}{2}$	$\frac{1}{2}$

Y	-1	0	1
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

且 $P(XY=0)=1$.

(1) 求 (X,Y) 的联合分布;

(2) 问 X,Y 是否独立.

解:

(1) $P(XY=0)=1 \Longrightarrow P(X=1,Y=-1)=P(X=1,Y=1)=0 \Longrightarrow P(X=0,Y=-1)=P(Y=-1)-P(X=1,Y=-1)=\frac{1}{4}$, $P(X=1,Y=0)=P(X=1)-P(X=1,Y=-1)-P(X=1,Y=1)=\frac{1}{2} \Longrightarrow P(X=0,Y=0)=P(Y=0)-P(X=1,Y=0)=0$, $P(X=0,Y=1)=P(Y=1)-P(X=1,Y=1)=\frac{1}{4}$. 即:

$P(X,Y)$ \ Y / X	-1	0	1
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{2}$	0

(2) $P(X=1,Y=-1)=0\neq P(X=1)P(Y=-1)\Longrightarrow X,\ Y$ 不独立.

33. 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 $X,\ Y$ 的边缘分布;
(2) 问 X 和 Y 是否相互独立.

解:
(1) 联合分布为:

$$P(X=x,Y=y)=\left(\frac{1}{3}\right)^{x-1}\cdot\frac{1}{6},\ x\geq 1, y=3,4,5,6.$$

直接得到边缘分布为 (也可以通过联合分布求和验证):

$$P(X=x)=\left(\frac{1}{3}\right)^{x-1}\cdot\frac{2}{3},\qquad P(Y=y)=\frac{1}{4}.$$

(2) $P(X=x,Y=y)=P(X=x)P(Y=y)\Longrightarrow X,\ Y$ 相互独立.

34. 令 X 和 Y 为独立的离散型随机变量, 且在 $\{0,1,\cdots,K-1\}$ 上均匀分布, 考虑

$$Z_n=X+nY\ (\mathrm{mod}\ K),$$

研究 Z_n 的独立性和两两独立性.

解:

$$P(Z_n=z)=P(X+nY\equiv z\mod K)=\sum_{y=0}^{K-1}P(X\equiv z-ny\mod K)P(Y=y)=\sum_{y=0}^{K-1}\frac{1}{K}\cdot\frac{1}{K}=\frac{1}{K},\quad \forall z\in\{0,1,\cdots,K-1\}.$$

两两独立性: 对 $\forall m\neq n$, $Z_m,\ Z_n$ 相互独立 $\Longleftrightarrow P(Z_m=z_m,Z_n=z_n)=P(Z_m=z_m)P(Z_n=z_n)=\dfrac{1}{K^2}$ $\Longleftrightarrow X,\ Y$ 满足下列方程组且取值唯一:

$$\begin{cases} X + mY \equiv z_m \mod K \\ X + nY \equiv z_n \mod K \end{cases} \Longrightarrow (m-n)Y \equiv z_m - z_n \mod K \Longrightarrow Y \equiv (z_m - z_n)(m-n)^{-1} \mod K$$

$(m-n)^{-1}$ 唯一, 要求 $(m-n,K)=1$, 从而此时 Y 唯一 $\Longrightarrow X$ 也唯一.
综上所述, Z_n 具有两两独立性 $\Longleftrightarrow (m-n,K)=1$.
一般独立性: 对 $\forall k\geq 2$, $\forall 1\leq n_1\leq n_2\cdots\leq n_k$, 有 $Z_{n_1}, Z_{n_2},\cdots, Z_{n_k}$ 相互独立, 显然整个序列/方程组的自由度为 2, 即不妨选取 Z_{n_1}, Z_{n_2}, 则对 $\forall 1\leq i\leq k$, Z_{n_i} 均可被 Z_{n_1}, Z_{n_2} 线性表示. 故易知 $k\geq 3$ 时, $\{Z_n\}$ 不具有相互独立性.

35. 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 $P(X>0.5|X+Y=1)$.

解:
(1)

$$\int_0^{+\infty}\int_0^{+\infty} Ae^{-(3x+4y)}\mathrm{d}x\mathrm{d}y = \frac{A}{12} = 1 \Longrightarrow A = 12.$$

(2) 先求边缘密度函数:

$$f_X(x) = \int_0^{+\infty} f(x,y)\mathrm{d}y = 3e^{-3x} \qquad f_Y(y) = \int_0^{+\infty} f(x,y)\mathrm{d}x = 4e^{-4y}.$$

显然 $f_{X,Y}(x,y) = f_X(x)f_Y(y) \Longrightarrow X, Y$ 独立.
(3) 由 (2) 知 X, Y 独立, 故由卷积公式:

$$f_Z(z) = \int_0^z f_X(x)f_Y(z-x)\mathrm{d}x = 12e^{-4z}(e^z-1),\ z\geq 0.$$

(4) 先求条件密度函数, 再积分:

$$f_{X|Z}(x|z) = \frac{f_{X,Y}(x,z-x)}{f_Z(z)} = \frac{e^x}{e-1},\ 0<x<1 \Longrightarrow P(X>0.5|X+Y=1) = \int_{0.5}^{1} f_{X|Z}(x|1)\mathrm{d}x = \frac{\sqrt{e}}{\sqrt{e}+1}.$$

➢ 卷积, 见课本 P106 例 3.19.

36. 设随机向量 (X,Y) 服从 $\{(x,y): |x+y|\le 1, |x-y|\le 1\}$ 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.

解:
(1) $|x+y|\le 1, |x-y|\le 1 \Longleftrightarrow -1\le x+y\le 1, -1\le x-y\le 1 \Longrightarrow f_{X,Y}(x,y)=\frac{1}{S}=\frac{1}{2}. \Longrightarrow$

$$f_X(x) = \int_{\max\{-1-x,-1+x\}}^{\min\{1-x,1+x\}} f_{X,Y}(x,y)\mathrm{d}y = \begin{cases} 1+x, & -1\le x\le 0 \\ 1-x, & 0<x\le 1 \\ 0, & x<-1\text{或}x>1 \end{cases} \overset{\text{对称性}}{\Longrightarrow} f_Y(y) = \begin{cases} 1+y, & -1\le y\le 0 \\ 1-y, & 0<y\le 1 \\ 0, & y<-1\text{或}y>1 \end{cases}$$

(2) $f_X(x)f_Y(y)\ne f_{X,Y}(x,y) \Longrightarrow X,\ Y$ 不相互独立.
(3)

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1}{2(1-x)},\quad 0<x<1,\ y\in[x-1,-x+1].$$

37. 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 $X=\dfrac{1}{2}$ 时, Y 的条件密度函数;
(2) 证明 X^2 和 Y^2 相互独立.

解:

(1)

$$f_X(x) = \int_{-1}^{1} f(x,y)\mathrm{d}y = 0.5 \Longrightarrow f_{Y|X}(y|x=\frac{1}{2}) = \frac{f(\frac{1}{2},y)}{f_X(\frac{1}{2})} = \frac{1}{2}+\frac{y}{4}.$$

(2) 由密度函数变换公式:

$$f_{X^2}(u) = f_U(u) = f_X(\sqrt{u})\cdot\frac{1}{2\sqrt{u}} + f_X(-\sqrt{u})\cdot\frac{1}{2\sqrt{u}} = \frac{1}{2\sqrt{u}} \overset{\text{对称性}}{\Longrightarrow} f_{Y^2}(v) = \frac{1}{2\sqrt{v}}.$$

$$J_{(x,y)\to(u,v)} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 0 \\ 0 & 2y \end{bmatrix} \Longrightarrow |\det J| = |2x\cdot 2y| = 4|xy| = 4\sqrt{uv}$$

$$\Longrightarrow f_{X^2,Y^2}(u,v) = \sum_{x^2=u,y^2=v} f_{X,Y}(x,y)|\det J_{(u,v)\to(x,y)}| = \frac{1}{4\sqrt{uv}}.$$

综上, $f_{X^2,Y^2}(u,v) = f_{X^2}(u)f_{Y^2}(v) \Longrightarrow X^2$ 与 Y^2 相互独立.

➢ 一元函数的密度函数变换公式, 见课本 P73; 第二题求 X^2, Y^2 的联合分布函数时用的是多元函数的密度函数变换公式, 即将一元函数的密度函数变换公式中反函数的导数换为反函数的 Jacobian 行列式, 当然不熟悉的同学也可以先求 X^2, Y^2 的联合分布函数, 再求二阶偏导得到: $f_{U,V}(u,v) = \frac{\partial^2}{\partial u\partial v}F_{U,V}(u,v)$, 这也是密度函数变换公式的本质.

38. 设随机变量 X, Y 的分布律分别为

X	0	1
P	$\frac{1}{3}$	$\frac{2}{3}$

Y	-1	0	1
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

且 $P(X^2=Y^2)=1$. 求

(1) (X,Y) 的联合分布;

(2) $Z=XY$ 的概率分布.

解:

(1) $P(X^2=Y^2)=1 \Longrightarrow P(X=0,Y=-1)=P(X=0,Y=1)=P(X=1,Y=0)=0 \Longrightarrow P(X=0,Y=0)=P(Y=0)-P(X=1,Y=0)=\frac{1}{3}$, $P(X=1,Y=-1)=P(Y=-1)-P(X=0,Y=-1)=\frac{1}{3}$, $P(X=1,Y=1)=P(Y=1)-P(X=0,Y=1)=\frac{1}{3}$, 即:

$P(X,Y)$ \ Y / X	-1	0	1
0	0	$\frac{1}{3}$	0
1	$\frac{1}{3}$	0	$\frac{1}{3}$

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

39. 设 (X,Y) 服从正方形 $\{(x,y):|x|+|y|\le 1\}$ 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?

解: 这道题和第 36 题是一样的, 答案见第 36 题 (1), (2).

40. 设随机变量 (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 $P(X+Y\le 1)$.

解:
(1)

$$f_X(x)=\int_0^x f(x,y)\mathrm{d}y=3x^2,\ 0<x<1 \qquad f_Y(y)=\int_y^1 f(x,y)\mathrm{d}x=\frac{3}{2}(1-y^2),\ 0<y<1.$$

(2) $f_X(x)f_Y(y)\ne f_{X,Y}(x,y)\Longrightarrow X$ 与 Y 不独立.
(3)

$$P(X+Y\le 1)=\int_0^1\int_0^{\min\{x,1-x\}}3x\mathrm{d}y\mathrm{d}x=\int_0^{\frac{1}{2}}\int_0^x 3x\mathrm{d}y\mathrm{d}x+\int_{\frac{1}{2}}^1\int_0^{1-x}3x\mathrm{d}y\mathrm{d}x=\frac{3}{8}.$$

41. 设 (X,Y) 的联合分布函数为

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

(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.

解:
(1) 易知 $|1-(x+1)e^{-x}|=1-(x+1)e^{-x}$, 故:

$$F_X(x)=\lim_{y\to\infty}F(x,y)=1-(x+1)e^{-x},\ x>0\qquad F_Y(y)=\lim_{X\to\infty}=\frac{y}{1+y},\ y>0.$$

(2)

$$f(x,y)=\frac{\partial^2F(x,y)}{\partial x\partial y}=\frac{xe^{-x}}{(1+y)^2},\ x>0,y>0$$

$$f_X(x)=F_X^{'}(x)=xe^{-x},\ x>0\qquad f_Y(y)=F_Y^{'}(y)=\frac{1}{(1+y)^2},\ y>0.$$

(3) $f_X(x)f_Y(y)=f_{X,Y}(x,y)\Longrightarrow X,\ Y$ 相互独立.

42. 设随机向量 (X,Y,Z) 的联合密度函数为

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

证明: X, Y, Z 两两独立但不相互独立.

解:

$$f_X(x)=\int_0^{2\pi}\int_0^{2\pi}f(x,y,z)\mathrm{d}y\mathrm{d}z=\frac{1}{2\pi},\ 0\le x\le 2\pi\overset{\text{对称性}}{\Longrightarrow}f_Y(y)=\frac{1}{2\pi},\ 0\le y\le 2\pi;\ f_Z(z)=\frac{1}{2},\ 0\le z\le 2\pi.$$

$$f_{X,Y}(x,y)=\int_0^{2\pi}f(x,y,z)\mathrm{d}z=\frac{1}{4\pi^2}=f_X(x)f_Y(y)\Longrightarrow X,Y\text{相互独立}\overset{\text{对称性}}{\Longrightarrow}X,Y,Z\text{两两独立}.$$

$f(x,y,x)\ne f_X(x)f_Y(y)f_Z(z)\Longrightarrow X,\ Y,\ Z$ 不相互独立.

43. 设随机变量 (X,Y) 的密度函数为

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

证明: X, Y 不独立但是 X^2, Y^2 是相互独立的.

解:

$$f_X(x)=\int_{-1}^{1}f(x,y)\mathrm{d}y=\frac{1}{2},\ |x|<1\overset{\text{对称性}}{\Longrightarrow}f_Y(y)=\frac{1}{2},\ |y|<1.\Longrightarrow f(x,y)\neq f_X(x)f_Y(y)\Longrightarrow X,Y\text{不独立}.$$

由密度函数变换公式:

$$f_{X^2}(u)=f_U(u)=f_X(\sqrt{u})\cdot\frac{1}{2\sqrt{u}}+f_X(-\sqrt{u})\cdot\frac{1}{2\sqrt{u}}=\frac{1}{2\sqrt{u}}\overset{\text{对称性}}{\Longrightarrow}f_{Y^2}(v)=\frac{1}{2\sqrt{v}}.$$

$$J_{(x,y)\to(u,v)}=\begin{bmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{bmatrix}=\begin{bmatrix}2x & 0\\ 0 & 2y\end{bmatrix}\Longrightarrow|\det J|=|2x\cdot 2y|=4|xy|=4\sqrt{uv}$$

$$\Longrightarrow f_{X^2,Y^2}(u,v)=\sum_{x^2=u,y^2=v}f_{X,Y}(x,y)|\det J_{(u,v)\to(x,y)}|=\frac{1}{4\sqrt{uv}}.$$

综上, $f_{X^2,Y^2}(u,v)=f_{X^2}(u)f_{Y^2}(v)\Longrightarrow X^2$ 与 Y^2 相互独立.

