习题1

1.1 将下列事件用事件A, B, C表示:

(1) 只有事件A 发生; (2) 3个事件中至少有2个发生; (3) 3个事件中恰有2个事件发生; (4) 3个事件中不多于1个事件发生.

解：(1) $A\overline{B}\,\overline{C}$;

(2)$AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC \cup ABC = AB \cup AC \cup BC$;

(3)$AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC$;

(4)$\overline{A}\,\overline{B}\,\overline{C} \cup A\overline{B}\,\overline{C} \cup \overline{A}B\overline{C} \cup \overline{A}\,\overline{B}C = \overline{A}\,\overline{B} \cup \overline{B}\,\overline{C} \cup \overline{A}\,\overline{C}$.

1.2 已知$P(A) = \dfrac{1}{2}$,

(1)若A, B互不相容, 求$P(A\overline{B})$,

(2)若$P(AB) = \dfrac{1}{8}$, 求$P(A\overline{B})$.

解：(1)$P(A\overline{B}) = P(A) - P(AB) = P(A) = \dfrac{1}{2}$.

(2)$P(A\overline{B}) = P(A) - P(AB) = \dfrac{1}{2} - \dfrac{1}{8} = \dfrac{3}{8}$.

1.3 抛一枚质地均匀的硬币 5 次,求即出现正面又出现反面的概率.

解：设$A =${即出现正面又出现反面},则

$P(A) = C_5^1\left(\dfrac{1}{2}\right)^5 + C_5^2\left(\dfrac{1}{2}\right)^5 + C_5^3\left(\dfrac{1}{2}\right)^5 + C_5^4\left(\dfrac{1}{2}\right)^5 = \dfrac{15}{16}$.

1.4 设A, B是两个事件,

(1) 已知$A\,\overline{B} = \overline{A}\,B$, 验证$A = B$.

(2) 验证事件A和事件B恰有一个发生的概率为$P(A) + P(B) - 2P(AB)$.

解：(1) 因为$A\,\overline{B} = \overline{A}\,B$ 所以$A\,\overline{B} \cup AB = \overline{A}\,B \cup AB$,即$A = B$.

(2) “事件A和事件B恰有一个发生”可以用事件$A\overline{B} \cup \overline{A}B$来表示。而

$P(A\overline{B} \cup \overline{A}B)$

$= P(A\overline{B}) + P(\overline{A}B)$

$= P(A) - P(AB) + P(B) - P(AB) = P(A) + P(B) - 2P(AB).$

1.5 试分析互不相容事件是对立事件吗?如果是,请给出理由;如果不是,请举例说明.

解：互不相容事件未必是对立事件。例如：抛掷一枚均匀的骰子，设事件$A=${掷出的点数恰好是奇数}，$B=${掷出的点数恰好是2},则事件A与B互不相容，但A与B不是对立事件。

1.6 试举例说明概率为零的事件未必是不可能事件; 概率为1的事件未必是必然事件.

解：在单位园内等可能的掷点，假设点落在任意子区域内的概率仅与该区域的面积成正比，而与它的位置和形状无关，设$A=${点恰好落在圆心}，则事件A不是不可能事件，但概率为0；而$\Omega - A$不是必然事件，但概率为1.

1.7 已知$P(A) = P(B) = P(C) = \dfrac{1}{4}$, $P(AB) = 0$,$P(AC) = P(BC) = \dfrac{1}{16}$. 求(1)$A, B, C$中至少有一个发生的概率; (2) A, B, C都不发生的概率.

解：(1)“A, B, C中至少有一个发生”表示为$A \cup B \cup C$, 因为$ABC \subseteq AB$,由概率的不等式$0 \le P(ABC) \le P(AB) = 0$得：$P(ABC) = 0$.

$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC)$

$= \dfrac{1}{4} + \dfrac{1}{4} + \dfrac{1}{4} - \dfrac{1}{16} - \dfrac{1}{16} = \dfrac{5}{8}.$

(2)$P(\overline{A}\,\overline{B}\,\overline{C}) = 1 - P(A \cup B \cup C) = \dfrac{3}{8}.$

1.8 从5双不同的鞋子中任取4只,求这4只鞋子中至少有两只配成一双的概率.

解：设$A=${4只鞋子中至少有两只配成一双},由于样本空间所含的样本点总数为C_{10}^4,不利于事件A的样本点总数为$C_5^4 2^4$,故

$P(A) = 1 - P(\overline{A}) = 1 - \dfrac{C_5^4 2^4}{C_{10}^4} = \dfrac{13}{21}.$

1.9 50只铆钉随机地取来用在10个部件上, 其中有三只铆钉强度太弱. 每个部件

用3只铆钉. 若3只强度太弱的铆钉都装在一个部件上, 则这个部件强度就太弱. 问发生一个部件强度太弱的概率是多少?

解：设$A=\{$有一个部件强度太弱$\}$，则A所含的样本点数为$C_{10}^1C_3^3C_{47}^3C_{44}^3\cdots C_{23}^3$, 所含样本点总数为$C_{50}^3C_{47}^3C_{44}^3\cdots C_{23}^3$, 所以

$$P(A)=\frac{C_{10}^1C_3^3C_{47}^3C_{44}^3\cdots C_{23}^3}{C_{50}^3C_{47}^3C_{44}^3\cdots C_{23}^3}=\frac{1}{1960}.$$

1.10 盒子中有3只白球、5只黑球和4只红球. 现从盒子中一个接一个地取出所有球, 试求红球比白球出现早的概率.

解：设$B=\{$红球比白球出现早$\}$，$A_i=\{$前$i-1$次抽到的都是黑球，而第i次首次抽到的不是黑球$\}, i=1,\cdots,6.$则$\sum\limits_{i=1}^{6}P(A_i)=1$,

$$P(A_1)=\frac{7}{12},\ P(A_2)=\frac{5}{12}\cdot\frac{7}{11},\ P(A_3)=\frac{C_5^2}{C_{12}^2}\cdot\frac{7}{10},\ P(A_4)=\frac{C_5^3}{C_{12}^3}\cdot\frac{7}{9},$$

$$P(A_5)=\frac{C_5^4}{C_{12}^4}\cdot\frac{7}{8},\ P(A_6)=\frac{C_5^5}{C_{12}^5}\cdot 1,\ P(B|A_i)=\frac{4}{7}, i=1,\cdots,6.$$

故$P(B)=\sum\limits_{i=1}^{6}P(A_iB)=\sum\limits_{i=1}^{6}P(A_i)P(B|A_i)=\dfrac{4}{7}.$

1.11 假设新购进了一批仪器,共有100件, 其中5件有质量问题. 抽样验收时从中任取5件, 假如均无质量问题, 则接收这批仪器, 否则拒收. 求这批仪器被拒收的概率.

解：设$A=\{$这批仪器被拒收$\}$，样本空间包含的样本点总数为C_{100}^5, $\overline{A}$中所包含的样本点总数为C_{95}^5, 故

$$P(A)=1-P(\overline{A})=1-\frac{C_{95}^5}{C_{100}^5}=0.2304.$$

这批仪器被拒收的概率为0.2304.

1.12 在一张打上方格的纸上投一枚直径为1的硬币, 方格的边长为多少才能使硬币与格线不相交的概率小于0.01.

解：设方格的边长为a,$A=\{$硬币与格线不想交$\}$. 考虑硬币的中心，它可以落在方格纸内任一个位置，要使得硬币与格线不想交，当且仅当硬币的中心与格线最短边

的距离大于$1/2$，故

$$P(A)=\frac{(a-1)^2}{a^2},$$

由此知，要使$P(A)<0.01$,只要$0<a<\frac{10}{9}$.

1.13 设有任意两数x,y, 满足$0<x<1,0<y<1$, 在此条件下, 试求满足条件$0<xy<\frac{1}{3}$的概率.

解：设$A=\{(x,y):0<xy<\frac{1}{3}\}$,则$P(A)=\frac{1}{3}+\int_{\frac{1}{3}}^{1}\frac{1}{x}\mathrm{d}x=\frac{1+\ln 3}{3}$.

1.14 设10片药片中有5片是安慰剂.

(1) 从中任意抽取5片, 求其中至少有2片安慰剂的概率;

(2) 从中每次取一片,做不放回抽样,求前3次都取到安慰剂的概率.

解：设$A=$\{至少有2片安慰剂\},$A_i=$\{恰有i片安慰剂\}$i=0,1,\cdots,5$,$B=$\{前3次都取到安慰剂\},则

(1) $P(A)=P(A_2\cup A_3\cup A_4\cup A_5)=1-P(A_0)-P(A_1)=1-\frac{C_5^0C_5^5}{C_{10}^5}-\frac{C_5^1C_5^4}{C_{10}^5}=\frac{113}{126}$.

(2)$P(B)=\frac{5}{10}\cdot\frac{4}{9}\cdot\frac{3}{8}=\frac{1}{12}$.

1.15 (1) 已知$P(\overline{A})=0.3$, $P(B)=0.4$, $P(A\overline{B})=0.5$, 求条件概率$P(B|A\cup\overline{B})$.

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

解：(1)因为$P(\overline{A})=0.3$, $P(A\overline{B})=0.5$, $P(A\overline{B})=P(A)-P(AB)$,

所以,$P(AB)=P(A)-P(A\overline{B})=1-P(A)-P(A\overline{B})=0.2$,

$$P(B|A\cup\overline{B})=\frac{P(B\cap(A\cup\overline{B}))}{P(A\cup\overline{B})}=\frac{P(AB)}{P(AB)+P(\overline{B})}=0.25.$$

(2) 因为$P(A)=\frac{1}{4}$, $P(B|A)=\frac{1}{3}$, $P(A|B)=\frac{1}{2}$,

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

所以，$P(AB)=P(A)P(B|A)=\frac{1}{12}$,$P(B)=\frac{P(AB)}{P(A|B)}=\frac{1}{6}$,

故$P(B|A\cup B)=P(A)+P(B)-P(AB)=\frac{1}{3}$.

1.16 设甲袋中有3只白球,2只黑球; 乙袋中有4只白球,5只黑球;先从甲袋中任取2只

球放入乙袋中, 再从乙袋中任取一只球, 求该球是白球的概率.

解：设A_i ={从甲袋中取出的2只球中恰有i只白球}，$i=0,1,2.B$ ={从乙袋中取出的一只球是白球}，根据已知条件得$P(A_i)=\dfrac{C_3^iC_2^{2-i}}{C_5^2}, P(B|A_i)=\dfrac{C_{4+i}^1}{C_{11}^1}=\dfrac{4+i}{11}$. 由全概率公式得：

$P(B)=\sum\limits_{i=0}^{2}(P(A_i)\cdot P(B|A_i))=\dfrac{26}{55}$.

1.17 有两箱同种类的零件,第一箱装50只,其中有10只一等品;第二箱装30只,其中有18只一等品, 今从两箱中任意挑出一箱,然后从该箱中依次取两只零件.试求

(1)第一次取到的零件是一等品的概率;

(2)在第一次取到一等品的条件下,第二次取到的零件也是一等品的概率.

解：设A_i ={第i次取到的零件是一等品},$i=1,2.B$ ={零件取自第一箱}，根据已知条件得$P(B)=\dfrac{1}{2}, P(A_1|B)=\dfrac{10}{50}, P(A_1|\overline{B})=\dfrac{18}{30}$.

$P(A_1A_2|B)=\dfrac{P_{10}^2}{P_{50}^2}=\dfrac{9}{245}, P(A_1A_2|\overline{B})=\dfrac{p_{18}^2}{P_{30}^2}=\dfrac{102}{290}$. 则

(1)$P(A_1)=P(B)P(A_1|B)+P(\overline{B})P(A_1|\overline{B})=\dfrac{2}{5}$,

(2)$P(A_1A_2)=P(B)P(A_1A_2|B)+P(\overline{B})P(A_1A_2|\overline{B})=\dfrac{276}{1421}$.

$P(A_2|A_1)=\dfrac{P(A_1A_2)}{P(A_1)}=\dfrac{690}{1421}=0.4856$.

1.18 假设一批100台液晶显示器中有80台优质品. 现在接连任取3台, 求

(1)第一台不是优质品而第二台是优质品的概率;

(2)最后才抽到一台优质品的概率.

解：设A_i ={第i次取到的显示器是优质品},$i=1,2,3$.

(1)$P(\overline{A_1}A_2)=P(\overline{A_1})P(A_2|\overline{A_1})=\dfrac{20}{100}\cdot\dfrac{80}{99}=\dfrac{16}{99}$.

(2)$P(\overline{A_1}\,\overline{A_2}A_3)=P(\overline{A_1})P(\overline{A_2}|\overline{A_1})P(A_3|\overline{A_1}\,\overline{A_2})=\dfrac{20}{100}\cdot\dfrac{19}{99}\cdot\dfrac{80}{98}=\dfrac{152}{4851}$.

1.19 假设在某条公路上载重汽车与其他汽车的数量之比为$3:2$, 前者中途停车修理的概率为0.02, 后者中途停车修理的概率为0.01. 现有一辆汽车中途停车修理, 求这辆汽

车是载重汽车的概率.

解：设$A=${汽车是载重汽车}，$B=${汽车中途停车修理}，则

$P(A)=\frac{3}{5}, P(B|A)=0.02, P(B|\overline{A})=0.01, P(\overline{A})=\frac{2}{5}.$

$P(A|B)=\frac{P(AB)}{P(B)}=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\overline{A})P(B|\overline{A})}=\frac{3}{4}.$

1.20 无线电通信中需要不断发出信号0和1, 大量统计资料表明, 发出信号0的概率为0.6, 而发出信号1的概率为0.4. 由于随机干扰, 当发出信号0时,分别以概率0.7和0.1收到0和1, 以0.2的概率收到模糊信号“x”; 发出信号1 时,分别以概率0.05和0.85收到0和1, 以0.1的概率收到模糊信号“x”. 问收到模糊信号“x”时,应翻译成哪个信号为好, 为什么?

解：设$A_i=${发出的信号为i},$i=0,1$, $B_j=${收到的信号为j},$j=0,1,x$. 则

$P(A_0)=0.6, P(A_1)=0.4, P(B_0|A_0)=0.7, P(B_1|A_0)=0.1, P(B_x|A_0)=0.2,$

$P(B_0|A_1)=0.05, P(B_1|A_1)=0.85, P(B_x|A_1)=0.1,$

由全概率公式得：$P(B_x)=P(A_0)P(B_x|A_0)+P(A_1)P(B_x|A_0)=0.16,$

由贝叶斯公式得：$P(A_0|B_x)=\frac{P(A_0B_x)}{P(B_x)}=\frac{P(A_0)P(B_x|A_0)}{P(B_x)}=\frac{3}{4}.$

因为$P(A_0|B_x)>0.5$, 所以,该信号应翻译成信号”0”.

1.21 普通人群中男女比例为51:49, 已知男性中有2%是色盲患者, 女性中有0.25%是色盲患者, 今从人群中随机挑选一人, 发现是色盲, 问此人是男性的概率是多少?

解：设$A=${此人是男性}，$B=${此人是色盲}，则

$P(A)=0.51, P(B|A)=2\%, P(B|\overline{A})=0.25\%,$

$P(A|B)=\frac{P(AB)}{P(B)}=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\overline{A})P(B|\overline{A})}=0.893,$

故此人是色盲的概率是0.893.

1.22 假定一架失事飞机在三个地区坠毁是等可能的. 用$1-\beta_i$表示该飞机在第i个地区发现的概率, 其中$i=1,2,3$(常数β_i称为未发现概率, 表示没有注意到坠机的概率,

通常可归因于该地区的地理和环境条件所致). 问在第1个地区未找到飞机,而飞机在第$i(i=1,2,3)$个地区坠毁的条件概率是多少?

解：设$A_i=${飞机在第$i(i=1,2,3)$个地区坠毁},$B=${在第1个地区未找到飞机},则

$$P(A_i)=\tfrac{1}{3}, i=1,2,3. P(B|A_1)=\beta_1, P(B|A_i)=1, i=2,3,$$

$$P(B)=\sum_{i=1}^{3}P(A_i)P(B|A_i)=\frac{\beta_1+2}{3},$$

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{\beta_1}{\beta_1+2},$$

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{P(B)}=\frac{1}{\beta_1+2},$$

$$P(A_3|B)=\frac{P(A_3)P(B|A_3)}{P(B)}=\frac{1}{\beta_1+2}.$$

故在第1个地区未找到飞机,而飞机在第$i(i=1,2,3)$个地区坠毁的条件概率分别是$\dfrac{\beta_1}{\beta_1+2},\dfrac{1}{\beta_1+2},\dfrac{1}{\beta_1+2}$.

1.23 某口袋中装有一球, 此球可能是白球, 也可能是黑球. 现在放一白球到袋中去, 然后再从袋中任取一球. 若已知取出的球是白球, 求剩下的球也是白球的概率.

解：设$A=${取出的球是白球}，$B=${口袋中的球是白球}，则

$$P(B)=0.5, P(A|B)=1, P(A|\overline{B})=0.5.$$

$$P(B|A)=\frac{P(B)P(A|B)}{P(B)P(A|B)+P(\overline{B})P(A|\overline{B})}=\frac{2}{3}.$$

1.24 根据报道,美国人血型的分布近似地为:A型占37%, O型占44%,B型占13%,AB型占6%. 夫妻拥有的血型相互独立.

(1)B型的人只有输入B、O两种血型才安全. 若妻为B型, 夫为何种血型未知, 求夫是妻的安全输血者的概率;

(2)随机取一对夫妻, 求妻为B型,夫为A型的概率;

(3)随机取一对夫妻, 求其中一人为A型, 另一人为B型的概率;

(4)随机取一对夫妻, 求其中至少有一人为O型的概率.

解：设$C_i=${妻子的血型是i型},$i=A,O,B,AB$,

$D_i=$\{丈夫的血型是j型\}，$j=A,O,B,AB,$

(1)$P(D_B\cup D_O)=P(D_B)+P(D_O)=0.13+0.44=0.57.$

(2)$P(C_BD_A)=P(C_B)P(D_A)=0.13\cdot 0.37=0.0481.$

(3)$P(C_AD_B\cup C_BD_A)=P(C_A)P(D_B)+p(C_B)P(D_A)=0.37\cdot 0.13\cdot 2=0.0962.$

(4)$P(C_O\cup D_O)=P(C_O)+P(D_O)-P(C_O)P(D_O)=0.44\cdot 2-0.44^2=0.6864.$

1.25 若干人独立地向一游动目标射击, 每人击中目标的概率都是0.6, 问至少需要多少人才能以0.99以上的概率击中目标?

解：设需要n个人才能以0.99以上的概率击中目标，$A_i=$\{第i个人击中目标\}，$i=1,\cdots,n.$

则至少有1个人击中目标的概率为

$P(\cup_{i=1}^n A_i)=1-\prod\limits_{i=1}^n P(\overline{A_i})=1-0.4^n\geq 0.99$,解之得：$n\geq 5.03.$

所以至少需要6个人才能以0.99以上的概率击中目标。

1.26 三个人独立地去破译一份密码, 已知3个人能破译出的概率分别为$1/5,1/3,1/4.$ 问三人中至少有一人能将此密码破译出的概率是多少?

解：设$A_i=$\{第i能破译出密码\},$i=1,2,3.$ $B=$\{至少有一人能破译出密码\}，则

$P(A_1)=1/5,P(A_2)=1/3,P(A_3)=1/4.$

$P(B)=P(A_1\cup A_2\cup A_3)$

$=P(A_1)+P(A_2)+P(A_3)-P(A_1)P(A_2)-P(A_1)P(A_3)-P(A_2)P(A_3)+P(A_1)P(A_2)P(A_3)$

$=0.6.$

1.27 设事件A,B满足$P(A)>0,P(B)>0$, 证明事件A,B相互独立与事件A,B互不相容不能同时成立.

证明：如果事件A,B相互独立，则$P(AB)=P(A)P(B)$,

又因为$P(A)>0,P(B)>0$,所以$P(AB)>0$, 从而A,B不是互不相容事件。

反之，如果A,B是互不相容事件，则$AB=\phi,P(AB)=0$,

又因为$P(A)>0,P(B)>0$,所以$P(AB)\neq P(A)P(B)$, 因此A,B相互不独立。

1.28 试分析任意两个互斥事件与独立事件之间的关系.

解：假设事件A,B互斥，则当$P(A)>0,P(B)>0$时，A,B一定不独立。

当$P(A)=0$或$P(B)=0$时，由于$AB\subseteq A,AB\subseteq B$,,所以$P(AB)=0$,因此$A,B$ 相互独立。

1.29 在如图**??**所示的开关电路中,开关S_1,S_2,S_3,S_4开或关的概率均为0.5,且是否开关相互独立.

(1)求灯亮的概率;

(2)已见灯亮,求开关S_1与S_2同时闭合的概率.

解：设A_i={开关S_i闭合}，$i=1,\cdots,4$. B={灯亮},则$P(A_i)=0.5,i=1,2,3,4$.

(1)$P(B)=P(A_1A_2\cup A_3\cup A_4)$

$=P(A_1A_2)+P(A_3)+P(A_4)-P(A_1A_2A_3)-P(A_1A_2A_4)-P(A_3A_4)+P(A_1A_2A_3A_4)$

$=P(A_1)P(A_2)+P(A_3)+P(A_4)-P(A_1)P(A_2)P(A_3)-P(A_1)P(A_2)P(A_4)$

$-P(A_3)P(A_4)+P(A_1)P(A_2)P(A_3)P(A_4)$

$=0.5^2+2\cdot 0.5-2\cdot 0.5^3-0.5^2+0.5^4=\dfrac{13}{16}=0.8125.$

(2)$P(A_1A_2|B)=\dfrac{P(A_1A_2)P(B|A_1A_2)}{P(B)}=\dfrac{0.5^2}{0.8125}=\dfrac{4}{13}=0.3077.$

1.30 某猎人在距离100m处射击一只兔子, 击中的概率为0.6,如果第一次未击中, 则进行第二次射击, 此时兔子跑到150m处. 如果第二次又未击中, 则进行第三次射击, 此时兔子跑到200m处. 超出200m则超出猎枪的射程.假定击中的概率与距离成反比, 求猎人击中兔子的概率.

解：设A_i={猎人与兔子的距离为$100+i$米}，$i=0,1,2$. B={猎人击中兔子}。

依题意得$P(A_0)=0.6,P(A_1)=\dfrac{2}{5},P(A_2)=\dfrac{3}{10},$

$P(B) = P(A_0) + P(\overline{A_0}A_1) + P(\overline{A_0}\,\overline{A_1}A_2)$

$= P(A_0) + P(\overline{A_0})P(A_1) + P(\overline{A_0})P(\overline{A_1})P(A_2) = 0.832.$

所以猎人击中兔子的概率是0.832.

1.31 设甲、乙、丙3人各自独立射击靶子一次,命中率分别是0.4,0.5和0.7.求

(1)恰有1人击中靶子的概率;

(2)至少有1人击中靶子的概率.

解：设A_i ={第i人击中靶子}，$i = 1, 2, 3.$ B ={恰有1人击中靶子},C ={至少有1人击中靶子},

则$P(A_1) = 0.4, P(A_2) = 0.5, P(A_3) = 0.7.$

(1)$P(B) = P(A_1\overline{A_2}\,\overline{A_3}) + P(\overline{A_1}A_2\overline{A_3}) + P(\overline{A_1}\,\overline{A_2}A_3)$

$$= P(A_1)P(\overline{A_2})P(\overline{A_3}) + P(\overline{A_1})P(A_2)P(\overline{A_3}) + P(\overline{A_1})P(\overline{A_2})P(A_3) = 0.36.$$

(2)$P(C) = P(A_1 \cup A_2 \cup A_3) = 1 - P(\overline{A_1}\,\overline{A_2}\,\overline{A_3}) = 1 - 0.6 \cdot 0.5 \cdot 0.3 = 0.91.$

1.32 甲、乙两选手进行兵乓球单打比赛, 已知在每局中甲胜的概率为0.6, 乙胜的概率为0.4. 比赛可采取三局二胜制或五局三胜制, 问哪一种比赛制度对甲更有利?

解：设A_i ={甲在第i中获胜}，$i = 1, \cdots, 5.$ B ={甲在三局二胜制中获胜}，C ={甲在五局三胜制中获胜}.

则$P(A_i) = 0.6, i = 1, \cdots, 5.$

$P(B) = P(A_1A_2) + P(A_1\overline{A_2}A_3) + P(\overline{A_1}A_2A_3) = 0.6^2 + 2 \cdot 0.6^2 \cdot 0.4 = 0.648.$

$P(C) = P(A_1A_2A_3) + P(A_1A_2\overline{A_3}A_4) + P(A_1\overline{A_2}A_3A_4) + P(\overline{A_1}A_2A_3A_4)$

$+P(A_1A_2\overline{A_3}\,\overline{A_4}A_5) + P(A_1\overline{A_2}A_3\overline{A_4}A_5) + P(A_1\overline{A_2}\,\overline{A_3}A_4A_5)$

$+P(\overline{A_1}A_2A_3\overline{A_4}A_5) + P(\overline{A_1}A_2\overline{A_3}A_4A_5) + P(\overline{A_1}\,\overline{A_2}A_3A_4A_5)$

$= 0.68256.$

因为$P(C) > P(B)$,所以，五局三胜制对甲更有利。

习题2

2.1 设随机变量X的分布函数为

$$F(x)=\begin{cases}0, & x<2;\\ \dfrac{1}{8}, & 2\leq x<4;\\ \dfrac{3}{8}, & 4\leq x<6;\\ 1, & x\geq 6.\end{cases}$$

求(1)X的概率分布律;　(2)$P\{X>2\}$, $P\{2<X<6\}$, $P\{2\leq X<6\}$.

解：(1)因为$P\{X=2\}=F(2)-F(2-0)=\dfrac{1}{8}$,

$P\{X=4\}=F(4)-F(4-0)=\dfrac{1}{4}$,

$P\{X=6\}=F(6)-F(6-0)=1-\dfrac{3}{8}=\dfrac{5}{8}$,

且随机变量X在其他各点的概率均为0.

所以随机变量X的概率分布律为：

X	2	4	6
P_X	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{5}{8}$

(2) $P\{X>2\}=1-P\{X\leq 2\}=1-F(2)=\dfrac{7}{8}$.

$P\{2<X<6\}=P\{X<6\}-P\{X\leq 2\}=F(6-0)-F(2)=\dfrac{3}{8}-\dfrac{1}{8}=\dfrac{1}{4}$.

$P\{2\leq X<6\}=P\{X<6\}-P\{X<2\}=F(6-0)-F(2-0)=\dfrac{3}{8}-0=\dfrac{3}{8}$.

2.2 以X表示某商店从早晨开始营业起直到第一个顾客到达的等待时间(以分计),

X的分布函数为

$$F(x)=\begin{cases}1-\mathrm{e}^{-0.4x}, & x\geq 0;\\ 0, & x<0.\end{cases}$$

求(1)$P\{$至多等待3分钟$\}$;

(2)$P\{$至少等待4分钟$\}$;

(3)$P\{$等待时间在3分钟至4分钟之间$\}$;

(4)$P\{$恰好等待2.5分钟$\}$.

解：(1)$P\{$至多等待3分钟$\}$=$P\{X \leq 3\} = F(3) = 1 - \mathrm{e}^{-1.2}$.

(2)$P\{$至少等待4分钟$\}$=$P\{X \geq 4\} = 1 - F(4) = \mathrm{e}^{-1.6}$.

(3)$P\{$等待时间在3分钟至4分钟之间$\}$=$P\{3 \leq X \leq 4\}$

$= F(4) - F(3) = \mathrm{e}^{-1.2} - \mathrm{e}^{-1.6}$.

(4)$P\{$恰好等待2.5分钟$\}$=$P\{X = 2.5\} = F(2.5) - F(2.5 - 0) = 0$.

2.3 设随机变量X的分布函数为

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

且$F(1-0) = \frac{3}{4}$, 求常数a, b, $P\left\{X \leq \frac{1}{2}\right\}$.

解：因为$\begin{cases} F(1-0) = \frac{3}{4}, \\ F(-1+0) = F(-1), \end{cases}$ 即$\begin{cases} a + b = \frac{3}{4}, \\ b - a = \frac{1}{8}, \end{cases}$

解之得$\begin{cases} a = \frac{5}{16}, \\ b = \frac{7}{16}. \end{cases}$

$P\left\{X \leq \frac{1}{2}\right\} = F\left(\frac{1}{2}\right) = \frac{5}{16} \cdot \frac{1}{2} + \frac{7}{16} = \frac{19}{32}$.

2.4 设离散型随机变量 X 的概率分布律为

X	-1	0	1
P_X	$\frac{1}{4}$	a	b

分布函数为

$$F(x) = P\{X \leq x\} = \begin{cases} c, & x < -1; \\ d, & -1 \leq x < 0; \\ \frac{3}{4}, & 0 \leq x < 1; \\ 1, & x \geq 1. \end{cases}$$

求常数a, b, c, d的值.

解：因为$\begin{cases} P\{X=-1\} = F(-1) - F(-1-0), \\ P\{X=0\} = F(0) - F(0-0), \\ P\{X=1\} = F(1) - F(1-0), \\ P\{X=-1\} + P\{X=0\} + P\{X=1\} = 1. \end{cases}$ 即$\begin{cases} \frac{1}{4} = d - c, \\ a = \frac{3}{4} - d, \\ b = 1 - \frac{3}{4}, \\ \frac{1}{4} + a + b = 1. \end{cases}$

解之得$\begin{cases} a = \frac{1}{2}, \\ b = \frac{1}{4}, \\ c = 0, \\ d = \frac{1}{4}. \end{cases}$

2.5 在一个盒子中有8个二极管,已知其中有4个正品,4个次品. 从盒子中依次取出进行测试, 直到取得一个正品为止.求在停止取二极管时,已取出的二极管个数X 的概率分布律.

解：据题意，随机变量X的所有可能取值为1,2,3,4,5.

$P\{X=1\}=\dfrac{4}{8}=\dfrac{1}{2},$

$P\{X=2\}=\dfrac{4}{8}\cdot\dfrac{4}{7}=\dfrac{2}{7},$

$P\{X=3\}=\dfrac{4}{8}\cdot\dfrac{3}{7}\cdot\dfrac{4}{6}=\dfrac{1}{7},$

$P\{X=4\}=\dfrac{4}{8}\cdot\dfrac{3}{7}\cdot\dfrac{3}{6}\cdot\dfrac{4}{5}=\dfrac{2}{35},$

$P\{X=5\}=\dfrac{4}{8}\cdot\dfrac{3}{7}\cdot\dfrac{2}{6}\cdot\dfrac{1}{5}\cdot 1=\dfrac{1}{70}.$

所以，随机变量X的概率分布律为

X	1	2	3	4	5
P_X	$\dfrac{1}{2}$	$\dfrac{2}{7}$	$\dfrac{1}{7}$	$\dfrac{2}{35}$	$\dfrac{1}{70}$

2.6 设箱中有20件产品,每次从箱中任取一件,取后放回,已知在3次抽取中至少出现1件正品的概率为0.936.

(1)问原来箱中有多少件正品?

(2)若以X表示在3次抽取中出现的正品数,求X的概率分布律及分布函数.

解：设原来箱中有n件正品,用X表示3次抽取中取到的正品数，则$X\sim B\left(3,\dfrac{n}{20}\right)$,

(1)$P\{X\geq 1\}=1-P\{X=0\}=1-\left(1-\dfrac{n}{20}\right)^3=0.936$, 解之得$n=12$.

所以，原来箱中有12件正品。

(2)随机变量X的概率分布律为$P\{X=k\}=C_3^k\left(\dfrac{3}{5}\right)^k\left(\dfrac{2}{5}\right)^2,k=0,1,2,3.$

分布函数为

$$F(x)=P\{X\le x\}=\begin{cases}0, & x<0;\\ 0.064, & 0\le x<1;\\ 0.352, & 1\le x<2;\\ 0.784, & 2\le x<3;\\ 1, & x\ge 3.\end{cases}$$

2.7 一袋中装有5只球, 编号分别为1,2,3,4,5. 在袋中同时取出3只球, 以X表示取出的3只球中的最大号码. 试写出X 的概率分布律, 求其分布函数并画出图像.

解：随机变量X的所有可能取值为3,4,5.

$P\{X=3\}=\dfrac{1}{C_5^3}=\dfrac{1}{10}=0.1,$

$P\{X=4\}=\dfrac{C_3^2}{C_5^3}=\dfrac{3}{10}=0.3,$

$P\{X=5\}=\dfrac{C_4^2}{C_5^3}=\dfrac{6}{10}=0.6.$

所以随机变量X的概率分布律为

X	3	4	5
P_X	0.1	0.3	0.6

分布函数为

$$F(x)=P\{X\le x\}=\begin{cases}0, & x<3;\\ 0.1, & 3\le x<4;\\ 0.4, & 4\le x<5;\\ 1, & x\ge 5.\end{cases}$$

2.8 一大楼装有5个同类型的供水设备,调查表明:在任一时刻每个设备被使用的概率

为0.1,各台设备是否被使用相互独立,问在同一时刻,

(1)恰有2个设备被使用的概率是多少?

(2)至少有3个设备被使用的概率是多少?

(3)至多有3个设备被使用的概率是多少?

(4)至少有1个设备被使用的概率是多少?

解：设X表示正在使用的设备数，则$X \sim B(5, 0.1)$.

(1)$P\{X = 2\} = C_5^2 0.1^2 0.9^3 = 0.0729.$

(2)$P\{X \geq 3\} = \sum\limits_{i=3}^{5} C_5^i 0.1^i 0.9^{5-i} = 0.00856.$

(3)$P\{X \leq 3\} = \sum\limits_{i=0}^{3} C_5^i 0.1^i 0.9^{5-i} = 0.99954.$

(4)$P\{X \geq 1\} = 1 - P\{X = 0\} = 1 - 0.9^5 = 0.40951.$

2.9 有一大批产品, 其验收方案如下, 先做第一次检验: 从中任取10件, 经检验若无次品,则接受这批产品;若次品数大于2,则拒收; 否则作第二次检验, 其做法是再从中任取5件, 仅当5件产品中无次品时才接受这批产品. 若产品的次品率为10%, 求

(1)这批产品经第一次检验就能接受的概率.

(2)需作第二次检验的概率.

(3)这批产品按第二次检验的标准被接受的概率.

(4)这批产品在第一次检验未能做决定且第二次检验时被通过的概率.

(5)这批产品被接受的概率.

解：设X表示第一次抽取10件产品中的次品数，则$X \sim B(10, 0.1)$,Y表示第二次抽取5件产品中的次品数，则$Y \sim B(10, 0.1)$.

(1)$P\{X = 0\} = 0.9^{10} = 0.3487.$

(2)$P\{1 \leq X \leq 2\} = P\{X = 1\} + P\{X = 2\}$

$= C_{10}^1 0.1 \cdot 0.9^9 + C_{10}^2 0.1^2 \cdot 0.9^8 = 0.5811.$

(3)$P\{Y=0\}=0.9^5=0.59049.$

(4)$P\{1\le X\le 2,Y=0\}=P\{1\le X\le 2\}P\{Y=0\}=0.3431.$

(5)$P\{X=0\}+P\{1\le X\le 2,Y=0\}=0.6918.$

2.10 某药厂声称某种药对某种疾病的治愈率为80%, 现有10名患者同时服用这种药物, 问这10人中至少有6人治愈的概率有多大?如果治愈者不超过5人, 其药效达80%的说法可靠吗?

解：设X表示10名患者中治愈的人数，则$X\sim B(10,0.8)$,

$P\{X\ge 6\}=\sum\limits_{i=6}^{10}C_{10}^{i}0.8^{i}\cdot 0.2^{10-i}=0.967.$

因为治愈者不超过5人的概率

$$P\{X\le 5\}=1-P\{X\ge 6\}=0.033$$

是小概率事件，在一次试验中不应该发生。如果发生了，就有理由认为药效达80%的说法不可靠。

2.11 设X 服从泊松分布,已知$P\{X=1\}=P\{X=2\}$,求$P\{X=4\}$.

解：因为 $P\{X=1\}=P\{X=2\}$,所以 $\lambda\mathrm{e}^{-\lambda}=\dfrac{\lambda^2}{2}\mathrm{e}^{-\lambda}$,

解之得$\lambda=2.$

故$P\{X=4\}=\dfrac{2^4}{4!}\mathrm{e}^{-2}=0.0902.$

2.12 设X 服从参数为λ的泊松分布, 求X的最可能取值.

解：因为$X\sim P(\lambda)$,所以$P\{X=k\}=\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda},k=0,1,2,\cdots.$

又因为$\dfrac{P\{X=k+1\}}{P\{X=k\}}=\dfrac{\lambda}{k+1}$,

所以当$k\le\lambda-1$时，$P\{X=k\}$单调递增，当$k<\lambda-1$时，$P\{X=k\}$单调递减。由于k取自然数，故

当λ是整数时，X的最可能取值为λ或$\lambda-1$。

当λ不是整数时，X的最可能取值为$[\lambda]$。

2.13 某一公安局在长度为t的时间间隔内收到紧急呼救的次数X 服从参数为$\frac{1}{2}t$泊松分布, 而与时间间隔的起点无关(时间以小时计).

(1)求某一天中午12时至下午3时未收到紧急呼救的概率.

(2)求某一天中午12时至下午5时至少收到1次紧急呼救的概率.

解：因为在长度为t的时间间隔内收到紧急呼救的次数$X \sim P\left(\frac{t}{2}\right)$，所以

(1)在中午12时至下午3时收到的紧急呼救次数$X \sim P(1.5)$,未收到紧急呼救的概率为 $P\{X=0\}=\mathrm{e}^{-1.5}=0.2231$.

(2)在中午12时至下午5时收到的紧急呼救次数$X \sim P(2.5)$,至少收到1次紧急呼救的概率为 $P\{X \geq 1\}=1-P\{X=0\}=1-\mathrm{e}^{-2.5}=0.9179$.

2.14 已知在一定工序下,生产某种产品的次品率为0.001. 今在此同一工序下,独立生产5000件这种产品,求至少有2件次品的概率.

解：因为次品率$p=0.001$很小，产品数量$n=5000$较大，所以随机变量X近似服从参数$\lambda=0.001\cdot 5000=5$的泊松分布.至少有2件次品的概率为

$P\{X \geq 2\}=1-P\{X=0\}-P\{X=1\}=1-\mathrm{e}^{-5}-5\cdot\mathrm{e}^{-5}=0.9596$.

2.15 某商店出售某种高档商品,根据以往经验,每月需求量X 服从参数$\lambda=3$ 的泊松分布,问在月初进货时要库存多少件这种商品,才能以99%以上的概率满足顾客的需要.

解：假设月初进货时要库存m件商品，才能以99%以上的概率满足顾客的需要.

据题意得$P\{X \leq m\} \geq 0.99$,

又因为$X \sim P(3)$,查表得$m \geq 8$.

2.16 假设某种昆虫产k个卵的概率为$\dfrac{\lambda^k \mathrm{e}^{-\lambda}}{k!}$,而一个卵孵成昆虫的概率为$p$. 假设各个卵是否孵化成昆虫是相互独立的,试求一只昆虫恰有l只后代的概率.

解：设A_k={昆虫产k个卵}，$k=1,2,\cdots$, B ={一只昆虫恰有l只后代}，

$P(A_k)=\dfrac{\lambda^k \mathrm{e}^{-\lambda}}{k!}, k=1,2,\cdots.$

$P(B)=\sum\limits_{k=l}^{\infty} P(A_k)P(B|A_k)=\sum\limits_{k=l}^{\infty}\left(\dfrac{\lambda^k \mathrm{e}^{-\lambda}}{k!}C_k^l p^l(1-p)^{k-l}\right)$

$=\sum\limits_{k=l}^{\infty}\left(\dfrac{\lambda^k \mathrm{e}^{-\lambda}}{k!}\cdot\dfrac{k!}{l!(k-l)!}\cdot^l(1-p)^{k-l}\right)$

$=\dfrac{\mathrm{e}^{-\lambda}}{l!}\left(\dfrac{p}{1-p}\right)^l\sum\limits_{k=l}^{\infty}\left(\dfrac{[\lambda(1-p))]^{k-l}}{(k-l)!}\cdot[\lambda(1-p)]^l\right)$

$=\dfrac{\mathrm{e}^{-\lambda}}{l!}(\lambda p)^l\mathrm{e}^{\lambda(1-p)}=\dfrac{(\lambda p)^l}{l!}\mathrm{e}^{-\lambda p}.$

2.17 自动生产线在调整之后出现不合格品的概率为p. 若在生产过程中出现不合格品,则立即再进行调整. 求

(1)在两次调整之间生产的合格品数X的概率分布律;

(2)在两次调整之间生产的合格品数X不小于5的概率.