➢ 一元函数的密度函数变换公式, 见课本 P73; 第二题求 X^2, Y^2 的联合分布函数时用的是多元函数的密度函数变换公式, 即将一元函数的密度函数变换公式中反函数的导数换为反函数的 Jacobian 行列式, 当然不熟悉的同学也可以先求 X^2, Y^2 的联合分布函数, 再求二阶偏导得到: $f_{U,V}(u,v)=\frac{\partial^2}{\partial u\partial v}F_{U,V}(u,v)$, 这也是密度函数变换公式的本质.

44. * 设连续型随机变量 $X\sim f(x)$, Y 为取有限值的离散型随机变量, X, Y 相互独立.

(1) 求 $Z=X+Y$ 的分布. 由此回答随机变量 Z 是否为连续型的?

(2) 求 $W=XY$ 的分布, 问 W 是不是连续型随机变量? 以 $X\sim N(\mu,\sigma^2)$, $Y\sim B(1,p)$ 为例求出 W 具体的分布.

解:

(1) X 是连续型随机变量 $\Longleftrightarrow$ F_X 可以表示为某个 f_X 的积分 $\Longleftrightarrow$ $F_X(x) \in \mathrm{AC}(\mathbb{R}) \Longrightarrow F_Z(z) = \sum_{i=1}^{n} F_X(z-y_i)p_i \in \mathrm{AC}(\mathbb{R}) \Longrightarrow Z$ 是连续型随机变量.

(2)

对于每个 y_i:

若 $y_i > 0$, 则 $P(Xy_i \le w) = P(X \le w/y_i) = F_X(w/y_i)$.

若 $y_i < 0$, 则 $P(Xy_i \le w) = P(X \ge w/y_i) = 1 - F_X(w/y_i)$.

若 $y_i = 0$, 则 $P(0 \le w) = I_{[0,+\infty)}$.

因此:

$$F_W(w) = \sum_{y_i>0} F_X\left(\frac{w}{y_i}\right)p_i + \sum_{y_i<0}\left(1 - F_X\left(\frac{w}{y_i}\right)\right)p_i + \sum_{y_i=0} I(w \ge 0)p_i$$

若 $P(Y=0)=0$, 则可以对 $F_W(w)$ 求导:

$$f_W(w) = \sum_{y_i>0} \frac{f_X(w/y_i)}{y_i}p_i + \sum_{y_i<0} \frac{-f_X(w/y_i)}{y_i}p_i = \sum_{y_i\neq 0} \frac{f_X(w/y_i)}{|y_i|}p_i \Longrightarrow F_W(w) \in \mathrm{AC}(\mathbb{R})$$

故 W 为连续型随机变量.

若 $P(Y=0) \neq 0$, 则 $w=0$ 处有一个非负的点概率, 从而 $F_w(w)$ 无法表示为一般 lebesgue 意义下的积分 $\Longrightarrow W$ 不是连续型随机变量.

➢ 点概率、概率与概率密度的区别与联系, 见 28 题注释, 其中对于这道题, 虽然离散型随机变量的分布函数可以表示为狄拉克函数的积分, 但狄拉克函数代表的是离散测度, 而非 Lebesgue 测度, 故不能推出是连续型随机变量.

45. 在长度为 1 的线段上任取两点, 把线段分为 3 段, 求这 3 条线段能组成三角形的概率.

解:

取 $\forall x < y$, $x, y \in [0,1]$, 则要使 $x < 1-x$, $y-x < 1-y+x$, $1-y < y$, 即:

$$P(\max\{x, y-x, 1-y\} < \frac{1}{2}) = P(x < \frac{1}{2}, y-x < \frac{1}{2}, 1-y < \frac{1}{2}) = \frac{1}{4}.$$

46. * 设随机变量 X 的密度函数为 $f(x)$, 令

$$Y=\begin{cases} a, & X<a, \\ X, & a\le X<b, \\ b, & X\ge b. \end{cases}$$

(1) 求随机变量 Y 的分布函数;
(2) 随机变量 Y 的分布有什么特点?

解:
(1) 显然, Y 的分布函数为:

$$F_Y(y)=\begin{cases} 0, & y<a \\ F_X(y), & a\le y<b \\ 1, & y\ge b \end{cases}$$

(2)
① Y 是由 X 在 $[a,b]$ 上的截断得来的.
② Y 的分布是一个 a,b 的离散部分与 (a,b) 的连续部分的混合分布.

47. * 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 证明: 存在常数 b, 使 $X+bY$, $X-bY$ 相互独立.

解:
由协方差的线性性:

$$\mathrm{Cov}(X+bY,X-bY)=\mathrm{Var}(X)-b^2\mathrm{Var}(Y)=\sigma_1^2-b^2\sigma_2^2=0\Longrightarrow b=\pm\frac{\sigma_1}{\sigma_2}.$$

➢ 二元正态分布的定义, 见课本 P93. 课本上关于二元和多元正态分布的解释比较少, 我会在习题课及讲义里进行详细的补充和推导, 不然这部分内容不好理解. 其中 ρ 代表 X 与 Y 的相关系数, 定义见课本 P146; Cov 代表协方差, 定义及性质见课本 P144. 两者的关系是 $Cox(X,Y)=\rho_{X,Y}\sigma_X\sigma_Y$. 而独立可以推出协方差等于 0, 但反之未必成立. 这里协方差等于 0 能推出独立是由多元正态分布的定义决定的.

48. * 设 $(X,Y)\sim N(0,0,1,1,\rho)$, 证明

(1) 随机变量 X 与随机变量 $Z=\dfrac{Y-\rho X}{\sqrt{1-\rho^2}}$ 相互独立;

(2) 利用 (1) 的结论证明

$$P(XY<0)=1-2P(X>0,Y>0)=\pi^{-1}\arccos\rho.$$

解:

(1)

$$\mathrm{Cov}(X,Z)=\mathrm{Cov}\left(X,\frac{Y-\rho X}{\sqrt{1-\rho^2}}\right)=\frac{1}{\sqrt{1-\rho^2}}\big(\mathrm{Cov}(X,Y)-\rho\,\mathrm{Var}(X)\big)=\frac{1}{\sqrt{1-\rho^2}}\big(\rho-\rho\cdot 1\big)=0$$

(2) 由对称性, $P(XY<0)=2P(X>0,Y<0)=1-2P(X>0,Y>0)$.

$Z=\frac{Y-\rho X}{\sqrt{1-\rho^2}}$ 为 X, Y 的线性组合, 且 $E[Z]=\frac{1}{\sqrt{1-\rho^2}}E[Y]-\frac{\rho}{\sqrt{1-\rho^2}}E[X]=0$, $\mathrm{Var}(Z)=\frac{1}{\sqrt{1-\rho^2}}\mathrm{Var}(Y)-\frac{\rho^2}{\sqrt{1-\rho^2}}\mathrm{Var}(x)=1\implies Z\sim N(0,1)$. 故:

$$\begin{aligned}P(X>0,Y>0)=P(X>0,\rho X+\sqrt{1-\rho^2}Z>0)\quad&=P(X>0,Z>-\tfrac{\rho X}{\sqrt{1-\rho^2}})\\&\overset{\text{独立}}{=\!=}\int_0^{+\infty}\phi(x)\int_{-\frac{\rho X}{\sqrt{1-\rho^2}}}^{+\infty}\phi(z)\mathrm{d}z\mathrm{d}x\\&=\pi^{-1}\arccos\rho.\end{aligned}$$

➢ 二元正态分布的定义, 见课本 P93. 课本上关于二元和多元正态分布的解释比较少, 我会在习题课及讲义里进行详细的补充和推导, 不然这部分内容不好理解. 其中 ρ 代表 X 与 Y 的相关系数, 定义见课本 P146; Cov 代表协方差, 定义及性质见课本 P144. 两者的关系是 $Cox(X,Y)=\rho_{X,Y}\sigma_X\sigma_Y$. 而独立可以推出协方差等于 0, 但反之未必成立. 这里协方差等于 0 能推出独立是由多元正态分布的定义决定的. 这里的期望 (E) 以及方差 (Var) 都是后面的内容, 包括直接判断 Z 服从的分布类型等, 放在这里还是太超纲了, 想要在这一章仅凭定义就做出来不太现实, 建议可以后面学完了再回头看.

49. * (多项分布) 设 $A_1, A_2, \cdots, A_n$ 是样本空间上的一个完备事件群, 即 $A_iA_j = \emptyset, i \neq j$, 其和为必然事件, 记 $P(A_i) = p_i > 0, i = 1, 2, \cdots, n$, 则 $p_1 + p_2 + \cdots + p_n = 1$. 由上面知, 在一次试验中, 事件 $A_1, A_2, \cdots, A_n$ 必然发生一个且仅发生一个. 现在把该试验独立重复 N 次, 以 X_i 记事件 A_i 出现的次数, $i = 1, 2, \cdots, n$, 则称 $(X_1, X_2, \cdots, X_n)$ 服从多项分布, 记为

$$(X_1, X_2, \cdots, X_n) \sim M(N, p_1, p_2, \cdots, p_n),$$

其中 $X_1 + X_2 + \cdots + X_n = N$, 证明:
(1) (X_i, X_j, Z) 在给定条件 $X_l = N_l$ 的条件下服从三项分布 $M(N - N_l, \dfrac{p_i}{1-p_l}, \dfrac{p_j}{1-p_l}, 1 - \dfrac{p_i + p_j}{1-p_l})$, 其中 $1 \leq i \neq j \neq l \leq n$, $Z = N - N_l - X_i - X_j$;
(2) 设 $1 \leq j < n$, 则

$$X_1 + X_2 + \cdots + X_j \sim B(N, p_1 + p_2 + \cdots + p_j).$$

解:
(1) 这实际上就是一种迭代的想法, 总试验数变为 $N' = N - N_l$, 相对概率变为 $\frac{p_i}{1-p_l}$, 对一般的 $n \geq 2$, 严格计算如下:

$$\begin{cases} P(X_1 = N_1, \ldots, X_n = N_n) = \dbinom{N}{N_1}\dbinom{N-N_1}{N_2}\cdots\dbinom{N - \sum\limits_{k=1}^{n-1} N_k}{N_n} p_1^{N_1}\cdots p_n^{N_n} = \dfrac{N!}{N_1!\cdots N_n!} p_1^{N_1}\cdots p_n^{N_n} \\ X_l \sim B(N, N_l) \Longrightarrow P(X_l = N_l) = \binom{N}{N_l} p_l^{N_l}(1-p_l)^{N-N_l} \end{cases}$$

$$\begin{aligned} \Longrightarrow P(X_1 = N_1, \ldots, X_n = N_n | X_l = N_l) &= \frac{P(X_1 = N_1, \ldots, X_n = N_n)}{P(X_l = N_l)} \\ &= \tfrac{(N-N_l)!}{N_1!\cdots N_{l-1}!\cdots N_{l+1}!\cdots N_n!}\left(\tfrac{p_1}{1-p_l}\right)^{N_1}\cdots\left(\tfrac{p_{l-1}}{1-p_l}\right)^{N_{l-1}}\left(\tfrac{p_{l+1}}{1-p_l}\right)^{N_{l+1}}\cdots\left(\tfrac{p_n}{1-p_l}\right)^{N_n} \end{aligned}$$

故 $(X_1, X_2, \cdots, N - N_l - \sum\limits_{1\leq k\leq n, k\neq l} X_k, \cdots, X_n) \sim M(N - N_l, \dfrac{p_1}{1-p_l}, \cdots, 1 - \dfrac{\sum\limits_{1\leq k\leq n, k\neq l} p_k}{1-p_l}, \cdots, \dfrac{p_n}{1-p_l})$. 对 $n = 3$, 自然成立.
(2) 将 $A_1, A_2, \ldots, A_j$ 合并为一个新事件 B, 其概率为 $p = p_1 + p_2 + \cdots + p_j$. 其余事件 $A_{j+1}, \ldots, A_n$ 合并为 B^c, 其概率为 $1 - p$.
这样, 每次试验可以看作是一个伯努利试验, 结果为 B 或 B^c.
原始的多项分布 $(X_1, X_2, \ldots, X_n)$ 中, $X_1 + X_2 + \cdots + X_j$ 表示 N 次独立试验中事件 B 发生的次数.
因此, $X_1 + X_2 + \cdots + X_j$ 服从参数为 N 和 $p = p_1 + p_2 + \cdots + p_j$ 的二项分布 $B(N, p)$.

第 4 章　随机变量的数字特征和极限定理

1. 篮球联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 $E(X)$. 若 A 队每场获胜的概率均为 $p = 0.6$ 呢?

解: 记 A 队获胜的概率为 p, 则 B 队获胜的概率为 $1-p$.
类似于负二项分布, 对 $X = x$, 某一队在前 $x-1$ 场胜 3 场且最后一场获胜:

$$P(X=x)=\binom{x-1}{3}\left[p^4(1-p)^{x-4}+(1-p)^4p^{x-4}\right]\Longrightarrow E(X)=\sum_{x=4}^{7}xP(X=x)=\begin{cases}\frac{93}{16}=5.8125,\ p=0.5\\ \frac{89088}{15625}\approx 5.6973,\ p=0.6\end{cases}$$

2. 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

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

(2) 若 X 为非负连续型随机变量, 且分布函数为 $F(x)$, 则

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

(3) 若 X 为非负随机变量, 则 (2) 中的结论依然成立.

解:

(1)

$$\begin{aligned}E(X)=\sum_{k=1}^{\infty}kP(X=k)=\sum_{k=1}^{\infty}\sum_{n=1}^{k}P(X=k)\xlongequal{Tonelli}\sum_{n=1}^{\infty}\sum_{k=n}^{\infty}P(X=k)\quad&=\sum_{n=1}^{\infty}P(X\geq n)\\&\xlongequal[n'=n-1]{P(X\geq n)=P(X>n-1)}\sum_{n=0}^{\infty}P(X>n)\end{aligned}$$

(2)

$$E(X)=\int_0^{\infty}xf(x)\mathrm{d}x=\int_0^{\infty}\int_0^{x}f(x)\mathrm{d}t\mathrm{d}x=\int_0^{\infty}\int_t^{\infty}f(x)\mathrm{d}x\mathrm{d}t=\int_0^{\infty}P(X>t)\mathrm{d}t=\int_0^{\infty}(1-F(t))\mathrm{d}t.$$

(3)

由于一般随机变量的密度函数不一定存在, 故记数学期望为 (Lebesgue-Stieltjes 积分):

$$E(X)=\int_{\Omega}X(\omega)\mathrm{d}P(\omega)$$

若随机变量非负, 再利用示性函数:

$$E(X)=\int_{\Omega}\left(\int_0^{\infty}I_{\{X(\omega)>t\}}\mathrm{d}t\right)\mathrm{d}P(\omega)\xlongequal{Tonelli}\int_0^{\infty}\left(\int_{\Omega}I_{\{X(\omega>t)\}}\mathrm{d}P(\omega)\right)\mathrm{d}t=\int_0^{\infty}P(X>t)\mathrm{d}t=\int_0^{\infty}(1-F(t))\mathrm{d}t$$

➢ Tonelli 定理:

设 $(X,\mathscr{A},\mu)$ 和 $(Y,\mathscr{B},\nu)$ 是 σ 有限的测度空间, 且函数 $f:X\times Y\to[0,\infty]$ 在 $\mathscr{A}\otimes\mathscr{B}$ 上非负可测, 则有:

$$\int_{X\times Y}f(x,y)\,d(\mu\times\nu)(x,y)=\int_X\left(\int_Y f(x,y)\,d\nu(y)\right)d\mu(x)=\int_Y\left(\int_X f(x,y)\,d\mu(x)\right)d\nu(y)$$

离散形式的 Tonelli 定理:

设 (X,μ) 和 (Y,ν) 是计数测度空间, 且函数 $f:X\times Y\to[0,\infty]$, 则 Tonelli 定理变成求和次序可交换的问题:

$$\sum_{x\in X}\sum_{y\in Y}f(x,y)=\sum_{(x,y)\in X\times Y}f(x,y)=\sum_{y\in Y}\sum_{x\in X}f(x,y)$$

➢事实上, 对于一般的连续型随机变量, 我们有:

$$\begin{aligned}E(X)&=\int_0^{\infty}tf(t)\,dt+\int_{-\infty}^0 tf(t)\,dt\\&=\int_0^{\infty}\left(\int_0^t 1\,dx\right)f(t)\,dt-\int_{-\infty}^0\left(\int_t^0 1\,dx\right)f(t)\,dt\\&\xlongequal{Fubini}\int_0^{\infty}\left(\int_x^{\infty}f(t)\,dt\right)dx-\int_{-\infty}^0\left(\int_{-\infty}^x f(t)\,dt\right)dx\\&=\int_0^{\infty}(1-F(x))\,dx-\int_{-\infty}^0 F(x)\,dx.\end{aligned}$$

同理, 对于一般的离散型随机变量, 我们有:

$$E(X)=\sum_{k=1}^{\infty}P(X\geq k)-\sum_{k=-\infty}^{-1}P(X\leq k)$$

3. 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, 求 $E(X)$.

解:

记 $X=0.5X_1+0.5X_2$, 其中 $X_1\sim N(0,1)$, $X_2\sim N(4,4)$. 故由期望的线性性:

$$E(X)=0.5E(X_1)+0.5E(X_2)=0.5\times 0+0.5\times 4=2$$

➢ 标准化变换: $X\sim N(\mu,\sigma^2)\Longleftrightarrow F(x)=\Phi(\frac{x-\mu}{\sigma})$, 见课本 P71.

➢ 数学期望的性质, 见课本 P127.

4. 设 X 为一个连续型随机变量, 试对下列各种情形, 分别求 $E(X)$ 和 $\mathrm{Var}(X)$.

(1) 若 X 的条件密度函数为

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

其中 $\sigma>0$ 为常数, 则称 X 服从瑞利 (Rayleigh) 分布;

(2) 若 X 的条件密度函数为

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

其中 $\alpha,\beta>0$ 为常数, $\Gamma(x)$ 为 Γ 函数, 则称 X 服从 β 分布.

(3) 若 X 的条件密度函数为

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

其中 $k,\lambda>0$ 为常数, 则称 X 服从韦布尔分布.

解:

(1)

期望为:

$$E(X)=\int_0^{\infty}x\cdot\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x\xlongequal[\mathrm{d}x=\sqrt{2}\sigma\mathrm{d}t]{t=\frac{x}{\sqrt{2}\sigma}}2\sqrt{2}\sigma\int_0^{\infty}t^2e^{-t^2}\mathrm{d}t=\sqrt{\frac{\pi}{2}}\sigma$$

先计算二阶原点矩:

$$E(X^2)=\int_0^{\infty}x^2\cdot\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}\mathrm{d}x=2\sigma^2$$

再求方差:

$$\mathrm{Var}(X)=E(X^2)-[E(X)]^2=\left(2-\frac{\pi}{2}\right)\sigma^2$$

(2)

期望为:

$$E(X)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{\alpha}(1-x)^{\beta-1}\mathrm{d}x=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot B(\alpha+1,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}=\frac{\alpha}{\alpha+\beta}$$

先计算二阶原点矩:

$$E(X^2)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 x^{\alpha+1}(1-x)^{\beta-1}\mathrm{d}x=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot B(\alpha+2,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot\frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)}=\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

再求方差:

$$\mathrm{Var}(X)=E(X^2)-[E(X)]^2=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

(3)

期望为:

$$E(X)=k\int_0^{\infty}\left(\frac{x}{\lambda}\right)^k e^{-\left(\frac{x}{\lambda}\right)^k}\mathrm{d}x\xlongequal[\mathrm{d}x=\frac{\lambda}{k}t^{\frac{1}{k}-1}\mathrm{d}t]{t=\left(\frac{x}{\lambda}\right)^k}\lambda\int_0^{\infty}t^{\frac{1}{k}}e^{-t}\mathrm{d}t=\lambda\Gamma\left(1+\frac{1}{k}\right)$$

同理计算二阶原点矩得:

$$E(X^2)=\lambda^2\Gamma\left(1+\frac{2}{k}\right)$$

从而, 方差为:

$$\mathrm{Var}(X)=E(X^2)-[E(X)]^2=\lambda^2\left[\Gamma\left(1+\frac{2}{k}\right)-\Gamma^2\left(1+\frac{1}{k}\right)\right]$$

➢ 方差的性质, 见课本 P136.

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

解:

归一化条件:

$$\int_0^1 ax^2+bx+c\mathrm{d}x=\frac{a}{3}+\frac{b}{2}+c=1$$

计算期望:

$$E(X)=\int_0^1 x(ax^2+bx+c)\mathrm{d}x=\frac{a}{4}+\frac{b}{3}+\frac{c}{2}=0.5$$

先计算二阶原点矩:

$$E(X^2) = \int_0^1 x^2(ax^2+bx+c)\mathrm{d}x = \frac{a}{5}+\frac{b}{4}+\frac{c}{3}$$

再计算方差:

$$\mathrm{Var}(X) = E(X^2)-[E(X)]^2 = \frac{a}{5}+\frac{b}{4}+\frac{c}{3}-0.5^2 = 0.15$$

综上所述, 根据有关 a, b, c 的三元一次方程组解得:

$$\begin{cases} \frac{a}{3}+\frac{b}{2}+c=1 \\ \frac{a}{4}+\frac{b}{3}+\frac{c}{2}=0.5 \\ \frac{a}{5}+\frac{b}{4}+\frac{c}{3}=0.4 \end{cases} \Longrightarrow \begin{cases} a=12 \\ b=-12 \\ c=3 \end{cases}$$

➢ 方差的性质, 见课本 P136.

6. 盲盒营销已成功地用于玩具销售, 现在某快餐店也希望引入这种营销方式. 设按传统模式销售, 每天按每份 15 元计价可以售出 1000 份, 毛利润为 20%, 现在改为盲盒营销, 盲盒内食品价格和出现概率构成如下:

价格/元	5	10	15	20	50
概率	0.2	0.3	0.3	0.1	0.1

按盲盒营销, 平均每盒毛利润为 18%. 在此模式下, 设每天能售出 1500 盒, 问商家平均每天毛利润有多少? 比传统营销每天增加多少利润?

解:

记 X 为单个盲盒内食品的价格 (成本), 则盲盒的平均价格 (成本) 为:

$$E(X) = 0.2\times5+0.3\times10+0.3\times15+0.1\times20+0.1\times50 = 15.5$$

从而盲盒的售价为 $\frac{x-15.5}{x}=18\% \Longrightarrow x\approx 18.90$ (元). 平均每天的毛利润为 $1500\times18.90\times0.18=5103$ (元). 每天增加的利润为 $\Delta = 5103-1000\times15\times0.2 = 2103$ (元).

7. 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.

解:

记 A_i={第 i 个盒子为空}, 空盒子的个数为 X, 则 $X=\sum_{i=1}^{n} I_{A_i}$. 易知 $P(A_i)=\left(1-\frac{1}{n}\right)^n$, 故由期望的线性性:

$$E(X)=\sum_{i=1}^{n} E(I_{A_i})=\sum_{i=1}^{n} P(A_i)=n\left(1-\frac{1}{n}\right)^n.$$

$n\to\infty$ 时空盒子的平均比例为:

$$\lim_{n\to\infty}\frac{E(X)}{n}=\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n=\frac{1}{e}$$

➢ 数学期望的性质, 见课本 P127.

8. 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. 某人想集齐这套卡片, 设他一共需要买 X_n 袋该零食. 试求:

(1) $E(X_n)$; (2) $\lim_{n\to\infty} E\left(\frac{X_n}{n\ln n}\right)$.

解:

(1) 记 Y_i 为已经收集到 $i-1$ 张不同卡片的情况在, 再收集到一张新的卡片所需的次数. 则 $X_n=\sum_{i=1}^{n} Y_i$, 且 $Y_i\sim Ge(\frac{n-i+1}{n})\Longrightarrow E(Y_i)=\frac{n}{n-i+1}$. 故由期望的线性性:

$$E(X_n)=\sum_{i=1}^{n}E(Y_i)=\sum_{i=1}^{n}\frac{n}{n-i+1}=n\sum_{i=1}^{n}\frac{1}{i}$$

这里 $n=12\Longrightarrow E(X_{12})\approx 37.24$.

(2)

$$\lim_{n\to\infty}E\left(\frac{X_n}{n\ln n}\right)=\lim_{n\to\infty}\frac{1}{n\ln n}E(X_n)=\lim_{n\to\infty}\frac{\sum_{i=1}^{n}\frac{1}{i}}{\ln n}=\lim_{n\to\infty}\frac{\ln n+\gamma+\varepsilon_n}{\ln n}=1$$

➢ 数学期望的性质, 见课本 P127.

9. 现有 $n(n\ge 1)$ 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中随机摸出一球放入第三个袋子中, 照此做法依次进行下去, 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 $E(W_n)$.

解:

记 A_i={第 i 次摸出白球}, 则 $W_n=\sum\limits_{i=1}^{n} I_{A_i}$. 故由期望的线性性:

$$E(W_n)=\sum_{i=1}^{n} E(I_{A_i})=\sum_{i=1}^{n} P(A_i)$$

数学归纳法证明 $P(A_i)=\frac{a}{a+b}$:

$n=1$ 时, $P(A_1)=\frac{a}{a+b}$; 假设 $n=k$ 时, $P(A_k)=\frac{a}{a+b}$,

则 $n=k+1$ 时, $P(A_{k+1})=\frac{a+1}{a+b+1}P(A_k)+\frac{a}{a+b+1}P(\overline{A_k})=\frac{a}{a+b}$.

故 $E(W_n)=\frac{a}{a+b}n$.

➢ 数学期望的性质, 见课本 P127.

10. 设随机变量 X 的概率质量函数为 $P(X=k)=C/k!$, $k=0,1,2,\cdots$, 求 $E(X^2)$.

解:

归一化条件解得 C:

$$\sum_{k=1}^{\infty} P(X=k)=C\sum_{k=1}^{\infty}\frac{1}{k!}=Ce=1\Longrightarrow C=\frac{1}{e}$$

再求二阶原点矩:

$$E(X^2)=\frac{1}{e}\sum_{k=1}^{\infty}\frac{k^2}{k!}=\frac{1}{e}\sum_{k=1}^{\infty}\frac{k+k(k-1)}{k!}=\frac{1}{e}\sum_{k=1}^{\infty}\left(\frac{1}{(k-1)!}+\frac{1}{(k-2)!}\right)=2$$

11. 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:

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

解:

记 $p_i=P(X=x_i)>0$, $x_j=\max\limits_{1\leq i\leq k}\{x_i\}$, 则:

$$\lim_{n\to\infty}\frac{E(X^{n+1})}{E(X^n)}=\lim_{n\to\infty}\frac{\sum\limits_{i=1}^{k} p_i x_i^{n+1}}{\sum\limits_{i=1}^{k} p_i x_i^{n}}=\lim_{n\to\infty}\frac{p_j x_j^{n+1}\sum\limits_{i=1}^{k}\frac{p_i}{p_j}\left(\frac{x_i}{x_j}\right)^{n+1}}{p_j x_j^{n}\sum\limits_{i=1}^{k}\frac{p_i}{p_j}\left(\frac{x_i}{x_j}\right)^{n}}=x_j=\max_{1\leq i\leq k}\{x_i\}$$

12. 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 $E(X^n)$.

解:
(1) 记 $Y=cX$, 则由密度函数变换公式:

$$f_X(x)=\lambda e^{-\lambda x}\Longrightarrow f_Y(y)=f_X(\frac{y}{c})\cdot\frac{1}{c}=\frac{\lambda}{c}e^{-(\lambda/c)y}\Longrightarrow Y\sim Exp(\lambda/c)$$

(2)
$$E(X^n)=\int_0^\infty x^n\lambda e^{-\lambda x}\mathrm{d}x\xlongequal{t=\lambda x}\frac{1}{\lambda^n}\int_0^\infty t^n e^{-t}\mathrm{d}t=\frac{\Gamma(n+1)}{\lambda^n}=\frac{n!}{\lambda^n}$$

➢ 密度函数变换公式, 见课本 P73.

13. 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 $Z=1/X$ 的数学期望.

解:
由佚名统计学家公式:

$$E(Y)=E(e^X)=\int_1^2 e^x\cdot 2(x-1)\mathrm{d}x\xlongequal{t=x-1}\int_0^1 e^{t+1}\cdot 2t\mathrm{d}t=2e$$

$$E(Z)=E(\frac{1}{X})=\int_1^2\frac{1}{x}\cdot 2(x-1)\mathrm{d}x=2-2\ln 2$$

➢ 佚名统计学家公式, 即课本上数学期望的第三条性质, 见课本 P129.