解：(1) 随机变量X的所有可能取值为$0,1,2,\cdots$,依题意得

$P\{X=k\}=(1-p)^k p$，所以，X的概率分布律为

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

(2)$P\{X\geq 5\}=\sum\limits_{k=5}^{\infty}P\{X=k\}=\sum\limits_{k=5}^{\infty}(1-p)^k p=(1-p)^5.$

2.18 设随机变量X的概率密度函数为

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

试确定常数C,并求X 的分布函数.

解：因为$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$, 即$\int_1^2 Cx^2\mathrm{d}x+\int_2^3 Cx\mathrm{d}x=1$,

解之得$C=\dfrac{6}{29}$.

当$x<1$时，$F(x)=P\{X\leq x\}=\int_{-\infty}^{x}f(x)\mathrm{d}x=0;$

当$1 \le x < 2$时，$F(x) = P\{X \le x\} = \int_{-\infty}^{x} f(x)\mathrm{d}x = \int_{1}^{x} \frac{6}{29} x^2 \mathrm{d}x = \frac{2}{29}(x^3 - 1)$;

当$2 \le x < 3$时，$F(x) = P\{X \le x\} = \int_{-\infty}^{x} f(x)\mathrm{d}x$

$$= \int_{1}^{2} \frac{6}{29} x^2 \mathrm{d}x \int_{2}^{x} \frac{6}{29} x \mathrm{d}x = \frac{1}{29}(3x^2 + 2);$$

当$2 \le x < 3$时，$F(x) = P\{X \le x\} = \int_{-\infty}^{x} f(x)\mathrm{d}x$

$$= \int_{1}^{2} \frac{6}{29} x^2 \mathrm{d}x \int_{2}^{3} \frac{6}{29} x \mathrm{d}x = 1.$$

综上得随机变量X的分布函数为

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{2}{29}(x^3 - 1), & 1 \le x < 2, \\ \frac{1}{29}(3x^2 + 2), & 2 \le x < 3, \\ 1, & x \ge 3. \end{cases}$$

2.19 设随机变量X的概率密度函数为

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

求常数a,使得$P\{X > a\} = P\{X < a\}$.

解：要求常数a,使得$P\{X > a\} = P\{X < a\}$,

只要$\int_{a}^{1} 2x \mathrm{d}x = \int_{0}^{a} 2x \mathrm{d}x$, 解之得$a = \frac{\sqrt{2}}{2}$.

2.20 设随机变量X的概率密度函数为$f(x) = A\mathrm{e}^{-|x|}, x \in \mathrm{R}$, 求(1)常数$A$; (2)$X$ 的分布函数; (3) $P\{-1 < X < 2\}$.

解：(1)因为$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$, 即$\int_{-\infty}^{+\infty} A\mathrm{e}-|x|\mathrm{d}x = 1$,

解之得$A = \frac{1}{2}$.

(2)当$x < 0$时，$F(x) = P\{X \le x\} = \int_{-\infty}^{x} f(x)\mathrm{d}x = \int_{-\infty}^{x} \frac{1}{2}\mathrm{e}^{x} \mathrm{d}x = \frac{1}{2}\mathrm{e}^{x}$;

当$x \ge 0$时，$F(x) = P\{X \le x\} = \int_{-\infty}^{0} \frac{1}{2}\mathrm{e}^{x} \mathrm{d}x + \int_{0}^{x} \frac{1}{2}\mathrm{e}^{-x} \mathrm{d}x = 1 - \frac{1}{2}\mathrm{e}^{x}$.

综上得随机变量X的分布函数为

$$F(x)=\begin{cases}\dfrac{1}{2}\mathrm{e}^{x}, & x<0;\\ 1-\dfrac{1}{2}\mathrm{e}^{-x}, & x\geq 0.\end{cases}$$

(3) $P\{-1<X<2\}=F(2)-F(-1)=1-\dfrac{1}{2}\left(\mathrm{e}^{-2}+\mathrm{e}^{-1}\right).$

2.21 设连续型随机变量X的分布函数为

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

求(1)常数A 、B、C; (2)随机变量X的概率密度函数; (3)$P\left\{X>\dfrac{1}{2}\right\}$.

解：(1)根据分布函数的连续性质得

$$\begin{cases}F(0-0) &= F(0+0)\\ F(1-0) &= F(0+0)\\ F(2-0) &= F(2)\end{cases} \quad 即 \quad \begin{cases}A=0\\ B=C-\dfrac{3}{2}\\ 2C-3=1\end{cases}$$

解之得$\begin{cases}A=0,\\ B=\dfrac{1}{2},\\ C=2.\end{cases}$

(2)当$x<0$或$x>2$时，$f(x)=F'(x)=0$;

当$0<x<1$时，$f(x)=F'(x)=x$;

当$1<x<2$时，$f(x)=F'(x)=2-x$;

$f(x)$在分界点处的导数可以任意赋值。所以，随机变量X的概率密度函数为

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

(3)$P\left\{X>\frac{1}{2}\right\}=1-P\{X\leq\frac{1}{2}\}=F\left(\frac{1}{2}\right)=\frac{7}{8}.$

2.22 设X在区间$[-2,5]$上服从均匀分布,求方程

$$4u^2+4Xu+X+2=0$$

有实根的概率.

解：方程$4u^2+4Xu+X+2=0$有实根,当且仅当$16X^2-16(X+2)\geq 0$,

即$X<-1$或$X>2$.

又因为X在区间$[-2,5]$上服从均匀分布，所以

$P\{X<-1\,or\,X>2\}=P\{-2\leq X<-1\}+P\{2<X\leq 5\}=\frac{4}{7}.$

2.23 设某种元件的寿命X（单位:h）服从参数$\lambda=1/600$的指数分布. 现有3个这种元件在独立地工作,以Y表示在最初使用的200小时内这3个元件损坏的个数,

(1)写出随机变量Y的概率分布律;

(2)求至少有1个元件损坏的概率.

解：(1)据题意，随机变量Y服从二项分布$B(3,p)$, 其中$p=P\{X\leq 200\}$.

因为随机变量$X\sim E\,(1/600)$, 所以$p=P\{X\leq 200\}=1-\mathrm{e}^{-1/3}$.

故随机变量Y的概率分布律为

$$P\{Y=k\}=C_3^k\left(1-\mathrm{e}^{-1/3}\right)^k\mathrm{e}^{-\frac{3-k}{3}},\ k=0,1,2,3.$$

(2)$P\{Y\geq 1\}=1-P\{Y=0\}=1-\mathrm{e}^{-1}.$

2.24 设带有3个炸弹的轰炸机向敌人铁路投弹,若炸弹落在铁路两旁40 m以内,便可破坏铁路交通,弹落点与铁路距离记为X(单位:m ,在铁路一侧为正,在另一侧为负),且X的概率密度函数为

$$f(x)=\begin{cases}\dfrac{100+x}{10000}, & -100\le x<0;\\[2ex] \dfrac{100-x}{10000}, & 0\le x\le 100;\\[2ex] 0, & \text{其他}.\end{cases}$$

若3个炸弹全部使用,问敌人铁路交通受到破坏的概率是多少?

解:铁路受到破坏当且仅当$|X|\le 40$,所以铁路受到破坏的概率为

$$p=P\{|X|\le 40\}=\int_{-40}^{0}\frac{100+x}{10000}\mathrm{d}x+\int_{0}^{40}\frac{100-x}{10000}\mathrm{d}x=0.64.$$

用Y表示3个炸弹中能破坏铁路的炸弹数，则$Y\sim B(3,p)$,若3个炸弹全部使用,则敌人铁路交通受到破坏的概率是

$P\{Y\ge 1\}=1-P\{Y=0\}=1-0.36^3=0.9533.$

2.25 设随机变量X 的概率密度函数为

$$f(x)=a\mathrm{e}^{-(x+1)^2},\ x\in \mathrm{R},$$

(1)求常数a;　(2)已知$X\sim N(\mu,\sigma^2)$, 问μ,σ各取何值?

解：(1)因为$f(x)=a\mathrm{e}^{-(x+1)^2}=\sqrt{\pi}a\cdot\dfrac{1}{\sqrt{2\pi}\cdot\sqrt{1/2}}\mathrm{e}^{-\left(\frac{x+1}{2\cdot 1/2}\right)^2}$,

根据正态分布概率密度函数的定义知，$\sqrt{\pi}a=1$,所以$a=1/\sqrt{\pi}$.

(2)由正态分布的概率密度函数的标准表达式$f(x)=\dfrac{1}{\sqrt{2\pi}\cdot\sqrt{1/2}}\mathrm{e}^{-\left(\frac{x+1}{2\cdot 1/2}\right)^2}$ 知

$$\mu=-1,\ \sigma=1/\sqrt{2}.$$

2.26 设随机变量$X\sim N(0,\sigma^2)$, 若$P\{|X|>k\}=0.1$, 求$P\{X<k\}$.

解：因为$X \sim N(0,\sigma^2)$,且$P\{|X|>k\}=0.1$,

所以 $P\{X>k\}=0.05$,

故$P\{X<k\}=1-P\{X>k\}=0.95$.

2.27 一工厂生产的某种元件的寿命X(以小时计)服从参数$\mu=160,\sigma^2(\sigma>0)$的正态分布. 若要求$P\{120<X\le 200\}\ge 0.80$, 允许$\sigma$最大为多少?

解：因为 $X\sim N(160,\sigma^2)$, 所以

$$P\{120<X\le 200\}=\Phi\left(\frac{200-160}{\sigma}\right)-\Phi\left(\frac{120-160}{\sigma}\right)=2\Phi\left(\frac{40}{\sigma}\right)-1,$$

要使$P\{120<X\le 200\}\ge 0.80$, 只要$2\Phi\left(\frac{40}{\sigma}\right)-1\ge 0.8$, 即 $\Phi\left(\frac{40}{\sigma}\right)\ge 0.9$,

查正态分布表得：$\frac{40}{\sigma}\ge 1.29$, 由此得$\sigma\le 31$.

2.28 某人上班, 自家里去办公楼要经过一交通指示灯, 这一指示灯有80%时间亮红灯, 此时他在指示灯旁等待直至绿灯亮. 等待时间在区间$[0,30]$(以秒计)服从均匀分布. 以X表示他的等待时间, 求X的分布函数$F(x)$. 问X是否为连续型随机变量, 是否为离散型随机变量, 为什么?

解：设$A=${指示灯亮红灯}，则$P(A)=0.80$.

当$x<0$时，$F(x)=P\{X\le x\}=0$;

当$0\le x<30$时，$F(x)=P\{X\le x\}=P(A\cup\{X\le x\})+P(\overline{A}\cup\{X\le x\})$

$$=P(A)P\{X\le x|A\}+P(\overline{A})P\{X\le x|\overline{A}\}=\frac{0.8x}{30}+0.2;$$

当$x\ge 30$时，$F(x)=P\{X\le x\}=1$.

综上得随机变量X的分布函数为

$$F(x)=\begin{cases}0, & x<0;\\ 0.2+\dfrac{2x}{75}, & 0\le x<30;\\ 1, & x\ge 30.\end{cases}$$

2.29 设随机变量X 的概率分布律为

X	-2	-1	0	1	2
P_X	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{11}{30}$

(1)令$Y_1 = 1 - X$, 求Y_1的概率分布律;

(2)令$Y_2 = X^2$, 求Y_2的概率分布律.

解：(1)首先确定随机变量Y_1的所有可能取值为：-1,0,1,2,3.

然后计算Y_1取每个可能值的概率。

$P\{Y_1 = -1\} = P\{1 - X = -1\} = P\{X = 2\} = 11/30;$

$P\{Y_1 = 0\} = P\{1 - X = 0\} = P\{X = 1\} = 1/15;$

$P\{Y_1 = 1\} = P\{1 - X = 1\} = P\{X = 0\} = 1/5;$

$P\{Y_1 = 2\} = P\{1 - X = 2\} = P\{X = -1\} = 1/6;$

$P\{Y_1 = 3\} = P\{1 - X = 3\} = P\{X = -2\} = 1/5.$

综上得随机变量Y_1的概率分布律为

Y_1	-1	0	1	2	3
P_{Y_1}	$\frac{11}{30}$	$\frac{1}{15}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$

(2)首先确定随机变量Y_2的所有可能取值为：0,1,4.

然后计算Y_2取每个可能值的概率。

$P\{Y_2 = 0\} = P\{X^2 = 0\} = P\{X = 0\} = 1/5;$

$P\{Y_2 = 1\} = P\{X^2 = 1\} = P\{X = -1\} + P\{X = 1\} = 7/30;$

$P\{Y_2 = 4\} = P\{X^2 = 4\} = P\{X = -2\} + P\{X = 2\} = 17/30.$

综上得随机变量Y_2的概率分布律为

Y_2	0	1	4
P_{Y_2}	$\frac{1}{5}$	$\frac{7}{30}$	$\frac{17}{30}$

3.30 设随机变量X服从参数为1的指数分布,令

$$Y=\begin{cases}-1, & X<1;\\ 0, & 1\le X<5;\\ 1, & X\ge 5,\end{cases}$$

求随机变量Y的分布.

解：因为$X\sim E(1)$,所以X的分布函数为$F(x)=\begin{cases}0 & x<0;\\ 1-\mathrm{e}^{-x} & x\ge 0.\end{cases}$

$P\{Y=-1\}=P\{X<1\}=F(1)=1-\mathrm{e}^{-1}$;

$P\{Y=0\}=P\{1\le X<5\}=F(5)-F(1)=\mathrm{e}^{-1}-\mathrm{e}^{-5}$;

$P\{Y=1\}=P\{X\ge 5\}=1-F(5)=\mathrm{e}^{-5}$.

所以，随机变量Y的概率分布律为

Y	-1	0	1
P_Y	$1-\mathrm{e}^{-1}$	$\mathrm{e}^{-1}-\mathrm{e}^{-5}$	e^{-5}

2.31 设随机变量X的概率密度函数为

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

现对X进行n次独立重复观测,用Y表示观测值不大于0.2的个数,试求Y的分布.

解：根据题意得$p=P\{X\le 0.2\}=\int_{-\infty}^{x}f(x)\mathrm{d}x=\int_{0}^{0.2}2x\mathrm{d}x=0.04$,

则$Y\sim B(n,0.04)$.

2.32 设随机变量$X\sim N(0,1)$. 令$Y_1=\mathrm{e}^X,Y_2=\sqrt{|X|}$, 分别求$Y_1,Y_2$的概率密度函数.

解：因为$X\sim N(0,1)$,所以$f_X(x)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$, $x\in\mathrm{R}$.

当$y\le 0$时，$F_{Y_1}(y)=P\{Y_1\le y\}=0$;

当$y>0$时，$F_{Y_1}(y)=P\{Y_1\le y\}=P\{\mathrm{e}^X\le y\}=P\{X\le \ln y\}=F_X(\ln y)$.

对分布函数求导数得随机变量Y_1的概率密度函数.

当$y<0$时，$f_{Y_1}(y)=0$;

当$y>0$时，$f_{Y_1}(y)==f_X(\ln y)\cdot\frac{1}{y}=\dfrac{1}{y\sqrt{2\pi}}\mathrm{e}^{-\frac{(\ln y)^2}{2}}$.

综上得Y_1的概率密度函数为

$$f_{Y_1}(y)=\begin{cases}\dfrac{1}{y\sqrt{2\pi}}\mathrm{e}^{-\dfrac{(\ln y)^2}{2}}, & y>0;\\ 0, & y\le 0.\end{cases}$$

下面计算随机变量Y_2的概率密度函数。

当$y\le 0$时，$F_{Y_2}(y)=P\{Y_2\le y\}=P\{\sqrt{|X|}\le y\}=0$;

当$y>0$时，$F_{Y_2}(y)=P\{Y_2\le y\}=P\{\sqrt{|X|}\le y\}$

$$=P\{|X|\le y^2\}=F_X(y^2)-F_X(-y^2).$$

对分布函数求导数得随机变量Y_2的概率密度函数.

当$y<0$时，$f_{Y_2}(y)=0$;

当$y>0$时，$f_{Y_2}(y)==f_X(y^2)\cdot(2y)-f_X(-y^2)\cdot(-2y)=\dfrac{4y}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^4}{2}}$.

综上得Y_2的概率密度函数为

$$f_{Y_2}(y)=\begin{cases}\dfrac{4y}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^4}{2}}, & y>0;\\ 0, & y\le 0.\end{cases}$$

2.33 设随机变量X的概率密度函数为

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

令$Y=(X-1)^2$,求Y的概率密度函数.

解：当$y<0$时，$F_Y(y)=P\{Y\le y\}=P\{(X-1)^2\le y\}=0$;

当$0 \le y < 1$时，$F_Y(y) = P\{(X-1)^2 \le y\} = P\{1-\sqrt{y} \le X \le 1+\sqrt{y}\}$

$$= \int_{1-\sqrt{y}}^{1+\sqrt{y}} \tfrac{3}{8} x^2 \mathrm{d}x = \frac{1}{4}\left(3y^{1/2} + y^{3/2}\right);$$

当$y \ge 1$时，$F_Y(y) = P\{(X-1)^2 \le y\} = \int_0^2 \tfrac{3}{8} x^2 \mathrm{d}x = 1.$

对分布函数求导数得Y的概率密度函数为

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

2.34 设随机变量X的概率密度函数为

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

令$Y = X^2$, 求Y的概率密度函数.

解：当$y < 0$时，$F_Y(y) = P\{Y \le y\} = P\{X^2 \le y\} = 0;$

当$0 \le y < 1$时，$F_Y(y) = P\{X^2 \le y\} = P\{-\sqrt{y} \le X \le \sqrt{y}\}$

$$= \int_{-\sqrt{y}}^{0} \tfrac{1}{2} \mathrm{d}x + \int_0^{\sqrt{y}} \tfrac{1}{4} \mathrm{d}x = \frac{3\sqrt{y}}{4};$$

当$1 \le y < 4$时，$F_Y(y) = P\{X^2 \le y\} = P\{-\sqrt{y} \le X \le \sqrt{y}\}$

$$= \int_{-1}^{0} \tfrac{1}{2} \mathrm{d}x + \int_0^{\sqrt{y}} \tfrac{1}{4} \mathrm{d}x = \frac{2+\sqrt{y}}{4};$$

当$y \ge 4$时，$F_Y(y) = P\{X^2 \le y\} = P\{-\sqrt{y} \le X \le \sqrt{y}\} = \int_{-1}^{0} \tfrac{1}{2} \mathrm{d}x + \int_0^{2} \tfrac{1}{4} \mathrm{d}x = 1.$

对分布函数求导数得Y的概率密度函数为

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

2.35 设某设备在任何长为 t 的时间 $[0,t]$ 内发生故障的次数 $N(t)$ 服从参数为 λt 的泊松分布, 求相继两次故障之间的时间间隔 T的分布.

解：当$t \le 0$时，$F_T(t) = P\{T \le t\} = 0$;

当$t > 0$时，$F_T(t) = P\{T \le t\} = 1 - P\{T > t\} = 1 - P\{N(t) = 0\} = 1 - \mathrm{e}^{-\lambda t}$.

综上得$T \sim E(\lambda)$.

2.36 设随机变量X的概率密度函数为

$$f(x) = \begin{cases} \dfrac{2x}{\pi^2}, & 0 < x < \pi; \\ 0, & \text{其他}. \end{cases}$$

令$Y = \sin X$, 求Y的概率密度函数.

解：当$0 < x < \pi$时，$0 < y < 1$.

当$y < 0$时，$F_Y(y) = P\{Y \le y\} = P\{\sin X \le y\} = 0$;

当$0 \le y < 1$时，$F_Y(y) = P\{Y \le y\} = P\{\sin X \le y\}$

$$\begin{aligned} &= P\{0 \le X \le \arcsin y\} + P\{\pi - \arcsin y \le X \le \pi\} \\ &= \int_0^{\arcsin y} \frac{2x}{\pi^2}\mathrm{d}x + \int_{\pi-arcsiny}^{\pi} \frac{2x}{\pi^2}\mathrm{d}x = \frac{2\arcsin y}{\pi}; \end{aligned}$$

当$y \ge 1$时，$F_Y(y) = P\{Y \le y\} = P\{\sin X \le y\} = 1$.

对分布函数求导数得Y的概率密度函数为

$$f_Y(y) = \begin{cases} \dfrac{2}{\pi\sqrt{1-y^2}}, & 0 < y < 1; \\ 0, & \text{其他}. \end{cases}$$

2.37 设电流I是一个随机变量, 它均匀分布在9A 至11A 之间,若此电流通过2Ω 的电阻,功率$W = 2I^2$,求W的概率密度函数.

解：依题意随机变量$I \sim U(9,11)$,所以I的概率密度函数为

$$f_I(x)=\begin{cases}\dfrac{1}{2}, & 9\le x\le 11;\\[2ex] 0, & 其他.\end{cases}$$
当$9\le I\le 11$时，$162\le W\le 242$.

当$w<162$时，$F_W(w)=P\{W\le w\}=P\{2I^2\le w\}=0,\ f_W(w)=0;$

当$162\le w<242$时，$F_W(w)=P\{W\le w\}=P\{2I^2\le w\}$

$$=P\{0\le I\le \sqrt{w/2}\}=F_I\left(\sqrt{w/2}\right),$$

求导数得 $f_W(w)=f_I\left(\sqrt{w/2}\right)\cdot\dfrac{1}{4}\sqrt{\dfrac{2}{w}}=\dfrac{1}{8}\sqrt{\dfrac{2}{w}}.$

当$w\ge 242$时，$F_W(w)=P\{W\le w\}=P\{2I^2\le w\}=1,\ f_W(w)=0.$

综上得W的概率密度函数为

$$f_W(w)=\begin{cases}\dfrac{1}{8}\sqrt{\dfrac{2}{w}}, & 162\le w\le 242;\\[2ex] 0, & 其他.\end{cases}$$

2.38 设随机变量X 在区间$\left[-\dfrac{\pi}{2},\dfrac{\pi}{2}\right]$上服从均匀分布,令$Y=\cos X$,求$Y$的分布函数.

解：当$-\dfrac{\pi}{2}\le X\le\dfrac{\pi}{2}$时，$0\le Y\le 1$.

当$y<0$时，$F_Y(y)=P\{Y\le y\}=P\{\cos X\le y\}=0;$

当$0\le y<1$时，$F_Y(y)=P\{Y\le y\}=P\{\cos X\le y\}=1-P\{\cos X>y\}$

$$=1-P\{-\arcsin y\le X\le \arccos y\}$$

$$=1-\textstyle\int_{-\arccos y}^{\arccos y}\dfrac{1}{\pi}\mathrm{d}x=1-\dfrac{2\arccos y}{\pi};$$

当$y\ge 1$时，$F_Y(y)=P\{Y\le y\}=P\{\sin X\le y\}=1.$

综上得随机变量Y的分布函数为

$$F_Y(y)=\begin{cases}0, & y<0;\\[2ex] 1-\dfrac{2\arccos y}{\pi}, & 0\le y<1;\\[2ex] 1, & y\ge 1.\end{cases}$$

2.39 设测量某距离的误差$X\sim N(-20,40^2)$（单位:mm），

(1)求在3次独立测量中,至少有2次误差的绝对值不超过30(mm) 的概率;

(2)令$Y=|X|$, 求Y 的概率密度函数.

解：因为$X\sim N(-20,40^2)$,所以

$p=P\{|X|\le 30\}=P\{-30\le X\le 30\}=\Phi(1.25)+\Phi(0.25)-1=0.493.$

(1)用Y表示3次独立测量中误差绝对值不超过30mm的次数，则$Y\sim B(3,p)$,

$$P\{Y\ge 2\}=\sum_{k=2}^{3}C_3^k p^k(1-p)^{3-k}=0.4895.$$

(2)当$y<0$时，$F_Y(y)=P\{Y\le Y\}=P\{|X|\le y\}=0,\ f_Y(y)=0;$

当$y\ge 0$时，$F_Y(y)=P\{|X|\le y\}=P\{-y\le X\le y\}=\Phi\left(\frac{20+y}{40}\right)-\Phi\left(\frac{20-y}{40}\right)$,

$$f_Y(y)=\varphi\left(\frac{20+y}{40}\right)/40+\varphi\left(\frac{20-y}{40}\right)/40=\frac{1}{40\sqrt{2\pi}}\left(\mathrm{e}^{-\dfrac{(y+20)^2}{3200}}+\mathrm{e}^{-\dfrac{(y-20)^2}{3200}}\right).$$

综上得Y 的概率密度函数为

$$f_Y(y)=\begin{cases}\dfrac{1}{40\sqrt{2\pi}}\left(\mathrm{e}^{-\dfrac{(y+20)^2}{3200}}+\mathrm{e}^{-\dfrac{(y-20)^2}{3200}}\right), & y\ge 0;\\ 0, & y<0.\end{cases}$$

2.40 假设一设备开机后无故障工作的时间X(h)服从参数为1/5 的指数分布.设备定时开机,出现故障时自动关机;无故障时工作2h便关机.求该设备每次开机无故障工作时间Y的分布函数.

解：依题意$Y=\begin{cases}X, & X<2;\\ 2, & X\ge 2.\end{cases}$

当$y<0$时，$F_Y(y)=P\{Y\le y\}=0;$

当$0\le y<2$时，$F_Y(y)=P\{Y\le y\}=P\{X\le y\}=1-\mathrm{e}^{-\dfrac{y}{5}};$

当$y\ge 2$时，$F_Y(y)=P\{Y\le y\}=1.$

综上得随机变量Y的分布函数为

$$F_Y(y)=\begin{cases}0, & y<0;\\ 1-\mathrm{e}^{-\dfrac{y}{5}}, & 0\le y<2;\\ 1, & y\ge 2.\end{cases}$$

习题3

3.1 设二维随机变量 (X,Y) 的联合分布函数为

$$F(x,y)=\begin{cases}\left(1-\mathrm{e}^{-\frac{1}{2}x}\right)\left(1-\mathrm{e}^{-\frac{1}{2}y}\right), & x\ge 0, y\ge 0;\\ 0, & \text{其他}.\end{cases}$$

求（1）随机变量X,Y 的边缘分布函数;

（2）$P\{1\le X\le 2, -1\le Y\le 3\}$;

（3）$P\{X>2, Y>3\}$.

解：(1) $F_X(x)=\lim\limits_{y\to+\infty}F(x,y)=\begin{cases}\lim\limits_{y\to+\infty}\left(1-\mathrm{e}^{-\frac{1}{2}x}\right)\left(1-\mathrm{e}^{-\frac{1}{2}y}\right), & x\ge 0\\ 0, & x<0,\end{cases}$

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

$F_Y(y)=\lim\limits_{x\to+\infty}F(x,y)=\begin{cases}\lim\limits_{x\to+\infty}\left(1-\mathrm{e}^{-\frac{1}{2}x}\right)\left(1-\mathrm{e}^{-\frac{1}{2}y}\right), & y\ge 0\\ 0, & y<0,\end{cases}$

$$=\begin{cases}1-\mathrm{e}^{-\frac{1}{2}y}, & y\ge 0\\ 0, & y<0,\end{cases}$$

(2)$P\{1\le X\le 2, -1\le Y\le 3\}$

$= F(2,3) - F(1,3) - F(2,-1) + F(1,-1)$

$= 1 - \mathrm{e}^{-\frac{1}{2}} - \mathrm{e}^{-2} + \mathrm{e}^{-\frac{5}{2}}$.（与答案不同）

(3) $P\{X > 2, Y > 3\} = 1 - P\{X \le 2\} - P\{Y \le 3\} + P\{X \le 2, Y \le 3\} = \mathrm{e}^{-\frac{5}{2}}$.

3.2 设二维随机变量 (X, Y) 的联合概率分布函数为

$$F(x,y) = A\left(B + \arctan\frac{x}{2}\right)\left(C + \arctan\frac{y}{2}\right), \quad (x,y) \in \mathrm{R}^2,$$

其中 A, B, C 为常数.

（1）确定常数 A, B, C;

（2）求 X , Y 的边缘分布函数;

（3）求 $P\{X > 2\}$.

解：(1)根据联合分布函数的性质得

$$\begin{cases} \lim\limits_{x\to+\infty, y\to+\infty} F(x,y) = 1, \\ \lim\limits_{x\to-\infty} F(x,y) = 0, \\ \lim\limits_{y\to-\infty} F(x,y) = 0. \end{cases} \quad 即 \begin{cases} A\left(B + \dfrac{\pi}{2}\right)\left(C + \dfrac{\pi}{2}\right) = 1, \\ A\left(B - \dfrac{\pi}{2}\right)\left(C + \arctan\dfrac{y}{2}\right) = 0, \\ A\left(B + \arctan\dfrac{x}{2}\right)\left(C - \dfrac{\pi}{2}\right) = 0. \end{cases}$$

解之得 $\begin{cases} A = \dfrac{1}{\pi^2}, \\ B = \dfrac{\pi}{2}, \\ C = \dfrac{\pi}{2}. \end{cases}$

(2) $F_X(x) = \lim\limits_{y\to+\infty} F(x,y)$

$$= \lim_{y\to+\infty} \frac{1}{\pi^2}\left(\frac{\pi}{2} + \arctan\frac{x}{2}\right)\left(\frac{\pi}{2} + \arctan\frac{y}{2}\right)$$

$$= \frac{1}{\pi}\left(\frac{\pi}{2} + \arctan\frac{x}{2}\right).$$

$F_Y(y) = \lim\limits_{x\to+\infty} F(x,y)$

$$= \lim_{x\to+\infty} \frac{1}{\pi^2}\left(\frac{\pi}{2} + \arctan\frac{x}{2}\right)\left(\frac{\pi}{2} + \arctan\frac{y}{2}\right)$$

$$= \frac{1}{\pi}\left(\frac{\pi}{2} + \arctan\frac{y}{2}\right).$$

(3) $P\{X > 2\} = 1 - P\{X \le 2\} = 1 - F_X(2) = \frac{1}{4}$.

3.3 将一均匀硬币连掷 3 次,用 X 表示在 3 次中出现正面的次数, Y 表示 3 次中出现正面次数与出现反面次数之差的绝对值,试求 (X, Y) 的联合概率分布律.

解：首先随机变量X的所有可能取值为0,1,2,3；Y的所有可能取值为1,3.

由于$P\{X=0,Y=1\}=0,\ \ P\{X=0,Y=3\}=P\{X=0\}=\frac{1}{8}$,

$P\{X=1,Y=1\}=P\{X=1\}=\frac{3}{8},\ \ P\{X=1,Y=3\}=0$,

$P\{X=2,Y=1\}=P\{X=2\}=\frac{3}{8},\ \ P\{X=2,Y=3\}=0$,

$P\{X=3,Y=1\}=0,\ \ P\{X=3,Y=3\}=P\{X=3\}=\frac{1}{8}$,

故(X,Y)的联合概率分布律为

Y \ X	0	1	2	3
1	0	$\frac{3}{8}$	$\frac{3}{8}$	0
3	$\frac{1}{8}$	0	0	$\frac{1}{8}$

3.4 设随机变量 X 服从参数为 1 的指数分布.定义随机变量 $Y_k(k=1,2)$ 如下:

$$Y_k = \begin{cases} 1, & 若X > k; \\ 0, & 若X \le k, \end{cases} \quad k = 1, 2.$$

求 (Y_1, Y_2) 的联合概率分布律.

解：因为$X \sim E(1)$,所以X的分布函数为

$$F_X(x) = \begin{cases} 1 - \mathrm{e}^{-x}, & x \ge 0; \\ 0, & x < 0. \end{cases}$$

由于$P\{Y_1=0,Y_2=0\}=P\{X\le 1, X\le 2\}=P\{X\le 1\}=F_X(1)=1-\mathrm{e}^{-1}$;

$P\{Y_1=0,Y_2=1\}=P\{X\le 1,X>2\}=0;$

$P\{Y_1=1,Y_2=0\}=P\{X>1,X\le 2\}=F_X(2)-F_X(1)=\mathrm{e}^{-1}-\mathrm{e}^{-2};$

$P\{Y_1=1,Y_2=1\}=P\{X>1,X>2\}=P\{X>2\}=1-F_X(2)=\mathrm{e}^{-2}.$

故(Y_1,Y_2)的联合概率分布律为

Y_2 \ Y_1	0	1
0	$1-\mathrm{e}^{-1}$	$\mathrm{e}^{-1}-\mathrm{e}^{-2}$
1	0	e^{-2}

3.5 设整数 X 随机地在 1,2,3 三个整数中任取一值,另一个整数 Y 随机地在 1 到 X 中任取一值.试求（X,Y）的联合概率分布律及边缘概率分布律.

解：依题意,随机变量X,Y的可能取值均为1,2,3，且$P\{X=i\}=\dfrac{1}{3},i=1,2,3.$

$$P\{X=i,Y=j\}=P\{X=i\}P\{Y=j|X=i\}=\begin{cases}\dfrac{1}{3}\cdot\dfrac{1}{i}, & j\le i;\\ 0, & j>i.\end{cases}$$

故(X,Y)的联合概率分布律为

Y \ X	1	2	3
1	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{9}$
2	0	$\frac{1}{6}$	$\frac{1}{9}$
3	0	0	$\frac{1}{9}$

又因为$P\{X=1\}=P\{X=1,Y=1\}+P\{X=1,Y=2\}+P\{X=1,Y=3\}=\dfrac{1}{3},$

$P\{X=2\}=P\{X=2,Y=1\}+P\{X=2,Y=2\}+P\{X=2,Y=3\}=\dfrac{1}{3},$

$P\{X=3\}=P\{X=3,Y=1\}+P\{X=3,Y=2\}+P\{X=3,Y=3\}=\frac{1}{3}$,

所以X的边缘概率分布律为

X	1	2	3
P_X	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

又因为

$P\{Y=1\}=P\{X=1,Y=1\}+P\{X=2,Y=1\}+P\{X=3,Y=1\}=\frac{11}{18}$,

$P\{Y=2\}=P\{X=1,Y=2\}+P\{X=2,Y=2\}+P\{X=3,Y=2\}=\frac{5}{18}$,

$P\{Y=3\}=P\{X=1,Y=3\}+P\{X=2,Y=3\}+P\{X=3,Y=3\}=\frac{1}{9}$,

所以Y的边缘概率分布律为

X	1	2	3
P_X	$\frac{11}{18}$	$\frac{5}{18}$	$\frac{1}{9}$

3.6 现有 5 件产品,其中有 2 件次品.现从中不放回地依次取两件,分别用 X, Y 表示第一次和第二次取到的次品数,求 (X,Y) 的联合概率分布律和边缘概率分布律;若抽取方式为有放回的,再求 (X,Y) 的联合概率分布律和边缘概率分布律.

解：(1)抽取方式为无放回

因为$P\{X=0,Y=0\}=\frac{A_3^2}{A_5^2}=\frac{3}{10}$;

$P\{X=0,Y=1\}=\frac{A_3^1A_2^1}{A_5^2}=\frac{3}{10}$;

$P\{X=1,Y=0\}=\frac{A_2^1A_3^1}{A_5^2}=\frac{3}{10}$;

$P\{X=1,Y=1\}=\frac{A_2^2}{A_5^2}=\frac{1}{10}$.

所以(X,Y)的联合概率分布律为

$X \backslash Y$	0	1
0	$\frac{3}{10}$	$\frac{3}{10}$
1	$\frac{3}{10}$	$\frac{1}{10}$

又因为$P\{X=0\}=P\{X=0,Y=0\}+P\{X=0,Y=1\}=\frac{3}{5}$,

$P\{X=1\}=P\{X=1,Y=0\}+P\{X=1,Y=1\}=\frac{2}{5}$.

所以X的边缘概率分布律为

X	0	1
P_X	$\frac{3}{5}$	$\frac{2}{5}$

又因为$P\{Y=0\}=P\{X=0,Y=0\}+P\{X=1,Y=0\}=\frac{3}{5}$,

$P\{Y=1\}=P\{X=0,Y=1\}+P\{X=1,Y=1\}=\frac{2}{5}$.

所以Y的边缘概率分布律为

Y	0	1
P_Y	$\frac{3}{5}$	$\frac{2}{5}$

(2)抽取方式为有放回

因为$P\{X=0,Y=0\}=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25}$;

$P\{X=0,Y=1\}=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25}$;

$P\{X=1,Y=0\}=\frac{2}{5}\cdot\frac{3}{5}=\frac{6}{25}$;

$P\{X=1,Y=1\}=\frac{2}{5}\cdot\frac{2}{5}=\frac{4}{25}$.

所以(X,Y)的联合概率分布律为

$X \backslash Y$	0	1
0	$\frac{9}{25}$	$\frac{6}{25}$
1	$\frac{6}{25}$	$\frac{4}{25}$

又因为$P\{X=0\}=P\{X=0,Y=0\}+P\{X=0,Y=1\}=\frac{3}{5}$,

$P\{X=1\}=P\{X=1,Y=0\}+P\{X=1,Y=1\}=\frac{2}{5}$.

所以X的边缘概率分布律为

X	0	1
P_X	$\frac{3}{5}$	$\frac{2}{5}$

又因为$P\{Y=0\}=P\{X=0,Y=0\}+P\{X=1,Y=0\}=\frac{3}{5}$,

$P\{Y=1\}=P\{X=0,Y=1\}+P\{X=1,Y=1\}=\frac{2}{5}$.

所以Y的边缘概率分布律为

Y	0	1
P_Y	$\frac{3}{5}$	$\frac{2}{5}$

3.7 设二维离散型随机变量 (X,Y) 的联合概率分布律为

Y \ X	1	2	3
1	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$
2	$\frac{1}{3}$	α	β

问 α,β 为何值时,X 与 Y 相互独立?

解：要使X 与 Y 相互独立,当且仅当

$$\begin{cases} P\{X=2,Y=2\}=P\{X=2\}P\{Y=2\}; \\ P\{X=3,Y=2\}=P\{X=3\}P\{Y=2\}, \end{cases} \quad 即 \begin{cases} \alpha=\left(\frac{1}{9}+\alpha\right)\left(\frac{1}{3}+\alpha+\beta\right); \\ \beta=\left(\frac{1}{18}+\beta\right)\left(\frac{1}{3}+\alpha+\beta\right), \end{cases}$$

解之得 $\begin{cases} \alpha=\frac{2}{9}, \\ \beta=\frac{1}{9}. \end{cases}$

3.8 设随机变量$X\sim B(1,0.2)$. 在$X=0$及$X=1$的条件下随机变量Y的条件概率分布律为

Y	1	2	3
$P_{Y\|X=0}$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{4}$
$P_{Y\|X=1}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

求(1)(X,Y)的联合概率分布律;

(2) 随机变量Y的边缘概率分布律;

(3) $P\{Y<3\}$;

(4) 在$Y<3$的条件下,随机变量X的条件概率分布律.

解：(1) 因为$P\{X=0,Y=1\}=P\{X=0\}P\{Y=1|X=0\}=0.8\cdot\frac{1}{8}=\frac{1}{10}$;

$P\{X=0,Y=2\}=P\{X=0\}P\{Y=2|X=0\}=0.8\cdot\frac{5}{8}=\frac{1}{2}$;

$P\{X=0,Y=3\}=P\{X=0\}P\{Y=3|X=0\}=0.8\cdot\frac{1}{4}=\frac{1}{5}$;

$P\{X=1,Y=1\}=P\{X=1\}P\{Y=1|X=1\}=0.2\cdot\frac{1}{2}=\frac{1}{10}$;

$P\{X=1,Y=2\}=P\{X=1\}P\{Y=2|X=1\}=0.2\cdot\frac{1}{3}=\frac{1}{15}$;

$P\{X=1,Y=3\}=P\{X=1\}P\{Y=3|X=1\}=0.2\cdot\frac{1}{6}=\frac{1}{30}$;

X \ Y	1	2	3
0	$\frac{1}{10}$	$\frac{1}{2}$	$\frac{1}{5}$
1	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{30}$

(2) $P\{Y=1\}=P\{X=0,Y=1\}+P\{X=1,Y=1\}=\dfrac{1}{5}$;

$P\{Y=2\}=P\{X=0,Y=2\}+P\{X=1,Y=2\}=\dfrac{17}{30}$;

$P\{Y=3\}=P\{X=0,Y=3\}+P\{X=1,Y=3\}=\dfrac{7}{30}$;

所以Y的边缘概率分布律为

Y	1	2	3
P_Y	$\frac{2}{10}$	$\frac{17}{30}$	$\frac{7}{30}$

(3)$P\{Y<3\}=P\{Y=1\}+P\{Y=2\}=\dfrac{23}{30}$.