14. 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 $E(X)$ 和方差 $\mathrm{Var}(X)$.

解:
记 $Y=\ln X$, 则由密度函数变换公式:

$$p(x)=f_X(x)=f_Y(\ln X)\cdot\frac{1}{x}=\frac{1}{\sqrt{2\pi}\sigma x}e^{-(\ln x-\mu)^2/(2\sigma^2)},\quad x>0$$

由佚名统计学家公式:

$$E(X)=E(e^Y)=\int_{-\infty}^{+\infty}e^y\cdot\frac{1}{\sqrt{2\pi}\sigma}e^{-(y-\mu)^2/(2\sigma^2)}\mathrm{d}y=e^{\mu+\sigma^2/2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-(y-\mu-\sigma^2)^2/(2\sigma^2)}\mathrm{d}y=e^{\mu+\sigma^2/2}$$

同理, 二阶原点矩为 $E(X^2)=e^{2\mu+2\sigma^2}$.

从而方差为:

$$\mathrm{Var}(X)=E(X^2)-[E(X)]^2=e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$$

➢ 密度函数变换公式, 见课本 P73.

➢ 佚名统计学家公式, 即课本上数学期望的第三条性质, 见课本 P129.

➢ 方差的性质, 见课本 P136.

15. 设随机变量 X 服从区间 $(-\pi/2,\pi/2)$ 上的均匀分布. 试求期望 $E(\sin X)$, $E(\cos X)$ 及 $E(X\cos X)$.

解:

由佚名统计学家公式:

$$E(\sin X)=\int_{-\pi/2}^{\pi/2}\sin x\cdot\frac{1}{\pi}\mathrm{d}x=0\quad E(\cos X)=\int_{-\pi/2}^{\pi/2}\cos x\cdot\frac{1}{\pi}\mathrm{d}x=\frac{2}{\pi}\quad E(X\cos X)=\int_{-\pi/2}^{\pi/2}x\cos x\cdot\frac{1}{\pi}\mathrm{d}x=0$$

➢ 佚名统计学家公式, 即课本上数学期望的第三条性质, 见课本 P129.

16. 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 $E[\mathrm{sgn}(X)\cdot X]$.

解:

(1) $P(\mathrm{sgn}=1)=P(0<X\le 1)=\dfrac{1}{3}$, $P(\mathrm{sgn}=0)=P(X=0)=0$, $P(\mathrm{sgn}=-1)=P(-2\le x<0)=\dfrac{2}{3}\Longrightarrow$ $E(\mathrm{sgn}(X))=-\dfrac{1}{3}$, $E([\mathrm{sgn}(X)]^2)=1$.

故方差为:

$$\mathrm{Var}(X)=E(X^2)-[E(X)]^2=\frac{8}{9}$$

(2)

注意到 $\mathrm{sgn}(X)\cdot X=|X|$, 则由佚名统计学家公式:

$$E[\mathrm{sgn}(X)\cdot X]=E(|X|)=\int_{-\infty}^{\infty}|x|\cdot\phi(x)\mathrm{d}x=2\int_0^{\infty}x\cdot\phi(x)\mathrm{d}x=\sqrt{\frac{2}{\pi}}$$

➢ 方差的性质, 见课本 P136.

➢ 佚名统计学家公式, 即课本上数学期望的第三条性质, 见课本 P129.

17. 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 $E(\min\{|X|,1\})$.

解:

由佚名统计学家公式:

$$\begin{aligned}E(\min\{|X|,1\})=\int_{-\infty}^{+\infty}\min\{|X|,1\}\cdot\frac{1}{\pi(1+x^2)}\mathrm{d}x&=2\int_0^{+\infty}\min\{X,1\}\cdot\frac{1}{\pi(1+x^2)}\mathrm{d}x\\&=2\left[\int_0^{+\infty}x\cdot\frac{1}{\pi(1+x^2)}\mathrm{d}x+\int_0^{+\infty}1\cdot\frac{1}{\pi(1+x^2)}\mathrm{d}x\right]\\&=\frac{1}{2}+\frac{\ln 2}{\pi}.\end{aligned}$$

➢ 佚名统计学家公式, 即课本上数学期望的第三条性质, 见课本 P129.

18. 设随机变量 X 的分布律为 $P(X=1)=P(X=2)=1/2$, 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 $E(Y)$.

解:

(1) 根据题意: $Y|X=1\sim U(0,1)$, $Y|X=2\sim U(0,2)$.

从而:

$$f_Y(y)=f_X(1)\cdot f_{Y|X=1}(y)+f_X(2)\cdot f_{Y|X=2}(y)=\frac{1}{2}I_{(0,1)}+\frac{1}{4}I_{(0,2)}\Longrightarrow f_Y(y)=\frac{3}{4}I_{(0,1)}+\frac{1}{4}I_{[1,2)}$$

$$\Longrightarrow F_Y(y)=\left(\frac{3}{4}y\right)I_{(0,1)}+\left(\frac{1}{2}+\frac{1}{4}y\right)I_{[1,2)}$$

(2)

$$E(Y)=\int_0^1 y\cdot\frac{3}{4}\mathrm{d}y+\int_1^2 y\cdot\frac{1}{4}\mathrm{d}y=\frac{3}{4}$$

19. 设二维随机变量 (X,Y) 服从二元正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 其中 $\mu_1=\mu_2=1$, $\sigma_1^2=\sigma_2^2=0.5$, $\rho=0.5$, 记

$$Z=|X-Y|,\qquad U=\max\{X,Y\},\qquad V=\min\{X,Y\}.$$

(1) 求 Z 的密度函数与期望 $E(Z)$;

(2) 分别求数学期望 $E(U)$ 和 $E(V)$.

解:

(1) 由期望和方差的性质, $E(X-Y)=E(X)-E(Y)=\mu_1-\mu_2=0$, $\mathrm{Var}(X-Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)=\sigma_1^2-\sigma_2^2=0.5$. 由正态组合的线性封闭性: $X-Y\sim N(0,0.5)$. 从而由推广的密度函数变换公式, Z 的密度函数为:

$$f_Z(z)=f_{X-Y}(z)+f_{X-Y}(-z)=2f_{X-Y}(z)=\frac{2}{\sqrt{\pi}}e^{-z^2},\quad z>0$$

进而 Z 的期望为:

$$E(Z)=\int_0^\infty z\cdot f_Z(z)\mathrm{d}z=\frac{1}{\sqrt{\pi}}\int_0^\infty 2ze^{-z^2}\mathrm{d}z=\frac{1}{\sqrt{\pi}}$$

(2) 易知

$$U=\frac{X+Y+|X-Y|}{2}=\frac{X+Y+Z}{2}\qquad V=\frac{X+Y-|X-Y|}{2}=\frac{X+Y-Z}{2}$$

故由期望的线性性:

$$E(U)=\frac{1}{2}[E(X)+E(Y)+E(Z)]=1+\frac{1}{2\sqrt{\pi}}\qquad E(V)=\frac{1}{2}[E(X)+E(Y)-E(Z)]=1-\frac{1}{2\sqrt{\pi}}$$

➢ 数学期望的性质, 见课本 P127; 方差的性质, 见课本 P136.

➢ 正态组合的线性封闭性: 正态分布的线性组合仍为正态分布 (无论独立与否), 课本上似乎只有独立条件下的证明, 见课本 P196.

➢ 推广的密度函数变换公式, 见课本 P74 的注.

20. 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) $E(X_2|X_1=k)$ 和 $\mathrm{Var}(X_2|X_1=k)$;

(2) $E(X_1+X_2)$ 和 $\mathrm{Var}(X_1+X_2+\cdots+X_k),\ k=1,2,\cdots,n$.

解:

(1) 显然有 $X_2|X_1=k\sim B(m-k,\dfrac{p_2}{1-p_1}),\ k\geq 2\Longrightarrow E(X_2|X_1=k)=(m-k)\dfrac{p_2}{1-p_1}$, $\mathrm{Var}(X_2|X_1=k)=(m-k)\dfrac{p_2}{1-p_1}\left(1-\dfrac{p_2}{1-p_1}\right)$.

(2) 由期望的线性性: $E(X_1+X_2)=E(X_1)+E(X_2)=m(p_1+p_2)$. 显然 $\sum\limits_{i=1}^{k}X_i\sim B(m,\sum\limits_{i=1}^{k}p_i)$, 故 $\mathrm{Var}(\sum\limits_{i=1}^{k}X_i)=m\sum\limits_{i=1}^{k}p_i\left(1-\sum\limits_{i=1}^{k}p_i\right)$.

➢ 二项分布的期望, 见课本 P128 例 4.7; 二项分布的方差, 见课本 P138 例 4.20.

➢ 数学期望的性质, 见课本 P127.

21. (1) 设随机变量 X 与 Y 相互独立, 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 $k\leq m$, 求 $P(X=k|X+Y=m)$ 及 $E(X|X+Y=m)$;

(2) 设随机变量 X 与 Y 相互独立, 均服从二项分布 $B(n,p)$, 对任何给定的非负整数 $k\leq m$, 求 $P(X=k|X+Y=m)$ 及 $E(X|X+Y=m)$.

解:

(1) $X\sim P(\lambda),\ Y\sim P(\mu)\Longrightarrow P(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda},\ P(Y=m-k)=\dfrac{\mu^{m-k}}{(m-k)!}e^{-\mu}$. 由独立性:

$$\begin{aligned}P(X+Y=m)=\sum_{k=0}^{m}P(X=k,Y=m-k)=\sum_{k=0}^{m}\frac{\lambda^k}{k!}e^{-\lambda}\cdot\frac{\mu^{m-k}}{(m-k)!}e^{-\mu}&=\frac{e^{-(\lambda+\mu)}}{m!}\sum_{k=0}^{m}\binom{m}{k}\lambda^k\mu^{m-k}\\&=\frac{(\lambda+\mu)^m}{m!}e^{-(\lambda+\mu)}\end{aligned}$$

即 $X+Y\sim P(\lambda+\mu)$. 则 $P(X=k|X+Y=m)=\dfrac{P(X=k,Y=m-k)}{P(X+Y=m)}=\dbinom{m}{k}\left(\dfrac{\lambda}{\lambda+\mu}\right)^k\left(\dfrac{\mu}{\lambda+\mu}\right)^{m-k}$, 即 $X|X+Y=m\sim B(m,\dfrac{\lambda}{\lambda+\mu})$. 故 $E(X|X+Y=m)=\dfrac{\lambda}{\lambda+\mu}m$.

(2)

① 根据二项分布的定义, X 与 Y 均为 n 次成功概率为 p 的伯努利试验得到的成功次数, 又由于 X 与 Y 相互独立, 故 $X+Y\sim B(2n,p)$. 即得 $P(X+Y=m)=\binom{2n}{m}p^m(1-p)^{2n-m}$.

② 根据概率卷积证明 $X+Y\sim B(2n,p)$:

$$\begin{aligned}P(X+Y=m)=\sum_{k=1}^{m}P(X=k)P(Y=m-k) &=\sum_{k=0}^{m}\binom{n}{k}p^k(1-p)^{n-k}\binom{n}{m-k}p^{m-k}(1-p)^{n-(m-k)}\\ &=p^m(1-p)^{2n-m}\sum_{k=0}^{m}\binom{n}{k}\binom{n}{m-k}\\ &\xlongequal{Vandermonde}p^m(1-p)^{2n-m}\binom{2n}{m}\end{aligned}$$

综上所述: $P(X=k|X+Y=m)=\dfrac{P(X=k,Y=m-k)}{P(X+Y=m)}=\dfrac{\binom{n}{k}\binom{n}{m-k}}{\binom{2n}{m}}$.

①$'$ $X|X+Y=m$ 服从超几何分布, 即 $X|X+Y=m\sim H(2n,n,m)$, 故 $E(X|X+Y=m)=\dfrac{mn}{2n}=\dfrac{m}{2}$.

②$'$ 根据 X 与 Y 的对称性, $E(X|X+Y=m)=E(Y|X+Y=m)$. 又由期望的线性性, $E(X+Y|X+Y=m)=E(X|X+Y=m)+E(Y|X+Y=m)=E(m)=m\Longrightarrow E(X|X+Y=m)=\dfrac{m}{2}$.

➢ 二项分布的期望, 见课本 P128 例 4.7.

➢ Vandermonde 恒等式: $\sum_{i=0}^{n}\binom{a}{i}\binom{b}{n-i}=\binom{a+b}{n}$, 证明略

➢ 超几何分布的期望, 见课本 P166 例 4.43.

➢ 数学期望的性质, 见课本 P127.

22. 假设随机变量 X 有分布律 $P(X=0)=P(X=1)=P(X=2)=1/3$, 随机变量 Y 在 $X=k$ 的条件下服从均值为 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 的协方差.

解:

(1)

$$f_Y(y)=\sum_{k=0}^{2}P(X=k)\cdot f_{Y|X=k}(y)=\frac{1}{3}\left[\frac{1}{\sqrt{2\pi}}e^{-(y-0)^2/2}+\frac{1}{\sqrt{2\pi}}e^{-(y-1)^2/2}+\frac{1}{\sqrt{2\pi}}e^{-(y-2)^2/2}\right]$$

$Y|X=k\sim N(k,1)\Longrightarrow E(Y|X=k)=k$, 即 $E(Y|X)=X$. 故由全期望公式: $E(Y)=E(E(Y|X))=E(X)=1$.

(2)

$$F_{X+Y}(z) = P(X+Y \le z) = \sum_{k=0}^{2} P(Y \le z-k|X=k)P(X=k) \xlongequal{\text{标准化}} \sum_{k=0}^{2} \Phi(\frac{z-k-k}{1})P(X=k)$$
$$= \frac{1}{3}[\Phi(z)+\Phi(z-2)+\Phi(z-4)]$$

(3) 由协方差的性质:

$$\left.\begin{aligned} E(XY) = E(E(XY|X)) = E(X \cdot E(Y|X)) = E(X^2) = \frac{5}{3} \\ E(X) = 1,\ E(Y) = 1 \end{aligned}\right\} \Longrightarrow \mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{2}{3}$$

➢ 全期望公式, 见课本 P131.

➢ 标准化变换: $X \sim N(\mu, \sigma^2) \Longleftrightarrow F(x) = \Phi(\frac{x-\mu}{\sigma})$, 见课本 P71.

➢ 协方差的性质, 见课本 P144.

23. 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. 试求最优的投资组合, 即找 $\omega \in [0,1]$ 使得 $\omega X + (1-\omega)Y$ 的夏普比率 (Sharpe ratio)

$$R(\omega) = \frac{E[\omega X + (1-\omega)Y]}{\sqrt{\mathrm{Var}(\omega X + (1-\omega)Y)}}$$

达到最大.

解:

$$E[\omega X + (1-\omega)Y] = \omega E(X) + (1-\omega)E(Y) = \mu(2-\omega)$$

$$\mathrm{Var}(\omega X + (1-\omega)Y) = \omega^2 \mathrm{Var}(X) + (1-\omega)^2 \mathrm{Var}(Y) + 2\omega(1-\omega)\rho\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}$$

$$\Longrightarrow R(\omega) = \frac{\mu(2-\omega)}{\sigma\sqrt{(4-\sqrt{3})\omega^2 - (6-\sqrt{3})\omega + 3}}$$

求导计算得 $\omega^* = \dfrac{42-2\sqrt{3}}{73}$.

➢ 数学期望的性质, 见课本 P127; 方差与协方差, 见课本 P144.

24. 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个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) 求到期收益的期望 $E((X+Y-3\mu)_+)$;
(2) 求到期收益的方差 $\mathrm{Var}((X+Y-3\mu)_+)$.

解:
$E(X+Y)=E(X)+E(Y)=3\mu$, $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y)\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}=4\sigma^2\Longrightarrow$ $X+Y\sim N(3\mu,4\sigma^2)$.

(1)
$$E((X+Y-3\mu)_+)=\int_0^\infty z\cdot\frac{1}{\sqrt{2\pi}\cdot 2\sigma}e^{-z^2/(8\sigma^2)}\mathrm{d}z=\sigma\sqrt{\frac{2}{\pi}}$$

(2)
$$E(((X+Y-3\mu)_+)^2)=\int_0^\infty z^2\cdot\frac{1}{\sqrt{2\pi}\cdot 2\sigma}e^{-z^2/(8\sigma^2)}\mathrm{d}z=2\sigma^2$$

$$\Longrightarrow \mathrm{Var}((X+Y-3\mu)_+)=E(((X+Y-3\mu)_+)^2)-E((X+Y-3\mu)_+)^2=2\sigma^2\left(1-\frac{1}{\pi}\right)$$

➢ 数学期望的性质, 见课本 P127; 方差与协方差, 见课本 P144.
➢ 正态组合的线性封闭性: 正态分布的线性组合仍为正态分布 (无论独立与否), 课本上似乎只有独立条件下的证明, 见课本 P196.
➢ 方差的性质, 见课本 P136.

25. 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 $E(\min\{X_1,X_2\})$ 和 $E(\max\{X_1,X_2\})$.

解:

$$P(\min\{X_1,X_2\}\le t)=1-P(\min\{X_1,X_2\}>t)=1-P(X_1>t,X_2>t)\xlongequal{\text{独立}}1-P(X_1>t)P(X_2>t)=1-e^{-4t}$$

$$P(\max\{X_1,X_2\}\le t)=P(X_1\le t,X_2\le t)=P(X_1\le t)P(X_2\le t)=\left(1-e^{-2t}\right)^2$$

$$\Longrightarrow f_{\min\{X_1,X_2\}}(t)=4e^{-4t}\qquad f_{\max\{X_1,X_2\}}(t)=4e^{-2t}\left(1-e^{-2t}\right)$$

$$\Longrightarrow E(\min\{X_1,X_2\})=\int_0^\infty 4te^{-4t}\mathrm{d}t=\frac{1}{4}\qquad E(\max\{X_1,X_2\})=\int_0^\infty 4te^{-2t}\left(1-e^{-2t}\right)\mathrm{d}t=\frac{3}{4}$$

(或算出 $E(\min\{X_1,X_2\})=\dfrac{1}{4}$ 后，根据期望的线性性：$E(\max\{X_1,X_2\})=E(X_1)+E(X_2)-E(\min\{X_1,X_2\})=\dfrac{3}{4}$.)

➢ 数学期望的性质，见课本 P127.

26. 设 X_1, X_2, X_3 服从球面 $\{(x_1,x_2,x_3)\in\mathbb{R}^3: x_1^2+x_2^2+x_3^2=1\}$ 上的均匀分布，求 $X_1+X_2+X_3$ 的期望 $E(X_1+X_2+X_3)$ 和方差 $\mathrm{Var}(X_1+X_2+X_3)$.

解:

由对称性和期望的线性性:

$$E(X_1)=E(X_2)=E(X_3)=0\Longrightarrow E(X_1+X_2+X_3)=0$$

$$E(X_1^2)=E(X_2^2)=E(X_3^2),\ E(X_1^2+X_2^2+X_3^2)=E(1)=1\Longrightarrow E(X_1^2)=E(X_2^2)=E(X_3^2)=\frac{1}{3}$$

$$\mathrm{Var}(X_1+X_2+X_3)=\mathrm{Var}(X_1)+\mathrm{Var}(X_2)+\mathrm{Var}(X_3)+2\mathrm{Cov}(X_1,X_2)+2\mathrm{Cov}(X_1,X_3)+2\mathrm{Cov}(X_2,X_3)$$

同样由对称性:

$$E(X_iX_j)=0\Longrightarrow \mathrm{Cov}(X_i,X_j)=E(X_iX_j)-E(X_i)E(X_j)=0,\quad i,j=1,2,3$$

综上, $\mathrm{Var}(X_1+X_2+X_3)=0$.

➢ 数学期望的性质，见课本 P127; 方差与协方差，见课本 P144.

27. 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 的泊松分布;

(3) 参数为 λ 的指数分布; (4) 正态分布 $N(\mu,\sigma^2)$.

解:

(1)

$$M_X(s)=E(e^{sX})=\sum_{k=0}^{n}e^{sk}\binom{n}{k}p^k(1-p)^{n-k}=\sum_{k=0}^{n}\binom{n}{k}\left(pe^t\right)^k(1-p)^{n-k}=\left(1-p+pe^t\right)^n$$

(2)

$$M_X(s)=E(e^{sX})=\sum_{k=0}^{\infty}e^{sk}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^s)^k}{k!}=e^{\lambda(e^s-1)}$$

(3)

$$M_X(s)=E(e^{sX})=\int_0^{\infty}e^{sx}\cdot\lambda e^{-\lambda x}\mathrm{d}x=\int_0^{\infty}\lambda e^{(s-\lambda)x}\mathrm{d}x=\frac{\lambda}{\lambda-s},\quad s<\lambda$$

(这里 $s<\lambda$ 是为了积分收敛, 当然从积分后的结果非负也可以看出.)

(4)

$$M_X(s)=E(e^{sX})=\int_{-\infty}^{+\infty}e^{sx}\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-\mu)^2/(2\sigma^2)}\mathrm{d}x=e^{\mu s+\sigma^2s^2/2}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-[x-(\mu+\sigma^2s)]^2/(2\sigma^2)}=e^{\mu s+\sigma^2s^2/2}$$

28. 设某人连续独立地投掷一枚均匀的骰子 (即投出点数 1, 2, 3, 4, 5, 6 的概率均为 1/6), 直到点数和大于或等于 10 停止.

(1) 求投掷次数的期望;

(2) 求停止时点数和的期望.

解:

(1) 记 e_n 为当前点数和为 n 时, 停止所需的投掷次数的期望. 则 $e_n=0,\ n\geq 10$ 且 $e_n=1+\frac{1}{6}\sum_{k=1}^{6}e_{n+k}$.

计算可得 $e_0=\frac{33495175}{10077696}\approx 3.32$.

(2) 同理, 记 S_n 为当前点数和为 n 时, 最终停止时点数和的期望. 则 $S_n=n,\ n\geq 10$ 且 $S_n=\frac{1}{6}\sum_{k=1}^{6}S_{n+k}$.

计算可得 $S_0=\frac{234466225}{20155392}\approx 11.63$.

29. 设 X_1,X_2 是相互独立的指数分布随机变量, 期望分别为 1 和 2. 定义

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

求 (1) $E(Y)$ 和 $E(Z)$; (2) $\mathrm{Var}(Y)$ 和 $\mathrm{Var}(Z)$.

解: $E(X_1)=1,E(X_2)=2\Longrightarrow X_1\sim Exp(1),\ X_2\sim Exp(\frac{1}{2})$.

(1)

$$P(Y\leq t)=1-P(Y>t)=1-P(X_1>t,X_2>t)\overset{\text{独立}}{=\!=}1-P(X_1>t)P(X_2>t)=1-e^{-3t/2}$$

$$P(Z\leq t)=P(X_1\leq t,X_2\leq t)=P(X_1\leq t)P(X_2\leq t)=\left(1-e^{-t}\right)\left(1-e^{-t/2}\right)$$

$$\Longrightarrow f_Y(t)=\frac{3}{2}e^{-3t/2}\qquad f_Z(t)=e^{-t}\left(1-e^{-t/2}\right)+\frac{1}{2}e^{-t/2}\left(1-e^{-t}\right)$$

$$\Longrightarrow E(Y)=\int_0^{\infty}tf_Y(t)\mathrm{d}t=\frac{2}{3}\qquad E(Z)=\int_0^{\infty}tf_Z(t)\mathrm{d}t=\frac{7}{3}$$

(或算出 $E(Y)=\frac{2}{3}$ 后, 根据期望的线性性: $E(Z)=E(X_1)+E(X_2)-E(Y)=\frac{7}{3}$.)

(2) 由 (1) 可得 $Y\sim Exp(\frac{3}{2})$, 故 $\mathrm{Var}(Y)=\frac{1}{(\frac{3}{2})^2}=\frac{4}{9}$. (根据概率密度函数和定义计算也可).

关于 $\mathrm{Var}(Z)$, 先计算 Z 的二阶原点矩:

$$Y+Z=X_1+X_2,\quad YZ=X_1X_2\Longrightarrow Z^2+Y^2=X_1^2+X_2^2\Longrightarrow E(Z^2)=E(X_1^2)+E(X_2^2)-E(Y^2)=\frac{82}{9}$$

其中 $E(N^2)=\mathrm{Var}(N)+[E(N)]^2=\frac{2}{\lambda^2}$, $N\sim Exp(\lambda)$.

故 $\mathrm{Var}(Z)=E(Z^2)-[E(Z)]^2=\frac{11}{3}$.

➢ 数学期望的性质, 见课本 P127; 方差的性质, 见课本 P136.

30. 已知二维随机变量 (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 $\mathrm{Cov}(X^2,Y^2)$.

解:

先求边际分布:

X	-1	0	1
P	0.5	0.3	0.2

Y	-1	0	1
P	0.2	0.35	0.45

X^2	0	1
P	0.3	0.7

Y^2	0	1
P	0.35	0.65

再计算期望:

$E(X)$	$E(Y)$	$E(XY)$	$E(X^2)$	$E(Y^2)$	$E(X^2Y^2)$
-0.3	0.25	-0.05	0.7	0.65	0.45

故

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0.025 \qquad \mathrm{Cov}(X^2,Y^2)=E(X^2Y^2)-E(X^2)E(Y^2)=-0.005$$

31. 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 $\mathrm{Cov}(X,Y)$.

解:
(1) 关于 X, 显然有: $E(X)=\frac{7}{2}$, $\mathrm{Var}(X)=\frac{35}{12}$. 关于 Y:

$$\begin{aligned}P(Y=k)=P(Y\le k)-P(Y\le k-1) &= P(\max\{X_1,X_2\}\le k)-P(\max\{X_1,X_2\}\le k-1)\\ &= P(X_1\le k)P(X_2\le k)-P(X_1\le k-1)P(X_2\le k-1)\\ &= \frac{2k-1}{36}\end{aligned}$$

$$\Longrightarrow E(Y)=\frac{161}{36} \qquad \mathrm{Var}(Y)=\frac{2555}{1296}$$

(2)

$$E(XY)=\frac{1}{36}\left(\sum_{1\le X_2\le X_1\le 6}X_1^2+\sum_{1\le X_1<X_2\le 6}X_1X_2\right)=\frac{1127}{72}\Longrightarrow \mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{35}{24}$$

32. 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?

解:
(1) 由协方差的性质:

$$\mathrm{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)=\alpha^2\mathrm{Var}(X)-\beta^2\mathrm{Var}(Y)=\left(\alpha^2-\beta^2\right)\sigma^2$$

(2) X 与 Y 均服从正态分布 $\Longrightarrow \alpha X+\beta Y$ 与 $\alpha X-\beta Y$ 也服从正态分布 $\Longrightarrow \alpha X+\beta Y$ 与 $\alpha X-\beta Y$ 相互独立当且仅当 $\mathrm{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)=0$, 即 $\alpha^2=\beta^2$.

➢ 协方差的性质, 见课本 P144.

➢ 正态组合的线性封闭性: 正态分布的线性组合仍为正态分布 (无论独立与否), 课本上似乎只有独立条件下的证明, 见课本 P196.

33. 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 $\mathrm{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)=0$? 如果存在请求出, 否则请说明原因.

解:

由协方差的性质:

$$\mathrm{Cov}(\alpha X+\beta Y,\alpha X-\beta Y)=\alpha^2\mathrm{Var}(X)-\beta^2\mathrm{Var}(Y)=\left(\alpha^2-\beta^2\right)\sigma^2=0 \Longleftrightarrow \alpha^2=\beta^2$$

➢ 协方差的性质, 见课本 P144.

34. 设随机变量 (X,Y) 服从区域 $G=\{(x,y):|x|+|y|\le 1\}$ 中的均匀分布.

(1) 求 $\mathrm{Cov}(X,Y)$; (2) X 与 Y 是否相互独立?

解:

(1) 由对称性, $E(X)=E(Y)=0$, $E(XY)=0$. 故 $\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0$.

(2)

$$S=2 \Longrightarrow f_{X,Y}(x,y)=\frac{1}{2} \Longrightarrow f_X(x)=\int_{-(1-|x|)}^{1-|x|} f_{X,Y}(x,y)\mathrm{d}y=1-|x| \overset{\text{对称性}}{\Longrightarrow} f_Y(y)=1-|y|$$

故 $f_{X,Y}(x,y)\neq f_X(x)f_Y(y) \Longleftrightarrow X$ 与 Y 不相互独立.

35. 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期望为 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期望和标准差.