(4)$P\{X=0|Y<3\}=\dfrac{P\{X=0,Y<3\}}{P\{Y<3\}}$

$=\dfrac{P\{X=0,Y=1\}+P\{X=0,Y=2\}}{P\{Y<3\}}=\dfrac{18}{23}$;

$P\{X=1|Y<3\}=\dfrac{P\{X=1,Y<3\}}{P\{Y<3\}}$

$=\dfrac{P\{X=1,Y=1\}+P\{X=1,Y=2\}}{P\{Y<3\}}=\dfrac{5}{23}$.

故在$Y<3$的条件下,随机变量X的条件概率分布律为

X	0	1
$P_{X\|Y<3}$	$\frac{18}{23}$	$\frac{5}{23}$

3.9 设二维随机变量 (X,Y) 的联合概率密度函数为

$$f(x,y)=\begin{cases} Ay\mathrm{e}^{-x}, & x>0,0<y<1; \\ 0, & \text{其他}. \end{cases}$$

（1）求常数 A;

（2）求 (X,Y) 的联合概率分布函数 $F(x,y)$.

解：(1)因为 $\iint\limits_{R^2} f(x,y)\mathrm{d}x\mathrm{d}y = 1$,　　所以 $\int_0^{+\infty}\mathrm{d}x\int_0^1 Ay\mathrm{e}^{-x}\mathrm{d}y = 1$,

解之得$A = 2$.

(2)当$X < 0$或$y < 0$时，$F(x,y) = P\{X \le x, Y \le y\} = 0$;

当$x \ge 0, 0 \le y < 1$时，$F(x,y) = \int_0^x \mathrm{d}x\int_0^y 2y\mathrm{e}^{-x}\mathrm{d}y = y^2\left(1-\mathrm{e}^{-x}\right)$;

当$x \ge 0, y \ge 1$时，$F(x,y) = \int_0^x \mathrm{d}x\int_0^1 2y\mathrm{e}^{-x}\mathrm{d}y = 1-\mathrm{e}^{-x}$;

综上得(X,Y)的联合分布函数为

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

3.10 设二维随机变量 (X,Y) 的联合概率密度函数为

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

求（1） $P\{(X,Y) \in D\}$,其中 $D = \{(x,y) : x+y \le 1\}$;

（2） $P\{X > 1\}$;

（3） $P\{X^2 \ge Y\}$.

解：(1) $P\{(X,Y) \in D\} = \iint\limits_D f(x,y)\mathrm{d}x\mathrm{d}y = \int_0^1 \mathrm{d}x\int_0^{1-x}\mathrm{e}^{-(x+y)}\mathrm{d}y = 1-2\mathrm{e}^{-1}$.

(2) $P\{X > 1\} = \int_1^{+\infty}\mathrm{d}x\int_0^{+\infty}\mathrm{e}^{-(x+y)}\mathrm{d}y = \mathrm{e}^{-1}$.

(3) $P\{X^2 \ge Y\} = \int_0^{+\infty}\mathrm{d}x\int_0^{x^2}\mathrm{e}^{-(x+y)}\mathrm{d}y = 1-\mathrm{e}^{\frac{1}{4}}\sqrt{\pi}\left(1-\Phi\left(\frac{\sqrt{2}}{2}\right)\right)$.

3.11 设二维随机变量 (X,Y) 的联合分布函数为

$$F(x,y)=\begin{cases} x^2(1-\mathrm{e}^{-y}), & 0\le x\le 1, y\ge 0;\\ 1-\mathrm{e}^{-y}, & x\ge 1, y>0;\\ 0, & \text{其他}. \end{cases}$$

（1）求边缘分布函数 $F_X(x), F_Y(y)$;

（2）求 (X,Y) 的联合概率密度函数 $f(x,y)$.

解：(1) $F_X(x)=\lim\limits_{y\to+\infty}F(x,y)=\begin{cases} \lim\limits_{y\to+\infty}x^2(1-\mathrm{e}^{-y}), & 0\le x\le 1;\\ \lim\limits_{y\to+\infty}(1-\mathrm{e}^{-y}), & x\ge 1;\\ 0, & x<0. \end{cases}$

$$=\begin{cases} 0, & x<0;\\ x^2, & 0\le x<1;\\ 1, & x\ge 1. \end{cases}$$

$F_Y(y)=\lim\limits_{x\to+\infty}F(x,y)=\begin{cases} \lim\limits_{x\to+\infty}(1-\mathrm{e}^{-y}), & y>0;\\ 0, & y\le 0. \end{cases}$

$$=\begin{cases} 0, & y<0;\\ 1-\mathrm{e}^{-y}, & y\ge 0. \end{cases}$$

(2)当$0<x<1, y>0$时，$f(x,y)=\dfrac{\partial\partial F(x,Y)}{\partial x\partial y}=2x\mathrm{e}^{-y}$;

在其他点处，$f(x,y)=0$.

所以 (X,Y) 的联合概率密度函数为

$$f(x,y)=\begin{cases}2xe^{-y}, & 0\le x\le 1, y>0;\\ 0, & 其他.\end{cases}$$

3.12 设二维随机变量 (X,Y) 的联合概率密度函数为

$$f(x,y)=\begin{cases}x, & 0\le x\le 2, \max(0,x-1)\le y\le \min(1,x);\\ 0, & 其他.\end{cases}$$

求X,Y的边缘概率密度函数.

解：根据$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$，$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$ 知，

当$x<0$或$x\ge 2$时，$f_X(x)=0$;

当$0\le x<1$时，$f_X(x)=\int_0^x x\mathrm{d}y=x^2$;

当$1\le x<2$时，$f_X(x)=\int_{x-1}^1 x\mathrm{d}y=x(2-x)$;

综上得X的边缘概率密度函数为

$$f_X(x)=\begin{cases}x^2, & 0\le x<1;\\ x(2-x), & 1\le x<2;\\ 0, & 其他.\end{cases}$$

当$y<0$或$y\ge 1$时，$f_Y(y)=0$;

当$0\le y<1$时，$f_Y(y)=\int_y^{y+1}x\mathrm{d}x=y+\dfrac{1}{2}$.

综上得Y的边缘概率密度函数为

$$f_Y(y)=\begin{cases}y+\dfrac{1}{2}, & 0\le y<1;\\ 0, & 其他.\end{cases}$$

3.13 设二维随机变量 (X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases} ce^{-2(x+y)}, & x>0,y>0; \\ 0, & 其他. \end{cases}$$

求（1）常数 c;（2）(X,Y)的联合概率分布函数 $F(x,y)$;（3）X的边缘概率分布函数 $F_X(x)$ 及 Y的边缘概率密度函数 $f_Y(y)$.

解：(1)因为 $\iint\limits_{R^2} f(x,y)\mathrm{d}x\mathrm{d}y=1,$ 所以 $\int_0^{+\infty}\mathrm{d}x\int_0^{+\infty}ce^{-2(x+y)}\mathrm{d}y=1,$

解之得$c=4.$

(2)当$x<0$或$y<0$时，$F(x,y)=\int_{-\infty}^{x}\mathrm{d}x\int_{-\infty}^{y}f(x,y)\mathrm{d}y=0;$

当$x\geq 0,y\geq 0$时，$F(x,y)=\int_0^x\mathrm{d}x\int_0^y 4e^{-2(x+y)}\mathrm{d}y=\left(1-e^{-2x}\right)\left(1-e^{-2y}\right);$

综上得(X,Y)的联合概率分布函数为

$$F(x,y)=\begin{cases} \left(1-e^{-2x}\right)\left(1-e^{-2y}\right), & x\geq 0,y\geq 0; \\ 0, & x<0或y<0. \end{cases}$$

(3)$F_X(x)=\lim\limits_{y\to+\infty}F(x,y)$

$$=\begin{cases} \lim\limits_{y\to+\infty}\left(1-e^{-2x}\right)\left(1-e^{-2y}\right), & x\geq 0; \\ 0, & x<0. \end{cases}=\begin{cases} 1-e^{-2x}, & x\geq 0; \\ 0, & x<0. \end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases} \int_0^{+\infty}4e^{-2(x+y)}\mathrm{d}x, & y\geq 0; \\ 0, & y<0. \end{cases}=\begin{cases} 2e^{-2y}, & y\geq 0; \\ 0, & y<0. \end{cases}$$

3.14 设二维随机变量 (X,Y) 的联合概率密度函数为

$$f(x,y)=\begin{cases} \dfrac{2e^{-y}}{x^3}, & x>1,y>0; \\ 0, & 其他, \end{cases}$$

证明 X 与 Y 相互独立.

证明：因为$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_0^{+\infty}\dfrac{2\mathrm{e}^{-y}}{x^3}\mathrm{d}y, & x>1;\\ 0, & x\le 1,\end{cases}=\begin{cases}\dfrac{2}{x^3}, & x>1;\\ 0, & x\le 1.\end{cases}$

$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_1^{+\infty}\dfrac{2\mathrm{e}^{-y}}{x^3}\mathrm{d}x, & y>0;\\ 0, & y\le 0,\end{cases}=\begin{cases}\mathrm{e}^{-y}, & y>0;\\ 0, & y\le 0.\end{cases}$

显然$f_X(x)f_Y(y)=f(x,y)$,所以随机变量X 与 Y 相互独立.

3.15 设二维连续型随机变量 (X,Y) 的联合概率密度函数为

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

求(1) 常数 c; (2) 判断 X 与 Y 是否独立?

解：(1)因为 $\iint\limits_{R^2}f(x,y)\mathrm{d}x\mathrm{d}y=1,$ 所以 $\int_0^1\mathrm{d}x\int_0^x cy(1-x)\mathrm{d}y=1,$

解之得$c=24.$

(2)$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_0^x 24y(1-x)\mathrm{d}y, & 0<x<1;\\ 0, & \text{其他},\end{cases}=\begin{cases}12x^2(1-x), & 0<x<1;\\ 0, & \text{其他}.\end{cases}$

$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_y^1 24y(1-x)\mathrm{d}x, & 0<y<1;\\ 0, & \text{其他},\end{cases}=\begin{cases}12y(1-y)^2, & 0<y<1;\\ 0, & \text{其他}.\end{cases}$

显然$f_X(x)f_Y(y)\ne f(x,y)$,所以随机变量X 与 Y 不独立.

3.16 设二维随机变量 (X,Y) 在区域 G 上服从均匀分布,其中G 由直线 $x=0,y=0$ 和 $x-y=1$ 围成.求

（1） X,Y的边缘概率密度函数 $f_X(x),f_Y(y)$;

（2）条件概率密度函数 $f_{X|Y}(x|y),f_{Y|X}(y|x)$.

解：因为(X,Y) 在区域 G 上服从均匀分布，所以(X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}2, & (x,y)\in G;\\ 0, & (x,y)\notin G.\end{cases}$$

$$(1)f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{x-1}^{0}2\mathrm{d}y, & 0<x<1;\\ 0, & \text{其他},\end{cases}=\begin{cases}2(1-x), & 0<x<1;\\ 0, & \text{其他}.\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_{0}^{y+1}2\mathrm{d}x, & -1<y<0;\\ 0, & \text{其他},\end{cases}=\begin{cases}2(1+y), & -1<y<0;\\ 0, & \text{其他}.\end{cases}$$

(2)当$-1<y<0$时，$f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}=\begin{cases}\dfrac{1}{1+y}, & 0<x<1+y;\\ 0, & \text{其他}.\end{cases}$

当$y\notin(-1,0)$时，$f_{X|Y}(x|y)$无意义。

当$0<x<1$时，$f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}=\begin{cases}\dfrac{1}{1-x}, & x-1<y<0;\\ 0, & \text{其他}.\end{cases}$

当$x\notin(0,1)$时，$f_{Y|X}(y|x)$无意义。

3.17 设随机变量 Y 的概率密度函数为

$$f_Y(y)=\begin{cases}5y^4, & 0<y<1;\\ 0, & \text{其他}.\end{cases}$$

且在条件 $Y=y(0<y<1)$ 下, X 的条件概率密度函数为

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

求 随机变量X 的边缘概率密度函数 $f_X(x)$ 及 $P\left\{X\le\dfrac{1}{2}\right\}$.

解：依题意得(X,Y)的联合概率密度函数为

$$f(x,y)=f_Y(y)f_{X|Y}(x|y)=\begin{cases}15x^2y, & 0<x<y, 0<y<1;\\ 0, & \text{其他}.\end{cases}$$

所以$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_x^1 15x^2y\mathrm{d}y, & 0<x<1;\\ 0, & \text{其他},\end{cases}=\begin{cases}\frac{15}{2}x^2(1-x^2), & 0<x<1;\\ 0, & \text{其他}.\end{cases}$

$P\left\{X\le\frac{1}{2}\right\}=\int_0^{\frac{1}{2}}\frac{15}{2}x^2(1-x^2)\mathbf{d}x=\frac{17}{64}.$

3.18 设二维随机变量(X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}cx^2y, & x^2\le y\le 1;\\ 0, & \text{其他}.\end{cases}$$

(1) 确定常数c;

(2) 判断X与Y是否独立?

(3) 求条件概率密度函数$f_{Y|X}(y|x)$;

(4) 求概率$P\left\{Y\ge\frac{3}{4}|X=\frac{1}{2}\right\}$, $P\left\{Y\le\frac{1}{4}|X=\frac{1}{2}\right\}$.

解：(1)因为$\iint\limits_{R^2}f(x,y)\mathrm{d}x\mathrm{d}y=1,$ 所以 $\int_{-1}^1\mathrm{d}x\int_{x^2}^1 cx^2y\mathrm{d}y=1,$

解之得$c=\frac{21}{4}.$

(2) 因为$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{x^2}^1\frac{21}{4}x^2y\mathrm{d}y, & -1<x<1;\\ 0, & \text{其他},\end{cases}=\begin{cases}\frac{21}{8}x^2(1-x^4), & -1<x<1;\\ 0, & \text{其他}.\end{cases}$

$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_{-\sqrt{y}}^{\sqrt{y}}\frac{21}{4}x^2y\mathrm{d}x, & 0<y<1;\\ 0, & \text{其他},\end{cases}=\begin{cases}\frac{7}{2}y^{\frac{5}{2}}, & 0<y<1;\\ 0, & \text{其他}.\end{cases}$

显然$f_X(x)f_Y(y)\ne f(x,y)$,所以随机变量X 与 Y 不独立.

(3)当$-1<x<1$时，$f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}=\begin{cases}\dfrac{2y}{1-x^4}, & x^2<y<1;\\ 0, & \text{其他}.\end{cases}$

当$x\notin(-1,1)$时，$f_{Y|X}(y|x)$无意义。

(4)$P\left\{Y\geq\frac{3}{4}\middle|X=\frac{1}{2}\right\}=\displaystyle\int_{\frac{3}{4}}^{1}\frac{2y}{1-0.5^4}\mathrm{d}y=\frac{7}{15}.$

$P\left\{Y\leq\frac{1}{4}\middle|X=\frac{1}{2}\right\}=0.$

3.19 设随机变量$X\sim U[0,1]$, 当给定$X=x\in(0,1)$时, 随机变量Y的条件概率密度函数为

$$f_{Y|X}(y|x)=\begin{cases}x, & 0\leq y\leq\dfrac{1}{x};\\ 0, & \text{其他},\end{cases}$$

求(1) (X,Y)的联合概率密度函数$f(x,y)$;

(2) Y的边缘概率密度函数$f_Y(y)$;

(3) 概率$P\{X>Y\}$.

解：(1)因为$X\sim U(0,1)$,所以X的概率密度函数为

$f_X(x)=\begin{cases}1, & x\in[0,1];\\ 0, & x\notin[0,1].\end{cases}$

所以 (X,Y)的联合概率密度函数为

$f(x,y)=f_X(x)f_{Y|X}(y|x)=\begin{cases}x, & 0<x<1,0\leq y\leq\dfrac{1}{x};\\ 0, & \text{其他},\end{cases}$

(2)$f_Y(y)=\displaystyle\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}0, & y<0;\\ \displaystyle\int_0^1 x\mathrm{d}x, & 0\leq y<1;\\ \displaystyle\int_0^{\frac{1}{y}}x\mathrm{d}x, & y\geq1;\end{cases}=\begin{cases}0, & y<0;\\ \dfrac{1}{2}, & 0\leq y<1;\\ \dfrac{1}{2y^2}, & y\geq1.\end{cases}$

(3) $P\{X>Y\}=\int_0^1\mathrm{d}x\int_0^x x\mathrm{d}y=\dfrac{1}{3}.$

3.20 设 X 与 Y 分别表示甲、乙两个元件的寿命 (单位:kh), 概率密度函数分别为

$$f_X(x)=\begin{cases}\mathrm{e}^{-x}, & x\geq 0;\\ 0, & x<0.\end{cases}\qquad f_Y(y)=\begin{cases}2\mathrm{e}^{-2y}, & y\geq 0;\\ 0, & y<0.\end{cases}$$

且 X 与 Y 相互独立. 若两个元件同时开始使用,求甲比乙先坏的概率.

解：因为 X 与 Y 相互独立，所以(X,Y)的联合概率密度函数为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}2\mathrm{e}^{-(x+2y)}, & x\geq 0, y\geq 0;\\ 0, & \text{其他}.\end{cases}$$

据题意，所求事件的概率为

$$P\{X<Y\}=\int_0^{+\infty}\mathrm{d}y\int_0^{y}2\mathrm{e}^{-(x+2y)}\mathrm{d}x=\frac{1}{3}.$$

3.21 若随机变量X,Y相互独立，且服从相同的分布，问$X+Y$与$2X$的分布相同吗？若相同，请说明理由；若不同，请举例说明。

解：不同。例如$X\sim N(0,1),Y\sim N(0,1)$,且$X$与$Y$相互独立，则$X+Y\sim N(0,2)$,而$2X\sim N(0,4)$.

3.22 若 X_1 与 X_2 相互独立,且服从同一分布:

$$P\{X_i=-1\}=P\{X_i=1\}=\frac{1}{2},\quad i=1,2.$$

令 $Y=2X_1, Z=X_1+X_2$,求（1） Y 的概率分布律; （2） Z 的概率分布律.

解：(1)因为Y的所有可能取值为-2,2.且

$P\{Y=-2\}=P\{2X_1=-2\}=P\{X_1=-1\}=0.5;$

$P\{Y=2\}=P\{2X_1=2\}=P\{X_1=1\}=0.5.$

所以Y的概率分布律为

Y	-2	2
P_Y	0.5	0.5

(2) 因为Z的所有可能取值为-2,0,2.且

$P\{Z=-2\}=P\{X_1+X_2=-2\}$

$=P\{X_1=-1,X_2=-1\}=P\{X_1=-1\}P\{X_2=-1\}=0.25;$

$P\{Z=0\}=P\{X_1+X_2=0\}$

$=P\{X_1=-1,X_2=1\}+\{X_1=1,X_2=-1\}=0.5;$

$P\{Z=2\}=P\{X_1+X_2=2\}$

$=P\{X_1=1,X_2=1\}=P\{X_1=1\}P\{X_2=1\}=0.25.$

所以Z的概率分布律为

Y	-2	0	2
P_Y	0.25	0.5	0.25

3.23 设二维随机变量(X,Y)的联合概率分布律为

Y \ X	0	1	2	3	4	5
0	0.00	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

令$U=\max\{X,2-Y\},V=\min\{X,Y\},W=X+Y$,求

(1)$P\{X=2|Y=2\},P\{Y=3|X=0\};$

(2) U的概率分布律;

(3) V的概率分布律;

(4) W的概率分布律.

解：(1) $P\{X=2|Y=2\}=\dfrac{P\{X=2,Y=2\}}{P\{Y=2\}}=\dfrac{1}{5}$;

$P\{Y=3|X=0\}=\dfrac{P\{X=0,Y=3\}}{P\{X=0\}}=\dfrac{1}{3}$;

(2)因为U的所有可能取值为0,1,2,3,4,5.且

$$P\{U=0\}=P\{\max\{X,2-Y\}=0\}$$

$$=P\{X=0,Y=2\}+P\{X=0,Y=3\}=0.02;$$

$$P\{U=1\}=P\{\max\{X,2-Y\}=1\}$$

$$=P\{X=0,Y=1\}+\sum_{i=1}^{3}P\{X=1,Y=i\}=0.08;$$

$$P\{U=2\}=P\{\max\{X,2-Y\}=2\}$$

$$=P\{X=0,Y=0\}+P\{X=1,Y=0\}+\sum_{i=0}^{3}P\{X=2,Y=i\}=0.17;$$

$$P\{U=3\}=P\{\max\{X,2-Y\}=3\}=\sum_{i=0}^{3}P\{X=1,Y=i\}=0.21;$$

$$P\{U=4\}=P\{\max\{X,2-Y\}=4\}=\sum_{i=0}^{3}P\{X=1,Y=i\}=0.24;$$

$$P\{U=5\}=P\{\max\{X,2-Y\}=5\}=\sum_{i=0}^{3}P\{X=1,Y=i\}=0.28.$$

所以U的概率分布律为

U	0	1	2	3	4	5
P_U	0.02	0.08	0.17	0.21	0.24	0.28

(3) 因为V的所有可能取值为0,1,2,3.且

$$P\{V=0\}=P\{\min\{X,Y\}=0\}$$

$$=\sum_{i=0}^{5}P\{X=i,Y=0\}+\sum_{j=1}^{3}P\{X=0,Y=j\}=0.28;$$

$$P\{V=1\}=P\{\min\{X,Y\}=1\}$$

$$= \sum_{i=1}^{5} P\{X = i, Y = 1\} + \sum_{j=2}^{3} P\{X = 1, Y = j\} = 0.30;$$

$$P\{V = 2\} = P\{\min\{X, Y\} = 2\}$$

$$= \sum_{i=2}^{5} P\{X = i, Y = 2\} + P\{X = 2, Y = 3\} = 0.25;$$

$$P\{V = 3\} = P\{\min\{X, Y\} = 1\} = \sum_{i=3}^{5} P\{X = i, Y = 3\} = 0.17;$$

所以V的概率分布律为

V	0	1	2	3
P_V	0.28	0.30	0.25	0.17

(4) 因为W的所有可能取值为1,2,3,4,5,6,7,8.且

$$P\{W = 2\} = P\{X + Y = 2\}$$

$$= P\{X = 0, Y = 2\} + P\{X = 1, Y = 1\} + P\{X = 2, Y = 0\} = 0.06;$$

同理可求得其他各点的概率，故W的概率分布律为

W	1	2	3	4	5	6	7	8
P_W	0.02	0.06	0.13	0.19	0.24	0.19	0.12	0.05

3.24 设X和Y是相互独立且服从同一分布的两个随机变量. 已知X的概率分布律为$P\{X = i\} = 1/3, i = 1, 2, 3$. 又设$U = \max\{X, Y\}, V = \min\{X, Y\}$. 求(1)$U$和$V$的联合概率分布律; (2)判断$U, V$是否独立?

解：(1) 据题意得，U, V的所有可能取值均为1,2,3.且

$$P\{U = 1, V = 1\} = P\{\max\{X, Y\} = 1, \min\{X, Y\} = 1\}$$

$$= P\{X = 1, Y = 1\} = P\{X = 1\}P\{Y = 1\} = \frac{1}{9};$$

$$P\{U = 1, V = 2\} = P\{\max\{X, Y\} = 1, \min\{X, Y\} = 2\} = 0;$$

$$P\{U = 1, V = 3\} = P\{\max\{X, Y\} = 1, \min\{X, Y\} = 3\} = 0;$$

$$P\{U = 2, V = 1\} = P\{\max\{X, Y\} = 2, \min\{X, Y\} = 1\}$$

$$= P\{X = 2, Y = 1\} + P\{X = 1, Y = 2\} = \frac{2}{9};$$

$P\{U=2,V=2\}=P\{\max\{X,Y\}=2,\min\{X,Y\}=2\}$

$=P\{X=2,Y=2\}=\dfrac{1}{9};$

$P\{U=2,V=3\}=P\{\max\{X,Y\}=2,\min\{X,Y\}=3\}=0;$

$P\{U=3,V=1\}=P\{\max\{X,Y\}=3,\min\{X,Y\}=1\}$

$=P\{X=3,Y=1\}+P\{X=1,Y=3\}=\dfrac{2}{9};$

$P\{U=3,V=2\}=P\{\max\{X,Y\}=3,\min\{X,Y\}=2\}$

$=P\{X=3,Y=2\}+P\{X=2,Y=3\}=\dfrac{2}{9};$

$P\{U=3,V=3\}=P\{\max\{X,Y\}=3,\min\{X,Y\}$

$=3\}=P\{X=3,Y=3\}=\dfrac{1}{9};$

所以(U,V)的联合概率分布律为

V \ U	1	2	3
1	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$
2	0	$\frac{1}{9}$	$\frac{2}{9}$
3	0	0	$\frac{1}{9}$

(2) 因为$P\{U=1,V=2\}=0$, $P\{U=1\}=\dfrac{1}{9}$, $P\{V=2\}=\dfrac{3}{9};$

$P\{U=1,V=2\}\neq P\{U=1\}\cdot P\{V=2\},$

所以，随机变量U与V不独立。

3.25 设随机变量 $X\sim P(\lambda_1)$,$Y\sim P(\lambda_2)$, 且 X 与 Y 相互独立,令 $Z=X+Y$,证明随机变量 $Z\sim P(\lambda_1+\lambda_2)$.

解：因为$X\sim P(\lambda_1)$,$Y\sim P(\lambda_2)$, 且 X 与 Y 相互独立,所以

$P\{X=i\}=\dfrac{\lambda_1^i}{i!}\mathrm{e}^{-\lambda_1},\ i=0,1,2,\cdots$

$P\{Y=j\}=\dfrac{\lambda_2^j}{j!}\mathrm{e}^{-\lambda_2},\ j=0,1,2,\cdots$

据题意得随机变量Z的所有可能取值为0,1,2,$\cdots$.且

$$
\begin{aligned}
P\{Z=k\}&=P\{X+Y=k\}=\sum_{i=0}^{k}P\{X=i,Y=k-i\}\\
&=\sum_{i=0}^{k}P\{X=i\}P\{Y=k-i\}\\
&=\sum_{i=0}^{k}\left(\frac{\lambda_1^i}{i!}\mathrm{e}^{-\lambda_1}\cdot\frac{\lambda_2^{k-i}}{(k-i)!}\mathrm{e}^{-\lambda_2}\right)\\
&=\frac{1}{k!}\mathrm{e}^{-(\lambda_1+\lambda_2}\sum_{i=0}^{k}\frac{k!}{(k-i)!i!}\lambda_1^i\lambda_2^{k-i}\\
&=\frac{1}{k!}\mathrm{e}^{-(\lambda_1+\lambda_2}\sum_{i=0}^{k}C_k^i\lambda_1^i\lambda_2^{k-i}\\
&=\frac{(\lambda_1+\lambda_2)^k}{k!}\mathrm{e}^{-(\lambda_1+\lambda_2},
\end{aligned}
$$

故随机变量$Z\sim P(\lambda_1+\lambda_2)$.

3.26 设某种商品一周的需求量是随机变量,概率密度函数为

$$
f(t)=\begin{cases} t\mathrm{e}^{-t}, & t\geq 0;\\ 0, & t<0.\end{cases}
$$

并设各周的需求量相互独立. 试求

（1）两周需求量的概率密度函数;

（2）三周需求量的概率密度函数.

解：设随机变量X,Y,Z相互独立，概率密度函数均为

$$
f(t)=\begin{cases} t\mathrm{e}^{-t}, & t\geq 0;\\ 0, & t<0.\end{cases}
$$

(1)令$U=X+Y$,则U的概率密度函数即为两周需求量的概率密度函数,即

$f_U(u)=\int_{-\infty}^{+\infty}f_X(x)f_Y(u-x)\mathrm{d}x$

$$=\begin{cases} 0, & u<0; \\ \int_0^u x\mathrm{e}^{-x}\cdot(u-x)\mathrm{e}^{-(u-x)}\mathrm{d}x, & u\geq 0, \end{cases}$$

$$=\begin{cases} 0, & u<0; \\ \frac{1}{6}u^3\mathrm{e}^{-u}, & u\geq 0. \end{cases}$$

(2)令$V=U+Z$,则V的概率密度函数即为三周需求量的概率密度函数,即

$$f_V(v)=\int_{-\infty}^{+\infty} f_U(u)f_Z(v-u)\mathrm{d}u$$

$$=\begin{cases} 0, & v<0; \\ \int_0^v \frac{1}{6}u^3\mathrm{e}^{-u}\cdot(v-u)\mathrm{e}^{-(v-u)}\mathrm{d}u, & v\geq 0, \end{cases}$$

$$=\begin{cases} 0, & v<0; \\ \frac{1}{120}v^5\mathrm{e}^{-v}, & v\geq 0. \end{cases}$$

3.27 设 (X,Y) 的联合概率密度函数为

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

令 $Z=2X-3Y$,求Z 的分布函数 $F_Z(z)$.

解：因为$F_Z(z)=P\{Z\leq z\}=P\{2x-3Y\leq z\}$

当$z<-6$时，$F_Z(z)=0$;

当$-6\leq z<-2$时，$F_Z(z)=\int_0^{3+\frac{z}{2}}\mathrm{d}x\int_{\frac{2x-z}{3}}^2\frac{1}{4}\mathrm{d}y=\frac{1}{48}(z+6)^2$;

当$-2\leq z<0$时，$F_Z(z)=\int_0^2\mathrm{d}x\int_{\frac{2x-z}{3}}^2\frac{1}{4}\mathrm{d}y=\frac{z+4}{6}$;

当$0\leq z<4$时，$F_Z(z)=1-\int_{\frac{z}{2}}^2\mathrm{d}x\int_0^{\frac{2x-z}{3}}\frac{1}{4}\mathrm{d}y=1-\frac{(z-4)^2}{48}$;

当$z\geq 4$时，$F_Z(z)=1$.

综上得随机变量Z的分布函数为

$$F_Z(z)=\begin{cases}0, & z<-6;\\ \dfrac{(z+6)^2}{48}, & -6\le z<-2;\\ \dfrac{z+4}{6}, & -2\le z<0;\\ 1-\dfrac{(4-z)^2}{48}, & 0\le z<4;\\ 1, & z\ge 4.\end{cases}$$

3.28 设 随机变量X 与 Y 相互独立,概率密度函数分别为

$$f_X(x)=\begin{cases}\dfrac{1}{2}, & 0\le x\le 2;\\ 0, & 其他.\end{cases}\qquad f_Y(y)=\begin{cases}\mathrm{e}^{-y}, & y\ge 0;\\ 0, & y<0.\end{cases}$$

令 $Z=X+Y$,求随机变量 Z 的概率密度函数及 $P\{X+Y>1\}$.

解：$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x$

$$=\begin{cases}0, & z<0;\\ \int_0^z \dfrac{1}{2}\mathrm{e}^{-(z-x)}\mathrm{d}x, & 0\le z<2;\\ \int_0^2 \dfrac{1}{2}\mathrm{e}^{-(z-x)}\mathrm{d}x, & z\ge 2,\end{cases}$$

$$=\begin{cases}0, & z<0.\\ \dfrac{1}{2}(1-\mathrm{e}^{-z}), & 0\le z<2;\\ \dfrac{1}{2}(\mathrm{e}^{2-z}-\mathrm{e}^{-z}), & z\ge 2,\end{cases}$$

$P\{X+Y>1\}=1-P\{X+Y\le 1\}=1-\int_0^1 \mathrm{d}x\int_0^{1-x}\dfrac{1}{2}\mathrm{e}^{-y}\mathrm{d}y=1-\dfrac{1}{2}\mathrm{e}^{-1}.$

3.29 设二维随机变量(X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{2}(x+y)\mathrm{e}^{-(x+y)}, & x>0,y>0;\\ 0, & 其他.\end{cases}$$

(1) 判断X和Y是否独立？

(2) 令$Z=2X-Y$,求Z的概率密度函数.

解：(1) 因为$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$

$$=\begin{cases}0, & x<0;\\ \int_0^{+\infty}\dfrac{1}{2}(x+y)\mathrm{e}^{-(x+y)}\mathrm{d}y, & x\geq 0,\end{cases}$$

$$=\begin{cases}0, & x<0;\\ \dfrac{x+1}{2}\mathrm{e}^{-x}, & x\geq 0;\end{cases}$$

同理得$f_Y(y)=\begin{cases}0, & y<0;\\ \dfrac{y+1}{2}\mathrm{e}^{-y}, & y\geq 0.\end{cases}$

显然$f(x,y)\neq f_X(x)f_Y(y)$,所以随机变量X与Y不独立。

(2)因为$F_Z(z)=P\{Z\leq z\}=P\{2X-Y\leq z\}$, 所以

当$z<0$时，$F_Z(z)=\int_{-z}^{+\infty}\mathrm{d}y\int_0^{(y+z)/2}\dfrac{1}{2}(x+y)\mathrm{e}^{-(x+y)}\mathrm{d}x=\dfrac{1}{6}(2-z)\mathrm{e}^{z}$;

当$z\geq 0$时，$F_Z(z)=1-\int_{z/2}^{+\infty}\mathrm{d}x\int_0^{2x-z}\dfrac{1}{2}(x+y)\mathrm{e}^{-(x+y)}\mathrm{d}y=1-\dfrac{z+4}{6}\mathrm{e}^{-z/2}$;

故随机变量Z的概率密度函数为

$$f_Z(z)=\begin{cases}\dfrac{1-z}{6}\mathrm{e}^{z}, & z<0;\\ \dfrac{z+2}{12}\mathrm{e}^{-z/2}, & z\geq 0.\end{cases}$$

3.30 设随机变量 $X\sim U(0,2),Y\sim Exp(1)$,且$X$与$Y$相互独立，令

$$Z_1=X+Y,\quad Z_2=X-Y,\quad Z_3=\max\{X,Y\},\quad Z_4=\min\{X,Y\},$$

分别求$Z_i, i=1,2,3,4$的分布.

解：因为 $X \sim U(0,2), Y \sim Exp(1)$,且$X$与$Y$相互独立，所以$(X,Y)$的联合概率密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{2}\mathrm{e}^{-y}, & 0\le x\le 2, y\ge 0;\\ 0, & \text{其他}.\end{cases}$$

随机变量X的分布函数为

$$F_X(x)=\begin{cases}0, & x<0;\\ \dfrac{x}{2}, & 0\le x<2;\\ 1, & x\ge 2.\end{cases}$$

随机变量Y的分布函数为

$$F_Y(y)=\begin{cases}0, & y<0;\\ 1-\mathrm{e}^{-y}, & y\ge 0.\end{cases}$$

因为$F_{Z_1}(z_1)=P\{Z_1\le z_1\}=P\{X+Y\le z_1\}=\int\int_{x+y\le z_1} f(x,y)\mathrm{d}x\mathrm{d}y$, 所以

当$z_1<0$时，$F_{Z_1}(z_1)=0$;

当$0\le z_1<2$时，$F_{Z_1}(z_1)=\int_0^{z_1}\mathrm{d}x\int_0^{z_1-x}\dfrac{1}{2}\mathrm{e}^{-y}\mathrm{d}y=\dfrac{1}{2}\left(z_1+\mathrm{e}^{-z_1}-1\right)$;

当$z_1\ge 2$时，$F_{Z_1}(z_1)=\int_0^{2}\mathrm{d}x\int_0^{z_1-x}\dfrac{1}{2}\mathrm{e}^{-y}\mathrm{d}y=1-\dfrac{1}{2}\left(\mathrm{e}^2-1\right)\mathrm{e}^{-z_1}$;

综上得随机变量Z_1的分布函数为

$$F_{Z_1}(z_1)=\begin{cases}0, & z_1<0;\\ \dfrac{1}{2}\left(z_1+\mathrm{e}^{-z_1}-1\right), & 0\le z_1<2;\\ 1-\dfrac{1}{2}\left(\mathrm{e}^2-1\right)\mathrm{e}^{-z_1}, & z_1\ge 2.\end{cases}$$

因为$F_{Z_2}(z_2)=P\{Z_2\le z_2\}=P\{X-Y\le z_2\}=\int\int_{x-y\le z_2} f(x,y)\mathrm{d}x\mathrm{d}y$, 所以

当$z_2<0$时，$F_{Z_2}(z_2)=\int_0^{2}\mathrm{d}x\int_0^{x-z_2}\dfrac{1}{2}\mathrm{e}^{-y}\mathrm{d}y=\dfrac{1}{2}(1-\mathrm{e}^{-2})\mathrm{e}^{z_2}$;

当$0\le z_2<2$时，$F_{Z_2}(z_2)=1-\int_{z_2}^{2}\mathrm{d}x\int_0^{x-z_2}\dfrac{1}{2}\mathrm{e}^{-y}\mathrm{d}y=\dfrac{1}{2}\left(1+z_2-\mathrm{e}^{z_2-2}\right)$;

当$z_2 \geq 2$时，$F_{Z_2}(z_2) = 1.$

综上得随机变量Z_2的分布函数为

$$F_{Z_2}(z_2) = \begin{cases} \dfrac{1}{2}(1-\mathrm{e}^{-2})\mathrm{e}^{z_2}, & z_2 < 0; \\ \dfrac{1}{2}\left(1+z_2-\mathrm{e}^{z_2-2}\right), & 0 \leq z_2 < 2; \\ 1, & z_2 \geq 2. \end{cases}$$

$$F_{Z_3}(z_3) = P\{Z_3 \leq z_3\} = P\{\max\{X,Y\} \leq z_3\} = P\{X \leq z_3\}P\{Y \leq z_3\}$$

$$= F_X(z_3)F_Y(z_3)$$

$$= \begin{cases} 0, & z_3 < 0; \\ \dfrac{z_3}{2}\left(1-\mathrm{e}^{-z_3}\right), & 0 \leq z_3 < 2; \\ 1-\mathrm{e}^{-z_3}, & z_3 \geq 2. \end{cases}$$

$$F_{Z_4}(z_4) = P\{Z_4 \leq z_4\} = P\{\min\{X,Y\} \leq z_4\} = 1 - P\{\min\{X,Y\} > z_4\}$$

$$= 1 - P\{X > z_4\}P\{Y > z_4\}$$

$$= 1 - \left(1-F_X(z_4)\right)\left(1-F_Y(z_4)\right)$$

$$= \begin{cases} 0, & z_4 < 0; \\ 1-\dfrac{2-z_4}{2}\mathrm{e}^{-z_4}, & 0 \leq z_4 < 2; \\ 1, & z_4 \geq 2. \end{cases}$$

3.31 设二维随机变量(X,Y)的联合概率密度函数为

$$f(x,y) = \begin{cases} 8xy, & 0 < x < y < 1; \\ 0, & \text{其他}. \end{cases}$$

令

$$Z_1 = X+Y, \quad Z_2 = X-Y, \quad Z_3 = \max\{X,Y\}, \quad Z_4 = \min\{X,Y\},$$

分别求$Z_i, i=1,2,3,4$的分布.

解：因为$F_{Z_1}(z_1)=P\{Z_1\le z_1\}=P\{X+Y\le z_1\}=\int\int_{x+y\le z_1} f(x,y)\mathrm{d}x\mathrm{d}y$, 所以

当$z_1<0$时，$F_{Z_1}(z_1)=0$;

当$0\le z_1<1$时，$F_{Z_1}(z_1)=\int_0^{z_1/2}\mathrm{d}x\int_x 0^{z_1-x}8xy\mathrm{d}y=\frac{1}{6}z_1^4$;

当$1\le z_1<2$时，$F_{Z_1}(z_1)=1-\int_{z_1/2}^1\mathrm{d}y\int_{z_1-y}^y 8xy\mathrm{d}x=-\frac{1}{6}z_1^4+2z_1^2-\frac{8}{3}z_1+1$;

当$z_1\ge 2$时，$F_{Z_1}(z_1)=1$.