解:

记每次事故的索赔额为 X, 保险公司单次赔付的实际事故金额为 Y, 总赔付金额为 S. $E(X)=2000$, $X\sim Exp(\lambda) \Longrightarrow \lambda=\frac{1}{2000} \Longrightarrow f_X(x)=\frac{1}{2000}e^{-x/2000}$, $x>0 \Longrightarrow E(Y)=\int_{700}^{\infty}(x-700)\cdot\frac{1}{2000}e^{-x/2000}\mathrm{d}x\approx 1409.37$.

由全期望公式:

$$E(S)=E(\sum_{i=1}^{N}Y_i)=E(E(\sum_{i=1}^{N}Y_i|N))=E(N\cdot E(Y))=E(N)E(Y)\approx 1174.48$$

由条件方差公式:

$$\begin{aligned}\mathrm{Var}(S)=E(\mathrm{Var}(S|N))+\mathrm{Var}(E(S|N)) &= E(\mathrm{Var}(\sum_{i=1}^{N}Y_i|N))+\mathrm{Var}(E(\sum_{i=1}^{N}Y_i|N))\\ &= E(N\cdot\mathrm{Var}(Y))+\mathrm{Var}(N\cdot E(Y))\\ &= \mathrm{Var}(Y)E(N)+[E(Y)]^2\mathrm{Var}(N)\approx 3980747\end{aligned}$$

故 $\sigma_S=\sqrt{\mathrm{Var}(S)}\approx 1995.18$.

➢ 全期望公式, 见课本 P131.

➢ 条件方差公式: $\mathrm{Var}(X)=E[\mathrm{Var}(X\mid Y)]+\mathrm{Var}(E[X\mid Y])$. 证明如下:

$$\begin{aligned}\mathrm{Var}(X)&=E[X^2]-(E[X])^2\\ &=E[E[X^2\mid Y]]-(E[E[X\mid Y]])^2\\ &=E[\mathrm{Var}(X\mid Y)+(E[X\mid Y])^2]-(E[E[X\mid Y]])^2\\ &=E[\mathrm{Var}(X\mid Y)]+E[(E[X\mid Y])^2]-(E[E[X\mid Y]])^2\\ &=E[\mathrm{Var}(X\mid Y)]+\mathrm{Var}(E[X\mid Y])\end{aligned}$$

36. 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 $\rho_{X,Y}=-0.5$, 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.

解:

(1) $\mathrm{Cov}(X,Y)=\rho_{X,Y}\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}=-0.5\sigma^2\Longrightarrow$

$$\mathrm{Var}(Z)=\pi^2\mathrm{Var}(X)+(1-\pi)^2\mathrm{Var}(Y)+2\pi(1-\pi)\mathrm{Cov}(X,Y)=\sigma^2[1-3\pi(1-\pi)]<\sigma^2$$

(2) ① $\dfrac{\partial\mathrm{Var}(Z)(\pi)}{\partial\pi}=0\Longrightarrow\pi^*=0.5$. ② 由对称性, $\pi^*=0.5$.

37.

(1) 证明

$$\operatorname{Cov}(X_1, X_2) = \operatorname{Cov}(X_1, E(X_2|X_1));$$

(2) 假设存在常数 c, $E(X_2|X_1) = 1 + cX_1$, 证明

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

解:

(1)

$$\begin{aligned}
\operatorname{Cov}(X_1, X_2) &= E(X_1 X_2) - E(X_1) E(X_2) \\
&\xlongequal{\text{全期望公式}} E(E(X_1 X_2|X_1)) - E(X_1) E(E(X_2|X_1)) \\
&= E(X_1 E(X_2|X_1)) - E(X_1) E(E(X_2|X_1)) \\
&= \operatorname{Cov}(X_1, E(X_2|X_1))
\end{aligned}$$

(2)

$$\begin{aligned}
\operatorname{Cov}(X_1, X_2) &\xlongequal{(1)} \operatorname{Cov}(X_1, E(X_2|X_1)) \\
&= \operatorname{Cov}(X_1, 1 + cX_1) \\
&= \operatorname{Cov}(X_1, 1) + c\operatorname{Cov}(X_1, X_1) \\
&= 0 + c\operatorname{Var}(X_1) \\
&= c\operatorname{Var}(X_1) \Longrightarrow c = \frac{\operatorname{Cov}(X_1, X_2)}{\operatorname{Var}(X_1)}
\end{aligned}$$

➢ 全期望公式, 见课本 P131.

38. 若 $E(X_2|X_1) = 1$, 证明

$$\operatorname{Var}(X_1 X_2) \geq \operatorname{Var}(X_1).$$

解:

$$\begin{aligned}
\operatorname{Var}(X_1X_2) &= E\left[(X_1X_2)^2\right] - [E(X_1X_2)]^2 \\
&\xlongequal{\text{全期望公式}} E\left[E\left(X_1^2X_2^2 \mid X_1\right)\right] - [E[E(X_1X_2|X_1)]]^2 \\
&= E\left[X_1^2E\left(X_2^2 \mid X_1\right)\right] - [E[X_1E(X_2|X_1)]]^2 \\
&= E\left[X_1^2E\left(X_2^2 \mid X_1\right)\right] - [E(X_1)]^2 \\
&= E\left[X_1^2\left(\operatorname{Var}(X_2 \mid X_1) + [E(X_2 \mid X_1)]^2\right)\right] - [E(X_1)]^2 \\
&= E\left[X_1^2(\operatorname{Var}(X_2 \mid X_1) + 1)\right] - [E(X_1)]^2 \\
&= E\left[X_1^2\operatorname{Var}(X_2 \mid X_1)\right] + E\left[X_1^2\right] - [E(X_1)]^2 \\
&= E\left[X_1^2\operatorname{Var}(X_2 \mid X_1)\right] + \operatorname{Var}(X_1) \\
&\geq \operatorname{Var}(X_1)
\end{aligned}$$

➢ 全期望公式, 见课本 P131.

39. 设 $X_1, X_2, \cdots, X_n, \cdots$ 是独立同分布的随机变量序列, 其分布为几何分布 $Ge(p)$; 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 $S = \sum_{i=1}^{N} X_i$. 求 $E(S|N=n)$, $\operatorname{Var}(S|N=n)$ 和 $\operatorname{Var}(S)$.

解:

$$E(S|N=n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \frac{n}{p}$$

$$\operatorname{Var}(S|N=n) = \operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}(X_i) = n\frac{1-p}{p^2}$$

由条件方差公式:

$$\operatorname{Var}(S) = E[\operatorname{Var}(S|N)] + \operatorname{Var}(E[S|N]) = \operatorname{Var}(N/p) + E[N\frac{1-p}{p^2}] = \frac{1}{p^2}\operatorname{Var}(N) + \frac{1-p}{p^2}E(N) = \frac{nq}{p^2}(2-p-q)$$

➢ 数学期望的性质, 见课本 P127; 方差的性质, 见课本 P136.

➢ 条件方差公式: $\operatorname{Var}(X) = E[\operatorname{Var}(X \mid Y)] + \operatorname{Var}(E[X \mid Y])$. 证明见第 35 题注释.

40. 设 $N(t)$ 是一个依赖于变量 t 的随机变量, 对 $t>0$, $N(t)$ 的分布律为

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

设 T 是一个期望为 a, 方差为 $b>0$ 的非负随机变量.
求 (1) $\mathrm{Cov}(T,N(T))$; (2) $\mathrm{Var}(N(T))$.

解:
(1)

$$N(T)|T=t\sim P(\lambda t)\Longrightarrow E(N(T)|T=t)=\lambda t\overset{\text{全期望}}{\underset{\text{公式}}{\Longrightarrow}}E(N(T))=E(E(N(T)|T))=E(\lambda T)=\lambda E(T)=\lambda a$$

$$E(T\cdot N(T)|T=t)=tE(N(t))=\lambda t^2\Longrightarrow E(T\cdot N(T))=E(E(T\cdot N(T)|T))=E(\lambda T^2)=\lambda E(T^2)$$

$$E(T^2)=\mathrm{Var}(T)+[E(T)]^2=a^2+b\Longrightarrow E(T\cdot N(T))=\lambda(a^2+b)$$

$$\Longrightarrow \mathrm{Cov}(T,N(T))=E(T\cdot N(T))-E(T)E(N(T))=\lambda b$$

(2) 由条件方差公式:

$$\mathrm{Var}(N(T))=E[\mathrm{Var}(N(T)|T)]+\mathrm{Var}(E[N(T)|T])=E[\lambda T]+\mathrm{Var}(\lambda T)=\lambda a+\lambda^2 b$$

➢ 全期望公式, 见课本 P131.
➢ 条件方差公式: $\mathrm{Var}(X)=E[\mathrm{Var}(X\mid Y)]+\mathrm{Var}(E[X\mid Y])$. 证明见第 35 题注释.

41. 设二维随机变量 (X,Y) 服从二元正态分布 $N(1,2,4,9,0.3)$, 求 $E(X|Y=2)$ 与 $E(XY^2+Y|Y=1)$.

解:
二元正态分布的条件期望为:

$$E(X|Y=y)=\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\Longrightarrow\begin{cases}E(X|Y=2)=1\\ E(XY^2+Y|Y=1)=E(X+1|Y=1)=E(X|Y=1)+1=1.8\end{cases}$$

➢ 二元正态分布的条件期望, 见课本 P131 例 4.11 (或二元正态分布的条件分布, 见课本例 3.14).

42. 设 $\{X_n, n=1,2,\cdots\}$ 和 $\{Y_n, n=1,2,\cdots\}$ 为定义在同一样本空间上的两个随机变量序列, 如果 $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$. 证明 $X_n+Y_n \xrightarrow{P} X+Y$.

解:

$$
\begin{aligned}
P(|X_n+Y_n-(X+Y)|>\varepsilon) &= P(|(X_n-X)+(Y_n-Y)|>\varepsilon)\\
&\le P\left(\left\{|X_n-X|>\frac{\varepsilon}{2}\right\}\cup\left\{|Y_n-Y|>\frac{\varepsilon}{2}\right\}\right)\\
&\le P\left(|X_n-X|>\frac{\varepsilon}{2}\right)+P\left(|Y_n-Y|>\frac{\varepsilon}{2}\right)\\
&\to 0. \quad (\varepsilon\to 0)
\end{aligned}
$$

43. 设 $X_1, X_2, \cdots, X_n, \cdots$ 为一随机变量序列, 且

$$X_n \sim Ge\left(\frac{\lambda}{n}\right), \quad n=1,2,3,\cdots,$$

其中 $\lambda>0$ 为常数. 定义随机变量 Y_n 为

$$Y_n=\frac{1}{n}X_n, \quad n=1,2,3,\cdots.$$

证明 $\{Y_n\}$ 依分布收敛于 Y, 其中 $Y\sim Exp(\lambda)$.

解:

$y\le 0$ 时, $P(Y_n\le y)=0=P(Y\le y)$. $y>0$ 时:

$$
\begin{aligned}
\lim_{n\to\infty}P(Y_n\le y) &= \lim_{n\to\infty}P\left(\frac{X_n}{n}\le y\right)\\
&= \lim_{n\to\infty}P(X_n\le ny)\\
&= \lim_{n\to\infty}1-\left(1-\frac{\lambda}{n}\right)^{\lfloor ny\rfloor}\\
&= \lim_{n\to\infty}1-\left(1-\frac{\lambda}{n}\right)^{ny}\\
&= 1-e^{-\lambda y}\\
&= P(Y\le y)
\end{aligned}
$$

综上, $Y_n \xrightarrow{\mathscr{L}} Y$.

44. 设 $\{X_n, n=1,2,\cdots\}$ 和 $\{Y_n, n=1,2,\cdots\}$ 为定义在同一样本空间上的两个随机变量序列, 如果 $X_n \xrightarrow{\mathscr{L}} X$, $Y_n \xrightarrow{P} c$, 其中 c 为常数. 证明
(1) $X_n + Y_n \xrightarrow{\mathscr{L}} X + c$;
(2) $X_n Y_n \xrightarrow{\mathscr{L}} cX$;
(3) $X_n / Y_n \xrightarrow{\mathscr{L}} X/c$, 这里 c 非零.

解: 这道题其实就是 Slutsky 定理.
(1) 对 $\forall\ t \in \mathbb{C}(F_{X+c})$,

$$\begin{cases} P(X_n + Y_n \le t) \\ = P(X_n + Y_n \le t, |Y_n - c| \le \varepsilon) + P(X_n + Y_n \le t, |Y_n - c| > \varepsilon) \\ \le P(X_n + Y_n \le t, |Y_n - c| \le \varepsilon) + P(|Y_n - c| > \varepsilon) \\ \le P(X_n + c \le t + \varepsilon) + P(|Y_n - c| > \varepsilon) \\ P(X_n + c \le t - \varepsilon) \\ = P(X_n + c \le t - \varepsilon, |Y_n - c| \le \varepsilon) + P(X_n + c \le t - \varepsilon, |Y_n - c| > \varepsilon) \\ \le P(X_n + c \le t - \varepsilon, |Y_n - c| \le \varepsilon) + P(|Y_n - c| > \varepsilon) \\ \le P(X_n + Y_n \le t) + P(|Y_n - c| > \varepsilon) \end{cases}$$

$$\Longrightarrow P(X_n + c \le t - \varepsilon) - P(|Y_n - c| > \varepsilon) \le P(X_n + Y_n \le t) \le P(X_n + c \le t + \varepsilon) + P(|Y_n - c| > \varepsilon)$$

由于 $X_n \xrightarrow{\mathscr{L}} X$, $Y_n \xrightarrow{P} c$, 令 $n \to \infty$:

$$P(X \le t - c - \varepsilon) \le \varliminf_{n\to\infty} P(X_n + Y_n \le t) \le \varlimsup_{n\to\infty} P(X_n + Y_n \le t) \le P(X \le t - c + \varepsilon)$$

再令 $\varepsilon \to 0$:

$$t - c \in \mathbb{C}(F_X) \Longrightarrow \lim_{n\to\infty} P(X_n + Y_n \le t) = P(X \le t - c) = P(X + c \le t) \Longrightarrow X_n + Y_n \xrightarrow{\mathscr{L}} X + c$$

(2) 对 $\forall\ t\in\mathbb{C}(F_{cX})$, 不妨令 $c>0$,

$$\begin{cases} P(X_nY_n\le t) \\ = P(X_nY_n\le t, |Y_n-c|\le\varepsilon)+P(X_nY_n\le t,|Y_n-c|>\varepsilon) \\ \le P(X_nY_n\le t, |Y_n-c|\le\varepsilon)+P(|Y_n-c|>\varepsilon) \\ \le P(X_n\le \dfrac{t}{c-\varepsilon})+P(|Y_n-c|>\varepsilon) \\ P(X_n\le \dfrac{t}{c+\varepsilon}) \\ = P((c+\varepsilon)X_n\le t, |Y_n-c|\le\varepsilon)+P((c+\varepsilon)X_n\le t,|Y_n-c|>\varepsilon) \\ \le P((c+\varepsilon)X_n\le t, |Y_n-c|\le\varepsilon)+P(|Y_n-c|>\varepsilon) \\ \le P(X_nY_n\le t)+P(|Y_n-c|>\varepsilon) \end{cases}$$

$$\Longrightarrow P(X_n\le\frac{t}{c+\varepsilon})-P(|Y_n-c|>\varepsilon)\le P(X_nY_n\le t)\le P(X_n\le\frac{t}{c-\varepsilon})+P(|Y_n-c|>\varepsilon)$$

由于 $X_n\xrightarrow{\mathscr{L}}X$, $Y_n\xrightarrow{P}c$, 令 $n\to\infty$:

$$P(X\le\frac{t}{c+\varepsilon})\le\varliminf_{n\to\infty}P(X_nY_n\le t)\le\varlimsup_{n\to\infty}P(X_nY_n\le t)\le P(X\le\frac{t}{c-\varepsilon})$$

再令 $\varepsilon\to 0$:

$$\frac{t}{c}\in\mathbb{C}(F_X)\Longrightarrow\lim_{n\to\infty}P(X_nY_n\le t)=P(cX\le t)\Longrightarrow X_nY_n\xrightarrow{\mathscr{L}}cX$$

(3) 不妨令 $c>0$, 则对 $\forall\ \varepsilon>0$,

$$P\left(\left|\frac{1}{Y_n}-\frac{1}{c}\right|>\frac{\varepsilon}{c(c-\varepsilon)}\right)=P\left(|Y_n-c|>\frac{\varepsilon}{c(c-\varepsilon)}|cY_n|\right)\le P(|Y_n-c|>\varepsilon)\to 0\quad(\varepsilon\to 0)$$

即 $Y_n\xrightarrow{P}c\Longrightarrow\dfrac{1}{Y_n}\xrightarrow{P}\dfrac{1}{c}$, $c\neq 0$. 由 (2) 即得 $X_n/Y_n\xrightarrow{\mathscr{L}}X/c$.

45. 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.

解:

记发生次数为 S_n, 则 $S_n \sim B(n=500, p=0.2) \Longrightarrow E(S_n)=np=100, \mathrm{Var}(S_n)=np(1-p)=80.$

① 由切比雪夫不等式:

$$P(|S_n-E(S_n)|\geq 20)\leq \frac{\mathrm{Var}(S_n)}{20^2}=0.2 \Longrightarrow P(80\leq S_n\leq 120)=1-P(|S_n-E(S_n)|\geq 20)\geq 0.8$$

② $n=500>30$, 故可以使用中心极限定理近似估计:

$$P(80\leq S_n\leq 120)=P\left(\frac{80-E(S_n)}{\sqrt{\mathrm{Var}(S_n)}}\leq \frac{S_n-E(S_n)}{\sqrt{\mathrm{Var}(S_n)}}\leq \frac{120-E(S_n)}{\sqrt{\mathrm{Var}(S_n)}}\right)\approx 2\Phi\left(\sqrt{5}\right)-1\approx 0.975$$

➢ 切比雪夫不等式, 见课本 P140.

➢ 中心极限定理, 见课本 P160/4.27.

➢ 中心极限定理估计离散型分布也可以用 "连续性修正", 见课本 P161 注, 这里不再给出结果.

46. 设 $X_1, X_2, \cdots, X_n, \cdots$ 为一列独立同分布的随机变量, 满足 $E(X_i^k)=\alpha_k, k=1,2,3,4$, 利用中心极限定理说明 $\sum\limits_{i=1}^{n} X_i^2$ 的渐近分布是什么.

解:

$E(X_i^2)=\alpha_2, \mathrm{Var}(X_i^2)=E(X_i^4)-\left[E(X_i^2)\right]^2=\alpha_4-\alpha_2^2$ 由林德伯格-莱维中心极限定理:

$$\frac{\sum\limits_{i=1}^{n} X_i^2-nE(X_i^2)}{\sqrt{n\mathrm{Var}(X_i^2)}}=\frac{\sum\limits_{i=1}^{n} X_i^2-n\alpha_2}{\sqrt{n(\alpha_4-\alpha_2^2)}}\xrightarrow{\mathscr{L}} N(0,1)$$

➢ 方差的性质, 见课本 P136.

➢ 林德伯格-莱维中心极限定理, 见课本 P159.

47. 设各零件的质量都是随机变量, 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 只零件的总质量超过 2510kg 的概率是多少?

解:

记每个零件的质量为 X_i, $1\leq i\leq 5000$, 5000 个零件的总质量记为 S. 则 $E(S)=n\mu=2500$, $\mathrm{Var}(S)=n\sigma^2=50$. 由中心极限定理:

$$\frac{S-E(S)}{\sqrt{\mathrm{Var}(S)}} \xrightarrow{\mathscr{L}} N(0,1) \Longrightarrow P(S>2510)=P\left(\frac{S-E(S)}{\sqrt{\mathrm{Var}(S)}}>\frac{2510-2500}{\sqrt{50}}\right)\approx 1-\Phi\left(\sqrt{2}\right)\approx 0.079$$

➢ 中心极限定理, 见课本 P160/4.27.

48.
(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?

解:
(1) 记 X 为正常部件的数量, 则 $X\sim B(n=100,p=0.9)\Longrightarrow E(X)=np=90, \mathrm{Var}(X)=np(1-p)=9$. 由中心极限定理:

$$P(X\geq 85)=P\left(\frac{X-E(X)}{\sqrt{\mathrm{Var}(X)}}\geq\frac{85-90}{\sqrt{3}}\right)\approx 1-\Phi\left(-\frac{5}{3}\right)=\Phi\left(\frac{5}{3}\right)\approx 0.9525$$

(2) $X\sim B(n,p=0.9)\Longrightarrow E(X)=np=0.9n, \mathrm{Var}(X)=np(1-p)=0.09n$. 由中心极限定理:

$$P(X\geq 0.8n)=P\left(\frac{X-E(X)}{\sqrt{\mathrm{Var}(X)}}\geq\frac{0.8n-0.9n}{\sqrt{0.09n}}\right)\approx 1-\Phi\left(-\frac{\sqrt{n}}{3}\right)=\Phi\left(\frac{\sqrt{n}}{3}\right)\geq 0.95$$

$\Longrightarrow \frac{\sqrt{n}}{3}\geq 1.645\Longrightarrow n\geq 24.35\Longrightarrow n_{min}=25.$

➢ 中心极限定理, 见课本 P160/4.27.

➢ 中心极限定理估计离散型分布也可以用“连续性修正”, 见课本 P161 注, 这里不再给出结果.

49. 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

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

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

解:
记单次取款额为 X_i, $1\leq i\leq 200$, 则 $E(X_i)=\frac{1}{10}\sum_{n=1}^{10}n=5.5$, $\mathrm{Var}(X_i)=\frac{1}{10}\sum_{n=1}^{10}n^2-\left(\frac{1}{10}\sum_{n=1}^{10}n\right)^2=8.25.$

记总取款额为 S, 则 $E(S)=nE(X_i)=1100$, $\mathrm{Var}(S)=n\mathrm{Var}(X_i)=1650$. 由中心极限定理:

$$P(S\le s)=P\left(\frac{S-E(S)}{\sqrt{\mathrm{Var}(S)}}\le\frac{s-1100}{\sqrt{1650}}\right)\approx\Phi\left(\frac{s-1100}{\sqrt{1650}}\right)\ge 0.95$$

$\Longrightarrow \dfrac{s-1100}{\sqrt{1650}}\ge 1.645\Longrightarrow s\ge 1166.82\Longrightarrow s_{min}=1167$ (百元).

➢ 中心极限定理, 见课本 P160/4.27.

50. 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) 若现在要进行 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?

解:

(1) 记单次加法运算的误差为 X_i, $1\le i\le n=1500$, 则 $X_i\sim U(-0.5,0.5)\Longrightarrow E(X_i)=0$, $\mathrm{Var}(X_i)=\dfrac{1}{12}$.
记误差总和为 S, 则 $E(S)=nE(X_i)=0$, $\mathrm{Var}(S)=n\mathrm{Var}(X_i)=125$. 由中心极限定理:

$$P(|S|\ge 15)=P\left(\left|\frac{S-E(S)}{\sqrt{\mathrm{Var}(S)}}\right|\ge\frac{15-0}{\sqrt{125}}\right)\approx 2\left(1-\Phi\left(\frac{3\sqrt{5}}{5}\right)\right)\approx 0.18$$

(2)

$$P(|S|\le 10)=P\left(\left|\frac{S-E(S)}{\sqrt{\mathrm{Var}(S)}}\right|\le\frac{10-0}{\sqrt{n/12}}\right)\approx 2\Phi\left(\frac{10}{\sqrt{n/12}}\right)-1\ge 0.90$$

$\Longrightarrow \dfrac{10}{\sqrt{n/12}}\ge 1.645\Longrightarrow n\le 443.4\Longrightarrow n_{max}=443$.

➢ 中心极限定理, 见课本 P160/4.27.

51. 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 10min, 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 15h 至 20h 的概率;

(2) 保证有 95% 的可能性, 问 16h 内最多可以组装多少件产品?

解:

(1) 记组装每件产品的时间为 X_i, $1\le i\le n=100$, 则 $E(X_i)=\dfrac{1}{6}$ $(h)\Longrightarrow X_i\sim Exp(6)\Longrightarrow \mathrm{Var}(X_i)=\lambda^2=\dfrac{1}{36}$.

记组装 n 件产品的时间为 S, 则 $E(S)=nE(X_i)=\frac{50}{3}$, $\mathrm{Var}(S)=n\mathrm{Var}(X_i)=\frac{25}{9}$. 由中心极限定理:

$$P(15\le S\le 20)=P\left(\frac{15-50/3}{\sqrt{25/9}}\le\frac{S-E(S)}{\mathrm{Var}(S)}\le\frac{20-50/3}{\sqrt{25/9}}\right)\approx\Phi(2)-\Phi(-1)\approx 0.819$$

(2)

$$P(S\le 16)=P\left(\frac{S-E(S)}{\mathrm{Var}(S)}\le\frac{16-n/6}{\sqrt{36/n}}\right)\approx\Phi\left(\frac{16-n/6}{\sqrt{n}/6}\right)\ge 0.95$$

$\Longrightarrow \frac{16-n/6}{\sqrt{n}/6}\ge 1.645\Longrightarrow n_{max}=81.$

➢ 中心极限定理, 见课本 P160/4.27.

52. 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险费为 5000 元. 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

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

且每次事故的索赔额度 (元) 服从 $[1000,5000]$ 上的均匀分布. 求平均每年保险公司盈利 200 万元的概率.

解:

记每次索赔金额为 $Y_i\sim U[1000,5000]$, $1\le i\le X$, 每辆车总索赔次数为 $S=\sum_{i=1}^{2400}X_i$. 则 $E(X)=\lambda=2$, $\mathrm{Var}(X)=\lambda=2$, $E(Y)=3000$, $\mathrm{Var}(Y)=\frac{4000000}{3}$. 由全期望公式:

$$E(S)=E(E(S|X))=E(E(\sum_{i=1}^{X}Y_i))=E(X\cdot E(Y))=E(X)E(Y)=6000$$

由条件方差公式:

$$\begin{aligned}\mathrm{Var}(S)=E[\mathrm{Var}(S|X)]+\mathrm{Var}(E[S|X]) &=E[\mathrm{Var}(\sum_{i=1}^{X}Y_i)]+\mathrm{Var}(E[\sum_{i=1}^{X}Y_i])\\ &=E[X\cdot\mathrm{Var}(Y)]+\mathrm{Var}(X\cdot E[Y])\\ &=E(X)\mathrm{Var}(Y)+\mathrm{Var}(X)\,[E(X)]^2\approx 20666666.67\end{aligned}$$

记 n 辆车的总赔付额为 S_n, 则由中心极限定理: $\frac{S_n-nE(S)}{\sqrt{n\mathrm{Var}(S)}}\sim N(0,1)\Longrightarrow$

$$P(2400\times 5000-S\ge 2000000)=P(S\le 10000000)\approx 1-\Phi(19.75)\approx 0$$

即该公司几乎不可能平均每年盈利 200 万元.

➢ 全期望公式, 见课本 P131.

➢ 条件方差公式: $\mathrm{Var}(X)=E[\mathrm{Var}(X\mid Y)]+\mathrm{Var}(E[X\mid Y])$. 证明见第 35 题注释.

➢ 中心极限定理, 见课本 P160/4.27.