综上得随机变量Z_1的分布函数为

$$F_{Z_1}(z)=\begin{cases}0, & z<0;\\ \frac{1}{6}z^4, & 0\le z<1;\\ -\frac{1}{6}z^4+2z^2-\frac{8}{3}z+1, & 1\le z<2;\\ 1, & z\ge 2.\end{cases}$$

因为$F_{Z_2}(z_2)=P\{Z_2\le z_2\}=P\{X-Y\le z_2\}=\int\int_{x-y\le z_2} f(x,y)\mathrm{d}x\mathrm{d}y$, 所以

当$z_2<-1$时，$F_{Z_2}(z_2)=0$;

当$-1\le z_2<0$时，$F_{Z_2}(z_2)=\int_{-z_2}^1\mathrm{d}y\int_0^{y+z_2}8xy\mathrm{d}x=\frac{1}{3}(3-z_2)\left(1+z_2\right)^3$;

当$z_2\ge 0$时，$F_{Z_2}(z_2)=1$.

综上得随机变量Z_2的分布函数为

$$F_{Z_2}(z_2)=\begin{cases}0, & z_2<-1;\\ \frac{1}{3}(3-z_2)\left(1+z_2\right)^3, & -1\le z_2<0;\\ 1, & z_2\ge 0.\end{cases}$$

$F_{Z_3}(z_3)=P\{Z_3\le z_3\}=P\{\max\{X,Y\}\le z_3\}=P\{Y\le z_3\}$

$$=\begin{cases} 0, & z_3<0; \\ \int_0^{z_3}\mathrm{d}x\int_x^{z_3}8xy\mathrm{d}y, & 0\le z_3<1; \\ 1, & z_3\ge 1. \end{cases}$$

$$=\begin{cases} 0, & z_3<0; \\ z_3^4, & 0\le z_3<1; \\ 1, & z_3\ge 1. \end{cases}$$

$F_{Z_4}(z_4)=P\{Z_4\le z_4\}=P\{\min\{X,Y\}\le z_4\}=P\{X\le z_4\}$

$=F_X(z_4)$

$$=\begin{cases} 0, & z_4<0; \\ 2z_4^2(1-z_4^2), & 0\le z_4<1; \\ 1, & z_4\ge 1. \end{cases}$$

3.32 设随机变量$X\sim Exp(2)$,随机变量$Y\sim B(1,0.6)$,且X与Y相互独立. 令

$$U=\begin{cases} 0, & X>Y; \\ 1, & X\le Y, \end{cases}\qquad V=\begin{cases} 0, & X>2Y; \\ 1, & X\le 2Y, \end{cases}\qquad Z=2X+Y,$$

求（1）(U,V)的联合概率分布律；（2）U,V的边缘概率分布律；（3）Z 的概率分布.

解：(1)因为$X\sim Exp(2),Y\sim B(1,0.6)$，所以

$$F_X(x)=\begin{cases} 0, & x<0; \\ 1-\mathrm{e}^{-2x}, & x\ge 0. \end{cases}$$

$P\{Y=1\}=0.6,\ P\{Y=0\}=0.4.$

$P\{U=0,V=0\}=P\{X>Y,X>2Y\}=P\{X>2Y\}$

$= P\{Y=0\}P\{X>0\}+P\{Y=1\}P\{X>2\}=0.4+0.6\mathrm{e}^{-4};$

$P\{U=0,V=1\}=P\{X>Y,X\leq 2Y\}=P\{Y<X\leq 2Y\}$

$= P\{Y=0\}P\{X=0\}+P\{Y=1\}P\{1<X\leq 2\}=0.6\left(\mathrm{e}^{-2}-\mathrm{e}^{-4}\right);$

$P\{U=1,V=0\}=P\{X\leq Y,X>2Y\}=P\{X>2Y\}0;$

$P\{U=1,V=1\}=P\{X\leq Y,X\leq 2Y\}=P\{X\leq Y\}$

$= P\{Y=0\}P\{X<0\}+P\{Y=1\}P\{X<1\}=0.6(1-\mathrm{e}^{-2}),$

所以(U,V)的联合概率分布律为

V \ U	0	1
0	$0.4+0.6\mathrm{e}^{-4}$	0
1	$0.6(\mathrm{e}^{-2}-\mathrm{e}^{-4})$	$0.6(1-\mathrm{e}^{-2})$

(2)因为$P\{U=0\}=P\{U=0,V=0\}+P\{U=0,V=1\}=0.4+0.6\mathrm{e}^{-2};$

$P\{U=1\}=P\{U=1,V=0\}+P\{U=1,V=1\}=0.6(1-\mathrm{e}^{-2}),$

所以随机变量U的概率分布律为

U	0	1
P_U	$0.4+0.6\mathrm{e}^{-2}$	$0.6(1-\mathrm{e}^{-2})$

同理可得随机变量V的概率分布律为

V	0	1
P_V	$0.4+0.6\mathrm{e}^{-4}$	$0.6(1-\mathrm{e}^{-4})$

(3)根据分布函数的定义知

$F_Z(z)=P\{Z\leq z\}=P\{2X+Y\leq z\}$

$\quad = P\{Y=0\}P\{X\leq z/2\}+P\{Y=1\}P\{X\leq (z-1)/2\}$

$\quad = 0.4F_X(z/2)+0.6F_X((z-1)/2)$

$$
= \begin{cases} 0, & z < 0; \\ 0.4(1-\mathrm{e}^{-z}), & 0 \le z < 1; \\ 1 - 0.4\mathrm{e}^{-z} - 0.6\mathrm{e}^{1-z}, & z \ge 1. \end{cases}
$$

3.33 设随机变量 X 与 Y 的概率密度函数分别为 $f_X(x)$ 与 $f_Y(y)$,且 X 与 Y 相互独立,令 $Z = aX + bY$(a, b为非零常数), 证明 Z 的概率密度函数为

$$
f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{|a||b|} f_X\left(\frac{u}{a}\right) f_Y\left(\frac{z-u}{b}\right) \mathrm{d}u\,, \quad z \in \mathrm{R}.
$$

证明：不妨设$a > 0$,可以类似讨论$a < 0$的情况。先求Z的分布函数。

$F_Z(z) = P\{Z \le z\} = P\{aX + bY \le z\} = \int_{-\infty}^{+\infty} \mathrm{d}y \int_{-\infty}^{\frac{z-by}{a}} f_X(x) f_Y(y) \mathrm{d}x$

$= \int_{-\infty}^{+\infty} F_X\left(\frac{z-by}{a}\right) f_Y(y) \mathrm{d}y,$

两边对z求导数得随机变量Z的概率密度函数为

$f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{a} f_X\left(\frac{z-by}{a}\right) f_Y(y) \mathrm{d}y$

作变量替换$u = z - by$得，$f_Z(z) = \int_{-\infty}^{+\infty} \frac{1}{a|b|} f_X\left(\frac{u}{a}\right) f_Y\left(\frac{z-u}{b}\right) \mathrm{d}u$

$= \int_{-\infty}^{+\infty} \frac{1}{|ab|} f_X\left(\frac{u}{a}\right) f_Y\left(\frac{z-u}{b}\right) \mathrm{d}u$

3.34 设电子元件的寿命 X(单位:h) 的概率密度函数为

$$
f(x) = \begin{cases} 0.0015\mathrm{e}^{-0.0015x}, & x \ge 0; \\ 0, & x < 0. \end{cases}
$$

今测试 4 个元件,并记录它们各自的失效时间. 求

（1）到 800h 时没有一个元件失效的概率;

（2）到 3000h 时所有元件都失效的概率.

解：(1) 一个元件到 800h 时失效的概率

$p_1 = P\{X \le 800\} = 1 - \mathrm{e}^{-1.2},$

令Y表示4个元件中到800h时失效的元件个数，则$Y\sim B(4,p_1)$,

故到 800h 时没有一个元件失效的概率为$P\{Y=0\}=(1-p_1)^4=\mathrm{e}^{-4.8}$.

(2)一个元件到 3000h 时失效的概率为

$p_2=P\{X\le 3000\}=1-\mathrm{e}^{-4.5}$,

令Z表示4个元件中到3000h时失效的元件个数，则$Z\sim B(4,p_2)$,

故到 3000h 时所有元件都失效的概率为$P\{Z=4\}=p_2^4=\left(1-\mathrm{e}^{-4.5}\right)^4$.

3.35 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立,且都服从区间 $[0,a](a>0$为常数) 上的均匀分布,令 $Y_1=\max\{X_1,X_2,\cdots,X_n\},Y_2=\min\{X_1,X_2,\cdots,X_n\}$,分别求Y_1和Y_2的概率密度函数.

解：因为随机变量 $X_1,X_2,\cdots,X_n$ 相互独立,且都服从区间 $[0,a](a>0$为常数) 上的均匀分布,所以，他们的分布函数与概率密度函数均为

$$F(x)=\begin{cases}0, & x<0;\\ \dfrac{x}{a}, & 0\le x<a;\\ 1, & x\ge a.\end{cases}\qquad f(x)=\begin{cases}\dfrac{1}{a}, & x\in[0,a];\\ 0, & x\notin[0,a].\end{cases}$$

下面求Y_1的分布函数。

$F_{Y_1}(y)=P\{Y_1\le y\}=P\{\max\{X_1,X_2,\cdots,X_n\}\le y\}$

$=P\{X_1\le y,X_2\le y,\cdots,X_n\le y\}$

$=P\{X_1\le y\}P\{X_2\le y\}\cdots P\{X_n\le y\}$

$=F_{X_1}(y)F_{X_2}(y)\cdots F_{X_n}(y)=F^n(y)$,

两边对y求导数得Y_1的概率密度函数为

$$f_{Y_1}(y)=nF^{n-1}(y)f(y)=\begin{cases}\dfrac{ny^{n-1}}{a^n}, & y\in[0,a];\\ 0, & y\notin[0,a].\end{cases}$$

下面再求Y_2的分布函数。

$F_{Y_2}(y) = P\{Y_2 \le y\} = P\{\min\{X_1, X_2, \cdots, X_n\} \le y\}$

$= 1 - P\{\min\{X_1, X_2, \cdots, X_n\} > y\}$

$= 1 - P\{X_1 > y, X_2 > y, \cdots, X_n > y\}$

$= 1 - P\{X_1 > y\}P\{X_2 > y\} \cdots P\{X_n > y\}$

$= 1 - (1 - F_{X_1}(y))(1 - F_{X_2}(y)) \cdots (1 - F_{X_n}(y)) = 1 - (1 - F(y))^n,$

两边对y求导数得Y_2的概率密度函数为

$$f_{Y_2}(y) = n(1 - F(y))^{n-1} f(y) = \begin{cases} \dfrac{n(a-y)^{n-1}}{a^n}, & y \in [0, a]; \\ 0, & y \notin [0, a]. \end{cases}$$

习题4

4.1 设射击比赛中每人可打4发弹, 并规定: 全部不中得0分; 中1弹得15分; 中2弹得30分; 中3弹得55分; 4弹全中得100分. 某人每次射击的命中率均为3/4, 问他平均可得多少分?

解：设X表示射击4发子弹时的得分数，则$X \sim B(4, 3/4)$,概率分布律为

X	0	15	30	55	100
P_X	$\frac{1}{256}$	$\frac{12}{256}$	$\frac{54}{256}$	$\frac{108}{256}$	$\frac{81}{256}$

所以平均得分为$\mathrm{E}(X) = 15 \cdot \dfrac{12}{256} + 30 \cdot \dfrac{54}{256} + 55 \cdot \dfrac{108}{256} + 100 \cdot \dfrac{81}{256} = 61.875.$

4.2 设排球队A与B进行比赛(无平局),若有一队胜4局,则比赛结束. 假定A、B在每局比赛中获胜的概率都是0.5, 试求比赛结束时所需比赛局数的数学期望.

解：设X表示比赛结束时所需要的比赛局数，则

$P\{X=k\}=2C_{k-1}^{3}0.5^{k}$, $k=4,5,6,7$. 即

X	4	5	6	7
P_X	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{5}{16}$	$\frac{5}{16}$

故此赛结束时所需要的平均局数为

$\mathrm{E}(X)=\sum\limits_{k=4}^{7}k\cdot 2C_{k-1}^{3}0.5^{k}=5.8125.$

4.3 (闯关游戏) 设汽车需通过4个有红绿信号灯的十字路口才能到达目的地. 如果汽车通过每个路口(即遇到绿灯)的概率是0.6,X 是首次停止时通过的路口数. 假定游戏规定, 通过k道关口将获得价值$10k(0\le k\le 4)$元的奖品. 问一个参与者获得的奖励平均价值是多少?

解：设Y表示获得的奖品价值，则

$P\{Y=10k\}=P\{X=k\}=0.6^{k}\cdot 0.4$, $k=0,1,2,3,$

$P\{Y=40\}=P\{X=4\}=0.6^{4}.$

故一个参与者获得奖励的平均价值是

$\mathrm{E}(Y)=\sum\limits_{k=0}^{4}10k\cdot P\{Y=10k\}=13.056\,.$

4.4 设某人月收入服从参数为$\lambda(\lambda>0)$的指数分布,月平均收入为2800元,按规定月收入超过3000元,应交个人所得税. 假定人在一年内各月的收入相互独立,且每年有 X个月需交个人所得税.求(1)此人每月需交个人所得税的概率; (2)随机变量 X的概率分布; (3)每年平均有几个月需交个人所得税.

解：用Y表示月收入，则$Y\sim E(1/2800)$.

(1)$P\{Y>3000\}=1-P\{Y\le 3000\}=\mathrm{e}^{-\frac{15}{14}}$.

(2)依题意 $X\sim B(12,p)$, 其中$p=\mathrm{e}^{-\frac{15}{14}}$.

(3)$\mathrm{E}(X)=np=12\mathrm{e}^{-\frac{15}{14}}$.

4.5 将3个球随机地投入到4个盒子中,球与盒子均可区分,以X表示所余的空盒子数,求$\mathrm{E}(X)$.

解：依题意X的概率分布律为

X	1	2	3
P_X	$\frac{C_4^3 3!}{4^3}$	$\frac{C_4^2(C_3^1+C_3^2)}{4^3}$	$\frac{C_4^1}{4^3}$

$\mathrm{E}(X)=\sum\limits_{k=1}^{3}kP\{X=k\}=\frac{27}{16}=1.6875.$

4.6 按规定,某车站每天$8:00\sim 9:00$, $9:00\sim 10:00$都恰有一辆客车到站,但到站的时间是随机的,且两者到站的时间相互独立.其规律为

到站时刻	8 : 10/9 : 10	8 : 30/9 : 30	8 : 50/9 : 50
概率	1/6	3/6	2/6

(1) 旅客8:00到站,求他候车时间的数学期望;

(2)一旅客8:20到站,求他候车时间的数学期望.

解：(1)若旅客8:00到站，则候车时间Y的概率分布律为

Y	10	30	50
P_Y	1/6	3/6	2/6

所以他的平均候车时间为

$\mathrm{E}(Y)=10\cdot\frac{1}{6}+30\cdot\frac{3}{6}+50\cdot\frac{2}{6}=\frac{100}{3}=33.33$(分钟)。

4.7 十万张奖卷为一组, 每组设头等奖2名, 奖金10000元, 二等奖20名, 奖金1000元, 三等奖200名, 奖金100元, 四等奖2000名, 奖金10元, 五等奖10000名, 奖金2元, 对于一个买了一张奖卷等待开奖的人来说, 他“期望”的奖金数X的平均值是多少?

解：设X表示获奖的奖金数，则X的概率分布律为

X	0	2	10	100	1000	10000
P_X	$\frac{87778}{10^5}$	$\frac{1}{10}$	$\frac{1}{50}$	$\frac{1}{500}$	$\frac{1}{5000}$	$\frac{1}{50000}$

他“期望”的奖金数X的平均值是

$$\mathrm{E}(X)=2\cdot\frac{1}{10}+10\cdot\frac{1}{50}+100\cdot\frac{1}{500}+1000\cdot\frac{1}{5000}+10000\cdot\frac{1}{50000}=1.$$

4.8 长途汽车起点站于每时的10分、30分、55分发车,设乘客不知发车时间,于每小时的任意时刻随机地到达车站,求乘客的平均候车时间.

解：设X表示乘客的到达时刻，Y表示乘客的候车时间，则$X\sim U(0,60)$,

$$f_X(x)=\begin{cases}\frac{1}{60}, & 0\le x\le 60;\\ 0, & \text{其他}.\end{cases}$$

$$Y=g(X)=\begin{cases}10-X, & X<10;\\ 30-X, & 10\le X<30;\\ 55-X, & 30\le X<55;\\ 70-X, & 55\le X<60.\end{cases}$$

平均候车时间为

$$\mathrm{E}(Y)=\int_0^{60}g(x)f_X(x)\mathrm{d}x$$
$$=\frac{1}{60}\left[\int_0^{10}(10-x)\mathrm{d}x+\int_{10}^{30}(30-x)\mathrm{d}x+\int_{30}^{55}(55-x)\mathrm{d}x+\int_{55}^{60}(70-x)\mathrm{d}x\right]=10.41.$$

4.9 设商店对某种家用电器的销售采用先使用后付款的方式,商店规定凡使用寿命在1年以内付款1500元;1年到2年内付款2000元; 2年到3年内付款2500元;3年以上付款3000元．若该家用电器的使用寿命X(单位:年) 服从指数分布$E\,(0.1)$, 试写出该商店对此家用电器每台的收款数Y的概率分布律,并求Y的数学期望.

解：由题意得Y的所有可能取值为1500,2000,2500，3000.

又因为使用寿命$X\sim E(0.1)$,所以X的分布函数为

$$F_X(x)=\begin{cases}1-\mathrm{e}^{-0.x}, & x\geq 0;\\ 0, & x<0.\end{cases}$$

由此得$P\{Y=1500\}=P\{X<1\}=F_X(1)=1-\mathrm{e}^{-0.1}$;

$P\{Y=2000\}=P\{1\leq X<2\}=F_X(2)-F_X(1)=\mathrm{e}^{-0.1}-\mathrm{e}^{-0.2}$;

$P\{Y=2500\}=P\{2\leq X<3\}=F_X(3)-F_X(2)=\mathrm{e}^{-0.2}-\mathrm{e}^{-0.3}$;

$P\{Y=3000\}=P\{X\geq 3\}=1-F_X(3)=\mathrm{e}^{-0.3}$.

综上得Y的概率分布律为

Y	1500	2000	2500	3000
p_Y	$1-\mathrm{e}^{-0.1}$	$\mathrm{e}^{-0.1}-\mathrm{e}^{-0.2}$	$\mathrm{e}^{-0.2}-\mathrm{e}^{-0.3}$	$\mathrm{e}^{-0.3}$

$\mathrm{E}(Y)=1500\cdot\left(1-\mathrm{e}^{-0.1}\right)+2000\cdot\left(\mathrm{e}^{-0.1}-\mathrm{e}^{-0.2}\right)+2500\cdot\left(\mathrm{e}^{-0.2}-\mathrm{e}^{-0.3}\right)+3000\cdot\mathrm{e}^{-0.3}$

$=2732.19$.（答案不对）

4.10 统计资料表明强烈地震的间隔服从参数为1/430 的指数分布,问平均多长时间发生一次强震?（建议删除）

解：

4.11 对球的直径进行近似测量,其值均匀分布在区间$[a,b]$上,求球体积的数学期望.

解：设X表示球的直径，V表示球的体积，则$V=\dfrac{\pi X^3}{6}$,由于$X\sim U(a,b)$,

所以X的概率密度函数为$f_X(x)=\begin{cases}\dfrac{1}{b-a}, & a\leq x\leq b;\\ 0, & 其他.\end{cases}$

$\mathrm{E}(V)=\int_a^b\dfrac{\pi x^3}{6}\cdot\dfrac{1}{b-a}\mathrm{d}x=\dfrac{\pi(a+b)(a^2+b^2)}{24}$.

4.12 某水果商店,冬季每周购进一批苹果. 已知该店一周苹果销售量 X(单位:kg)服从$U(1000,2000)$. 若购进的苹果在一周内售出, 则每千克获纯利1.5元; 一周内没售出,

则每千克需付耗损、储藏等费用0.3元. 问一周应购进多少千克苹果, 商店才能获得最大的平均利润?

解：设Y表示水果店一周的利润，一周应购进苹果m千克,则

$$Y=g(X)=\begin{cases}1.5X-0.3(m-X), & X\le m;\\ \\ 1.5m, & X>m.\end{cases}$$

又因为销售量$X\sim U(1000,2000)$，所以X的概率密度函数为

$$f_X(x)=\begin{cases}\dfrac{1}{1000}, & 1000\le x\le 2000;\\ \\ 0, & 其他.\end{cases}$$

$\mathrm{E}(Y)=\int_{1000}^{2000}g(x)\cdot\dfrac{1}{1000}\mathrm{d}x=\dfrac{1}{1000}\left[\int_{1000}^{m}[1.5x-0.3(m-x)]\mathrm{d}x+\int_{m}^{2000}1.5m\mathrm{d}x\right]$

$=\dfrac{1}{1000}(-1.2m^2+3300m+600000)$,

要使得平均利润$\mathrm{E}(Y)$达到最大，只要取$m=1375$.(和答案不同。)即应购进1375千克苹果, 商店才能获得最大平均利润。

4.13 一个系统由两个子系统串联而成, 只要有一个子系统发生故障, 系统就不能正常工作. 设两个子系统的工作寿命分别为 X和Y, 且相互独立, 并服从相同的指数分布$E(\lambda)$, 其中$\lambda>0$为参数, 求系统工作寿命 T的数学期望. (建议删掉，与4.17重复)

解：因为系统由两个子系统串联组成，所以系统的寿命T取决于子系统的最短寿命，即$T=\min\{X,Y\}$.

又因为X,Y相互独立，且都服从指数分布$E(\lambda)$,概率密度函数和分布函数分别为

$$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x\ge 0;\\ \\ 0, & x<0.\end{cases}\qquad F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x\ge 0;\\ \\ 0, & x<0.\end{cases}$$

随机变量T的概率密度函数为

$$f_T(t)=2f(t)(1-F(t))=\begin{cases}2\lambda \mathrm{e}^{-2\lambda t}, & t\geq 0;\\ 0, & t<0.\end{cases}\qquad \text{即} T\sim E(2\lambda).$$

故系统的平均寿命为$\mathrm{E}(T)=\dfrac{1}{2\lambda}$.

4.14 一部机器在一天内发生故障的概率为0.2,且一旦发生故障全天停止工作. 一周5个工作日, 如果一周内不发生故障, 厂家可获利润10万元; 若只发生1次故障, 仍可获利润5万元; 若发生2次故障, 不获利也不亏损; 若发生故障3次或3次以上, 就要亏损2万元.假定机器每天发生故障与否相互独立。（修改）求一周内平均获利多少万元?

解：设X表示一周内机器发生的故障次数，则$X\sim B(5,0.2)$.

用Y表示厂家的利润，则$Y=g(X)=\begin{cases}10, & X=0;\\ 5, & X=1;\\ 0, & X=2;\\ -2, & X\geq 3.\end{cases}$

故厂家在一周内平均获利为$\mathrm{E}(Y)=\sum\limits_{k=0}^{5}g(k)P\{X=k\}=5.20896$(万元).

4.15 某人寿保险公司有3000个同一年龄段的人参加人身保险, 在一年内这些人的死亡率为0.1%. 每个参加保险的人在一年的第1天交付保险费, 若在该年内死亡, 其家属可从保险公司领取赔偿金1万元. 如果保险公司期望获利不少于20万元, 那么每个参加保险的人每年交的保险费至少为多少元?

解：设每个参加保险的人每年交的保险费至少为m元,X表示一年中的死亡人数，则$X\sim B(3000,0.001)$.

公司的利润(单位：万元)函数为$Y=0.3m-X$,平均利润为$\mathrm{E}(Y)=0.3m-\mathrm{E}(X)=0.3m-3$,

要使保险公司获利不少于20万元，只要$\mathrm{E}(Y) > 20$,解之得$m > 76.7$(元).

4.16 设二维随机变量 (X, Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{4}(1-x^3y+xy^3), & |x|<1,|y|<1;\\ 0, & \text{其他}.\end{cases}$$

试验证$\mathrm{E}(XY)=\mathrm{E}(X)\mathrm{E}(Y)$,但$X$与$Y$不独立.

解：因为$\mathrm{E}(X)=\int\int_{R^2} xf(x,y)\mathrm{d}x\mathrm{d}y=\int_{-1}^{1}\mathrm{d}x\int_{-1}^{1}x\cdot\dfrac{1}{4}(1-x^3y+xy^3)\mathrm{d}y=0$;

$\mathrm{E}(Y)=\int\int_{R^2} yf(x,y)\mathrm{d}x\mathrm{d}y=\int_{-1}^{1}\mathrm{d}x\int_{-1}^{1}y\cdot\dfrac{1}{4}(1-x^3y+xy^3)\mathrm{d}y=0$;

$\mathrm{E}(XY)=\int\int_{R^2} xyf(x,y)\mathrm{d}x\mathrm{d}y=\int_{-1}^{1}\mathrm{d}x\int_{-1}^{1}xy\cdot\dfrac{1}{4}(1-x^3y+xy^3)\mathrm{d}y=0$;

所以$\mathrm{E}(XY)=\mathrm{E}(X)\mathrm{E}(Y)$.

又因为$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{-1}^{1}\dfrac{1}{4}(1-x^3y+xy^3)\mathrm{d}y, & -1\le x\le 1;\\ 0, & \text{其他}.\end{cases}$

$$=\begin{cases}\dfrac{1}{2}, & -1\le x\le 1;\\ 0, & \text{其他}.\end{cases}$$

$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_{-1}^{1}\dfrac{1}{4}(1-x^3y+xy^3)\mathrm{d}x, & -1\le y\le 1;\\ 0, & \text{其他}.\end{cases}$

$$=\begin{cases}\dfrac{1}{2}, & -1\le y\le 1;\\ 0, & \text{其他}.\end{cases}$$

显然$f(x,y)\neq f_X(x)f_Y(y)$,所以X与Y不独立。

4.17 有5个相互独立工作的电子装置,它们的寿命$X_k(k=1,2,3,4,5)$均服从参数为λ的指数分布.(1)若将这5个电子装置串联连接组成整机,求整机寿命（以小时记）N的数学期望; (2)若将这5个电子装置并联连接组成整机,求整机寿命（以小时记）M的数学

期望.

解：因为$X_k(k=1,2,3,4,5)$均服从参数为λ的指数分布，所以$X_k(k=1,2,3,4,5)$的概率密度函数和分布函数分别为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0;\\ \\ 0, & x<0.\end{cases}\qquad F(x)=\begin{cases}1-e^{-\lambda x}, & x\geq 0;\\ \\ 0, & x<0.\end{cases}$$

(1)由于整机寿命$N=\min\{X_1,\cdots,X_5\}$，所以N的概率密度函数为

$$f_N(x)=5\left(1-F(x)\right)^4 f(x)=\begin{cases}5\lambda e^{-5\lambda x}, & x\geq 0;\\ \\ 0, & x<0.\end{cases}\quad \text{即} N\sim E(5\lambda).$$

故$\mathrm{E}(N)=\dfrac{1}{5\lambda}$.

(2)由于整机寿命$M=\max\{X_1,\cdots,X_5\}$，所以M的概率密度函数为

$$f_M(x)=5F^4(x)f(x)=\begin{cases}5\lambda\left(1-e^{-\lambda x}\right)^4 e^{-\lambda x}, & x\geq 0;\\ \\ 0, & x<0.\end{cases}$$

故$\mathrm{E}(M)=\int_0^{+\infty}5\lambda\left(1-e^{-\lambda x}\right)^4 e^{-\lambda x}\mathrm{d}x=\dfrac{137}{60\lambda}$.

4.18 理发店里有甲、乙、丙三个顾客,假定理发店对三个顾客的服务时间都服从参数为λ的指数分布, 对甲和乙立即开始服务,在对甲或乙服务结束后开始对丙服务,对每个人服务所需的时间是独立的. 求丙在理发店的等待时间与逗留时间（逗留时间等于等待时间与服务时间之和）的数学期望.

解：假设理发店对甲、乙、丙的服务时间分别为X,Y,Z.则丙在理发店的等待时间为$N=\min X,Y$,逗留时间为$M=N+Z$.

因为X,Y,Z均服从参数为λ的指数分布，所以X,Y,Z的概率密度函数和分布函数分别为

$$f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, & x\geq 0;\\ 0, & x<0.\end{cases}\qquad F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x\geq 0;\\ 0, & x<0.\end{cases}$$

所以N的概率密度函数为

$$f_N(x)=2\left(1-F(x)\right)f(x)=\begin{cases}2\lambda\mathrm{e}^{-2\lambda x}, & x\geq 0;\\ 0, & x<0.\end{cases}\quad \text{即}N\sim E(2\lambda).$$

故$\mathrm{E}(N)=\dfrac{1}{2\lambda}$.

$\mathrm{E}(M)=\mathrm{E}(N)+\mathrm{E}(Z)=\dfrac{1}{2\lambda}+\dfrac{1}{\lambda}=\dfrac{3}{2\lambda}$.

4.19 设$X\sim N(\mu,\sigma^2)$,求证:$\mathrm{E}\left(|X-\mu|\right)=\sqrt{\dfrac{2}{\pi}}\sigma$.

解：因为$X\sim N(\mu,\sigma^2)$，所以$X-\mu\sim N(0,\sigma^2)$,概率密度函数为

$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\dfrac{x^2}{2\sigma^2}\right\}$.

$$\text{故}\mathrm{E}|X-\mu|=\int_{-\infty}^{+\infty}|x|\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}\mathrm{d}x=2\int_0^{+\infty}x\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}\mathrm{d}x$$
$$=\sqrt{\frac{2}{\pi}}\sigma.$$

4.20 一超市购物车载有20位顾客自超市开出,沿途有10个站点可以下车,如到达一个站点没有顾客下车就不停车,以X表示停车的次数, 求$\mathrm{E}(X)$(设每位顾客在各个站点下车是等可能的并设各顾客是否下车相互独立).

解：设X_i表示在第i站停车的次数，则X_i的概率分布律为

X_i	0	1
P_{X_i}	$\left(\dfrac{9}{10}\right)^{20}$	$1-\left(\dfrac{9}{10}\right)^{20}$

由于$X=\sum\limits_{i=1}^{10}X_i$,

所以$\mathrm{E}(X)=\sum\limits_{i=1}^{10}\mathrm{E}(X_i)=10\cdot\left(1-\left(\dfrac{9}{10}\right)^{20}\right)=8.784$.

4.21 设某图书馆的读者借阅甲种图书的概率为p,借阅乙种图书的概率为α,且对每位读者而言, 他是否借阅甲种图书与是否借阅乙种图书相互独立,且设读者与读者之间

是否借阅甲、乙图书的行动也相互独立．设某天恰有n名读者,求(1)借阅甲种图书的平均人数, (2)至少借阅了一种图书的平均人数.

解：设借阅甲、乙图书的人数分别为X,Y,则$X\sim B(n,p),Y\sim B(n,\alpha)$,且$X$与$Y$相互独立。

(1)因为$X\sim B(n,p)$,所以$\mathrm{E}(X)=np$.

(2)设$Z_i=0$表示第i个人没有借书，$Z_i=1$表示第i个人至少借阅一种图书，则$Z_i(i=1,\cdots,n)$的概率分布律为

Z_i	0	1
P_{Z_i}	$(1-p)(1-\alpha)$	$p+\alpha-p\alpha$

至少借阅一种图书的总人数$Z=\sum\limits_{i=1}^{n}Z_i$,

故$\mathrm{E}(Z)=\sum\limits_{i=1}^{n}\mathrm{E}(Z_i)=n(p+\alpha-p\alpha)$.

4.22 将n只球(1至n号)随机地放进n只盒子(1至n号)中去,一只盒子装一只球,将一只球放入与球同号的盒子中称为一个配对, 设X为配对的个数,求$\mathrm{E}(X)$.

解：设X_i表示第i只盒子与第i只球的配对数，$i=1,\cdots,n$.则X_i的概率分布律为

X_i	0	1
P_{X_i}	$1-\dfrac{1}{n}$	$\dfrac{1}{n}$

由于$X=\sum\limits_{i=1}^{n}X_i$,所以

$\mathrm{E}(X)=\sum\limits_{i=1}^{n}\mathrm{E}(X_i)=1$.

4.23 一台设备由三大部件构成,在设备运转中各部件需要调整的概率分别为0.1,0.2和0.3. 假设各部件的状态相互独立, 以 X表示同时需要调整的部件数.试用下列两种方法求 X的数学期望和方差. (1)利用 X的概率分布律; (2)利用分解法.

解：(1)设$A_i=\{$第i个部件需要调整$\},i=1,2,3$.

$P(A_1)=0.1, P(A_2)=0.2, P(A_3)=0.3.$

由于随机变量X的所有可能取值为0,1,2,3.且

$P\{X=0\}=P(\overline{A_1}\,\overline{A_2}\,\overline{A_3})=P(\overline{A_1})P(\overline{A_2})P(\overline{A_3})=0.9\cdot 0.8\cdot 0.7=0.504;$

$P\{X=1\}=P(A_1\overline{A_2}\,\overline{A_3})+P(\overline{A_1}\,A_2\,\overline{A_3})+P(\overline{A_1}\,\overline{A_2}\,A_3)=0.398.$

$P\{X=2\}=P(A_1A_2\overline{A_3})+P(A_1\overline{A_2}A_3)+P(\overline{A_1}\,A_2A_3)=0.092.$

$P\{X=2\}=P(A_1A_2A_3)=0.006$

所以X的概率分布律为

X	0	1	2	3
P_X	0.504	0.398	0.092	0.006

故$\mathrm{E}(X)=0\cdot 0.504+1\cdot 0.398+2\cdot 0.092+3\cdot 0.006=0.6.$

$\mathrm{D}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)=0^2\cdot 0.504+1^2\cdot 0.398+2^2\cdot 0.092+3^3\cdot 0.006-0.6^2=0.46.$

(2)设X_i表示第i个部件需要调整的数量,$i=1,2,3$，则X_i的概率分布律为

X	0	1
P_X	$1-0.i$	$0.i$

$\mathrm{E}(X_i)=0.i, \qquad \mathrm{D}(X_i)=0.i\cdot(1-0.i),$

由于$X=X_1+X_2+X_3$,且X_1,X_2,X_3相互独立，所以

$\mathrm{E}(X)=\mathrm{E}(X_1)+\mathrm{E}(X_2)+\mathrm{E}(X_3)=0.1+0.2+0.3=0.6.$

$\mathrm{D}(X)=\mathrm{D}(X_1)+\mathrm{D}(X_2)+\mathrm{D}(X_3)=0.1\cdot 0.9+0.2\cdot 0.8+0.3\cdot 0.7=0.46.$

4.24(负二项分布随机变量的数学期望)对于一系列独立试验, 每次成功的概率为p, 求第r 次成功需要的试验次数的数学期望(提示:用随机变量的分解法,令X_i表示从第$i-1$次成功到第i次成功所需的试验次数).

解：设X_i表示从第$i-1$次成功到第i次成功所需的试验次数，$i=1,\cdots,r$.则X_i的概率分布律为

$P\{X_i = k\} = (1-p)^{k-1}p, k = 1, 2, \cdots,$

$\mathrm{E}(X_i) = \sum\limits_{k=1}^{\infty} k \cdot (1-p)^k p = \dfrac{1}{p},$

由于第r次成功需要的试验次数$X = \sum\limits_{i=1}^{r} X_i$,所以$\mathrm{E}(X) = \sum\limits_{i=1}^{r} \mathrm{E}(X_i) = \dfrac{r}{p}$.

4.25 设二维离散型随机变量 (X, Y)的联合概率分布律为

$Y \backslash X$	-2	0	1
0	0.20	0.05	0.10
1	0.05	0.10	0.25
2	0	0.15	0.10

求$\mathrm{E}(X), \mathrm{E}(Y), \mathrm{E}(X^2Y), \mathrm{D}(X), \mathrm{D}(Y)$.

解：$\mathrm{E}(X) = -2 \cdot 0.25 + 0 \cdot 0.3 + 1 \cdot 0.45 = -0.05.$

$\mathrm{E}(Y) = 0 \cdot 0.35 + 1 \cdot 0.4 + 2 \cdot 0.25 = 0.9.$

$\mathrm{E}(X^2Y) = (-2)^2 \cdot 1 \cdot 0.05 + 1^2 \cdot 1 \cdot 0.25 + 1^2 \cdot 2 \cdot 0.1 = 0.65.$

$\mathrm{D}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = (-2)^2 \cdot 0.25 + 0^2 \cdot 0.3 + 1^2 \cdot 0.45 - (-0.05)^2 = 1.4475.$

$\mathrm{D}(Y) = \mathrm{E}(Y^2) - \mathrm{E}^2(Y) = 0^2 \cdot 0.35 + 1^2 \cdot 0.4 + 2^2 \cdot 0.25 - 0.9^2 = 0.59.$

4.26 一批零件中有9件合格品和3件次品. 安装机器时从这批零件中任取一件, 若每次取出的次品不放回, 求在取得合格品以前已取出的次品数的数学期望和方差.

解：设X表示在取得合格品以前已取出的次品数，则X的所有可能取值为0,1,2,3.

$P\{X = 0\} = \dfrac{9}{12} = \dfrac{3}{4};$

$P\{X = 1\} = \dfrac{3}{12} \cdot \dfrac{9}{11} = \dfrac{9}{44};$

$P\{X = 2\} = \dfrac{3}{12} \cdot \dfrac{2}{11} \cdot \dfrac{9}{10} = \dfrac{9}{220};$

$P\{X = 3\} = \dfrac{3}{12} \cdot \dfrac{2}{11} \cdot \dfrac{1}{10} = \dfrac{1}{220};$

$\mathrm{E}(X)=1\cdot\dfrac{9}{44}+2\cdot\dfrac{9}{220}+3\cdot\dfrac{1}{220}=0.3.$

$\mathrm{D}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)=1^2\cdot\dfrac{9}{44}+2^2\cdot\dfrac{9}{220}+3^2\cdot\dfrac{1}{220}-0.3^2=\dfrac{351}{1100}.$

4.27 某种商品每件表面上的疵点数 X服从泊松分布,平均每件上有0.8个疵点.若规定表面不超过一个疵点的为一等品; 价值10元,表面疵点数大于1不多于4的为二等品,价值8元;表面疵点数有4个以上者为废品,求产品价值的均值和方差.

解：设随机变量Y表示产品的价值，则

$$Y=\begin{cases}0, & X>4;\\ 8, & 1<X\le 4;\\ 10, & X\le 1.\end{cases}$$

因为$X\sim P(0.8)$,所以

$P\{Y=0\}=P\{X>4\}=1-P\{X\le 4\}=0.0014.$

$P\{Y=8\}=P\{1<X\le 4\}=P\{X\le 4\}-P\{X\le 1\}=0.1898.$

$P\{Y=10\}=P\{X\le 1\}=0.8088.$

故$\mathrm{E}(Y)=8\cdot 0.1898+10\cdot 0.8088=9.6064.$

$\mathrm{D}(Y)=\mathrm{E}(Y^2)-\mathrm{E}^2(Y)=8^2\cdot 0.1898+10^2\cdot 0.8088-9.6064^2=0.7443.$

4.28 某流水生产线上每个产品不合格的概率为$p(0<p<1)$,各产品合格与否相互独立,当出现一个不合格品时立即停机检修, 求开机到第一次检修时已生产的产品件数 X的期望和方差.

解：因为X的概率分布律为$P\{X=k\}=(1-p)^{k-1}p, k=1,2,\cdots.$

$\mathrm{E}(X)=\sum\limits_{k=1}^{+\infty}k\cdot(1-p)^{k-1}p=\dfrac{1}{p}.$

$\mathrm{D}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)=\sum\limits_{k=1}^{+\infty}k^2\cdot(1-p)^{k-1}p-\dfrac{1}{p^2}=\dfrac{1-p}{p^2}.$

4.29 为了开发某种新产品,某厂需要对生产设备的投资规模作一次决策．设计部门

提出3种可供选择的方案:(I)购买大型设备;(Ⅱ)购买中型设备;(III)购买小型设备．一个末确定的因素是市场对这种新产品的需求量．预计新产品投放市场后,市场对这种产品的需求量为高、中和低的概率分别为0.3, 0.4和0.3．厂方估计三种方案的利润(单位:万元)如下表所示．

需求量状态 \ 方案	I	II	III
高	50	30	10
中	20	25	10
低	-20	-10	10

(1)计算3种方案的期望利润,哪一个方案的期望利润最大?

(2)计算3种方案利润的方差,哪一个方案风险最小?

解：(1)方案I：平均利润为：$0.3 \cdot 50 + 0.4 \cdot 20 + 0.3 \cdot (-20) = 17$(万元)；

方案II：平均利润为：$0.3 \cdot 30 + 0.4 \cdot 25 + 0.3 \cdot (-10) = 16$(万元)；

方案III：平均利润为：$0.3 \cdot 10 + 0.4 \cdot 10 + 0.3 \cdot 10 = 10$(万元)；

由上知，方案I的期望利润最大。

(2)方案I：方差为：$0.3 \cdot 50^2 + 0.4 \cdot 20^2 + 0.3 \cdot (-20)^2 - 17^2 = 741$(万元)；

方案II：方差为：$0.3 \cdot 30^2 + 0.4 \cdot 25^2 + 0.3 \cdot (-10)^2 - 16^2 = 294$(万元)；

方案III：方差为：$0.3 \cdot 10^2 + 0.4 \cdot 10^2 + 0.3 \cdot 10^2 - 10^2 = 0$(万元)；

由上知，方案III的风险最小。

4.30 设$X \sim N(-1,3)$, $Y \sim N(1,4)$,且 X与Y相互独立,令$Z = 4X - 2Y + 1$.求$P\{Z > -5\}$, $\mathrm{E}(Z^2)$.

解：因为$X \sim N(-1,3)$, $Y \sim N(1,4)$,且 X与Y相互独立,

所以$Z = 4X - 2Y + 1 \sim N(-5, 64)$, E(Z)=-5, D(Z)=64.

故$P\{Z > -5\} = 0.5$, $\mathrm{E}(Z^2) = \mathrm{D}(Z) + \mathrm{E}^2(Z) = 64 + (-5)^2 = 89$.

4.31 设随机变量$X \sim U(-2,\ 2)$,令

$$Y = \begin{cases} -1, & X \leq -1; \\ 1, & X > -1, \end{cases} \qquad Z = \begin{cases} -1, & X \leq 1; \\ 1, & X > 1. \end{cases}$$

试求(1) Y与Z的联合概率分布律; (2)D$(Y + Z)$

解：(1)因为$X \sim U(-2,\ 2)$,所以X的分布函数为

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

由于$P\{Y = -1, Z = -1\} = P\{X \leq -1, X \leq 1\} = P\{X \leq -1\} = \dfrac{1}{4}$;

$P\{Y = -1, Z = 1\} = P\{X \leq -1, X > 1\} = 0$;

$P\{Y = 1, Z = -1\} = P\{X > -1, X \leq 1\} = P\{-1 < X \leq 1\} = \dfrac{1}{2}$;

$P\{Y = 1, Z = 1\} = P\{X > -1, X > 1\} = P\{X > 1\} = 1 - P\{X \leq 1\} = \dfrac{1}{4}$;

故(Y, Z)的联合概率分布律为

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

(2)$\mathrm{D}(Y+Z) = \mathrm{E}[(Y+Z)^2] - \mathrm{E}^2(Y+Z) = (-2)^2 \cdot \dfrac{1}{4} + 2^2 \cdot \dfrac{1}{4} - \left[-2 \cdot \dfrac{1}{4} + 2 \cdot \dfrac{1}{4}\right]^2 = 2$.

4.32 设二维随机变量 (X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases} kx(1+3y^2), & 0<x<2, 0<y<1; \\ 0, & \text{其他}. \end{cases}$$

试确定常数k,并求$\mathrm{E}(X),\mathrm{E}(Y),\mathrm{E}(XY),\mathrm{D}(X),\mathrm{D}(Y),\mathrm{Cov}(X,Y),\rho_{X,Y}$.

解：因为$\int\int_{R^2} f(x,y)\mathrm{d}x\mathrm{d}y=1$,

所以$\int_0^2 \mathrm{d}x\int_0^1 kx(1+3y^2)\mathrm{d}y=1$,

解之得$k=\dfrac{1}{4}$.

$\mathrm{E}(X)=\int\int_{R^2} xf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^2 \mathrm{d}x\int_0^1 x\cdot\dfrac{1}{4}x(1+3y^2)\mathrm{d}y=\dfrac{4}{3}$.

$\mathrm{E}(Y)=\int\int_{R^2} xf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^2 \mathrm{d}x\int_0^1 y\cdot\dfrac{1}{4}x(1+3y^2)\mathrm{d}y=\dfrac{5}{8}$.

$\mathrm{E}(Y)=\int\int_{R^2} xf(x,y)\mathrm{d}x\mathrm{d}y=\int_0^2 \mathrm{d}x\int_0^1 xy\cdot\dfrac{1}{4}x(1+3y^2)\mathrm{d}y=\dfrac{5}{6}$.

$\mathrm{D}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)=\int_0^2 \mathrm{d}x\int_0^1 x^2\cdot\dfrac{1}{4}x(1+3y^2)\mathrm{d}y-\left(\dfrac{4}{3}\right)^2=\dfrac{2}{9}$.

$\mathrm{D}(Y)=\mathrm{E}(Y^2)-\mathrm{E}^2(Y)=\int_0^2 \mathrm{d}x\int_0^1 y^2\cdot\dfrac{1}{4}x(1+3y^2)\mathrm{d}y-\left(\dfrac{5}{8}\right)^2=\dfrac{73}{960}$.

$\mathrm{Cov}(X,Y)=\mathrm{E}(XY)-\mathrm{E}(X)\mathrm{E}(Y)=\int_0^2 \mathrm{d}x\int_0^1 xy\cdot\dfrac{1}{4}x(1+3y^2)\mathrm{d}y-\dfrac{4}{3}\cdot\dfrac{5}{8}=0$.

$\rho_{X,Y}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{D}(X)\mathrm{D}(Y)}}=0$.

4.33 设随机变量 X与Y相互独立, 且均服从正态分布$N(\mu,\sigma^2)$,记$U=aX+bY,\ V=aX-bY$ a,b为常数. 问a,b取何值时，U与V不相关?

解：因为随机变量 X与Y相互独立, 且均服从正态分布$N(\mu,\sigma^2)$，

所以$\mathrm{E}(X)=\mu,\ \mathrm{D}(X)=\sigma^2$.

$$\mathrm{Cov}(U,V)=\mathrm{E}(UV)-\mathrm{E}(U)\mathrm{E}(V)=\mathrm{E}(a^2X^2-b^2Y^2)-(a\mu+b\mu)(a\mu-b\mu)$$

$$=a^2(\mu^2+\sigma^2)-b^2(\mu^2+\sigma^2)-(a^2\mu^2-b^2\mu^2)=(a^2-b^2)\sigma^2,$$

要使U与V不相关，只要$(a^2-b^2)\sigma^2=0$, 即$a=\pm b$.

4.34 设随机变量 (X,Y)在区域$D=\left\{(x,y)|x^2+y^2\leq 1\right\}$ 服从均匀分布,试证

明: X与Y不相关也不相互独立.

解：因为(X,Y)在区域$D=\{(x,y)|x^2+y^2\leq 1\}$ 服从均匀分布，

所以(X,Y)的联合概率密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{\pi}, & (x,y)\in D;\\ 0, & (x,y)\notin D.\end{cases}$$

又因为$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y=\begin{cases}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\dfrac{1}{\pi}\mathrm{d}y, & -1<x<1;\\ 0, & \text{其他},\end{cases}=\begin{cases}\dfrac{2\sqrt{1-x^2}}{\pi}, & -1<x<1;\\ 0, & \text{其他}.\end{cases}$

$$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x=\begin{cases}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\dfrac{1}{\pi}\mathrm{d}x, & -1<y<1;\\ 0, & \text{其他},\end{cases}=\begin{cases}\dfrac{2\sqrt{1-y^2}}{\pi}, & -1<y<1;\\ 0, & \text{其他}.\end{cases}$$

显然$f(x,y)\neq f_X(x)f_Y(y)$, 随机变量X,Y不独立。

又因为$\mathrm{Cov}(X,Y)=\mathrm{E}(XY)-\mathrm{E}(X)\mathrm{E}(Y)$

$$=\int_{-1}^{1}\mathrm{d}x\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}xy\cdot\frac{1}{\pi}\mathrm{d}y-\int_{-1}^{1}\mathrm{d}x\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}x\cdot\frac{1}{\pi}\mathrm{d}y\cdot\int_{-1}^{1}\mathrm{d}x\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}y\cdot\frac{1}{\pi}\mathrm{d}y$$

$=0,$

所以，随机变量X,Y不相关。

4.35 设随机变量$X\sim B(4,0.5),Y\sim N(0,1),\mathrm{Cov}(X,Y)=-1$,求$V=4X+3Y+1$与$W=-2X+4Y$的方差与协方差.

解：因为$X\sim B(4,0.5),Y\sim N(0,1)$，所以

$\mathrm{E}(X)=2,\ \mathrm{D}(X)=1,\ \mathrm{E}(Y)=0,\ \mathrm{D}(Y)=1,$

故$\mathrm{D}(V)=\mathrm{D}(4X+3Y+1)=16\mathrm{D}(X)+24\mathrm{Cov}(X,Y)+9\mathrm{D}(Y)=1.$

$\mathrm{D}(W)=\mathrm{D}(-2X+4Y)=4\mathrm{D}(X)-16\mathrm{Cov}(X,Y)+16\mathrm{D}(Y)=36.$

$\mathrm{D}(V+W)=\mathrm{D}(2X+7Y+1)=4\mathrm{D}(X)+28\mathrm{Cov}(X,Y)+49\mathrm{D}(Y)=25.$

$\mathrm{Cov}(X,Y)=\dfrac{\mathrm{D}(V+W)-\mathrm{D}(V)-\mathrm{D}(W)}{2}=-6.$

4.36 设 X与Y都是标准化随机变量,它们的相关系数$\rho_{X,Y}=0.5$,令$Z_1=aX$, $Z_2=bX+cY$, 试确定a,b,c 的值,使$\mathrm{D}(Z_1)=\mathrm{D}(Z_2)=1$,且$Z_1$ 与Z_2不相关.

解：因为$\mathrm{D}(Z_1)=a^2$,

$\mathrm{D}(Z_2)=b^2\mathrm{D}(X)+c^2\mathrm{D}(Y)+2bc\mathrm{Cov}(X,Y)=b^2+bc+c^2$,

$\mathrm{Cov}(Z_1,Z_2)=\mathrm{Cov}(aX,bX+cY)=ab+0.5ac.$

要使$\mathrm{D}(Z_1)=\mathrm{D}(Z_2)=1$,且$Z_1$ 与Z_2不相关，只要

$$\begin{cases} a^2=1 \\ b^2+bc+c^2=1, \\ ab+0.5ac=0, \end{cases}$$

解之得：$\begin{cases} \text{a=1} \\ \text{b=}\dfrac{\sqrt{3}}{3} \\ \text{c=}\dfrac{-2\sqrt{3}}{3} \end{cases}$, $\begin{cases} \text{a=-1} \\ \text{b=}\dfrac{\sqrt{3}}{3} \\ \text{c=}\dfrac{-2\sqrt{3}}{3} \end{cases}$, $\begin{cases} \text{a=1} \\ \text{b=}\dfrac{-\sqrt{3}}{3} \\ \text{c=}\dfrac{2\sqrt{3}}{3} \end{cases}$, $\begin{cases} \text{a=-1} \\ \text{b=}\dfrac{-\sqrt{3}}{3} \\ \text{c=}\dfrac{2\sqrt{3}}{3} \end{cases}$.

4.37 已知随机变量$(X,Y)\sim N(1,0;9,16;-0.5)$,设$Z=X/3+Y/2$.求

(1)Z的数学期望和方差; (2) X与Z的相关系数.

解：因为$(X,Y)\sim N(1,0;9,16;-0.5)$,所以$\mathrm{E}(X)=1$, $\mathrm{E}(Y)=0\,\mathrm{D}(X)=9$, $\mathrm{D}(Y)=16$, $\rho(X,Y)=-0.5$.

$\mathrm{Cov}(X,Y)=\rho(X,Y)\sqrt{\mathrm{D}(X)\mathrm{D}(Y)}=-6.$

(1)$\mathrm{E}(Z)=\dfrac{1}{3}\mathrm{E}(X)+\dfrac{1}{2}\mathrm{E}(Y)=\dfrac{1}{3}.$

$\mathrm{D}(Z)=\dfrac{1}{9}\mathrm{D}(X)+\dfrac{1}{3}\mathrm{Cov}(X,Y)+\dfrac{1}{4}\mathrm{D}(Y)=3.$

(2)$\mathrm{Cov}(X,Z)=\mathrm{Cov}(X,X/3+Y/2)=\dfrac{1}{3}\mathrm{D}(X)+\dfrac{1}{2}\mathrm{Cov}(X,Y)=0.$

4.38 某箱子装有100件产品,其中一、二、三等品的数量分别为80、10和10件,现在

从中随机抽取一件,记

$$X_i = \begin{cases} 1, & \text{若抽到}i\text{等品;} \\ 0, & \text{其他.} \end{cases} \qquad i = 1,2,3.$$

试求(1)随机变量X_1与X_2的联合概率分布律;

(2)随机变量X_1与X_2的相关系数ρ_{X_1,X_2}.

解:(1)因为$P\{X_1 = 0, X_2 = 0\} = \dfrac{10}{100} = 0.1$;

$P\{X_1 = 0, X_2 = 1\} = \dfrac{10}{100} = 0.1$;

$P\{X_1 = 1, X_2 = 0\} = \dfrac{80}{100} = 0.8$;

$P\{X_1 = 1, X_2 = 1\} = 0$;

所以(X, Y)的联合概率分布律为

X_1 \ X_2	0	1
0	0.1	0.1
1	0.8	0

(2)$\mathrm{E}(X_1) = 0.8$, $\mathrm{E}(X_2) = 0.1$, $\mathrm{E}(X_1X_2) = 0$, $\mathrm{Cov}(X_1, X_2) = -0.08$,

$\mathrm{D}(X_1) = \mathrm{E}(X_1^2) - \mathrm{E}^2(X_1) = 0.16$, $\mathrm{D}(X_2) = \mathrm{E}(X_2^2) - \mathrm{E}^2(X_2) = 0.09$,

故$\rho(X_1, X_2) = \dfrac{\mathrm{Cov}(X_1, X_2)}{\sqrt{\mathrm{D}(X_1)\mathrm{D}(X_2)}} = -\dfrac{2}{3}$.

4.39 假设二维随机变量 (X, Y)在矩形$G = \{(x, y) : 0 \le x \le 2, 0 \le y \le 1\}$上服从均匀分布.记

$$U = \begin{cases} 0, & \text{若}X \le Y; \\ 1, & \text{若}X > Y, \end{cases} \qquad V = \begin{cases} 0, & \text{若}X \le 2Y; \\ 1, & \text{若}X > 2Y. \end{cases}$$

求U与V的相关系数$\rho_{U,V}$.

解：因为(X,Y)在矩形$G=\{(x,y):0\le x\le 2,0\le y\le 1\}$上服从均匀分布,所以$(X,Y)$的联合概率密度函数为

$$f(x,y)=\begin{cases}\dfrac{1}{2}, & (x,y)\in G;\\[2ex] 0, & (x,y)\notin G.\end{cases}$$

又因为$P\{U=0,V=0\}=P\{X\le Y,X\le 2Y\}=P\{X\le Y\}=\dfrac{1}{4}$;

$$P\{U=0,V=1\}=P\{X\le Y,X>2Y\}=0;$$

$$P\{U=1,V=0\}=P\{X>Y,X\le 2Y\}=P\{Y<X\le 2Y\}=\frac{1}{4};$$

$$P\{U=1,V=1\}=P\{X>Y,X>2Y\}=P\{X>2Y\}=\frac{1}{2};$$

所以(U,V)的联合概率分布律为

U \ V	0	1
0	0.25	0
1	0.25	0.5

又因为$\mathrm{E}(U)=0.75$, $\mathrm{E}(V)=0.5$,

$$\mathrm{Cov}(UV)=\mathrm{E}(UV)-\mathrm{E}(U)\mathrm{E}(V)=0.5-0.75\cdot 0.5=\frac{1}{8}.$$

$$\mathrm{D}(U)=\frac{3}{16},\ \mathrm{D}(V)=\frac{1}{4}.$$

故 $\rho(U,V)=\dfrac{\mathrm{Cov}(U,v)}{\sqrt{\mathrm{D}(U)\mathrm{D}(V)}}=\dfrac{\sqrt{3}}{3}$.

4.40 设 A和 B是随机试验E的两个随机事件,且$P(A)>0$, $P(B)>0$,并定义随机变量 X与Y如下:

$$X=\begin{cases}1 & 若A发生;\\[2ex] 0 & 若A不发生,\end{cases}\qquad Y=\begin{cases}1 & 若B发生;\\[2ex] 0 & 若B不发生.\end{cases}$$

证明若$\rho_{X,Y}=0$,则 X与Y 必定相互独立.

解：因为(X,Y)的联合概率分布律为

$U \backslash V$	0	1
0	$P(\overline{A}\,\overline{B})$	$P(\overline{A}B)$
1	$P(A\overline{B})$	$P(AB)$

所以$\mathrm{Cov}(X,Y)=\mathrm{E}(XY)-\mathrm{E}(X)\mathrm{E}(Y)=P(AB)-P(A)P(B)$,

若$\rho_{X,Y}=0$，则$\mathrm{Cov}(X,Y)=P(AB)-P(A)P(B)=0,$， 即$P(AB)=P(A)P(B)$,

故X与Y 相互独立.

4.41 已知随机变量 X与Y的协方差矩阵为$\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$,试求$U=X-2Y$与$V=2X-Y$的相关系数.

解：根据协方差矩阵的定义知，$\mathrm{D}(X)=1$, $\mathrm{D}(Y)=4$, $\mathrm{Cov}(X,Y)=1$,

所以$\mathrm{Cov}(U,V)=\mathrm{Cov}(X-2Y,2X-Y)=2\mathrm{D}(X)-5\mathrm{Cov}(X,Y)+2\mathrm{D}(Y)=5$,

$\mathrm{D}(U)=\mathrm{D}(X-2Y)=\mathrm{D}(X)-4\mathrm{Cov}(X,Y)+4\mathrm{D}(Y)=13$,

$\mathrm{D}(V)=\mathrm{D}(2X-Y)=4\mathrm{D}(X)-4\mathrm{Cov}(X,Y)+\mathrm{D}(Y)=4$,

故$\rho(U,V)=\dfrac{\mathrm{Cov}(U,V)}{\sqrt{\mathrm{D}(U)\mathrm{D}(V)}}=\dfrac{5\sqrt{13}}{26}$.

习题5

5.1 设随机变量序列$X_1,X_2,\cdots,X_n,\cdots$相互独立,且概率分布律为

$$P\{X_n=\pm 2^n\}=\frac{1}{2^{2n+1}},\quad P\{X_n=0\}=1-\frac{1}{2^{2n}},\ n=1,2,\cdots,$$

证明随机变量$\{X_n\}$服从大数定律.

解：因为相互独立随机变量序列X_n的数学期望和方差分别为

$E(X_n)=0,\ D(X_n)=2\cdot 2^{2n}\cdot\dfrac{1}{2^{2n+1}}=1,$

故随机变量$\{X_n\}$服从切比雪夫大数定律.

5.2 设连续型随机变量X的概率密度函数为$f(x),\ x\in \mathrm{R}$,若$\lambda>0$,$Y=\mathrm{e}^{\lambda X}$. 证明:对任意实数a,有$P\{X\geq a\}\leq \mathrm{e}^{-\lambda a}\mathrm{E}(Y)$.

证明：因为对任意实数a,

$\mathrm{E}(Y)=\int_{\mathrm{R}}\mathrm{e}^{\lambda x}f(x)\mathrm{d}x\geq\int_a^{+\infty}\mathrm{e}^{\lambda x}f(x)\mathrm{d}x\geq \mathrm{e}^{\lambda a}\int_a^{+\infty}f(x)\mathrm{d}x=\mathrm{e}^{\lambda a}P\{X\geq a\},$

所以对任意实数a,有$P\{X\geq a\}\leq \mathrm{e}^{-\lambda a}\mathrm{E}(Y)$.

5.3 设随机变量X的概率密度函数为

$$f(x)=\begin{cases}\dfrac{x^m}{m!}\mathrm{e}^{-x}, & x\geq 0;\\ \\ 0, & x<0.\end{cases}$$

其中m为正整数, 试用切比雪夫不等式估计概率$P\{0<X<2(m+1)\}$.

解：因为$\mathrm{E}(X)=\int_{\mathrm{R}}xf(x)\mathrm{d}x=\int_0^{+\infty}x\cdot\dfrac{x^m}{m!}\mathrm{e}^{-x}\mathrm{d}x=m+1,$

$\mathrm{D}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)=(m+2)(m+1)-(m+1)^1=m+1.$

根据切比雪夫不等式得

$P\{0<X<2(m+1)\}=P\{-(m+1)<X-\mathrm{E}(X)<m+1\}$

$\geq 1-\dfrac{\mathrm{D}(X)}{(m+1)^2}=\dfrac{m}{m+1}.$

5.4 设$g(x)$在区间$[0,1]$上连续,并记$I=\int_0^1 g(x)\mathrm{d}x$,设随机变量$X\sim U(0,1), X_1,X_2,\cdots,X_n,\cdots$相互独立且与$X$同分布,令$Y_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}g(X_i)$.

(1)求$\mathrm{E}(Y_n),\mathrm{D}(Y_n)$,并设$\mathrm{D}[g(X)]=\sigma^2<C$, 试证明当$n\to\infty$时,$Y_n$依概率收敛于$I$.

(2)对任意$\varepsilon>0$,利用中心极限定理估计$P\{|Y_n-I|\leq\varepsilon\}$.

解：(1)因为$X_1,X_2,\cdots,X_n,\cdots$相互独立且与$X$同分布，所以$g(X_1),g(X_2),\cdots,g(X_n),\cdots$也相互独立且与$g(X)$同分布,由于

$\mathrm{E}(Y_n)=\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}g(X_i)\right)=\mathrm{E}(g(X))=I,$
$\mathrm{D}(Y_n)=\mathrm{D}\left(\frac{1}{n}\sum_{i=1}^{n}g(X_i)\right)=\frac{\mathrm{D}(g(X))}{n}=\frac{\sigma^2}{n}<C,$
故根据切比雪夫大数定律知，当$n\to\infty$时,Y_n依概率收敛于I.

(2)对任意$\varepsilon>0$,由中心极限定理得$Y_n\dot{\sim}N(I,\frac{\sigma^2}{n})$, 因此

$P\{|Y_n-I|\le\varepsilon\}\simeq 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sigma}\right)-1.$

5.5 设$X_1,X_2,\cdots,X_n,\cdots$为一列独立同分布随机变量序列,且在区间$[0,\beta](\beta>0)$上服从均匀分布, 令$\xi_n=\max\{X_1,X_2,\cdots,X_n\}$, 证明: $\xi_n\xrightarrow{P}\beta$.

解：由3.35知，ξ_n的概率密度函数为

$$f_{\xi_n}(x)=\begin{cases}\frac{n}{\beta^n}x^{n-1}, & x\in[0,\beta];\\ 0, & y\notin[0,\beta].\end{cases}$$

因为$\mathrm{E}(\xi_n)=\int_0^\beta x\cdot\frac{nx^{n-1}}{\beta^n}\mathrm{d}x=\frac{n}{n+1}\beta,$
$\mathrm{D}(\xi_n)=\int_0^\beta x^2\cdot\frac{nx^{n-1}}{\beta^n}\mathrm{d}x-\left(\frac{n}{n+1}\beta\right)^2=\frac{n}{(n+1)^2(n+2)}\beta^2,$
对任意$\varepsilon>0$,$P\{|\xi_n-\beta|\ge\varepsilon\}=P\{\xi_n-\frac{n}{n+1}\beta-\frac{1}{n+1}\beta\ge\varepsilon\}$
$\le P\left\{|\xi_n-\frac{n}{n+1}\beta|+\frac{1}{n+1}\beta\ge\varepsilon\right\}$
$P\left\{|\xi_n-\frac{n}{n+1}\beta|\ge\varepsilon-\frac{1}{n+1}\beta\right\}$
$\le\frac{\mathrm{D}(\xi_n)}{\left(\varepsilon-\frac{1}{n+1}\beta\right)^2}=\frac{n\beta^2}{(n+1)^2(n+2)\left(\varepsilon-\frac{1}{n+1}\beta\right)^2}\xrightarrow{n\to+\infty}0,$
所以$\xi_n\xrightarrow{P}\beta$.

5.6 设$X_1,X_2,\cdots,X_n,\cdots$为一列独立同分布随机变量序列, $\mathrm{E}(X_i)=\mu,\mathrm{D}(X_i)=\sigma^2,i=1,2,\cdots$, 令$Y_n=\frac{2}{n(n+1)}\sum_{i=1}^{n}iX_i$,证明随机变量序列$\{Y_n\}$ 依概率收敛于μ.

证明：因为$X_1,X_2,\cdots,X_n,\cdots$为一列独立同分布随机变量序列, $\mathrm{E}(X_i)=\mu,\mathrm{D}(X_i)=\sigma^2,i=1,2,\cdots$,所以

$\mathrm{E}(Y_n) = \dfrac{2}{n(n+1)}\sum\limits_{i=1}^{n} i\mathrm{E}(X_i) = \dfrac{2}{n(n+1)}\sum\limits_{i=1}^{n} i\mu = \mu,$

$\mathrm{D}(Y_n) = \dfrac{4}{n^2(n+1)^2}\sum\limits_{i=1}^{n} i^2\mathrm{D}(X_i) = \dfrac{4}{n^2(n+1)^2}\sum\limits_{i=1}^{n} i^2\sigma^2 = \dfrac{2(2n+1)}{3n(n+1)}\sigma^2 \le \sigma^2,$

根据切比雪夫大数定律知，随机变量序列$\{Y_n\}$ 依概率收敛于μ.

5.7 对敌人的防御地段进行100次炮轰,每次炮轰命中的炮弹数服从同一分布,其数学期望为2,均方差为1.5. 若各次轰击命中的炮弹数相互独立,在100 次轰击中,求(1)有180到220发炮弹命中目标的概率; (2)至少命中180发炮弹的概率.

解：以X_i表示第i次命中的炮弹数，则$\mathrm{E}(X_i)=2, \mathrm{D}(X_i)=1.5^2, i=1,\cdots,100.$

令Y表示100次炮轰命中的炮弹数，则$Y=\sum\limits_{i=1}^{100} X_i, \mathrm{E}(Y)=100\cdot 2=200, \mathrm{D}(Y)=100\cdot 1.5^2=15^2,$

由中心极限定理知，$Y \dot\sim N(200,15^2)$。

$(1)P\{180 \le Y \le 220\} \simeq \Phi\left(\dfrac{220-200}{15}\right) - \Phi\left(\dfrac{180-200}{15}\right) = 0.8175.$

$(2)P\{X \ge 180\} = 1 - P\{X < 180\} \simeq 1 - \Phi\left(\dfrac{180-200}{15}\right) = 0.9088.$

5.8 根据以往经验,某种电器元件的寿命服从均值为100小时的指数分布.现随机抽取16只,假设它们的寿命相互独立. 求这16只元件寿命的总和大于1920小时的概率.

解：以X_i表示第i只元件的寿命，则$\mathrm{E}(X_i)=100, \mathrm{D}(X_i)=100^2, i=1,\cdots,16.$

令Y表示16只元件的寿命总和，则

$Y=\sum\limits_{i=1}^{16} X_i, \mathrm{E}(Y)=16\cdot 100=1600, \mathrm{D}(Y)=16\cdot 100^2=400^2,$

由中心极限定理知，$Y \dot\sim N(1600,400^2).$

$P\{Y>1920\} = 1-P\{Y\le 1920\} \simeq 1-\Phi\left(\dfrac{1920-1600}{400}\right) = 0.2119.$

5.9 检验员逐个检查某产品,每查一个需用10秒钟. 但有的产品需重复检查一次,再用去10秒钟.若产品需重复检查的概率为0.5, 求检验员在8 小时内检查的产品数多于1900个的概率.

解：以X_i表示第i个20秒检查的产品个数，则$X_i(i=1,\cdots,1440)$的概率分布律为

X_i	1	2
P_{X_i}	0.5	0.5

检验员在8 小时内检查的产品数为$Y=\sum\limits_{i=1}^{1440}X_i$,

因为$\mathrm{E}(Y)=1440\cdot(0.5+1)=2160,\ \mathrm{D}(Y)=1440\cdot(0.5+2-1.5^2)=360$,

所以$Y\dot{\sim}N(2160,360)$.

$P\{Y>1900\}=1-P\{Y\le 1900\}\simeq 1-\Phi\left(\dfrac{1900-2160}{\sqrt{360}}\right)=1.(?)$

5.10 一加法器同时收到20个噪声电压$V_i,(i=1,2,\cdots,20)$, 并设它们是相互独立的随机变量,且都在区间[0,10] 上服从均匀分布,记$V=\sum\limits_{i=1}^{20}V_i$, 求$P\{V>105\}$.

解：因为$\mathrm{E}(V_i)=5,\ \mathrm{D}(V_i)=\dfrac{25}{3},i=1,\cdots,20.$

而$V=\sum\limits_{i=1}^{20}V_i\dot{\sim}N\left(100,\dfrac{500}{3}\right)$,

故$P\{V>105\}=1-P\{V\le 105\}\simeq 1-\Phi\left(\dfrac{105-100}{\sqrt{500/3}}\right)=0.3492.$

5.11 计算机在进行加法运算时,对每个加数取整(取为最接近于它的整数).设所有的取整误差是相互独立的, 且都在$[-0.5,0.5]$上服从均匀分布.

(1)若取1 500个数相加,问误差总和绝对值超过15的概率是多少?

(2)可将几个数加在一起使得误差总和的绝对值小于10的概率为0.90?

解：以X_i表示第i个数的取整误差，则$X_i(i=1,\cdots,1440)\sim U(-0.5,0.5),i=1,\cdots,1500$,

(1)用随机变量Y表示1500个数的取整误差，则$Y=\sum\limits_{i=1}^{1500}X_i$,

因为$\mathrm{E}(Y)=1500*0=0,\ \mathrm{D}(Y)=1500*\dfrac{1}{12}=125$,

所以$Y\dot{\sim}N\left(0,125\right)$,

故$P\{|Y|>15\}=1-P\{|Y|\le 15\}\simeq 2-2\Phi\left(\frac{15}{\sqrt{125}}\right)=0.1797.$

(2)由于$Z=\sum\limits_{i=1}^{n}X_i\dot{\sim}N\left(0,\frac{n}{12}\right),$

要使$P\{|Y|<10\}=2\Phi\left(\frac{10}{\sqrt{n/12}}\right)-1=0.9,$

只要$\Phi\left(\frac{10}{\sqrt{n/12}}\right)=0.95$,即$\frac{10}{\sqrt{n/12}}=1.645,$

解之得$n=443.$

5.12 一批电子管,其中一等品占20%,,现用重复抽取的方式从中抽取10000只进行测试.

(1)求所抽的10000只电子管中一等品不少于1 900只且不多于2100只的概率.

(2)应抽取多少只,才能以95%的概率保证至少抽到2000只一等品?

解：设X_i表示第i只电子管是一等品，则$X_i\sim B(1,0.2),i=1,\cdots,10000.$

10000只电子管中含有的一等品数为$Y=\sum\limits_{i=1}^{10000}X_i\dot{\sim}N\left(2000,40^2\right),$

(1) $P\{1900\le Y\le 2100\}\simeq\Phi\left(\frac{2100-2000}{40}\right)-\Phi\left(\frac{1900-2000}{40}\right)=2\Phi(2.5)-1=0.9876.$

(2)要使$P\{Y\ge 2000\}=1-\Phi\left(\frac{2000-0.2n}{\sqrt{0.16n}}\right)\ge 0.95,$

只要$\Phi\left(\frac{2000-0.2n}{\sqrt{0.16n}}\right)\le 0.05$, 即$\frac{2000-0.2n}{\sqrt{0.16n}}\le -1.645,$

解之得$n\ge 10335.$

5.13 一份俄语试卷由100个单项选择题组成．每题有四项供选择,只有一项正确,答对得1分,答错得0分.求一个对俄语一无所知的人至少得60分的概率.

解：设X_i表示第i题答对数，则$X_i\sim B(1,0.25),i=1,\cdots,100.$ 则一个对俄语一无所知的人至少得60分的概率为

$$P\left\{\sum_{i=0}^{100}X_i\ge 60\right\}=1-\left\{\sum_{i=0}^{100}X_i<60\right\}\simeq 1-\Phi\left(\frac{60-100*0.25}{\sqrt{100*0.25*0.75}}\right)=0.$$

5.14 某电站供应10000户居民用电,设在高峰时每户用电的概率为0.8,且各户的用电

量相互独立.求

(1)同一时刻有8100户以上居民用电的概率;

(2)若每户用电功率为100kW,则电站至少需要供应多少功率才能保证以0.975的概率供应居民用电?

解：令 $X_i = \begin{cases} 1, & \text{若第}i\text{户居民用电;} \\ 0, & \text{否则,} \end{cases} \quad i = 1, 2, \cdots, 10000.$

同一时刻用电总户数可表示为$Y = \sum\limits_{i=1}^{10000} X_i$ 则$Y \dot{\sim} N(8000, 40^2)$

(1)$P\left\{Y \geq 8100\right\} \simeq 1 - \Phi\left(\dfrac{8100 - 8000}{40}\right) = 0.0062.$

(2)假设电站至少供应的功率为$m$kw,依题意得

$P\{100Y \leq m\} \geq 0.975,$ 即$\Phi\left(\dfrac{0.01m - 8000}{40}\right) \geq 0.975,$

等价于$\dfrac{0.01m - 8000}{40} \geq 1.96$，解之得$m \geq 807840(\text{kw}).$

5.15 从某工厂的产品中任取200件,检查结果发现其中有6件次品,问能否相信该厂的次品率不大于1%?（建议删除）

解：令X表示200件中的次品数，若该产品的次品率为1%,则$X \sim B(200, 1\%) \dot{\sim} P(2)$,

有6件次品的概率为$P\{X = 6\} \simeq \dfrac{2^6}{6!}\mathrm{e}^{-2} = 0.012,$

由于该事件发生的概率很小，故不能认为该工厂的次品率不大于1%.

5.16 质检部门抽样检查某种产品质量时,如果发现次品数多于10件,则认为该批产品不合格.设某批产品的次品率为10%,问至少应抽取多少件产品检查才能使该批产品不合格的概率达90%?

解：假设抽取m件产品检查，令

$$X_i=\begin{cases}1, & 若第i件产品是次品;\\ 0, & 否则,\end{cases}\quad i=1,2,\cdots,m.$$

则m件产品中的次品数为$Y=\sum\limits_{i=1}^{m}X_i\sim B(m,0.1)\dot{\sim}N(0.1m,0.09m)$,

该批产品为不合格品的概率为

$P\{Y>10\}=1-P\{Y\le 10\}\simeq 1-\Phi\left(\dfrac{10-0.1m}{0.3\sqrt{m}}\right)\ge 0.9,$

即$\Phi\left(\dfrac{10-0.1m}{0.3\sqrt{m}}\right)\le 0.1$, 等价的有$\dfrac{10-0.1m}{0.3\sqrt{m}}\le -1.28$, 解之得$m\ge 147$.

5.17 有一批建筑用木柱,其中80%长度不少于3m. 现从这批木柱中随机抽取100根,问其中至少有30根短于3m的概率是多少?

解：令

$$X_i=\begin{cases}1, & 若第i根木柱的长度小于3米;\\ 0, & 否则,\end{cases}\quad i=1,2,\cdots,100.$$

则100根木柱中长度小于3米的根数为$Y=\sum\limits_{i=1}^{100}X_i\sim B(100,0.2)\dot{\sim}N(20,4^2)$,

至少有30根短于3m的概率为

$P\{Y\ge 30\}=1-P\{Y30\}1-P\{Y\le 29\}\simeq 1-\Phi\left(\dfrac{29-20}{4}\right)=0.012.$

5.18 一个复杂系统,由100个相互独立的部件组成.在整个运行期间每个部件损坏的概率都为0.10. 为了使整个系统正常工作, 至少必须有85个部件工作,求整个系统正常工作的概率.

解：令

$$X_i=\begin{cases}1, & 若第i个部件正常工作;\\ 0, & 否则,\end{cases}\quad i=1,2,\cdots,100.$$

则100个部件中正常工作的部件数为$Y=\sum\limits_{i=1}^{100}X_i\sim B(100,0.9)\dot{\sim}N(90,3^2)$,

整个系统正常工作的概率为

$P\{Y \geq 85\} = 1 - P\{Y \leq 84\} = 1 - \Phi\left(\frac{84-90}{3}\right) = 0.9772.$

5.19 用切比雪夫不等式确定当掷一均匀铜币时,需投多少次才能保证使得正面出现的频率在0.4 至0.6之间的概率不小于90%, 再用中心极限定理计算同一问题并进行对照.

解：设需要投掷的次数为m,令

$$X_i = \begin{cases} 1, & \text{若第}i\text{次出现正面;} \\ 0, & \text{否则,} \end{cases} \quad i = 1, 2, \cdots, m.$$

则投掷m次铜币出现正面的总次数为$Y = \sum\limits_{i=1}^{m} X_i \sim B(m, 0.5) \dot{\sim} N(0.5m, 0.25m)$,

依题意，要使得$P\left\{0.4 \leq \frac{Y}{m} \leq 0.6\right\} \geq 0.9$,

由切比雪夫不等式知

$P\left\{0.4 \leq \frac{Y}{m} \leq 0.6\right\} = P\left\{-0.1m \leq Y - 0.5m \leq 0.1m\right\}$

$\geq 1 - \frac{\mathrm{D}(Y)}{0.01m^2} = 1 - \frac{25}{m}$,

故只要$1 - \frac{25}{m} \geq 0.9$, 解之得$m \geq 250$.

另外根据中心极限定理知

$P\left\{0.4 \leq \frac{Y}{m} \leq 0.6\right\} = P\{0.4m \leq Y \leq 0.6m\}$

$\simeq \Phi\left(\frac{0.6m - 0.5m}{0.5\sqrt{m}}\right) - \Phi\left(\frac{0.4m - 0.5m}{0.5\sqrt{m}}\right) = 2\Phi(0.2\sqrt{m}) - 1$,

故只要$2\Phi(0.2\sqrt{m}) - 1 \geq 0.9$, 等价的$\Phi(0.2\sqrt{m}) \geq 0.95$,查表得$0.2\sqrt{m} \geq 1.645$,

解之得$m \geq 68$.

5.20 在一个罐子中,装有10个编号分别为0～ 9的同样的球,从罐中有放回地抽取若干次,每次抽一个,并记下号码. 设

$$X_k = \begin{cases} 1, & \text{第}k\text{次取到号码0;} \\ 0, & \text{第}k\text{次取不到号码0.} \end{cases} \quad k = 1, 2, \cdots,$$

(1) 至少应抽取多少次才能使“0”出现的频率在0.09-0.11之间的概率至少是0.95?

(2)用中心极限定理计算在100次抽取中,数码“0”出现的次数在7和13之间的概率.

解：设应抽取m次才能使“0”出现的频率在0.09-0.11之间的概率至少是0.95，

依题意, 随机变量$Y=\sum\limits_{i=1}^{m}X_i$表示$m$次抽取中出现"0"的总次数，且

$Y\sim B(m,0.1)\dot{\sim}N(0.1m,0.09m)$.

(1) 由切比雪夫不等式知

$P\left\{0.09\leq\dfrac{Y}{m}\leq 0.11\right\}=P\left\{-0.01m\leq Y-0.1m\leq 0.01m\right\}$

$\geq 1-\dfrac{\mathrm{D}(Y)}{0.01^2m^2}=1-\dfrac{900}{m}$,

故要使$P\left\{0.09\leq\dfrac{Y}{m}\leq 0.11\right\}\geq 0.95$,

只要$1-\dfrac{900}{m}\geq 0.95$, 解之得$m\geq 18000$.（与答案相差很多）

若用中心极限定理，则

$P\left\{0.09\leq\dfrac{Y}{m}\leq 0.11\right\}=P\left\{-0.01m\leq Y-0.1m\leq 0.01m\right\}$

$\simeq 2\Phi\left(\dfrac{0.01m}{0.3\sqrt{m}}\right)-1$,

故要使$P\left\{0.09\leq\dfrac{Y}{m}\leq 0.11\right\}\geq 0.95$,

只要$2\Phi\left(\dfrac{0.01m}{0.3\sqrt{m}}\right)-1\geq 0.95$, 即$\Phi\left(\dfrac{0.01m}{0.3\sqrt{m}}\right)\geq 0.975$,

等价的$\dfrac{0.01m}{0.3\sqrt{m}}\geq 1.96$, 解之得$m\geq 3458$.

(2) $P\left\{7\leq Y\leq 13\right\}\simeq\Phi\left(\dfrac{13-10}{3}\right)-\Phi\left(\dfrac{7-10}{3}\right)=2\Phi(1)-1=0.6827$.

5.21 某保险公司的多年统计资料表明,在索赔客户中,被盗索赔户占20%. 以X表示在随机抽查的100个索赔户中, 因被盗向保险公司索赔的户数.

(1)写出X的概率分布;

(2)利用中心极限定理,求被盗索赔户不少于14户且不多于30户的概率.

解：(1) 因为$X\sim B(100,0.2)$, 所以随机变量X的概率分布律为

$P\{X=k\}=C_{100}^{k}0.2^{k}0.8^{100-k},k=0,1,\cdots,100$.

(2)$P\{14 \geq X \leq 30\} = P\{13 < X \leq 30\} \simeq \Phi\left(\dfrac{30-20}{4}\right) - \Phi\left(\dfrac{13-20}{4}\right) = 0.9537.$

5.22 甲地到乙地有两个汽车站各有一班公共汽车同时开出,假定每个发车时刻有100位乘客等可能地选乘其中一个汽车站乘车,为保证以95%的概率使旅客有座位,每车至少要设几个座位?

解：设有X个乘客选择第1个汽车站，则有$100-X$个乘客选择第2个汽车站，且$X \sim B(100, 0.5) \dot{\sim} N(50, 25)$,假定汽车的座位数为$m$,根据题意得：

$$\begin{cases} P\{X \leq m\} & \geq 0.95, \\ P\{100 - X \leq m\} & \geq 0.95. \end{cases} \quad \text{等价地} \quad \begin{cases} \Phi\left(\dfrac{m-50}{5}\right) & \geq 0.95, \\ 1 - \Phi\left(\dfrac{50-m}{5}\right) & \geq 0.95. \end{cases}$$

即 $\dfrac{m-50}{5} \geq 1.645$, 解之得 $m \geq 59.$

5.23 设某农贸市场某种商品每日价格的变化是相互独立且均值为0, 方差为$\sigma^2 = 2$的随机变量Y_n,并满足

$$X_{n+1} = X_n + Y_n, \quad n = 1, 2, \cdots,$$

其中X_n是第n天该商品的价格.如果今天的价格为100,求18天后该商品的价格在96与104之间的概率.

解：因为$X_{19} = X_1 + \sum\limits_{i=1}^{18} Y_i = 100 + \sum\limits_{i=1}^{18} Y_i$,而$\sum\limits_{i=1}^{18} Y_i \dot{\sim} N(0, 36)$.

所以$P\{96 \leq X_{19} \leq 104\} = P\{-4 \leq \sum\limits_{i=1}^{18} Y_i \leq 4\} = 2\Phi\left(\dfrac{4}{6}\right) - 1 = 0.495.$

习题6

6.1 某厂生产的电容器的使用寿命服从指数分布,但参数λ未知. 现随机抽取n只电容器,测得其实际使用寿命.试问本题中什么是总体,什么是样本?并求样本的联合概率分布.

解：所有电容器的使用寿命是总体，抽取的n只电容器的使用寿命是样本，联合概率分布为

$$f(x_1,\cdots,x_n)=\prod_{i=1}^{n}f(x_i)=\begin{cases}\prod\limits_{i=1}^{n}\left(\lambda e^{-\lambda x_i}\right), & x_i\geq 0, i=1,\cdots,n;\\ 0, & 否则,\end{cases}$$

$$=\begin{cases}\lambda^n e^{-\lambda\sum\limits_{i=1}^{n}x_i}, & x_i\geq 0, i=1,\cdots,n;\\ 0, & 否则.\end{cases}$$

6.2 设$(x_1,x_2,\cdots,x_5)=(2,7,3,5,8)$是来自总体$X$的一个容量为5的样本. 求样本均值$\overline{x}$,样本方差$s^2$及经验分布函数$F_n(x)$,并做出$F_n(x)$的图像.

解：$\overline{x}=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i=5,\qquad s^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\overline{x})=6.5.$

$$F_n(x)\begin{cases}0, & x<2;\\ \frac{1}{5}, & 2\leq x<3;\\ \frac{2}{5}, & 3\leq x<5;\\ \frac{3}{5}, & 5\leq x<7;\\ \frac{4}{5}, & 7\leq x<8;\\ 1, & x\geq 8.\end{cases}$$

6.3 设$X\sim B(1,p)$, p未知, $(X_1,X_2,\cdots X_n)$是来自X的样本.

(1) 写出样本的联合概率分布;

(2)指出$X_1+X_2, 2p+X_1, (pX_1+X_n)^2, \dfrac{X_1}{X_2}$中哪些是统计量,哪些不是统计量,为什么?

解：(1)因为$P\{X=x\}=p^x(1-p)^{1-x}, x=0,1.$ 所以

$$P\{X_1=x_1,\cdots,X_n=x_n\}=\prod_{i=1}^{n}P\{X_i=x_i\}=\prod_{i=1}^{n}\left(p^{x_i}(1-p)^{1-x_i}\right)$$

$$= p^{\sum\limits_{i=1}^{n} x_i}(1-p)^{n-\sum\limits_{i=1}^{n} x_i}.$$

(2)$X_1+X_2, \dfrac{X_1}{X_2}$是统计量，$2p+X_1, (pX_1+X_n)^2$不是统计量。

6.4 设总体X服从正态分布$N(\mu, 3^2)$, $(X_1, X_2, \cdots, X_n)$是取自总体X的样本，写出样本的联合概率密度函数.

解：因为$f(x;\mu) = \dfrac{1}{3\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{18}}$,

所以样本$(X_1, X_2, \cdots, X_n)$的联合概率密度函数为

$$f(x_1, \cdots, x_n; \mu) = \prod_{i=1}^{n} f(x_i) = \frac{1}{(18\pi)^{n/2}}\mathrm{e}^{-\frac{1}{18}\sum\limits_{i=1}^{n}(x_i-\mu)^2}.$$

6.5 设总体X的概率密度函数为

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

$(X_1, X_2, \cdots, X_n)$是来自总体X的样本,试求最小次序统计量$X_{(1)}$和最大次序统计量$X_{(n)}$的概率密度函数.

解：据题意得X的分布函数为

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \begin{cases} 0, & x < 0; \\ x^2, & 0 \le x < 1; \\ 1, & x \ge 1. \end{cases}$$

最小次序统计量$X_{(1)}$的概率密度函数为

$$f_1(x) = n\left(1-F(x)\right)^{n-1} f(x) = \begin{cases} 2nx(1-x^2)^{n-1}, & 0 < x < 1; \\ 0, & \text{其他}. \end{cases}$$

最大次序统计量$X_{(n)}$的概率密度函数为

$$f_n(x) = nF^{n-1}(x)f(x) = \begin{cases} 2nx^{2n-1}, & 0 < x < 1; \\ 0, & \text{其他}. \end{cases}$$

6.6 设$X \sim U(a,b)$, $(X_1, X_2, \cdots, X_n)$是来自X的样本, 试分别求最小次序统计量$X_{(1)}$和最大次序统计量$X_{(n)}$的概率密度函数.

解：因为$X \sim U(a,b)$,所以X的概率密度函数和分布函数分别为

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b]; \\ 0, & x \notin [a,b]. \end{cases} \qquad F(x) = \begin{cases} 0, & x < a; \\ \dfrac{x-a}{b-a}, & a \le x < b; \\ 0, & x \ge b. \end{cases}$$

最小次序统计量$X_{(1)}$的概率密度函数为

$$f_1(x) = n\left(1 - F(x)\right)^{n-1} f(x) = \begin{cases} \dfrac{n(b-x)^{n-1}}{(b-a)^n}, & a \le x \le b; \\ 0, & \text{其他}. \end{cases}$$

最大次序统计量$X_{(n)}$的概率密度函数为

$$f_n(x) = nF^{n-1}(x)f(x) = \begin{cases} \dfrac{n(x-a)^{n-1}}{(b-a)^n}, & a \le x \le b; \\ 0, & \text{其他}. \end{cases}$$

6.7 设总体X服从正态分布$N(52, 6.3^2)$, $(X_1, \cdots, X_{36})$是来自总体X的样本，样本均值为$\overline{X}$,求$P\{50.8 < \overline{X} < 53.8\}$.

解：因为$X \sim N(52, 6.3^2)$,所以$\overline{X} \sim N\left(52, \dfrac{6.3^2}{36}\right)$, 故

$$P\{50.8 < \overline{X} < 53.8\} = \Phi\left(\frac{53.8-52}{1.05}\right) - \Phi\left(\frac{50.8-52}{1.05}\right) = 0.833.$$

6.8 设$X \sim N(\mu, 1)$, $(X_1, X_2, \cdots, X_n)$是来自X的样本,$\overline{X}$为样本均值,试问样本容量n 应取多大,才能使

(1) $\mathrm{E}(|\overline{X} - \mu|^2) \le 0.1$;

(2) $P(|\overline{X} - \mu| \le 0.1) \ge 0.95$.

解：因为$X \sim N(\mu, 1)$,所以$\overline{X} \sim N\left(\mu, \dfrac{1}{n}\right)$, $\overline{X} - \mu \sim N\left(0, \dfrac{1}{n}\right)$, $n(\overline{X} - \mu)^2 \sim \chi_1^2$.

(1)由于$\mathrm{E}(|\overline{X} - \mu|^2) = \dfrac{1}{n}\mathrm{E}(n|\overline{X} - \mu|^2) = \dfrac{1}{n}$,

故要使$\mathrm{E}(|\overline{X}-\mu|^2)\leq 0.1$，只要$\dfrac{1}{n}\leq 0.1$,解之得$n\geq 10$.

(2)由于$P(|\overline{X}-\mu|\leq 0.1)=P(\sqrt{n}|\overline{X}-\mu|\leq 0.1\sqrt{n})=2\Phi(0.1\sqrt{n})-1$,

故要使$P(|\overline{X}-\mu|\leq 0.1)\geq 0.95$,只要$2\Phi(0.1\sqrt{n})-1\geq 0.95$,

即$0.1\sqrt{n}\geq 1.96$,解之得$n\geq 385$.

6.9 设S^2是从正态总体$N(\mu,\sigma^2)$抽取的容量为16的样本方差,其中μ,σ^2未知,求$P\left\{\dfrac{S^2}{\sigma^2}\leq 2.04\right\}$，$\mathrm{D}(S^2)$.

解：因为$\dfrac{15S^2}{\sigma^2}\sim\chi^2_{15}$, 所以

$P\left\{\dfrac{S^2}{\sigma^2}\leq 2.04\right\}=P\left\{\dfrac{15S^2}{\sigma^2}\leq 30.6\right\}=1-P\left\{\dfrac{15S^2}{\sigma^2}\geq 30.6\right\}=0.99.$

$\mathrm{D}(S^2)=\dfrac{\sigma^4}{15^2}\mathrm{D}\left(\dfrac{15S^2}{\sigma^2}\right)=\dfrac{2\sigma^4}{15}.$

6.10 设$(X_1,\cdots,X_{2n})$来自正态总体$N(\mu,\sigma^2)$的样本,令$\overline{X}=\dfrac{1}{2n}\sum\limits_{i=1}^{2n}X_i$, $Y=\sum\limits_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2$,求$\mathrm{E}(Y)$.

解：令$Z_i=X_i+X_{n+i},i=1,\cdots,n$, 则$\overline{Z}=\dfrac{1}{n}\sum\limits_{i=1}^{n}Z_i=2\overline{X}$, $(n-1)S_Z^2=Y$.

因为$X_i\sim N(\mu,\sigma^2)$, 所以$Z_i\sim N(2\mu,2\sigma^2)$,

根据抽样分布定理知 $\dfrac{Y}{2\sigma^2}=\dfrac{(n-1)S_Z^2}{2\sigma^2}\sim\chi^2_{n-1}$,

故$\mathrm{E}(Y)=2\sigma^2\mathrm{E}\left(\dfrac{Y}{2\sigma^2}\right)=2(n-1)\sigma^2$.

6.11 设(X_1,X_2,X_3,X_4)是来自正态总体$N(0,2^2)$的样本,$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$,求常数$a,b$使得统计量$X$服从$\chi^2$分布, 自由度是多少？

解：因为$X_i\sim N(0,2^2),i=1,2,3,4$,

所以 $\sqrt{a}(X_1-2X_2)\sim N(0,20a)$, $\sqrt{b}(3X_3-4X_4)\sim N(0,100b)$,

由于$\sqrt{a}(X_1-2X_2)$与$\sqrt{b}(3X_3-4X_4)$相互独立，故

要使$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$服从$\chi^2$分布，

只要$\sqrt{a}(X_1-2X_2)\sim N(0,1)$, $\sqrt{b}(3X_3-4X_4)\sim N(0,1)$,

即 $20a=1, 100b=1$, 解之得 $a=\dfrac{1}{20}, b=\dfrac{1}{100}$.

因为独立平方和的项数为2，所以自由度是2.

6.12 设$(X_1, X_2, \cdots, X_6)$是来自正态总体$N(0,1)$的样本,试确定常数c, 使随机变量$Y=c[(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2]$服从$\chi^2$分布, 并求其自由度.

解：因为$X_i \sim N(0,1), i=1,\cdots,6$,

所以$\sqrt{c}(X_1+X_2+X_3) \sim N(0,3c)$, $\sqrt{c}(X_4+X_5+X_6) \sim N(0,3c)$,

由于$\sqrt{c}(X_1+X_2+X_3)$与$\sqrt{c}(X_4+X_5+X_6)$相互独立，故

要使$Y=c[(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2]$服从$\chi^2$分布，

只要$3c=1$, 即$c=\dfrac{1}{3}$.

因为独立平方和的项数为2，所以自由度是2.

6.13 设总体$X \sim N(0,\sigma^2)$, $(X_1, X_2, \cdots, X_n, X_{n+1}, \cdots, X_{n+m})$是来自总体$X$的样本,试求统计量$Y=\dfrac{\sqrt{m}\sum\limits_{i=1}^{n} X_i}{\sqrt{n}\sqrt{\sum\limits_{i=n+1}^{n+m} X_i^2}}$的分布.

解：因为$X_i \sim N(0,\sigma^2), i=1,\cdots,n+m$,

所以 $\dfrac{X_i}{\sigma} \sim N(0,1), i=1,\cdots,n+m$,

从而$\dfrac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}\dfrac{X_i}{\sigma} \sim N(0,1)$, $\sum\limits_{i=n+1}^{n+m}\left(\dfrac{X_i}{\sigma}\right)^2 \sim \chi_m^2$,

由于$X_1, \cdots, X_n, X_{n+1}, \cdots, X_{n+m}$相互独立，所以$\dfrac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}\dfrac{X_i}{\sigma}$与$\sum\limits_{i=n+1}^{n+m}\left(\dfrac{X_i}{\sigma}\right)^2$相互独立，

根据t分布的定义知,

$$Y=\frac{\sqrt{m}\sum\limits_{i=1}^{n} X_i}{\sqrt{n}\sqrt{\sum\limits_{i=n+1}^{n+m} X_i^2}}=\frac{\frac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}\dfrac{X_i}{\sigma}}{\sqrt{\sum\limits_{i=n+1}^{n+m}\left(\dfrac{X_i}{\sigma}\right)^2/m}} \sim t_m.$$

6.14 设总体$X \sim N(\mu,\sigma^2)$, $(X_1, X_2, \cdots, X_n, X_{n+1})$是来自总体$X$的样本,令$\overline{X_n}=\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i, S_n^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X_n})^2$, 试确定常数$c$,使随机变量$T=c\dfrac{X_{n+1}-\overline{X_n}}{S_n}$服从$t$分

布,其自由度为多少?

解：因为$X_i \sim N(\mu,\sigma^2), i=1,\cdots,n+1$,且相互独立，

所以X_{n+1}与$\overline{X_n}$独立，且二者也与S_n相互独立，

又因为$X_{n+1}-\overline{X_n} \sim N\left(0,\dfrac{n+1}{n}\sigma^2\right)$，$\dfrac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1}$，

且$X_{n+1}-\overline{X_n}$与$\dfrac{(n-1)S_n^2}{\sigma^2}$相互独立，

根据t分布的定义知，

$$\frac{\frac{\sqrt{n}}{\sqrt{n+1}\sigma}(X_{n+1}-\overline{X_n})}{\sqrt{\frac{(n-1)S_n^2}{\sigma^2}/(n-1)}}=\frac{\sqrt{n}}{\sqrt{n+1}}\frac{X_{n+1}-\overline{X_n}}{S_n}\sim t_{n-1},$$

故取$c=\dfrac{\sqrt{n}}{\sqrt{n+1}}$,可使随机变量$T=c\dfrac{X_{n+1}-\overline{X_n}}{S_n}$服从自由度是$n-1$的$t$分布。

6.15 设随机变量$X \sim N(\mu_1,\sigma^2)$,随机变量$Y \sim N(\mu_2,\sigma^2)$,且X与Y相互独立,$(X_1,X_2,\cdots,X_n)$与$(Y_1,Y_2,\cdots,$

分别来自总体X与Y的样本, 求随机变量

$$(1)\quad \frac{(n-1)(S_1^2+S_2^2)}{\sigma^2};\qquad (2)\quad \frac{n\left[(\overline{X}-\overline{Y})-(\mu_1-\mu_2)\right]^2}{S_1^2+S_2^2}$$

的分布.

解：(1) 因为$X \sim N(\mu_1,\sigma^2)$，$Y \sim N(\mu_2,\sigma^2)$,

根据抽样分布定理知，$\dfrac{(n-1)S_1^2}{\sigma^2} \sim \chi^2_{n-1}$，$\dfrac{(n-1)S_2^2}{\sigma^2} \sim \chi^2_{n-1}$,

又因为S_1^2与S_2^2相互独立，所以

$$\frac{(n-1)(S_1^2+S_2^2)}{\sigma^2}=\frac{(n-1)S_1^2}{\sigma^2}+\frac{(n-1)S_2^2}{\sigma^2}\sim\chi^2_{2n-2}.$$

(2)

6.16 已知X_1,X_2,X_3相互独立且都服从$N(0,\sigma^2)$.证明：$Y=\sqrt{\dfrac{2}{3}}\dfrac{X_1+X_2+X_3}{|X_2-X_3|}$ 服从自由度为1的t分布。

6.17 设$X \sim N(\mu_1,\sigma^2), Y \sim N(\mu_2,\sigma^2), (X_1,X_2,\cdots,X_m)$与$(Y_1,Y_2,\cdots,Y_n)$分别是来自总体$X$和$Y$的样本,它们对应的样本均值分别为$\overline{X}$和$\overline{Y}$,样本方差分别为$S_1^2$和$S_2^2$,令$S^2=$

$\dfrac{m-1}{m+n-2}S_1^2+\dfrac{n-1}{m+n-2}S_2^2, \alpha, \beta$为两个已知常数,试求统计量

$$Z=\frac{\alpha(\overline{X}-\mu_1)+\beta(\overline{Y}-\mu_2)}{\sqrt{\dfrac{\alpha^2}{m}+\dfrac{\beta^2}{n}}S}$$

的分布。

7.1　设总体X的概率密度函数为

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

其中$\alpha>-1$是未知参数,$(X_1,X_2,\ldots,X_n)$是来自X的样本,试求参数α的矩估计和极大似然估计. 现有样本观测值(0.1,0.2,0.9,0.8,0.7,0.7),求参数α的估计值.

解： 由于$EX=\int_0^1 x(\alpha+1)x^{\alpha}\mathrm{d}x=\frac{\alpha+1}{\alpha+2}$, 故令$\bar{X}=\frac{\alpha+1}{\alpha+2}$, 解得矩估计$\hat{\alpha}=\frac{2\bar{X}-1}{1-\bar{X}}$.

似然函数$L(x,\alpha)=\prod_{i=1}^n f(x_i,\alpha)=(\alpha+1)^n\prod_{i=1}^n x_i^{\alpha}$;

取对数$\ln L(x,\alpha)=n\ln(\alpha+1)+\alpha\sum_{i=1}^n\ln x_i$;

令$\frac{\mathrm{d}\ln L}{\mathrm{d}\alpha}=\frac{n}{\alpha+1}+\sum_{i=1}^n\ln x_i=0$

解得最大似然估计$\hat{\alpha}=-1-\frac{n}{\sum_{i=1}^n\ln X_i}$.

将所给观测值代入两个表达式，得$\bar{X}=\frac{3.4}{6}$, 于是α的矩估计值$\hat{\alpha}=\frac{2\bar{X}-1}{1-\bar{X}}=\frac{4}{13}=0.3077$;

α的最大似然估计$\hat{\alpha}=-1-\frac{n}{\sum_{i=1}^n\ln x_i}=0.2112$

7.2　设总体$X\sim B(m,p), 0<p<1,(X_1,X_2,\ldots,X_n)$是取自总体$X$的样本,试求参数$p$的矩估计和极大似然估计.

解： 由于$EX=mp$, 故令$\bar{X}=mp$, 解得矩估计$\hat{p}=\frac{\bar{X}}{m}$.

似然函数

$$L(x,p)=\prod_{i=1}^n P(X_i=x_i)=\prod_{i=1}^n \mathrm{C}_m^{x_i}p^{x_i}(1-p)^{m-x_i}=\prod_{i=1}^n \mathrm{C}_m^{x_i}p^{\sum_{i=1}^n x_i}(1-p)^{nm-\sum_{i=1}^n x_i}$$

取对数$\ln L(x,p)=\ln\prod_{i=1}^n\mathrm{C}_m^{x_i}+\sum_{i=1}^n x_i\ln p+(nm-\sum_{i=1}^n x_i)\ln(1-p)$;

令$\frac{\mathrm{d}\ln L}{\mathrm{d}p}=\frac{\sum_{i=1}^n x_i}{p}-\frac{nm-\sum_{i=1}^n x_i}{1-p}=\frac{n\bar{X}}{p}-\frac{nm-n\bar{X}}{1-p}=0$

解得最大似然估计$\hat{p}=\frac{\bar{X}}{m}$.

7.3 设X具有概率密度函数

$$f(x,\theta)=\begin{cases}\dfrac{2\theta^2}{(\theta^2-1)x^3}, & 1<x<\theta,\\ 0, & \text{其他}.\end{cases}$$

$(X_1,X_2,\ldots,X_n)$是取自总体X的样本,试求参数θ的矩估计.

解由于$EX=\int_1^\theta x\frac{2\theta^2}{(\theta^2-1)x^3}\mathrm{d}x=-\frac{2\theta^2}{(\theta^2-1)}x^{-1}\Big|_1^\theta=\frac{2\theta}{1+\theta}$, 故令$\bar{X}=\frac{2\theta}{1+\theta}$, 解得矩估计$\hat{\theta}=\frac{\bar{X}}{2-\bar{X}}$.

7.4 设总体X服从Gamma分布,其概率密度函数为

$$f(x,\beta)=\begin{cases}\dfrac{\beta^{-(\alpha+1)}}{\Gamma(\alpha+1)}x^\alpha e^{-\frac{x}{\beta}}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

$(X_1,X_2,\ldots,X_n)$是来自总体X的样本,若α已知,试求参数β的极大似然估计量.

解: 似然函数

$$L(x,\beta)=\prod_{i=1}^n f(x_i,\beta)=\prod_{i=1}^n\frac{\beta^{-(\alpha+1)}}{\Gamma(\alpha+1)}x_i^\alpha e^{-\frac{x}{\beta}}=\frac{\beta^{-(n\alpha+n)}}{\Gamma^n(\alpha+1)}\prod_{i=1}^n x_i^\alpha e^{-\frac{\sum_{i=1}^n x_i}{\beta}}$$

取对数$\ln L(x,\beta)=-n(\alpha+1)\ln\beta-\ln\Gamma^n(\alpha+1)+\alpha\sum_{i=1}^n\ln x_i-\frac{1}{\beta}\sum_{i=1}^n x_i$;

令$\frac{\mathrm{d}\ln L}{\mathrm{d}\beta}=\frac{-n(\alpha+1)}{\beta}+\frac{n}{\beta^2}\bar{X}=0$

解得最大似然估计$\hat{\beta}=\frac{\bar{X}}{\alpha+1}$.

7.5 设X的概率密度函数为

$$f(x,\beta)=\begin{cases}\beta\alpha x^{\alpha-1}e^{-\beta x^\alpha}, & x>0,\\ 0, & x\le 0,\end{cases}$$

其中$\alpha>0$已知,$\beta>0$未知,$(X_1,X_2,\ldots,X_n)$是来自X的样本,试求$\dfrac{1}{\beta}$的极大似然估计量.

解: 似然函数

$$L(x,\beta)=\prod_{i=1}^n f(x_i,\beta)=\prod_{i=1}^n\beta\alpha x_i^{\alpha-1}e^{-\beta x_i^\alpha}=\beta^n\alpha^n\prod_{i=1}^n x_i^{\alpha-1}e^{-\beta\sum_{i=1}^n x_i^\alpha}$$

取对数$\ln L(x,\beta)=n\ln\beta+\ln\alpha^n+\alpha\sum_{i=1}^n\ln x_i-\beta\sum_{i=1}^n x_i^\alpha$;

令$\frac{\mathrm{d}\ln L}{\mathrm{d}\beta}=\frac{n}{\beta}-\sum_{i=1}^n x_i^\alpha=0$

解得$\frac{1}{\beta}$的最大似然估计$\hat{\frac{1}{\beta}}=\frac{1}{n}\sum_{i=1}^n x_i^\alpha$.

7.6　设$X\sim N(\tan\mu+5,\sigma_0^2)$, 其中参数$\mu$未知, 且满足$-\dfrac{\pi}{2}<\mu<\dfrac{\pi}{2}$, σ_0^2已知, $(X_1,X_2,\ldots,X_n)$是来自X的样本, 试求μ的极大似然估计量.

解： 似然函数

$$L(x,\mu)=\prod_{i=1}^n f(x_i,\mu)=\prod_{i=1}^n\frac{1}{\sqrt{2\pi}|\sigma_0|}e^{-\frac{[x_i-(\tan\mu+5)]^2}{2\sigma_0^2}}=\left(\frac{1}{\sqrt{2\pi}|\sigma_0|}\right)^n e^{-\frac{\sum_{i=1}^n(x_i-5-\tan\mu)^2}{2\sigma_0^2}}$$

取对数$\ln L(x,\mu)=-n\ln(\sqrt{2\pi}|\sigma_0|)-\frac{\sum_{i=1}^n(x_i-5-\tan\mu)^2}{2\sigma_0^2}$;

令$\frac{\mathrm{d}\ln L}{\mathrm{d}\mu}=-\frac{\sum_{i=1}^n(x_i-5-\tan\mu)}{\sigma_0^2}\frac{1}{1+\mu^2}=0$

解得μ的最大似然估计$\hat{\mu}=\arctan\frac{n\bar{X}-5n}{n}=\arctan(\bar{X}-5)$.

7.7 设总体X服从几何分布, 概率分布律为

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

$(X_1,X_2,\cdots,X_n)$是取自总体X的样本, 试求参数p的最大似然估计.

解： 似然函数为

$$L(p)=\prod_{i=1}^n P\{X_i=x_i\}=p^n(1-p)^{\sum\limits_{i=1}^n x_i-n}$$

对数似然函数为

$$l(p)=\ln L(p)=n\ln p+(\sum_{i=1}^n x_i-n)\ln(1-p),$$

令

$$\frac{\mathrm{d}l(p)}{\mathrm{d}p}=0,$$

即

$$\frac{n}{p}-\frac{\sum\limits_{i=1}^n x_i-n}{1-p}=0$$

解得p的最大似然估计为$\hat{p}=\frac{1}{\bar{X}}$.

7.8 设总体X的概率密度函数为,

$$f(x,\theta)=\begin{cases}5e^{-5(x-\theta)}, & x\geq\theta,\\ 0, & x<\theta,\end{cases}$$

$(X_1,X_2,\cdots,X_n)$是取自X的样本, 试求参数θ的最大似然估计.

解: 当 $x_i\geq\theta$ 时, 似然函数为

$$L(\theta)=\prod_{i=1}^{n}f(x_i,\theta)=5^n e^{-5\sum\limits_{i=1}^{n}(x_i-\theta)}$$

对数似然函数为

$$l(\theta)=\ln L(\theta)=n\ln 5-5(\sum_{i=1}^{n}x_i-n\theta),$$

由于$\dfrac{\mathrm{d}l(\theta)}{\mathrm{d}\theta}=5n>0$, 知$l(\theta)$为$\theta$的单调递增函数, 故$\theta$取最大值即可, 由$x_i\geq\theta$, 故$\theta$的最大似然估计为$\hat{\theta}=\min\limits_{1\leq i\leq n\{X_i\}}=X_{(1)}$.

7.9 设总体X的概率分律为

X	-1	0	1	2
$f(x,\theta)$	θ^3	$3\theta^2(1-\theta)$	$3\theta(1-\theta)^2$	$(1-\theta)^3$

$(X_1,X_2,\cdots,X_n)$是取自总体X的样本, 试求参数θ的矩估计和最大似然估计. 现有样本观测值$(-1,1,0,2,2,-1,0,-1)$, 求参数θ的矩估计和最大似然估计值.

解:

因为

$$\mathrm{E}(X)=-\theta^3+3\theta(1-\theta)^2+2(1-\theta)^3=-3\theta+2,$$

令

$$-3\theta+2=\bar{X}$$

得参数θ的矩估计量为

$$\hat{\theta}_{\text{矩}}=\frac{2-\bar{X}}{3}.$$

设$(X_1, X_2, \cdots, X_n)$中取$-1, 0, 1$的个数分别为k_1, k_2, k_3, 则取2的个数为$n-k_1-k_2-k_3$, 似然函数为

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} P\{X_i = x_i\} \\ &= (\theta^3)^{k_1}[3\theta^2(1-\theta)]^{k_2}[3\theta(1-\theta)^2]^{k_3}[(1-\theta)^3]^{n-k_1-k_2-k_3} \\ &= 3^{k_2+k_3}\theta^{3k_1+2k_2+k_3}[(1-\theta)]^{3n-3k_1-2k_2-k_3} \end{aligned}$$

$$\begin{aligned} l(\theta) &= \ln(L(\theta)) \\ &= (k_2+k_3)\ln 3 + (3k_1+2k_2+k_3)\ln\theta + (3n-3k_1-2k_2-k_3)\ln(1-\theta) \end{aligned}$$

令

$$\frac{\mathrm{d}l(\theta)}{\mathrm{d}\theta} = 0,$$

即

$$\frac{3k_1+2k_2+k_3}{\theta} - \frac{3n-3k_1-2k_2-k_3}{1-\theta} = 0$$

解得θ的最大似然估计为

$$\hat{\theta}_{\text{最}} = \frac{3k_1+2k_2+k_3}{3n}.$$

将样本观测值代入得$\hat{\theta}_{\text{矩}} = \frac{7}{12}$, $\hat{\theta}_{\text{最}} = \frac{7}{12}$.

事实上, 此题目$\hat{\theta}_{\text{矩}} = \hat{\theta}_{\text{最}}$.

7.10 设总体$X \sim U(\theta, 2\theta)$, $(X_1, X_2, \cdots, X_n)$是取自总体X的样本, 试求参数θ的矩估计和最大似然估计.

解:

(1) 因为

$$\mathrm{E}(X) = \frac{3\theta}{2},$$

令

$$\frac{3\theta}{2} = \bar{X}$$

得参数θ的矩估计量为

$$\hat{\theta} = \frac{2}{3}\bar{X}.$$

(2)似然函数为

$$L(\theta)=\prod_{i=1}^{n}f(x_i,\theta)=\begin{cases}\frac{1}{\theta^n}, & \theta\le x_i\le 2\theta,\\ 0, & 其他,\end{cases}$$

由$L(\theta)=\frac{1}{\theta^n}$ 知, 要使$L(\theta)$达到最大, 只要θ达到最小值即可, 而

$$\theta\le x_i\le 2\theta,$$

即

$$\frac{x_i}{2}\le\theta\le x_i,$$

故应满足$\hat{\theta}_{最}\in[\frac{X_{(n)}}{2},X_{(1)}]$, 即区间$[\frac{X_{(n)}}{2},X_{(1)}]$中的最小值为$\theta$的最大似然估计, 即有$\hat{\theta}_{最}=\frac{X_{(n)}}{2}$.

7.11 设总体$X\sim N(\sigma,\sigma)$, $(X_1,X_2,\cdots,X_n)$是取自X的样本, 试求参数$\sigma(>0)$的最大似然估计.

解: 总体X的概率密度函数为

$$f(x,\sigma)=\frac{1}{2\pi\sigma}e^{-\frac{1}{2\sigma}(x-\sigma)^2}$$

似然函数为

$$L(\sigma)=\prod_{i=1}^{n}f(x_i,\sigma)=(2\pi\sigma)^{-\frac{n}{2}}e^{-\frac{1}{2\sigma}\sum\limits_{i=1}^{n}(x_i-\sigma)^2}$$

对数似然函数为

$$l(\sigma)=\ln L(\sigma)=-\frac{n}{2}\ln(2\pi\sigma)-\frac{1}{2\sigma}\sum_{i=1}^{n}(x_i-\sigma)^2,$$

$$\begin{aligned}\frac{\mathrm{d}l(\sigma)}{\mathrm{d}\sigma} &= -\frac{n}{2\sigma}+\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\sigma)^2+\frac{1}{2\sigma}\sum_{i=1}^{n}2(x_i-\sigma)\\ &= \frac{1}{2\sigma}[\frac{1}{\sigma}\sum_{i=1}^{n}(x_i-\sigma)^2+\sum_{i=1}^{n}2(x_i-\sigma)-n]\\ &= \frac{1}{2\sigma}[\frac{1}{\sigma}\sum_{i=1}^{n}x_i^2-n\sigma-n]\end{aligned}$$

令

$$\frac{\mathrm{d}l(\sigma)}{\mathrm{d}\sigma}=0,$$

即

$$n\sigma^2+n\sigma-\sum_{i=1}^{n}x_i^2=0$$

解得σ的最大似然估计为$\hat{\sigma}=\sqrt{\frac{1}{4}+\frac{1}{n}\sum\limits_{i=1}^{n}x_i^2}-\frac{1}{2}$.

7.12 设总体X的概率密度函数为,

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

, $(X_1,X_2,\cdots,X_n)$是取自X的样本, 试求参数θ的矩估计.

解： 因为

$$\mathrm{E}(X)=\int_0^{\theta}\frac{2x}{\theta^2}\mathrm{d}x=\frac{\theta}{3}.$$

令

$$\frac{\theta}{3}=\bar{X},$$

得参数θ的矩估计量为

$$\hat{\theta}=3\bar{X}.$$

7.13 一个盒子里装有白球和黑球，有放回地取出一个容量为n的样本，其中有k个白球，求盒子里黑球数与白球数之比R的最大似然估计值。

解： 因盒子里黑球与白球个数之比为R，则每次取到白球的概率为$\frac{1}{R+1}$。所以似然函数为

$$L(R)=\left(\frac{1}{R+1}\right)^k\left(\frac{R}{R+1}\right)^{n-k}.$$

$$\begin{aligned}l(R)=\ln L(R) &= -k\ln(R+1)+(n-k)\left[\ln(R)-\ln(R+1)\right]\\ &= (n-k)\ln(R)-n\ln(R+1)\end{aligned}$$

令

$$\frac{dl(R)}{dR}=0,$$

即
$$\frac{n-k}{R}-\frac{n}{R+1}=0,$$
解得R的最大似然估计值为
$$\hat{R}=\frac{n-k}{k}.$$

7.14 设总体X的概率密度函数为
$$f(x;\theta)=\begin{cases}\frac{1}{\theta}e^{-\frac{x}{\theta}}, & x\geq 0,\\ 0, & x<0.\end{cases}$$
从该总体中抽取样本(X_1,X_2,X_3)，考虑θ的如下4种估计量：
$$\hat{\theta}_1=X_1,\qquad \hat{\theta}_2=\frac{1}{3}\left(X_1+2X_2+X_3\right),$$
$$\hat{\theta}_3=\frac{1}{2}\left(X_1+X_2\right),\qquad \hat{\theta}_2=\frac{1}{3}\left(X_1+X_2+X_3\right).$$

(1)试判断这些估计量哪些是θ的无偏估计？

(2)试比较这些估计量的方差。

解： 事实上，$X\sim Exp(\frac{1}{\theta})$，所以$EX=\theta$，$DX=\theta^2$。又因为$(X_1,X_2,X_3)$是来自$X$的样本，故
$$E(X_1)=E(X_2)=E(X_3)=EX=\theta,$$
$$D(X_1)=D(X_2)=D(X_3)=DX=\theta^2.$$

(1) 因为$E(\hat{\theta}_1)=E(X_1)=\theta$，所以$\hat{\theta}_1$是$\theta$的无偏估计。因为
$$E(\hat{\theta}_2)=E\left[\frac{1}{3}\left(X_1+2X_2+X_3\right)\right]=\frac{1}{3}EX_1+\frac{2}{3}EX_2+\frac{1}{3}EX_3=\frac{4}{3}\theta\neq\theta,$$
所以$\hat{\theta}_2$不是θ的无偏估计；

类似地，因为$E(\hat{\theta}_3)=\theta$，$E(\hat{\theta}_4)=\theta$，所以$\hat{\theta}_3$与$\hat{\theta}_4$是θ的无偏估计。

(2) 因为

$$\begin{aligned}
&D(\hat{\theta}_1)=D(X_1)=DX=\theta^2,\\
&D(\hat{\theta}_2)=D\left[\frac{1}{3}(X_1+2X_2+X_3)\right]=\frac{1}{9}(DX_1+4DX_2+DX_3)=\frac{6}{9}DX=\frac{2}{3}\theta^2,\\
&D(\hat{\theta}_3)=D\left[\frac{1}{2}(X_1+X_2)\right]=\frac{1}{4}(DX_1+DX_2)=\frac{2}{4}DX=\frac{1}{2}\theta^2,\\
&D(\hat{\theta}_4)=D\left[\frac{1}{3}(X_1+X_2+X_3)\right]=\frac{1}{9}(DX_1+DX_2+DX_3)=\frac{3}{9}DX=\frac{1}{3}\theta^2,
\end{aligned}$$

所以$D(\hat{\theta}_4)<D(\hat{\theta}_3)<D(\hat{\theta}_2)<D(\hat{\theta}_1)$。

7.15 设总体$X\sim U(0,\theta)$, $\theta>0$是未知参数，$(X_1,X_2,\ldots,X_n)$是取自X的样本。试证明：

$$\hat{\theta}_1=\frac{n+1}{n}X_{(n)},\ \hat{\theta}_2=(n+1)X_{(1)}$$

都是θ的无偏估计量，且$\hat{\theta}_1$较$\hat{\theta}_2$有效。

解：因为

$$f(x)=\begin{cases}\frac{1}{\theta}, & 0\le x\le\theta\\ 0, & else\end{cases},\ F(x)=\begin{cases}0, & x<0\\ \frac{x}{\theta}, & 0\le x<\theta\\ 1, & x\ge\theta\end{cases},$$

所以$X_{(n)}$的概率密度函数为

$$f_n(x)=n\left[F(x)\right]^{n-1}f(x)=\begin{cases}nx^{n-1}/\theta^n, & 0\le x\le\theta,\\ 0, & else.\end{cases}$$

$$\begin{aligned}
E(X_{(n)})&=\int_0^\theta\frac{n}{\theta^n}x^n dx=\frac{n}{n+1}\theta,\\
E(X_{(n)}^2)&=\int_0^\theta\frac{n}{\theta^n}x^{n+1}dx=\frac{n}{n+2}\theta^2,\\
D(X_{(n)})=E(X_{(n)}^2)-\left[E(X_{(n)})\right]^2&=\frac{n}{(n+2)(n+1)^2}\theta^2.
\end{aligned}$$

$X_{(1)}$的概率密度函数为

$$f_1(x)=n\left[1-F(x)\right]^{n-1}f(x)=\begin{cases}n(\theta-x)^{n-1}/\theta^n, & 0\le x\le\theta,\\ 0, & else.\end{cases}$$

$$E(X_{(1)}) = \frac{n}{\theta^n}\int_0^\theta x(\theta-x)^{n-1}dx \xlongequal{令x=\theta-t} \frac{n}{\theta^n}\int_0^\theta (\theta-t)t^{n-1}dt = \frac{\theta}{n+1},$$

$$E(X_{(1)}^2) = \frac{n}{\theta^n}\int_0^\theta x^2(\theta-x)^{n-1}dx \xlongequal{令x=\theta-t} \frac{n}{\theta^n}\int_0^\theta (\theta-t)^2t^{n-1}dt = \frac{2\theta^2}{(n+2)(n+1)},$$

$$D(X_{(1)}) = E(X_{(1)}^2) - \left[E(X_{(1)})\right]^2 = \frac{n}{(n+2)(n+1)^2}\theta^2.$$

则有

$$E(\hat{\theta}_1) = E\left[\frac{n+1}{n}X_{(n)}\right] = \frac{n+1}{n}E(X_{(n)}) = \theta,$$

$$E(\hat{\theta}_2) = E\left[(n+1)X_{(1)}\right] = (n+1)E(X_{(1)}) = \theta,$$

即$\hat{\theta}_1$，$\hat{\theta}_2$都是θ的无偏估计量。

$$D(\hat{\theta}_1) = D\left[\frac{n+1}{n}X_{(n)}\right] = \frac{(n+1)^2}{n^2}D(X_{(n)}) = \frac{1}{n(n+2)}\theta^2,$$

$$D(\hat{\theta}_2) = D\left[(n+1)X_{(1)}\right] = (n+1)^2D(X_{(1)}) = \frac{n}{n+2}\theta^2.$$

因为$D(\hat{\theta}_1) < D(\hat{\theta}_2)$，所以$\hat{\theta}_1$较$\hat{\theta}_2$有效。

7.16 设总体X的概率密度函数为

$$f(x;\lambda) = \begin{cases} \frac{1}{2\lambda}e^{-\frac{x}{2\lambda}}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

其中$\lambda > 0$是未知参数，$(X_1, X_2, \ldots, X_n)$是取自X的样本。令

$$\hat{\lambda}_1 = \frac{1}{2}\overline{X}, \quad \hat{\lambda}_2 = \frac{n}{2}X_{(1)}.$$

(1)试验证$\hat{\lambda}_1$和$\hat{\lambda}_2$都是λ的无偏估计量；

(2)试判断估计量$\hat{\lambda}_1$和$\hat{\lambda}_2$哪个更有效。

解：事实上，$X \sim Exp\left(\frac{1}{2\lambda}\right)$，所以$EX = 2\lambda$，$DX = 4\lambda^2$，

$$E(\hat{\lambda}_1) = E\left(\frac{1}{2}\overline{X}\right) = \frac{1}{2}E\overline{X} = \frac{1}{2}EX = \frac{1}{2}\cdot 2\lambda = \lambda,$$

$$D(\hat{\lambda}_1) = D\left(\frac{1}{2}\overline{X}\right) = \frac{1}{4}D\overline{X} = \frac{1}{4}\cdot\frac{4\lambda^2}{n} = \frac{\lambda^2}{n}.$$

因为

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

所以$X_{(1)}$的概率密度函数为

$$f_1(x)=n[1-F(x)]^{n-1}f(x)=\begin{cases}\frac{n}{2\lambda}e^{-\frac{n}{2\lambda}x}, & x\geq 0\\ 0, & x<0\end{cases}.$$

可看出$X_{(1)}\sim Exp\left(\frac{n}{2\lambda}\right)$，所以$E[X_{(1)}]=\frac{2\lambda}{n}$，$D[X_{(1)}]=\frac{4\lambda^2}{n^2}$。因此

$$E(\hat{\lambda}_2)=E\left[\frac{n}{2}X_{(1)}\right]=\frac{n}{2}E\left[X_{(1)}\right]=\lambda,$$
$$D(\hat{\lambda}_2)=D\left[\frac{n}{2}X_{(1)}\right]=\frac{n^2}{4}D\left[X_{(1)}\right]=\lambda^2.$$

故有$D(\hat{\lambda}_1)<D(\hat{\lambda}_2)$。

由上可知，$\hat{\lambda}_1$和$\hat{\lambda}_2$均为λ的无偏估计量，且$\hat{\lambda}_1$比$\hat{\lambda}_2$更有效。

7.17 设总体$X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$，且X与Y相互独立。从总体X, Y中分别抽取容量分别为n_1, n_2的两个样本，样本二阶中心距分别为S_1^2, S_2^2。试证明：对任何常数a, b，如果

$$\frac{n_1-1}{n_1}a+\frac{n_2-1}{n_2}b=1,$$

则$T=aS_1^2+bS_2^2$ 都是σ^2的无偏估计量，并确定a,b的值，使$D(T)$达到最小。

解：(1)（方法一）

$$S_1^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2\equiv\frac{1}{n_1}\sum_{i=1}^{n_1}X_i^2-\overline{X}^2,$$

$$S_2^2=\frac{1}{n_2}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\equiv\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i^2-\overline{Y}^2,$$

又

$$
\begin{aligned}
&E\left(X_i^2\right)=(EX_i)^2+DX_i=(EX)^2+DX=\mu_1^2+\sigma^2,\\
&E\left(\overline{X}^2\right)=\left(E\overline{X}\right)^2+D\overline{X}=(EX)^2+\frac{DX}{n_1}=\mu_1^2+\frac{\sigma^2}{n_1},\\
&E\left(Y_i^2\right)=(EY_i)^2+DY_i=(EY)^2+DY=\mu_2^2+\sigma^2,\\
&E\left(\overline{Y}^2\right)=\left(E\overline{Y}\right)^2+D\overline{Y}=(EY)^2+\frac{DY}{n_2}=\mu_2^2+\frac{\sigma^2}{n_2},
\end{aligned}
$$

所以

$$
\begin{aligned}
E(T) &= E(aS_1^2+bS_2^2)\\
&= aE(S_1^2)+bE(S_2^2)\\
&= aE\left(\frac{1}{n_1}\sum_{i=1}^{n_1}X_i^2-\overline{X}^2\right)+bE\left(\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i^2-\overline{Y}^2\right)\\
&= \left(\frac{n_1-1}{n_1}a+\frac{n_2-1}{n_2}b\right)\sigma^2
\end{aligned}
$$

若要使$E(T)=\sigma^2$，则

$$\frac{n_1-1}{n_1}a+\frac{n_2-1}{n_2}b=1.$$

(1)（方法二）因为

$$\frac{1}{\sigma^2}\sum_{i=1}^{n_1}(X_i-\overline{X})^2\sim\chi^2_{n_1-1},\quad \frac{1}{\sigma^2}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\sim\chi^2_{n_2-1},$$

则

$$
\begin{aligned}
&E\left[\frac{1}{\sigma^2}\sum_{i=1}^{n_1}(X_i-\overline{X})^2\right]=n_1-1, \qquad E\left[\frac{1}{\sigma^2}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\right]=n_2-1,\\
&D\left[\frac{1}{\sigma^2}\sum_{i=1}^{n_1}(X_i-\overline{X})^2\right]=2(n_1-1), \qquad D\left[\frac{1}{\sigma^2}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\right]=2(n_2-1),
\end{aligned}
$$

即

$$
\begin{aligned}
&E\left[\sum_{i=1}^{n_1}(X_i-\overline{X})^2\right]=(n_1-1)\sigma^2, \qquad E\left[\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\right]=(n_2-1)\sigma^2,\\
&D\left[\sum_{i=1}^{n_1}(X_i-\overline{X})^2\right]=2(n_1-1)\sigma^4, \qquad D\left[\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2\right]=2(n_2-1)\sigma^4,
\end{aligned}
$$

从而

$$E(S_1^2)=\frac{n_1-1}{n_1}\sigma^2,\quad E(S_2^2)=\frac{n_2-1}{n_2}\sigma^2,$$

若要使

$$E(T)=E(aS_1^2+bS_2^2)=\left(\frac{n_1-1}{n_1}a+\frac{n_2-1}{n_2}b\right)\sigma^2=\sigma^2,$$

则

$$\frac{n_1-1}{n_1}a+\frac{n_2-1}{n_2}b=1.$$

(2)

$$D(S_1^2)=\frac{2(n_1-1)}{n_1^2}\sigma^4,\ D(S_2^2)=\frac{2(n_2-1)}{n_2^2}\sigma^4,$$

$$D(T)=D(aS_1^2+bS_2^2)=2\left[(n_1-1)\frac{a^2}{n_1^2}+(n_2-1)\frac{b^2}{n_2^2}\right]\sigma^4.$$

若使$D(T)$最小，则令

$$m=\left[(n_1-1)\frac{a^2}{n_1^2}+(n_2-1)\frac{b^2}{n_2^2}\right]+\lambda\left(\frac{n_1-1}{n_1}a+\frac{n_2-1}{n_2}b-1\right),$$

$$\begin{cases}\dfrac{\partial m}{\partial a}=2a\dfrac{n_1-1}{n_1^2}+\lambda\dfrac{n_1-1}{n_1}=0, & (0.0.1)\\ \dfrac{\partial m}{\partial b}=2b\dfrac{n_2-1}{n_2^2}+\lambda\dfrac{n_2-1}{n_2}=0, & (0.0.2)\\ \dfrac{\partial m}{\partial \lambda}=\dfrac{n_1-1}{n_1}a+\dfrac{n_2-1}{n_2}b=1 & (0.0.3)\end{cases}$$

由(1)，(2)解得

$$\frac{a}{n_1}=\frac{b}{n_2},$$

代入(3)，得

$$a=\frac{n_1}{n_1+n_2-2},\quad b=\frac{n_2}{n_1+n_2-2}.$$

7.18 设总体X的概率密度函数为

$$f(x;\theta)=\begin{cases}1, & \theta-\frac{1}{2}\le x\le\theta+\frac{1}{2},\\ 0, & else.\end{cases}$$

$(X_1,X_2,\ldots,X_n)$是取自总体X的样本。

(1) 求参数θ 的矩估计和最大似然估计；

(2) 证明$\overline{X}$ 及$\frac{1}{2}\left[X_{(1)}+X_{(n)}\right]$ 都是θ的无偏估计量。

解：X的分布函数为

$$F(x;\theta)=\begin{cases}0, & x<\theta-\frac{1}{2}\\ x+\dfrac{1}{2}-\theta, & \theta-\frac{1}{2}\leq x<\theta+\frac{1}{2}\\ 1, & x\geq\theta+\frac{1}{2}\end{cases}$$

因为$E\overline{X}=EX=\frac{1}{2}\left(\theta-\frac{1}{2}+\theta+\frac{1}{2}\right)=\theta$，所以$\theta$的矩估计量$\hat{\theta}=\overline{X}$。

又$E\overline{X}=EX=\theta$，所以$\overline{X}$是θ的无偏估计量。

因为$L(\theta)=1$，其中$x_{(n)}-\frac{1}{2}\leq\theta\leq x_{(1)}+\frac{1}{2}$，所以$\theta$的最大似然估计量是$\left[x_{(n)}-\frac{1}{2},x_{(1)}+\frac{1}{2}\right]$里的任一值。

$X_{(1)}$的概率密度函数为

$$f_1(x)=n[1-F(x)]^{n-1}f(x)=\begin{cases}n\left(\dfrac{1}{2}+\theta-x\right)^{n-1}, & \theta-\frac{1}{2}\leq x\leq\theta+\frac{1}{2}\\ 0, & \text{else}\end{cases}$$

$X_{(n)}$的概率密度函数为

$$f_n(x)=nF^{n-1}(x)f(x)=\begin{cases}n\left(\dfrac{1}{2}-\theta+x\right)^{n-1}, & \theta-\frac{1}{2}\leq x\leq\theta+\frac{1}{2}\\ 0, & \text{else}\end{cases}$$

$$\begin{aligned}E\left[X_{(1)}\right] &= \int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} nx\left(\frac{1}{2}+\theta-x\right)^{n-1}dx\\ &= \int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} -xd\left(\frac{1}{2}+\theta-x\right)^{n}\\ &= -x\left(\frac{1}{2}+\theta-x\right)^{n}\Bigg|_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}}+\int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}}\left(\frac{1}{2}+\theta-x\right)^{n}dx\\ &= \theta-\frac{1}{2}+\frac{1}{n+1}.\end{aligned}$$

$$\begin{aligned} E\left[X_{(n)}\right] &= \int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} nx\left(\frac{1}{2}-\theta+x\right)^{n-1} dx \\ &= \int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} xd\left(\frac{1}{2}-\theta+x\right)^{n} \\ &= x\left(\frac{1}{2}-\theta+x\right)^{n}\Big|_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} - \int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} \left(\frac{1}{2}-\theta+x\right)^{n} dx \\ &= \theta+\frac{1}{2}-\frac{1}{n+1}. \end{aligned}$$

$$E\left\{\frac{1}{2}\left[X_{(1)}+X_{(n)}\right]\right\} = \frac{1}{2}\left\{E\left[X_{(1)}\right]+E\left[X_{(n)}\right]\right\} = \frac{1}{2}\cdot 2\theta = \theta.$$

所以$\frac{1}{2}\left[X_{(1)}+X_{(n)}\right]$是$\theta$的无偏估计量。

7.19 设总体$X\ P(\lambda), \lambda>0, (X_1, X_2, ..., X_n)$是取自总体$X$的样本.

(1)试证: $\bar{X}$和$S^2=\frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2$均为$\lambda$的无偏估计;

(2)试证: 对一切$a(0\le a\le 1), a\bar{X}+(1-a)S^2$都是$\lambda$的无偏估计量;

(3)试求λ^2的无偏估计量.

解：(1)

$$E(\bar{X}) = E(\frac{1}{n}\sum_{i=1}^{n}X_i) = \frac{1}{n}\sum_{i=1}^{n}EX_i = \lambda.$$

$$\begin{aligned} E(S^2) &= E[\tfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2] \\ &= \tfrac{1}{n-1}E[\tfrac{1}{n}\sum_{i=1}^{n}X_i^2-n\bar{X}^2] \\ &= \tfrac{1}{n-1}[\sum_{i=1}^{n}E(X_i^2)-nE(\bar{X}^2)] \\ &= \tfrac{1}{n-1}[\sum_{i=1}^{n}(\lambda+\lambda^2)-n(\lambda^2+\tfrac{\lambda}{n})] = \lambda. \end{aligned}$$

或$E(S^2)=D(X)=\lambda.$

(2)

$$\begin{aligned}
& E[a\bar{X}+(1-a)S^2] \\
= \ & aE(\bar{X})+(1-a)E(S^2) \\
= \ & a\lambda+(1-a)\lambda \\
= \ & \lambda
\end{aligned}$$

(3)

$\because E(\bar{X}^2)=[E(\bar{X})]^2+D(\bar{X})=\lambda^2+\lambda, \quad E(\bar{X})=\lambda$

$\therefore E(\bar{X}^2-\bar{X})=E(\bar{X}^2)-E(\bar{X})=\lambda^2$

$\therefore \bar{X}^2-\bar{X}$是$\lambda^2$的无偏估计.同理,$\bar{X}^2-S^2$也是$\lambda^2$的无偏估计.

7.20 设$\hat{\theta}_1,\hat{\theta}_2$是参数$\theta$的两个相互独立的无偏估计量，且$D(\hat{\theta}_1)=2D(\hat{\theta}_2)$,求常数$k_1,k_2$, 使$k_1\hat{\theta}_1+k_2\hat{\theta}_2$也是$\theta$的无偏估计量，并且使它在所有这种形式的估计量中方差最小.

解： 选k_1,k_2使得$E(k_1\hat{\theta}_1+k_2\hat{\theta}_2)=\theta$, $k_1E(\hat{\theta}_1)+k_2E(\hat{\theta}_2)=\theta$. $\because \hat{\theta}_1,\hat{\theta}_2$均为$\theta$的无偏估计, $\therefore k_1\theta+k_2\theta=\theta$, 即$k_1+k_2=1$.

$D(k_1\hat{\theta}_1+k_2\hat{\theta}_2)=k_1^2D(\hat{\theta}_1)+k_2^2D(\hat{\theta}_2)=(2k_1^2+k_2^2)D(\hat{\theta}_2)$

即求$\min(2k_1^2+k_2^2), s.t. k_1+k_2=1$.

令$L(\lambda,k_1,k_2)=2k_1^2+k_2^2-\lambda(k_1+k_2-1)$,

$$\begin{aligned}
\tfrac{\partial L}{\partial \lambda}&=-\lambda_1-\lambda_2+1=0\\
\tfrac{\partial L}{\partial k_1}&=4k_1-\lambda=0\\
\tfrac{\partial L}{\partial k_1}&=2k_2-\lambda=0
\end{aligned}$$

$\Rightarrow k_1=\frac{1}{3}, k_2=\frac{2}{3}$.

7.21 设总体$X\ N(\mu,\sigma^2)$, $(X_1,X_2,...,X_n)$是取自总体X的样本，

(1) 求k, 使$\hat{\sigma}=\frac{1}{k}\sum\limits_{i=1}^{n-1}(X_{i+1}-X_i)^2$为$\sigma^2$的无偏估计量;

(2) 求k, 使$\hat{\sigma}=\frac{1}{k}\sum\limits_{i=1}^{n}|X_i-\bar{X}|$为$\sigma$的无偏估计量.

解： (1) 选k, s.t.

$$\begin{aligned}
E[\tfrac{1}{k}\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2] &= \sigma^2\\
\tfrac{1}{k}\sum_{i=1}^{n-1}E[X_{i+1}^2+X_i^2-2X_iX_{i+1}] &= \sigma^2\\
\tfrac{1}{k}\sum_{i=1}^{n-1}[\mu^2+\sigma^2+\mu^2+\sigma^2-2E(X_i)E(X_{i+1})] &= \sigma^2\\
\tfrac{1}{k}\sum_{i=1}^{n-1}[2\mu^2+2\sigma^2-2\mu^2] &= \sigma^2\\
\tfrac{2(n-1)}{k}\sigma^2 &= \sigma^2
\end{aligned}$$

$\Rightarrow k=2(n-1)$.

(2) $X_i\sim N(\mu,\sigma^2), \bar{X}\sim N(\mu,\frac{\sigma^2}{n})$, 令$Y_i=X_i-\bar{X}$, $\because Cov(X_i,\frac{1}{n}\sum\limits_{i=1}^{n}X_i)=\frac{\sigma^2}{n}$,

$\therefore Var(Y_i)=Var(X_i)+Var(\bar{X})-2Cov(X_i,\bar{X})=\dfrac{n-1}{n}\sigma^2, Y_i\sim N(0,\dfrac{n-1}{n}\sigma^2)$,

$$\begin{aligned}
&|Y_i|\sim F(y)=P|Y_i|\le y=\textstyle\int_{-y}^{y}\frac{1}{\sqrt{2\pi}\sqrt{\frac{n-1}{n}}\sigma}\mathrm{e}^{-\frac{nx^2}{2(n-1)\sigma^2}}dx\\
&p(y)=F'(y)=\tfrac{\sqrt{2n}}{\sqrt{(n-1)\pi\sigma}}\mathrm{e}^{-\frac{ny^2}{2(n-1)\sigma^2}}, y>0\\
&E|X_i-\bar{X}|=\textstyle\int_0^{+\infty}\frac{\sqrt{2n}}{\sqrt{(n-1)\pi\sigma}}\mathrm{e}^{-\frac{ny^2}{2(n-1)\sigma^2}}ydy=\frac{\sqrt{2(n-1)}}{\sqrt{n\pi}}\sigma\\
&E\hat{\sigma}=\frac{1}{k}\sum_{i=1}^{n}\frac{\sqrt{2(n-1)}}{\sqrt{n\pi}}\sigma=\sigma, \therefore k=\frac{\sqrt{2n(n-1)}}{\sqrt{\pi}}.
\end{aligned}$$

7.22 设总体X的概率密度函数为

$$f(x,\theta)=\begin{cases}\frac{6x(\theta-x)}{\theta^3}, & 0<x<\theta,\\ 0, & \text{其他}.\end{cases}$$

$(X_1,X_2,...,X_n)$是取自总体X的样本.

(1)求θ的矩估计量$\hat{\theta}$; (2)$\hat{\theta}$是θ的无偏估计量吗？(3)求$\hat{\theta}$的方差$D(\hat{\theta})$.

解： (1)

$$E(X)=\textstyle\int_0^{\theta}x\cdot\frac{6x\theta-6x^2}{\theta^3}dx=\frac{\theta}{2}, \bar{X}=\frac{\theta}{2}\Rightarrow\hat{\theta}=2\bar{X}$$

(2)

$$E(\hat{\theta})=2E(\bar{X})=2E(X)=\theta, \therefore \hat{\theta}\text{是}\theta\text{的无偏估计}.$$

(3)

$$D(X)=E(X^2)-(E(X))^2=\int_0^\theta x^2\cdot\tfrac{6x\theta-6x^2}{\theta^3}dx-\tfrac{\theta^2}{4}=\tfrac{\theta^2}{20},$$

$$D(\hat{\theta})=D(2\bar{X})=4D(\bar{X})=\tfrac{4}{n}D(X)=\tfrac{\theta^2}{5n}.$$

7.23 从总体中抽取容量$n=50$的样本:

样本值x_i	2	5	7	10
频率	16	12	8	14

试求总体均值的无偏估计.

解： $2\cdot\frac{16}{50}+5\cdot\frac{12}{50}+7\cdot\frac{8}{50}+10\cdot\frac{14}{50}=\frac{144}{25}$.

7.24 设总体X的概率密度函数为

$$f(x;\lambda)=\begin{cases}\frac{a}{2}x\mathrm{e}^{-\frac{x^2}{\lambda}}, & x>0,\\ 0, & x\le 0,\end{cases}$$

其中λ为未知参数, $(X_1,X_2,...,X_n)$是取自总体X的样本.

(1) 试求常数a;

(2) 试求λ的矩估计量$\hat{\lambda}$;

(3) $\hat{\lambda}$是否为λ的无偏估计量? 若不是, 能否修正为λ的无偏估计量.

解： (1)

$$\int_0^{+\infty}\frac{a}{2}x\mathrm{e}^{-\frac{x^2}{\lambda}}dx=1,\Rightarrow a=\frac{4}{\lambda}$$

(2)

$$\begin{aligned}EX &= \textstyle\int_0^{+\infty}\frac{2x^2}{\lambda}\mathrm{e}^{-\frac{x^2}{\lambda}}dx\\ &= -x\mathrm{e}^{-\frac{x^2}{\lambda}}|_0^{+\infty}+\textstyle\int_0^{+\infty}\mathrm{e}^{-\frac{x^2}{\lambda}}dx\\ &= \sqrt{2\pi}\sqrt{\tfrac{\lambda}{2}}\textstyle\int_0^{\infty}\frac{1}{\sqrt{2\pi}\sqrt{\frac{\lambda}{2}}}\mathrm{e}^{-\frac{x^2}{2\cdot\frac{\lambda}{2}}}dx\\ &= \tfrac{\sqrt{\pi\lambda}}{2}.\end{aligned}$$

$$\bar{X}=\frac{\sqrt{\pi\lambda}}{2}\Rightarrow\hat{\lambda}=\frac{4}{\pi}\bar{X}^2.$$

(3)

$$\begin{aligned} D(X) &= \int_0^{+\infty} x^2 \frac{ax}{2} \mathrm{e}^{-\frac{x^2}{\lambda}} dx - (EX)^2 \\ &= \lambda - \frac{\pi\lambda}{4} \end{aligned}$$

$$\begin{aligned} &E(\hat{\lambda}) = \tfrac{4}{\pi} E(\bar{X}^2) = \tfrac{4}{\pi}[(E\bar{X})^2 + D\bar{X}] \\ &= \tfrac{4}{\pi}[\tfrac{\pi\lambda}{4} + \tfrac{1}{n}(\lambda - \tfrac{\pi\lambda}{4})] \\ &= \tfrac{(n-1)\pi+4}{n\pi}\lambda \end{aligned}$$

$\hat{\lambda}$的修正$\frac{n\pi}{(n-1)\pi+4}\hat{\lambda}$为$\lambda$的无偏估计.

7.25 设总体X的概率密度函数为

$$f(x;\mu) = \begin{cases} \dfrac{1}{\sqrt{2\pi}x} \mathrm{e}^{-\frac{(\ln x - \mu)^2}{2}}, & x > 0; \\ 0, & x \le 0, \end{cases}$$

其中μ为未知参数, $(X_1, X_2, \cdots, X_n)$是取自X的样本.

(1)试求$g(\mu) = 3\mu + 1$的最大似然估计量$\hat{g}(\mu)$;

(2)试验证$\hat{g}(\mu)$是$g(\mu)$的无偏估计量.

解：(1)因为总体X的概率密度函数为

$$f(x;\mu) = \begin{cases} \dfrac{1}{\sqrt{2\pi}x} \mathrm{e}^{-\frac{(\ln x - \mu)^2}{2}}, & x > 0; \\ 0, & x \le 0, \end{cases}$$

所以，当$x > 0$时，似然函数为

$$L(\mu) = \prod_{i=1}^n f(x_i;\mu) = \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi}x_i} \mathrm{e}^{-\frac{(\ln x_i - \mu)^2}{2}} \right) = (2\pi)^{-\dfrac{n}{2}} (\prod_{i=1}^n x_i)^{-1} \mathrm{e}^{-\sum\limits_{i=1}^n \frac{(\ln x_i - \mu)^2}{2}},$$

$$\ln L(\mu) = \ln \left((2\pi)^{-\dfrac{n}{2}} \right) - \ln(\prod_{i=1}^n x_i) - \sum_{i=1}^n \frac{(\ln x_i - \mu)^2}{2},$$

令$\dfrac{d \ln L(\mu)}{d\mu} = 0$得 $\quad \sum\limits_{i=1}^n (\ln x_i - \mu) = 0$,

解之得$\hat{\mu} = \dfrac{1}{n} \sum\limits_{i=1}^n \ln X_i$,

故$g(\mu) = 3\mu + 1$的最大似然估计量$\hat{g}(\mu) = \dfrac{3}{n} \sum\limits_{i=1}^n \ln X_i + 1$.

(2)因为$\mathrm{E}(\ln X)=\int_{-\infty}^{+\infty}\ln x f(x;\mu)\mathrm{d}x\int_0^{+\infty}\dfrac{\ln x}{\sqrt{2\pi}x}\mathrm{e}^{-\frac{(\ln x-\mu)^2}{2}}\mathrm{d}x=\mu,$

所以$\mathrm{E}(\hat{g}(\mu))=\dfrac{3}{n}\sum\limits_{i=1}^{n}\mathrm{E}(\ln X_i)+1=3\mu+1.$

故$\hat{g}(\mu)$是$g(\mu)$的无偏估计量.

7.26 设有一批产品,为估计废品率p,随机取一样本$(X_1,X_2,\cdots,X_n)$, 其中

$$X_i=\begin{cases}1, & \text{取得废品;}\\ 0, & \text{取得正品,}\end{cases}\quad i=1,2,\cdots,n.$$

试证明$\hat{p}=\overline{X}$是p的无偏和相合估计量.

证明：由题意得，$X_i\sim B(1,p),i=1,\cdots,n$,所以。$\mathrm{E}(X_i)=p,\mathrm{D}(X_i)=p(1-p),i=1,\cdots,n.$

因为$\mathrm{E}(\hat{p})=\mathrm{E}(\overline{X})=\dfrac{1}{n}\sum\limits_{i=1}^{n}\mathrm{E}(X_i)=p,$

所以$\hat{p}=\overline{X}$是p的无偏估计量。

又因为$\mathrm{D}(\hat{p})=\dfrac{1}{n^2}\sum\limits_{i=1}^{n}\mathrm{D}(X_i)=\dfrac{1}{n}p(1-p),$

由切比雪夫不等式，有

$P\left\{|\overline{X}-p|\geq\varepsilon\right\}\leq\dfrac{\mathrm{D}(\overline{X})}{\varepsilon^2}=\dfrac{p(1-p)}{n\varepsilon^2}\longrightarrow 0(n\to\infty)$，

故$\hat{p}=\overline{X}$是p的相合估计量.

7.27 设总体X的概率密度函数为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-\frac{x}{\theta}}, & x\geq 0;\\ 0, & x<0,\end{cases}$$

其中$\theta>0$是未知参数, $(X_1,X_2,\cdots,X_n)$是从X中抽取的样本, 试证明$\hat{\theta}=\overline{X}$是参数$\theta$的无偏和相合估计量.

证明：由题设得$X\sim E(1/\theta)$,　故$\mathrm{E}(X)=\theta,\mathrm{D}(X)=\theta^2,$

因为$\mathrm{E}(\hat{\theta})=\mathrm{E}(\overline{X})=\dfrac{1}{n}\sum\limits_{i=1}^{n}\mathrm{E}(X_i)=\theta,$

所以$\hat{\theta}=\overline{X}$是$\theta$的无偏估计量。

又因为$\mathrm{D}(\hat{\theta})=\mathrm{D}(\overline{X})=\dfrac{\mathrm{D}(X)}{n}=\dfrac{\theta^2}{n}$,

由切比雪夫不等式，有

$P\left\{|\overline{X}-\theta|\geq\varepsilon\right\}\leq\dfrac{\mathrm{D}(\overline{X})}{\varepsilon^2}=\dfrac{\theta}{n\varepsilon^2}\longrightarrow 0(n\to\infty)$，

故$\hat{\theta}=\overline{X}$是$\theta$的相合估计量.

7.28 设总体$X\sim U(\theta,\theta+1)$, $(X_1,X_2,\cdots,X_n)$是取自总体X的样本.

(1)求参数θ的矩估计量和最大似然估计量;

(2)试证明$\hat{\theta}_1=\overline{X}-\dfrac{1}{2},\hat{\theta}_2=X_{(n)}-\dfrac{n}{n+1},\hat{\theta}_3=X_{(1)}-\dfrac{1}{n+1}$ 都是θ的无偏估计量.

(3)试说明估计量$\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3$中,哪一个方差最小;

(4)$\hat{\theta}_1,\hat{\theta}_2$和$\hat{\theta}_3$是否为$\theta$的相合估计量?

解：(1)由于$X\sim U(\theta,\theta+1)$,所以$\mathrm{E}(X)=\dfrac{2\theta+1}{2},\mathrm{D}(X)=\dfrac{1}{12}$.

令$\mathrm{E}(X)=\overline{X}$, 得$\theta$的矩估计为$\hat{\theta}=\overline{X}-\dfrac{1}{2}$.

又$(X_1,\cdots,X_n)$的似然函数为

$$L(\theta)=\begin{cases}1, & X_{(n)}-1\leq\theta\leq X_{(1)};\\ 0, & \text{其他}.\end{cases}$$

由于似然函数是分段常量函数，所以要使得似然函数达到最大，只要

$X_{(n)}-1\leq\theta\leq X_{(1)}$

故区间$[X_{(n)}-1,X_{(1)}]$中的任一点都可以作为θ的极大似然估计值。

(2)因为$\mathrm{E}(\hat{\theta}_1)=\mathrm{E}\left(\overline{X}-\dfrac{1}{2}\right)=\mathrm{E}\left(X\right)-\dfrac{1}{2}=\theta$,

所以$\hat{\theta}_1$是θ的无偏估计.

由于最大次序统计量$X_{(n)}$的概率密度函数为

$$f_{(n)}(x;\theta)=\begin{cases}n(x-\theta)^{n-1}, & \theta\in[\theta,\theta+1];\\ 0, & x\notin[\theta,\theta+1],\end{cases}$$

最小次序统计量$X_{(1)}$的概率密度函数为

$$f_{(1)}(x;\theta)=\begin{cases} n(\theta+1-x)^{n-1}, & \theta\in[\theta,\theta+1]; \\ 0, & x\notin[\theta,\theta+1], \end{cases}$$

故$\mathrm{E}(X_{(n)})=\int_{-\infty}^{+\infty}xf_{(n)}(x;\theta)\mathrm{d}x=\int_{\theta}^{\theta+1}x\cdot n(x-\theta)^{n-1}\mathrm{d}x=\dfrac{n}{n+1}+\theta$,

$\mathrm{E}(X_{(1)})=\int_{-\infty}^{+\infty}xf_{(1)}(x;\theta)\mathrm{d}x=\int_{\theta}^{\theta+1}x\cdot n(\theta+1-x)^{n-1}\mathrm{d}x=\dfrac{1}{n+1}+\theta$,

从而$\mathrm{E}(\hat{\theta}_2)=\mathrm{E}(X_{(n)}-\dfrac{n}{n+1})=\theta$,

$\mathrm{E}(\hat{\theta}_3)=\mathrm{E}(X_{(1)}-\dfrac{1}{n+1})=\theta$,

因此所以$\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3$都是$\theta$的无偏估计.

(3)因为$\mathrm{D}(\hat{\theta}_1)=\mathrm{D}\left(\overline{X}-\dfrac{1}{2}\right)=\dfrac{\mathrm{D}(X)}{n}=\dfrac{1}{12n}$;

$$\mathrm{D}(\hat{\theta}_2)=\mathrm{D}\left(X_{(n)}-\frac{n}{n+1}\right)=\mathrm{D}(X_{(n)})=\mathrm{E}(X_{(n)}^2)-\mathrm{E}^2(X_{(n)})$$

$$=\int_{\theta}^{\theta+1}x^2\cdot n(x-\theta)^{n-1}\mathrm{d}x-\left(\frac{n}{n+1}+\theta\right)^2=\frac{n}{(n+2)(n+1)^2};$$

$$\mathrm{D}(\hat{\theta}_3)=\mathrm{D}\left(X_{(1)}-\frac{1}{n+1}\right)=\mathrm{D}(X_{(1)})=\mathrm{E}(X_{(1)}^2)-\mathrm{E}^2(X_{(1)})$$

$$=\int_{\theta}^{\theta+1}x^2\cdot n(\theta+1-x)^{n-1}\mathrm{d}x-\left(\frac{1}{n+1}+\theta\right)^2=\frac{n}{(n+2)(n+1)^2};$$

当$n\le 7$时，$\mathrm{D}(\hat{\theta}_1)\le\mathrm{D}(\hat{\theta}_2)=\mathrm{D}(\hat{\theta}_3)$;

当$n>7$时，$\mathrm{D}(\hat{\theta}_1)>\mathrm{D}(\hat{\theta}_2)=\mathrm{D}(\hat{\theta}_3)$.

(4)因为$\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3$都是$\theta$的无偏估计量，且当$n\to+\infty$时，他们方差的极限均为0，

故对$i=1,2,3$,由切比雪夫不等式得

$$P\left\{|\hat{\theta}_i-\theta|\ge\varepsilon\right\}\le\frac{\mathrm{D}(\hat{\theta}_i)}{\varepsilon^2}\longrightarrow 0(n\to\infty),$$

故$\hat{\theta}_i,i=1,2,3$都是θ的相合估计量.

7.29 从商店一年来的发票存根中随机抽取26张,计算得平均金额为78.5元,样本标准差为20元. 假设发票金额数服从正态分布$N(\mu,\sigma^2)$,其中μ,σ^2为未知参数.

试求该商店一年来发票平均金额数μ的置信度为0.90的置信区间.

解： 由于总体$X\sim N(\mu,\sigma^2)$,且μ,σ^2未知，故选择枢轴量

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}},$$

对于给定的置信水平$1-\alpha=$0.90,μ的置信度为0.90的置信区间为

$$\left[\overline{x}-t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}},\overline{x}+t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\right]$$

对自由度$n-1=25$,查表得$t_{0.05}(25)=1.7081$,将$\overline{x}=78.5,s=20$代入上式得μ的置信度为0.90的置信区间为[71.8,85.2].

7.30 某种零件尺寸偏差X服从正态分布$N(\mu,\sigma^2)$,其中μ,σ^2为未知参数. 今随机抽取10个零件测得尺寸偏差(单位：μm)为:$(1,2,-2,3,2,4,-2,5,3,4)$,试分别求μ和σ^2的置信度为0.99的置信区间.

解： 先求μ的置信度为0.99的置信区间.

由于总体$X\sim N(\mu,\sigma^2)$,且μ,σ^2未知，故选择枢轴量

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}},$$

对于给定的置信水平$1-\alpha=0.99$,μ的置信度为0.99的置信区间为

$$\left[\overline{x}-t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}},\overline{x}+t_{\alpha/2}(n-1)\frac{s}{\sqrt{n}}\right]$$

对自由度$n-1=9$,查表得$t_{0.005}(9)=3.2498$,根据数据计算得$\overline{x}=2,s=\frac{\sqrt{52}}{3}$,代入上式得$\mu$的置信度为0.99的置信区间为[-0.47,4.47].

再求σ^2的置信度为0.99的置信区间.

由于总体$X\sim N(\mu,\sigma^2)$,且μ,σ^2未知，故选择枢轴量

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2},$$

对于给定的置信水平$1-\alpha=0.99$,σ^2的置信度为0.99的置信区间为

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\right]$$

对自由度$n-1=9$,查表得$\chi^2_{0.005}(9)=23.587,\chi^2_{0.995}(9)=1.735$根据数据计算得$s=\frac{\sqrt{52}}{3}$, 代入上式得$\sigma^2$ 的置信度为0.99 的置信区间为[2.20,29.97].

7.31　为了比较两批灯泡的寿命，从标有商标A的一批灯泡中抽取150只灯泡组成一个样本，计算得样本均值$\bar{x}=1400\text{h}$,样本标准差$s_1=120\text{h}$,从标有商标B的一批灯泡中抽取100只灯泡组成一个样本，计算得样本均值$\bar{y}=1200\text{h}$,样本标准差$s_2=80\text{h}$.假设两批灯泡分别服从正态分布$N(\mu_1,\sigma_1^2)$,$N(\mu_2,\sigma_2^2)$.试求总体X和Y平均寿命之差$\mu_1-\mu_2$的置信度为0.99的置信区间.

解： $\mu_1-\mu_2$的置信区间为

$$[\bar{x}-\bar{y}-z_{\frac{\alpha}{2}}\sqrt{\frac{1}{m}s_1^2+\frac{1}{n}s_2^2}, \bar{x}-\bar{y}+z_{\frac{\alpha}{2}}\sqrt{\frac{1}{m}s_1^2+\frac{1}{n}s_2^2}]$$

,将$m=150$,$n=100$,$\bar{x}=1400$,$s_1=120$,$\bar{y}=1200$,$s_2=80$,$\alpha=0.01$, 查表得$z_{\frac{\alpha}{2}}=2.58$,代入置信区间得[167.365,232.635].

7.32　设A、B两批导线的电阻分别从无正态分布$N(\mu_1,\sigma_1^2)$,$N(\mu_2,\sigma_2^2)$，参数均未知，从两个总体分别独立地抽取容量$n_1=4$,$n_2=5$的样本，测得电阻（单位：Ω）为

A导线：0.143,0.142,0.143,0.137

B导线：0.140，0.142,0.136,0.138,0.140

试求$\mu_1-\mu_2$的置信度为0.95的置信区间.

解： 已知两正态总体的方差相等且未知，所以$\mu_1-\mu_2$的$1-\alpha$的置信区间为

$$[\bar{x}-\bar{y}-t_{m+n-2}(\frac{\alpha}{2})s_w\sqrt{\frac{1}{m}+\frac{1}{n}}, \bar{x}-\bar{y}+t_{m+n-2}(\frac{\alpha}{2})s_w\sqrt{\frac{1}{m}+\frac{1}{n}}]$$

其中$s_w^2=\frac{(m-1)s_1^2+(n-1)s_2^2}{m+n-2}$,由所给数据得，$\bar{x}=0.14125$,$\bar{y}=0.1392$,$s_1^2=8.25$,$s_2^2=5.2$,$\alpha=0.05$，查表得$t_7(0.025)=2.3646$,代入得置信区间[-0.002,0.006].

7.33　在某一地区，随机抽取100名成年居民进行民意测验，其中有80名居民支持粮食调价，试求在该地区所有居民中，支持粮食调价的居民比率p的置信度为0.95的置信区间.

解： 比率p的$1-\alpha$的置信区间为

$$[\frac{m}{n}-z_{\frac{\alpha}{2}}\sqrt{\frac{1}{n}\frac{m}{n}(1-\frac{m}{n})},\frac{m}{n}+z_{\frac{\alpha}{2}}\sqrt{\frac{1}{n}\frac{m}{n}(1-\frac{m}{n})}]$$

其中$n=100$,$m=80$,$z_{\frac{\alpha}{2}}=z_{0.025}=1.96$,代入得置信区间[0.7216,0.8784].

7.34　从两台机器所生产的滚珠轴承中分别抽取容量为10的样本，测得它们直径的标准差分别为0.042cm和0.035cm，假设直径服从正态分布.试求两个正态总体方差比的置信度为0.98的置信区间.

解： $\frac{\sigma_1^2}{\sigma_2^2}$的$1-\alpha$的置信区间为

$$[F_{n-1,m-1}(1-\frac{\alpha}{2})\frac{S_1^2}{S_2^2},F_{n-1,m-1}(\frac{\alpha}{2})\frac{S_1^2}{S_2^2}],$$

其中$n=m=10$,$s_1=0.042$,$s_2=0.035$,$\alpha=0.02$,查表得$F_{9,9}(0.01)=5.35$,$F_{9,9}(0.99)=\frac{1}{F_{9,9}(0.01)}$,代入得置信区间[0.269,7.704].

7.35　设$(X_1,X_2,\cdots,X_n)$是从正态总体$N(\mu,\sigma^2)$中抽取的样本，其中μ,σ^2均未知，用L表示μ的置信度为$1-\alpha$的置信区间长度. 对下列不同的置信水平$1-\alpha$和样本容量n，分别求$E(L^2)$.

(1) $1-\alpha=0.95,n=5$;　(2) $1-\alpha=0.90,n=8$.

解： σ未知时，μ的$1-\alpha$的置信区间为

$$[\bar{X}-t_{n-1}(\frac{\alpha}{2})\frac{S}{\sqrt{n}},\bar{X}+t_{n-1}(\frac{\alpha}{2})\frac{S}{\sqrt{n}}],$$

故$L=2t_{n-1}(\frac{\alpha}{2})\frac{S}{\sqrt{n}}$,$E(L^2)=\frac{4}{n}t_{n-1}^2(\frac{\alpha}{2})E(S^2)=\frac{4\sigma^2t_{n-1}^2(\frac{\alpha}{2})}{n}$

(1)当$1-\alpha=0.95,n=5$，查表得$t_4(0.025)=2.7764$,代入得$E(L^2)=6.1669\sigma^2$;

(2)当$1-\alpha=0.90,n=8$，查表得$t_7(0.05)=1.8946$,代入得$E(L^2)=1.7947\sigma^2$.

7.36　设总体$X\sim N(\mu,\sigma^2)$，$(X_1,X_2,\cdots,X_n)$是取自X的样本，a,b为常数，且$0<b<a$. 试求随机区间$(\sum_{i=1}^n\frac{(X_i-\mu)^2}{a},\sum_{i=1}^n\frac{(X_i-\mu)^2}{b})$的长度的数学期望和方差.

解： $L=\sum_{i=1}^n\frac{(X_i-\mu)^2}{a}-\sum_{i=1}^n\frac{(X_i-\mu)^2}{b}=(\frac{1}{b}-\frac{1}{a})\sum_{i=1}^n(X_i-\mu)^2$,由$\chi^2$分布性质有，$E(\frac{1}{\sigma^2}\sum_{i=1}^n(X_i-\mu)^2)=n,D(\frac{1}{\sigma^2}\sum_{i=1}^n(X_i-\mu)^2)=2n$,

$E(L)=(\frac{1}{b}-\frac{1}{a})\sum_{i=1}^n(X_i-\mu)^2=(\frac{1}{b}-\frac{1}{a})n\sigma^2=\frac{(a-b)n\sigma^2}{ab}$,

$D(L)=(\frac{(a-b)}{ab})^2D(\sum_{i=1}^n(X_i-\mu)^2)=2n\sigma^4(\frac{(a-b)}{ab})^2$.

7.37　设总体X的概率密度函数为

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

其中$0<\theta<1$,$(X_1,X_2,\cdots,X_n)$是来自总X的样本,

(1)求未知参数θ的矩估计量;

(2)若样本容量$n=400$,置信度为0.95，求θ的置信区间.

解： (1)$E(X)=\int_0^\infty-\theta^x x\ln\theta\mathrm{d}x=\frac{1}{\ln\theta}\int_0^\infty-e^{x\ln\theta}x\ln\theta\mathrm{d}(x\ln\theta)=-\frac{1}{\ln\theta}=\bar{X}$，所以$\hat{\theta}=e^{-\frac{1}{\bar{X}}}$;

(2)尽管总体X不服从正态分布，但由于$n=400$属于大样本，因此$U=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim N(0,1)$,即总体X的数学期望$\mu=-\frac{1}{\ln\theta}$的置信区间近似为

$[\bar{X}-z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}},\bar{X}+z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}]$. 因为$\alpha=0.05,z_{\frac{\alpha}{2}}=1.96$,故$-\frac{1}{\ln\theta}$的置信区间为

$[\bar{X}-0.098S,barX+0.098S]$

从而θ的置信区间为$[\exp(-\frac{1}{\bar{X}-0.098S}),\exp(-\frac{1}{\bar{X}-0.098S})]$.

习题8

8.1 判断下列命题是否正确，如不正确，试指出错误并改正之。

(1)在假设检验中，经检验接受原假设H_0，这就证明了H_0所述的命题是绝对正确的。

(2)检验的显著性水平α，是当H_0为真时接受H_0的最大概率。

(3)一般说来，当样本容量固定时，若要减少犯一类错误的概率，则犯另一类错误的概率往往增大。若要使两类错误的概率都减少，除非增加样本容量。

解：（1）错误，接受原假设，只是$1-\alpha$的概率认为原假设是正确的。

（2）错误，是当H_0为真时拒绝H_0的最大概率。

（3）正确。

8.2 在假设检验中，若检验结果是接受原假设，则检验可能犯哪一类错误？若检验结果是拒绝原假设，则又有可能犯哪一类错误？

解：第二类错误；第一类错误。

8.3 区间估计与假设检验提法是否相同？解决问题的途径相通吗？以未知方差关于期望的区间估计与假设检验为例说明(置信度$1-\alpha$，即检验水平α).

解：不同，但是解决的方法是相通的。

8.4 在产品检验时，原假设H_0 :产品合格。为了使“次品混入正品”的可能性很小，在样本容量n固定的条件下，显著水平α应取大些还是小些？

解：α应取大些.

8.5 设$X_1, X_2, \cdots, X_n$是来自$N(\mu, 1)$的样本，考虑如下假设检验问题

$$H_0: \mu = 2 \quad H_1: \mu = 3,$$

若检验由拒绝域为$W = \{\overline{X} \geq 2.6\}$确定.

(1) 当$n=20$时，求检验犯两类错误的概率；

(2) 如果要使得检验犯第二类错误的概率$\beta \leq 0.01$，n最小应取多少？

(3) 证明：当$n \to +\infty$时，$\alpha \to 0, \beta \to 0$。

解：$H_0: \ \mu = 2 \ \ H_1: \ \mu = 3$. $W = \{(X_1, X_2, \cdots X_n) : \overline{X} \geq 2.6\}$.

(1)当$n=20$,

$$
\begin{aligned}
\alpha &= P\{\overline{X} \geq 2.6 | \mu = 2\} \\
&= P\{\frac{\overline{X}-2}{1/\sqrt{20}} \geq \frac{2.6-2}{1/\sqrt{20}}\} \\
&= 1 - \Phi(\frac{0.6}{1/\sqrt{20}}); \\
\beta &= P\{\overline{X} < 2.6 | \mu = 3\} \\
&= P\{\frac{\overline{X}-3}{1/\sqrt{20}} < \frac{2.6-3}{1/\sqrt{20}}\} \\
&= \Phi(\frac{-0.4}{1/\sqrt{20}}).
\end{aligned}
$$

(2)$1-\Phi(\frac{0.4}{1/\sqrt{n}}) \leq 0.01$,解得$n \geq \frac{2.33^2}{0.4^2}$

(3)$n \to +\infty, \alpha = 1-\Phi(\frac{0.6}{1/\sqrt{n}}) \to 0$, $n \to +\infty, \beta = 1-\Phi(\frac{-0.4}{1/\sqrt{n}}) \to 0$.

8.6 设$X \sim U(0,\theta)$，其中$\theta > 0$是未知参数，$X_1, X_2, \cdots, X_n$是来自总体X的样本，记$X_{(n)} = max\{X_1, X_2, \cdots, X_n\}$.检验假设$H_0: \theta \geq 2 \quad H_1: \theta < 2$,取拒绝域$W = \{X_{(n)} < 1.5\}$，求犯第一类错误的概率的最大值？

解：$X_{(n)}$ 的密度函数为

$$
f(x) = \begin{cases} n(\frac{x}{\theta})^{n-1}\frac{1}{\theta}, & 0 < x < \theta, \\ 0, & \text{其它} \end{cases}
$$

$$
\begin{aligned}
\alpha &= P\{X_{(n)} < 1.5|\theta \geq 2\} \\
&= \int_0^{1.5} n(\frac{x}{\theta})^{n-1}\frac{1}{\theta}dx = (\frac{1.5}{\theta})^n, n \geq 2,
\end{aligned}
$$

要使α最大，取$\theta = 2, \alpha_{\max} = 0.75^n$

8.7 设总体$X \sim B(n,p)$，检验假设$H_0 : p = 0.6 \quad H_1 : p = 0.3$，假设$H_0$ 的拒绝域$W = \{X \leq c_1\} \bigcup \{X \geq c_2\}$,对

(1)$n = 10, c_1 = 1, c_2 = 9$;

(2)$n = 20, c_1 = 7, c_2 = 17$.

求犯第一类错误的概率α和犯第二类错误的概率β。

解：$X \sim B(n,p), H_0 : p = 0.6 \;\; H_1 : p = 0.3$,拒绝域为$W = \{X \leq c_1\} \bigcup \{X \geq c_2\}$，则第一类错误$\alpha = P\{X \leq c_1\} \bigcup \{X \geq c_2|p = 0.6\}$,而第二类错误$\beta = P\{c_1 < X < c_2|p = 0.3\}$

(1)$n = 10, c_1 = 1, c_2 = 9$则$\alpha = P\{X \leq 1 \bigcup X \geq 9|p = 0.6\} = 0.0480 \;\; \beta = \sum_{i=2}^{8} C_{10}^i 0.3^i 0.7^{10-i} = 0.8505$.

(2)$n = 20, c_1 = 7, c_2 = 17$,类似可得$\alpha = 0.0370, \beta = 0.2277$.

8.8 设总体$X \sim P(\lambda)$，X_1, X_2是来自总体X的样本，假设检验$H_0 : \lambda = 1/2 \quad H_1 : \lambda = 2$，取拒绝域$W = \{(X_1, X_2) : X_1 + X_2 \geq 3\}$，试求犯第一类和第二类错误的概率？

解: $X \sim P(\lambda) \; H_0 : \lambda = 0.5 \quad H_1 : \lambda = 2$, 拒绝域为$W = \{(X_1, X_2)|X_1 + X_2 \geq 3\}$，则第一类错误$\alpha = P\{X_1 + X_2 \geq 3|\lambda = 0.5\} = 1 - \sum_{i=0}^{2} \frac{(2\lambda)^i e^{-2\lambda}}{i!}|_{\lambda=0.5} = 0.0803$,而第二类错误$\beta = P\{X_1 + X_2 < 3|\lambda = 2\} = \sum_{i=0}^{2} \frac{(2\lambda)^i e^{-2\lambda}}{i!}|_{\lambda=2} = 0.2381$.

8.9 某厂商声称他们生产的某种型号的装潢材料抗断强度（单位：MPa）服从正态

分布，平均抗断强度为3.25，方差$\sigma^2=1.21$，今从中随机抽取9件进行检验，测得平均抗断强度为3.15，问能否接受厂商的说法（$\alpha=0.05$）？

解：已知正态总体方差为1.21, 检验$H_0:\mu=3.25\quad H_1:\mu\neq 3.25$ 由样本$\overline{X}=3.15, Z=\frac{\overline{X}-3.25}{\sigma_0/\sqrt{n}}=-0.2727$, $Z_{0.025}=1.96$,又$|Z|=0.027<1.96$, 所以不拒绝H_0,即可以认为平均抗断强度为3.25, 接受厂商说法。

8.10 某厂对废水进行处理，要求某种有毒物质的浓度小于$19(g/l)$。抽样检查得到10个数据，其样本均值$\overline{x}=17.1(g/l)$。设有毒物质含量服从正态分布，且已知方差$\sigma^2=8.5(g^2/l^2)$。问在显著性水平$\alpha=0.05$下处理后的废水是否合格？

解：$X\sim N(\mu,\sigma^2),\sigma^2=8.5$, 建立假设$H_0:\mu\geq 19,\quad H_1:\mu<19$.拒绝域为$W=\{\frac{\overline{X}-19}{\sqrt{8.5/10}}<-Z_0.05\}$, 因为$\frac{\overline{X}-19}{\sqrt{8.5/10}}=\frac{17.1-19}{\sqrt{8.5/10}}=-2.065<-1.645$, 所以拒绝$H_0$, 认为废水合格。

8.11 假定考生成绩服从正态分布，在某地一次数学统考中，随机抽取了36位考生的成绩，算得平均成绩为66.5分，标准差为15分，问在显著性水平0.05下，是否可以认为这次考试全体考生的平均成绩为70分？

解：$X\sim N(\mu,\sigma^2)$, 建立假设$H_0:\mu=70,\ \ H_1:\mu\neq 70$. 拒绝域为$W=\{|\frac{\overline{X}-70}{\sqrt{s^2/n}}|\geq t_{0.025}(35)\}$, 因为$|\frac{\overline{X}-70}{\sqrt{s^2/n}}|=|\frac{66.5-70}{\sqrt{15^2/36}}|=1.4<t_{0.025}(35)=2.0301$, 所以不拒绝$H_0$, 认为平均成绩为70。

8.12 电视台广告部称某类企业在该台黄金时段内播放电视广告后的平均受益量（平均利润增加量）至少为15万元，已知这类企业广告播出后的受益量近似服从正态分布，为此，某调查公司对该电视台广告播出后的此类企业进行了随机抽样调查，抽出容量为20的样本，得平均受益量为13.2万元，标准差为3.4万元，试在$\alpha=0.05$下判断该广告部的说法是否正确？

解：建立假设$H_0: \mu \geq 15$, $H_1: \mu < 15$. 拒绝域为$W = \{\frac{\overline{X}-15}{\sqrt{s^2/n}} < -t_{0.05}(19) = -1.7291\}$, 因为$T = \frac{13.2-15}{\sqrt{3.4^2/20}} = -2.3676 < -1.7291$, 所以拒绝$H_0$, 认为说法不正确。

8.13 下表分别给出两个文学家马克·吐温(Mark Twain)的8篇小品文以及斯诺特格拉斯(Snodgrass)的10篇小品文中个字母组成的单词的比例：

马克·吐温	0.225	0.262	0.217	0.240	0.230	0.229	0.235	0.217		
斯诺特格拉斯	0.209	0.205	0.196	0.210	0.202	0.207	0.224	0.223	0.220	0.201

设两组数据分别来自正态分布，且两总体方差相等，但参数均未知，两样本相互独立，问两作家所写的小品文中包含由三个字母组成的单词的比例是否有显著的差异(取$\alpha = 0.05$)？

解：设马克·吐温,斯诺特格拉斯小品文中包含由三个字母组成的单词的比例依次为$X \sim N(\mu_1, \sigma^2), Y \sim N(\mu_2, \sigma^2)$, 建立假设$H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$. ，检验统计量：$T = \frac{\overline{X}-\overline{Y}}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}$, 拒绝域为$W = \{\frac{\overline{X}-\overline{Y}}{S_w\sqrt{1/n+1/m}} \geq t_{m+n-2}(\frac{\alpha}{2}) = t_{16}(0.025)\}$, 计算$\overline{X} = \frac{1}{8}(0.225 + \cdots + 0.217) = 0.232, \overline{Y} = 0.2097, S_1^2 = 0.000212175, S_2^2 = 0.0000933, S_w^2 = 0.0001453$ 查表$t_{16}(0.025) = 2.1199$, 计算$|T| = 3.8781 > 2.1199$, 所以拒绝H_0, 认为有显著性差异。

8.14 设有甲、乙两种零件彼此可以代用，但乙零件比甲零件制造简单、造价低，经过试验获得它们的抗压强度如下表（单位：Kg/cm^2）

甲种零件	88	87	92	90	91
乙种零件	89	89	90	84	88

已知甲、乙两种零件的抗压强度分别服从正态分布$N(\mu_1, \sigma^2)$, $N(\mu_2, \sigma^2)$。问能否在保证抗压强度质量下，用乙种零件代替甲种零件（取$\alpha = 0.05$）？

解：建立假设$H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$. 检验统计量$T = \frac{\overline{X}-\overline{Y}}{S_w\sqrt{\frac{1}{5}+\frac{1}{5}}} \sim t_8$, 拒绝域为$W = \{|T| \geq t_{m+n-2}(\frac{\alpha}{2}) = t_8(0.025)\}$, 计算$\overline{X} = 89.6, \overline{Y} = 88, S_1^2 = 4.3, S_2^2 =$

$5.5, S_w^2=4.9$ 查表$t_8(0.025)=2.3060$, 计算$|T|=1.143<2.3060$, 所以不拒绝H_0, 乙可以代替甲。

8.15 在针织品漂白工艺过程中，要考察温度对针织品断裂强力（主要质量指标）的影响。为了比较$70°C$与$80°C$的影响有无差别，在这两个温度下，分别重复做了8次试验，得数据如下（单位：N）

$70°C$时的强力	20.5	18.8	19.8	20.9	21.5	19.5	21.0	21.2
$80°C$时的强力	17.7	20.3	20.0	18.8	19.0	20.1	20.0	19.1

根据经验，温度对针织品断裂强力的波动没有影响。问在$70°C$时的平均断裂强力与$80°C$时的平均断裂强力间是否有显著差别？（假定断裂强力服从正态分布，$\alpha=0.05.$）

解：$70°C$时的平均断裂强力记为$X\sim N(\mu_1,\sigma^2)$, $80°C$时的平均断裂强力记为$Y\sim N(\mu_2,\sigma^2)$. 建立假设$H_0:\mu_1=\mu_2,\quad H_1:\mu_1\neq\mu_2$. 检验统计量$T=\frac{\overline{X}-\overline{Y}}{S_w\sqrt{\frac{1}{8}+\frac{1}{8}}}\sim t_{14}$, 拒绝域为$W=\{|T|\geq t_{m+n-2}(\frac{\alpha}{2})=t_{14}(0.025)\}$, 计算$\overline{X}=20.4, \overline{Y}=19.375, S_1^2=0.8857, S_2^2=0.7879, S_w^2=0.8368 s_w=0.91$ 查表$t_{14}(0.025)=2.1448$, 计算$|T|=2.241>2.1448$, 所以拒绝H_0, 有显著性差异。

8.16 由于存在声音反射的原因，人们讲英语时在辅音识别上会遇到麻烦。有人随机选取了20个以英语为母语的人（记为A组）和20个以英语为外国语的人（记为B组）进行了试验，下面记录了他们正确反映的比例（%）：

A:	93	85	87	86	89	92	95	88	85	86
	90	91	89	88	91	85	84	87	89	88
B:	76	82	78	85	79	86	80	83	82	80
	78	79	76	75	80	82	83	82	85	86

假定这些数据来自正态分布，其具有公共方差，试在$\alpha = 0.05$下，检验这两组的反应是否有显著性差异？

解：A组人的反应记为$X \sim N(\mu_1, \sigma^2)$, B组人的反应记为$Y \sim N(\mu_2, \sigma^2)$. 建立假设$H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$. 检验统计量$T = \frac{\overline{X}-\overline{Y}}{S_w\sqrt{\frac{1}{20}+\frac{1}{20}}} \sim t_{38}$, 拒绝域为$W = \{|T| \geq t_{m+n-2}(\frac{\alpha}{2}) = t_{38}(0.025)\}$, 计算$\overline{X} = 88.4, \overline{Y} = 80.85, S_1^2 = 8.67, S_2^2 = 10.98, S_w^2 = 9.825$, 查表$t_{38}(0.025) = 2.0244$, 计算$|T| = 7.6169 > 2.0244$, 所以拒绝$H_0$, 有显著性差异。

8.17 随机的选了8个人,分别测量了他们在早晨起床时和晚上就寝时的身高(cm),得到以下数据:

序号	1	2	3	4	5	6	7	8
早上(x_i)	172	168	180	181	160	163	165	177
晚上(y_i)	172	167	177	179	159	161	166	175
$d_i = x_i - y_i$	0	1	3	2	1	2	-1	2

设各对数据的差$D_i = X_i - Y_i (i = 1, 2, \cdots, 8)$是来自正态总体$N(\mu_D, \sigma_D^2)$的样本,$\mu_D, \sigma_D^2$均未知. 问是否可以认为早晨的身高比晚上的身高要高(取$\alpha = 0.05$)?

解：建立假设$H_0: \mu \leq 0$, $H_1: \mu > 0$. $\overline{d} = \frac{1}{8}\sum_{i=1}^{8} D_i = 1.25, s_d^2 = \frac{1}{7}\sum_{i=1}^{8}(d_i - \overline{d})^2 = 1.642857$. 检验统计量$T = \frac{\overline{d}}{s_d/\sqrt{8}} \sim t_7$, 拒绝域为$W = \{T \geq t_7(0.05)\}$, 计算$T = 2.758386$, 查表$t_7(0.05) = 1.8946$, 计算$T = 2.758386 > 1.8946$, 所以拒绝$H_0$, 早上比晚上高。

8.18 用中草药青木香治疗高血压，记录了13个病例，所测定的舒张压(mmHg)数据如下：

序号	1	2	3	4	5	6	7	8	9	10	11	12	13
治疗前	110	115	133	133	126	108	110	110	140	104	160	120	120
治疗后	90	116	101	103	110	88	92	104	126	86	114	88	112

试检验该药是否具有降低血压的作用($\alpha = 0.01$)。

解：$X_1 \cdots X_{13}$ 为治疗前的舒张压（mmHg）,$Y_1 \cdots Y_{13}$为治疗后的舒张压（mmHg）.$D_i = X_i - Y_i, i = 1 \cdots 13, \sim N(\mu, \sigma^2)$ 建立假设$H_0 : \mu \le 0,\quad H_1 : \mu > 0$. $\overline{d} = \frac{1}{13}\sum_{i=1}^{13} D_i = 19.92308, s_d^2 = \frac{1}{7}\sum_{i=1}^{8}(d_i - \overline{d})^2 = 158.7436$. 检验统计量$T = \frac{\overline{d}}{s_d/\sqrt{13}} \sim t_{12}$, 拒绝域为$W = \{T \ge t_{12}(0.01)\}$, 计算$T = 5.70138$, 查表$t_{12}(0.01) = 2.6810$, 计算$T = 5.70138 > 2.6810$, 所以拒绝$H_0$, 认为该药有效。

8.19 根据设计要求，某零件的内径标准差不得超过0.30（单位：cm），现从该产品中随机抽验了25件，测得样本标准差为0.36，问该批产品是否合格？($\alpha = 0.05$)

解：设零件内径服从正态分布记为$X \sim N(\mu, \sigma^2)$, 建立假设$H_0 : \sigma^2 \le 0.3^2,\quad H_1 : \sigma^2 > 0.3^2$. 检验统计量$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{24}$, 拒绝域为$W = \{\frac{24S^2}{0.3^2} \ge \chi^2_{24}(0.05)\}$, 计算$S^2 = 0.36^2, \frac{24\times 0.36^2}{0.3^2} = 34.56$, 查表$\chi^2_{24}(0.05) = 36.415$ 所以不拒绝H_0, 产品是合格的。

8.20 A厂三车间生产铜丝的折断力服从正态分布，生产一直比较稳定，今从产品中随机抽出9根检验折断力，测得数据如下（单位；Kg）：

289 268 285 284 286 285 286 298 292

问是否可相信该车间的铜丝折断力的方差为20($\alpha = 0.05$)?

解：设A厂三车间生产铜丝的折断力服从正态分布记为$X \sim N(\mu, \sigma^2)$, 建立假设$H_0 : \sigma^2 = 20,\quad H_1 : \sigma^2 \ne 20$. 检验统计量$\frac{(n-1)S^2}{20} \sim \chi^2_8$, 拒绝域为$W = \{\frac{8S^2}{20} \ge \chi^2_8(0.025), \frac{8S^2}{20} \le \chi^2_8(0.975)\}$, 计算$S^2 = 64.86114, \frac{8\times S^2}{20} = 25.9444$, 查表$\chi^2_8(0.025) = 17.534, \chi^2_8(0.975) = 2.180$ 所以拒绝H_0, 不可相信铜丝折断力的方差为20。

8.21 某电工器材厂生产一种保险丝, 保险丝的熔化时间服从正态分布. 按规定, 熔化时间的方差不得超过400. 今从一批产品中随机抽取25件样品, 测得熔化时间的方差为410, 问在显著性水平$\alpha = 0.05$下, 能认为这批产品的方差显著偏大吗?

解：假设

$$H_0 : \delta_0^2 \leq 400, H_1 : \delta_0^2 > 400.$$

用χ^2检验, $n = 25$, $S^2 = 410$, 查χ^2分布表得

$$\chi_{n-1}^2(\alpha) = \chi_{24}^2(0.05) = 36.415,$$

又由样本值得

$$\chi^2 = \frac{(n-1)S^2}{\delta_0^2} = 24.6 < \chi_{n-1}^2(\alpha).$$

故不能拒绝H_0, 即不能认为产品方差显著偏大.

8.22 某电工器材厂生产的漆包线的抗拉强度服从均值$\mu = 25$, 方差 $\delta^2 = (1.8)^2$的正态分布. 现由于原材料发生了变化, 对现在生产的产品随机抽取12个进行测试, 的抗拉强度如下：

22.5	28.3	25.6	29.3	24.4	26.3	27.4	27.3	25.2	30.2	28.4	29.1

试在$\alpha = 0.01$下, 分别检验假设$H_0 : \mu = 25$和$H_0 : \delta^2 = 1.8^2$.

解：(1) 检验假设$H_{10} : \mu = 25, H_{11} : \mu \neq 25; \delta^2$ 未知, 令$T = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$, 则拒绝域为$W = \{|T| > t_{n-1}(\frac{\alpha}{2}\}$, 由样本观测值$\overline{x} = 27$, $S^2 = \frac{1}{11}\sum_{i=1}^{12}(x_i - \overline{x})^2 = 5.14$, 所以$T = 3.056$, 查$t$分布表得到$t_{n-1}(\frac{\alpha}{2}) = t_{11}(0.005) = 3.1058$. 因为$|T| < t_{n-1}(\frac{\alpha}{2})$, 故不能拒绝$H_0$;

(2) 检验假设$H_{20} : \delta^2 = (1.8)^2, H_{21} : \delta^2 \neq (1.8)^2; \delta^2$ 未知, 用χ^2检验法, 由样本观测值知$\chi^2 = \frac{(n-1)S^2}{\delta_0^2} = 17.45$, 查$\chi^2$分布表得

$\chi^2_{n-1}(\frac{\alpha}{2}) = \chi^2_{11}(0.005) = 26.757$, $\chi^2_{n-1}(1-\frac{\alpha}{2}) = \chi^2_{11}(0.995) = 2.603$, 因为$chi^2_{n-1}(1-\frac{\alpha}{2}) < \chi^2 < \chi^2_{n-1}(\frac{\alpha}{2})$,

故不能拒绝原假设, 即认为方差$\delta^2 = (1.8)^2$.

8.23 某教师同时给A, B两个班级讲授同一门课程, 先从A班随机抽取16名学生, 从B班随机抽取26名学生. 在同一次测试中, A班成绩的样本标准差为$s_1 = 9$, B 班成绩的样本标准差为$s_2 = 12$, 假设A, B两班测试成绩均服从正态分布. 在显著性水平$\alpha = 0.01$下, 能否认为B班成绩的标准差比A 班大?

解：假设

$$H_0 : \delta_1^2 \geq \delta_2^2, H_1 : \delta_1^2 < \delta_2^2.$$

用F检验法, $F = \frac{s_1^2}{s_2^2} \sim F_{m-1,n-1}$, 拒绝域为$W = \{F < F_{15,25}(1-\alpha)\}$, 计算得$F = \frac{s_1^2}{s_2^2} = \frac{81}{144} = 0.5625$, 查表得：$F_{15,25}(0.99) = 0.3050$. 因为$F > F_{15,25}(1-\alpha)$, 所以不能拒绝$H_0$, 即不能认为$B$班标准差比$A$班大.

8.24 用老工艺生产的机械零件方差较大, 抽查了25个, 得$s_1^2 = 6.37$. 现改用新工艺生产，随机抽查25个零件, 得$s_2^2 = 3.19$. 设这两种生产过程皆服从正态分布, 问新工艺的精度是否比老工艺的精度显著地高($\alpha = 0.05$)?

解：假设

$$H_0 : \delta_1^2 \leq \delta_2^2, H_1 : \delta_1^2 > \delta_2^2.$$

用F检验法. $m = n = 25$, 查F分布得到：$F_{24,24} = 1.98$, 由样本值计算得

$$F = \frac{s_1^2}{s_2^2} = 1.99 > F_{m-1,n-1}(\alpha),$$

故拒绝H_0, 即认为新工艺精度比老工艺好.

8.25 某厂使用两种不同的原料生产同一类型产品, 随机选取使用原料A生产的样品22件, 测得平均质量为$2.36kg$, 样本标准差为$0.57kg$. 选取使用原料B生产的样品24

件, 测得平均质量为$2.55kg$, 样本标准差为$0.48kg$. 假定产品质量服从正态分布, 两个样本独立. 问能否认为使用原料B生产的产品平均质量较使用原料A生产的显著大(取$\alpha = 0.05$)?

解：假设

$$H_0 : \mu_2 \leq \mu_1, H_1 : \mu_2 > \mu_1.$$

已知$n_1 = 22, \overline{X} = 2.36, \delta_1 = 0.57, n_2 = 24, \overline{Y} = 2.55, \delta_2 = 0.48$.选取检验统计量

$$U = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\delta_1^2}{n_1} + \frac{\delta_2^2}{n_2}}} \sim N(0,1)$$

则$U = 1.22$, 查表知$U_\alpha = U_{0.05} = 1.64$. $U < U_\alpha$, 故不能拒绝H_0, 即不能认为使用原料B生产的产品平均质量较使用原料A生产的显著大.

8.26 某化工厂为了提供某种化学药品独立的得率, 提出了两种工艺方案. 为了研究哪一种方案好, 分别用两种工艺进行试验, 数据如下：

甲得率(%)	68.1	62.4	64.3	64.7	68.4	66.0	65.5	66.7	67.3	66.2
乙得率(%)	69.1	71.0	69.1	70.0	69.1	69.1	67.3	70.2	72.1	67.3

假设甲、乙方案的得率分别服从正态分布, 问方案乙是否比方案甲能显著提高得率($\alpha = 0.01$)?

解：(1)两个正态总体数学期望的比较：

假设$H_0 : \mu_2 \leq \mu_1, H_1 : \mu_2 > \mu_1$.

$z_i = X_i - Y_i : -1 \quad -8.6 \quad -4.8 \quad -5.3 \quad -0.7 \quad -3.3 \quad -1.8 \quad -3.5 \quad -4.8 \quad -1.1$

采用T检验, 则

$\overline{z} = \frac{1}{n}\sum_{i=1}^{n} z_i = -2.716$, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \overline{z})^2 \approx 6.8222$, $S = 2.6119$.

$T = \frac{\overline{z}}{S}\sqrt{n} = 3.1196 \sim t(9)$. 固$3.1196 > t_0.99(9)$.

(2)两个正态总体方差比较：

$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} x_i = 65.96,$

$S_1^2 = \frac{1}{9}\sum^1 0_{i=1}(x_i - \overline{X})^2 = \frac{1}{9}(2.14^2 + 3.56^2 + 1.66^2 + 1.26^2 + 2.44^2 + 0.04^2 + 0.46^2 + 0.74^2 + 1.34^2 + 0.24^2) = 3.3516$

$\overline{Y} = \frac{1}{10}\sum_{i=1}^{10} y_i = 69.43,$

$S_1^2 = \frac{1}{9}\sum^1 0_{i=1}(y_i - \overline{Y})^2 = \frac{1}{9}(0.33^2 + 1.57^2 + 0.33^2 + 0.57^2 + 0.33^2 + 0.33^2 + 2.13^2 + 0.77^2 + 2.67^2 + 2.13^2) = 2.2246$

采用F检验, $F = \frac{S_1^2}{S_2^2} \sim F(9,9)$, $F = 1.5067 < 5.35$. 故方案乙能显著提高效率.

8.27 有人称某地成年人中大学生毕业生比例不低于30%. 为检验之, 随机调查该地15名成年人, 发现有3名大学毕业生. 取$\alpha = 0.05$, 问该人看法是否成立？并给出检验的pz值.

解：以X表示人的学历水平：

$$X = \begin{cases} 0 & \text{尚未大学毕业} \\ 1 & \text{大学毕业以上} \end{cases}$$

则$X \sim B(1,p)$, 其中p表示大学生比例. 问题归结为检验假设.

$$H_0 : p = p_0 = 30\%, H_1 : p \neq p_0.$$

$\overline{X}$是样本比例,$\hat{p} = \frac{3}{15} = 20\%$, $\hat{p}$的均值是p, 方差是$\frac{p(1-p)}{n}$. 则$\hat{P} \sim N(p, \frac{p(1-p)}{n})$.

因此当原假设成立时, 取检验统计量$Z = \frac{\hat{p}-p_0)}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1)$.

对于给定的显著水平$\alpha = 0.05$, $Z_{0.05} = 1,96$, 故拒绝域$W = |Z| > 1.96$.

统计量Z的实测值$Z = \frac{0.1}{\sqrt{0.014}} = 0.8451 < 1.96$.

故不能拒绝原假设. 可以认为该地成年人中大学生毕业生比例不低于30%.

此处检验p值, $p = 2p\{Z \geq 0.8451\} = 0.1991 > \alpha = 0.05$.

8.28 某厂的产品质量较好, 不合格品率不超过0.05. 最近, 该厂采用新工艺, 从新工艺生产的产品中随机抽取100件产品进行检查, 发现有6 个不合格品. 试问, 在下列两个情形下能否认为产品的质量没有显著变化?

(1)为了维护生产方的利益, 要求犯第一类错误的概率不超过0.1; (2)为了维护使用方的利益, 要求犯第一类错误的概率不超过0.1.

解: (1)检验假设$H_0: p \leq 0.05$, $H_1: p > 0.05$.

检验统计量$Z = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1)$.

其中$\hat{p} = \frac{6}{100}, p_0 = 0.05, n = 100, \alpha = 0.1$.

则$Z = \frac{0.06-0.05}{\sqrt{\frac{0.05\times 0.95}{100}}} = \frac{0.1}{\sqrt{0.0475}} = 0.459$, $Z_\alpha = 1.29$.

拒绝域: $W = Z > Z_\alpha$. 因为$Z < Z_\alpha$, 所以不能拒绝原假设.

(2)检验假设$H_0: p \geq 0.05$, $H_1: p < 0.05$.

检验统计量$Z = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1)$.

则$Z = 0.459$, $Z_\alpha = 1.29$.

拒绝域: $W = Z < -Z_\alpha$, 所以不能拒绝原假设.

故认为产品质量有显著变化.

8.29 某大学随机调查120名男同学, 发现有50人非常喜欢看武侠小说, 而随机调查的85名女同学中有23人喜欢, 在$\alpha = 0.05$下用大样本检验方法确认: 男女同学在喜欢武侠小说方面有无显著性差异? 并给出检验的p值.

解: 假设n_1名男同学中有$n_1\overline{X} = \sum_{i=1}^{n_1} x_i$人喜爱看武侠小说, 而$n_2$名女同学中有$n_2\overline{Y} = \sum_{i=1}^{n_2} y_i$人喜爱看武侠小说.

有$n_1\overline{X} \sim B(n_1, p_1)$, $n_2\overline{Y} \sim B(n_2, p_2)$大样本.

有$\overline{X} \sim N(p_1, \frac{p_1(1-p_1)}{n_1})$, $\overline{Y} \sim N(p_2, \frac{p_2(1-p_2)}{n_2})$.

则$\overline{X} - \overline{Y} \sim N(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2})$.

即$\frac{\overline{X}-\overline{Y}-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}} \sim N(0,1)$.

当P_1与p_2未知时, 用总频率$\hat{p}=\frac{n_1\overline{X}+n_2\overline{Y}}{n_1+n_2}$作为参数, p的点估计替换p. 大样本下有

$$U=\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\sqrt{\hat{p}(1-\hat{p})}} \sim N(0,1)$$

假设$H_0: p_1=p_2$, 拒绝域$H_1: p_1 \neq p_2$.

显著水平$\alpha=0.05$, $u_{1-\frac{\alpha}{2}}=1.96$, 双侧拒绝域$W=|u| \geq 1.96$.

因$n_1=120, n_2=85, n_1\overline{X}=50, n_2\overline{Y}=25$. 有$\hat{p}=0.3561$.

此时$u=\frac{\frac{50}{120}-\frac{23}{85}}{\sqrt{\frac{1}{120}+\frac{1}{85}}\sqrt{0.3561(1-0.3561)}}=2.1519 \in W$.

并且检验值$p=2pU \geq 2.1519=0.0314<\alpha=0.05$.

故拒绝H_0, 接受H_1, 认为男女同学在喜爱武侠小说上有差异.

8.30 假定电话总机在单位时间内接到呼叫次数服从泊松分布, 现观测了40个单位时间, 接到的呼叫次数如下：

0	2	3	2	3	2	1	0	2	2	1	2	2	1	3	1	1	4	1	1
5	1	2	2	3	3	1	3	1	3	4	0	6	1	1	1	4	0	1	3

问在显著性水平0.05下, 能否认为单位时间内接到的平均呼叫次数不低于2.5次？并给出检验的p值.

解：以X记电话总机在该单位时间内接到的呼叫次数, 设$X \sim P(\lambda)$.

检验假设$H_0: \lambda \geq 2.5$, $H_1: \lambda < 2.5$.

由于$n=40$较大, 可采用大样本检验, 检验统计量为$u=\frac{\sqrt{n}(\overline{X}-2.5)}{\sqrt{2.5}}$,

拒绝域为$W=u \leq u_{0.05}=\{u \leq -1.645\}$.

有样本计算得$u=-2.1<-1.645 \quad (\overline{X}=1.975)$, 拒绝原假设.

检验的p值为$p=P\{u \leq -2.1\}=-0.0179$.

所以不能认为单位时间内接到的呼叫次数不低于2.5次.

8.31 通常每平方米某种布上的瑕疵点数服从泊松分布, 现观测该种布$100m^2$, 发现有126个瑕疵点, 在显著性水平为0.05下, 能否认为该种布每平方米上平均瑕疵点数不超过1个？并给出检验的p值.

解：设每平方米该种布上的瑕疵点数$X \sim P(\lambda)$. 有$n\overline{X} = \sum_{i=1}^{n} x_i \sim P(n\lambda)$大样本,

有$\frac{n\overline{X}-n\lambda}{\sqrt{n\lambda}} = \frac{\overline{X}-\lambda}{\sqrt{\lambda/n}} \sim N(0,1)$.

假设$H_0 : \lambda = 1, H_1 : \lambda > 1$.

选取统计量$U = \frac{\overline{X}-\lambda}{\sqrt{\lambda/n}} \sim N(0,1)$.

显著水平$\alpha = 0.05, u_{1-\alpha} = 1.645$, 右侧拒绝域$W = u \geq 1.645$.

由于$\overline{X} = 1.26, n = 100, \lambda = 1$, 则$u = 2.6 \in W$.

并且检验p值, $p = Pu \geq 2.6 = 0.0047 < \alpha = 0.05$.

故拒绝域H_0, 接受H_1, 不能认为该种布每平方米上瑕疵点数不超过1个.

8.32 有人对$\pi = 3.1415926\cdots$的小数点后800位数字0, 1, 2,$\cdots$,9出现的次数进行了统计, 结果如下

数字	0	1	2	3	4	5	6