53. * 设随机变量 X, $0<a\le X\le b$, $E(|X|)<\infty$, 证明

$$1\le E(X)\cdot E\left(\frac{1}{X}\right)\le\frac{(a+b)^2}{4ab}.$$

解:

LHS:

① Cauchy-Schwarz 不等式:

$$1=\left(E(\sqrt{X}\cdot\frac{1}{\sqrt{X}})\right)^2\le E(X)E\left(\frac{1}{X}\right)$$

② Jensen 不等式: 令 $f(x)=\dfrac{1}{x}$ 为严格凸函数, 则:

$$E[f(X)]=E[\frac{1}{X}]\ge f(E[X])=\frac{1}{E[X]}\Longrightarrow E(X)\cdot E\left(\frac{1}{X}\right)\ge 1$$

③ 辅助函数法: 构造辅助函数 $\phi(t)=E\left(\dfrac{(X-t)^2}{X}\right)$, $t\in\mathbb{R}$.

$$\phi(t)=E(X)-2t+t^2E\left(\frac{1}{X}\right)\ge 0\Longrightarrow\phi_{min}(t)=E(X)-\frac{1}{E\left(\frac{1}{X}\right)}\ge 0\Longrightarrow E(X)E\left(\frac{1}{X}\right)\ge 1$$

RHS:

$$E\left(\frac{(X-a)(b-X)}{X}\right)\ge 0\Longrightarrow E(X)+abE\left(\frac{1}{X}\right)\le a+b\Longrightarrow E(X)\cdot E\left(\frac{1}{X}\right)\le\frac{E(X)\,(a+b-E(X))}{ab}\le\frac{(a+b)^2}{4ab}$$

➢ Cauchy-Schwarz 不等式:

$$(E[XY])^2 \le E[X^2]E[Y^2]$$

证明:

$$E[(Y-tX)^2] = t^2E[X^2] - 2tE[XY] + E[Y^2] \ge 0 \Longleftrightarrow \Delta = 4(E[XY])^2 - 4E[X^2]E[Y^2] \le 0$$

➢ Jensen 不等式: 对严格凸函数 f,

$$E[f(X)] \ge f(E[X])$$

证明:

$$f\text{严格凸} \Longrightarrow \exists c \in \mathbb{R},\ s.t.\ f(X) \ge f(E[X]) + c(X - E[X])$$

$$\text{两边取期望} \Longrightarrow E[f(x)] \ge E[f(E[x]) + c(X - E[X])] = f(E[X])$$

54. * 设随机变量 X_1, X_2 独立同分布, 都只取正值, 证明必有 $E(X_1/X_2) \ge 1$, 等号成立当且仅当 X_1, X_2 只取一个值. (提示: 后一结论是由于 X_1, X_2 地位平等, 所以也有 $E(X_2/X_1) \ge 1$, 故 $E(X_1/X_2)E(X_2/X_1) \ge 1$, 但是 $(X_1/X_2)(X_2/X_1) \equiv 1$.)

解:

① 由第 53 题的左半边不等式:

$$E\left(\frac{X_1}{X_2}\right) \xlongequal{\text{独立}} E(X_1)\cdot E\left(\frac{1}{X_2}\right) = E(X)\cdot E\left(\frac{1}{X}\right) \ge 1$$

② 由 Cauchy-Schwarz 不等式:

$$E\left(\frac{X_1}{X_2}\right)\cdot E\left(\frac{X_2}{X_1}\right) \ge \left[E\left(\sqrt{\frac{X_1}{X_2}}\cdot\sqrt{\frac{X_2}{X_1}}\right)\right]^2 = 1, \quad E\left(\frac{X_1}{X_2}\right) \xlongequal{\text{对称性}} E\left(\frac{X_2}{X_1}\right) \Longrightarrow E\left(\frac{X_1}{X_2}\right) \ge 1$$

根据第 53 题的证明过程或不等式等号成立条件, 等号成立当且仅当 $\exists a, b \ne 0 \in \mathbb{R}$, s.t. $P(aX_1 = bX_2) = 1$. 由于 X_1, X_2 独立, 故 X_1, X_2 只能取单个常数值; 又由同分布, X_1, X_2 只取同一个值.

➢ Cauchy-Schwarz 不等式:

$$(E[XY])^2 \le E[X^2]E[Y^2]$$

证明见第 53 题注释.

55. * 设随机变量 X 取值于 $[0,1]$, 证明 $\mathrm{Var}(X)\le 1/4$. 什么时候等号成立? 把该结果推广到 $0<a\le X\le b$ 的情况.

解:

$$X\in[0,1]\Longrightarrow X^2\le X\Longrightarrow \mathrm{Var}(X)=E(X^2)-[E(X)]^2\le E(X)-[E(X)]^2\le\frac{1}{4}$$

等号成立当且仅当 $E(X^2)=E(X)=\frac{1}{2}$, 即 X 服从 $\{0,1\}$ 上的两点分布.
推广到一般情况, 归一化即可:

$$X=a+(b-a)Y\Longrightarrow \mathrm{Var}(X)=(b-a)^2\mathrm{Var}(Y)\le\frac{(b-a)^2}{4}$$

等号成立当且仅当 Y 服从 $\{0,1\}$ 上的两点分布, 即 X 服从 $\{a,b\}$ 上的两点分布.

56. * 设随机变量 X 的期望 μ 存在, 中位数记为 m.
(1) 证明对任意实数 b, $E(|X-m|)\le E(|X-b|)$;
(2) 若 $\mathrm{Var}(X)=\sigma^2<\infty$, 则 $|m-\mu|\le\sigma$.

解:
(1)

$$g(b):=E(|X-b|)=\int_{-\infty}^{b}(b-x)f_X(x)\mathrm{d}x+\int_{b}^{+\infty}(x-b)f_X(x)\mathrm{d}x\Longrightarrow g'(b)=\int_{-\infty}^{b}f_X(x)\mathrm{d}x-\int_{b}^{+\infty}f_X(x)\mathrm{d}x$$

令 $g'(b)=0\Longrightarrow\int_{-\infty}^{b}f_X(x)\mathrm{d}x=\int_{b}^{+\infty}f_X(x)\mathrm{d}x$, 即中位数的定义. 且 $g''(b)=2f_X(b)\ge 0\Longrightarrow E(|X-m|)\le E(|X-b|)$.
(2) 由 (1) 及 Cauchy-Schwarz 不等式 (取 $X'=|X-\mu|$, $Y'=1$):

$$|m-\mu|=|E(X-m)|\le E(|X-m|)\le E(|X-\mu|)\le\sqrt{E(|X-\mu|^2)}=\sqrt{\mathrm{Var}(X)}=\sigma$$

➢ Cauchy-Schwarz 不等式:

$$(E[XY])^2\le E[X^2]E[Y^2]$$

证明见第 53 题注释.

57. * 某大型商场大体分为生活用品、首饰、家电手机、衣帽鞋类和饮食五大区域. 根据以往大数据资料的分析, 认为周末 11:00-12:00 在商场的人数服从参数为 600 的泊松分布, 到上述五个区域的人流比例为 2:1:2:4:3. 为简单起见, 假设每个人的行为独立. 求
(1) 到各个区域人数的期望和方差;
(2) 计算各区域人数之间的协方差和相关系数.

解:
(1) 记商场总人数为 $N\sim P(600)$, 到各区域的人数为 $X_i|N=n\sim B(n,p_i)$, $1\le i\le 5$. 对 $\forall 1\le i\le 5$,

$$P(X_i=k)=\sum_{n=k}^{\infty}P(N=n)\cdot P(X_i=k|N=n)=\sum_{n=k}^{\infty}\frac{e^{-\lambda}\lambda^n}{n!}\cdot\binom{n}{k}p_i^k(1-p_i)^{n-k}=\frac{e^{-\lambda p_i}(\lambda p_i)^k}{k!}$$

即 $X_i\sim P(\lambda p_i)\Longrightarrow E(X_i)=\lambda p_i, \mathrm{Var}(X_i)=\lambda p_i$.

i	1	2	3	4	5
$E(X_i)$	100	50	100	200	150
$\mathrm{Var}(X_i)$	100	50	100	200	150

(2) 记 $\{Y_a=b\}$={第 a 个人选择的区域为 b}, 则

$$E(X_iX_j|N=n)=E\left(\sum_{1\le p\ne q\le n}I_{\{Y_p=i\}}I_{\{Y_q=j\}}\right)=n(n-1)p_ip_j$$

$$\Longrightarrow E(X_iX_j)=E(E(X_iX_j|N))=E(N(N-1)p_ip_j)=\lambda^2p_ip_j$$

$$\Longrightarrow \mathrm{Cov}(X_i,X_j)=E(X_i,X_j)-E(X_i)E(X_j)=0\Longrightarrow\rho=\frac{\mathrm{Cov}(X_i,X_j)}{\sqrt{\mathrm{Var}(X_i)\mathrm{Var}(X_j)}}=0$$

58. * 设 $f(x)$, $g(x)$ 是 $[0,1]$ 上的非负连续函数, 且存在常数 c, 使得 $f(x)\le cg(x)$. 用大数定理证明下式成立:

$$\lim_{n\to\infty}\int_0^1\int_0^1\cdots\int_0^1\frac{f(x_1)f(x_2)\cdots f(x_n)}{g(x_1)g(x_2)\cdots g(x_n)}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n=\frac{\int_0^1 f(x)\mathrm{d}x}{\int_0^1 g(x)\mathrm{d}x}.$$

提示: 设 X_1, X_2, $\cdots$, X_n 为独立同分布的随机变量, 其公共分布为 $U(0,1)$, 记被积函数为 $h(x_1,x_2,\cdots,x_n)$, 则上述重积分为 $E[h(X_1,X_2,\cdots,X_n)]$, 因为被积函数非负且有上界, 所以求极限与求期望可交换次序.

解:

个人怀疑这道题是否正确, 被积函数表达为乘积的分式似乎得不到该等式. 故根据题目描述, 猜测等式的正确形式为 (若有同学证出来了题目中的等式, 欢迎联系):

$$\lim_{n\to\infty}\int_0^1\int_0^1\cdots\int_0^1 \frac{\sum\limits_{i=1}^n f(x_i)}{\sum\limits_{i=1}^n g(x_i)}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n = \frac{\int_0^1 f(x)\mathrm{d}x}{\int_0^1 g(x)\mathrm{d}x}$$

证明:

$$\frac{f(x)}{g(x)} \le c \Longrightarrow \frac{\sum\limits_{i=1}^n f(x_i)}{\sum\limits_{i=1}^n g(x_i)} \le c$$

由控制收敛定理:

$$LHS = \lim_{n\to\infty} E\left(\frac{\sum\limits_{i=1}^n f(x_i)}{\sum\limits_{i=1}^n g(x_i)}\right) = E\left(\lim_{n\to\infty}\frac{\sum\limits_{i=1}^n f(x_i)}{\sum\limits_{i=1}^n g(x_i)}\right) = E\left(\lim_{n\to\infty}\frac{\sum\limits_{i=1}^n f(x_i)/n}{\sum\limits_{i=1}^n g(x_i)/n}\right) \xlongequal{\text{大数定律}} E\left(\frac{E(f(X))}{E(g(X))}\right) = RHS$$

59. * 用中心极限定理证明: 当 $n\to\infty$ 时,

$$e^{-n}\sum_{k=0}^n \frac{n^k}{k!} \to \frac{1}{2}.$$

解:

记 $X_n \sim P(n) \Longrightarrow E(X_n)=n, \mathrm{Var}(X_n)=n$, 则由中心极限定理:

$$e^{-n}\sum_{k=0}^n \frac{n^k}{k!} = P(X_n\le n) = P(\frac{X_n-n}{\sqrt{n}}\le 0) \to \Phi(0) = \frac{1}{2}$$

➢ 中心极限定理, 见课本 P160/4.27.

60. 考虑独立同分布的随机变量序列 $X_1, X_2, \cdots, X_i, \cdots$, 其中 X_i 服从整数 $\{0,1,2,\cdots,9\}$ 上的离散均匀分布. 令

$$U_n = \sum_{i=1}^{n} \frac{1}{10^i} X_i = 0.1X_1 + 0.01X_2 + \cdots + 0.1^n X_n.$$

直观上, 这是区间 $[0,1]$ 上一个随机数的有理数展开前 n 位. 证明当 $n \to \infty$ 时, U_n 依分布收敛于 $[0,1]$ 上的均匀分布随机变量 U, 即 $P(U_n \le u) \to P(U \le u)$, $u \in \mathbb{R}$.

解:

$$u \le P(U_n \le u) = P(10^n U_n \le 10^n u) = P(10^n U_n \le \lfloor 10^n u \rfloor) = \frac{\lfloor 10^n u \rfloor + 1}{10^n} \le u + \frac{2}{10^n}$$

令 $n \to \infty$, 夹逼即得 $P(U_n \le u) \to u = P(U \le u)$.

第 5 章　统计学基本概念

1.

解:

2.

解:

3.

解:

4.

解:

5.

解:

6.

解:

7.

解:

8.

解:

9.

解:

10.

解:

11.

解:

12.

解:

13.

解:

14.

解:

15.

解:

16.

解:

17.

解:

18.

解:

19.

解:

20.

解:

21.

解:

22.

解:

第 6 章　参数点估计

1.

解:

2.

解:

3.

解:

4.

解:

5.

解:

6.

解:

7.

解:

8.

解:

9.

解:

10.

解:

11.

解:

12.

解:

13.

解:

14.

解:

15.

解:

16.

解:

17.

解:

18.

解:

19.

解:

20.

解:

21.

解:

22.

解:

23.

解:

24.

解:

25.

解:

26.

解:

27.

解:

28.

解:

29.

解:

30.

解:

31.

解:

32.

解:

33.

解:

34.

解:

35.

解:

36.

解:

37.

解:

38.

解:

39.

解:

40.

解:

41.

解:

42.

解:

43.

解:

44.

解:

45.

解:

46.

解:

47.

解:

48.

解:

49.

解:

50.

解:

51.

解:

52.

解:

53.

解:

54.

解:

55.

解:

56.

解:

57.

解:

58.

解:

59.

解:

60.

解:

61.

解:

第 7 章　区间估计

1.

解:

2.

解:

3.

解:

4.

解:

5.

解:

6.

解:

7.

解:

8.

解:

9.

解:

10.

解:

11.

解:

12.

解:

13.

解:

解:

14.

解:

15.

解:

16.

解:

17.

解:

18.

解:

19.

解:

20.

解:

21.

解:

22.

解:

23.

解:

24.

解:

25.

解:

26.

解:

27.

解:

28.

解:

29.

解:

30.

解:

31.

解:

32.

解:

33.

解:

34.

解:

第 8 章　假设检验

1.

解:

2.

解:

3.

解:

4.

解:

5.

解:

6.

解:

7.

解:

8.

解:

9.

解:

10.

解:

11.

解:

12.

解:

13.

解:

14.

解:

15.

解:

16.

解:

17.

解:

18.

解:

19.

解:

20.

解:

21.

解:

22.

解:

23.

解:

24.

解:

25.

解:

26.

解:

27.

解:

28.

解:

29.

解:

30.

解:

31.

解:

32.

解:

33.

解:

34.

解:

35.

解:

36.

解:

37.

解:

38.

解:

39.

解:

40.

解:

41.

解:

42.

解:

43.

解:

44.

解:

45.

解:

46.

解:

47.

解:

48.

解:

49.

解:

第 9 章　非参数假设检验

1.

解:

2.

解:

3.

解:

4.

解:

5.

解:

6.

解:

7.

解:

8.

解:

9.

解:

10.

解:

11.

解:

12.

解:

13.

解:

14.

解: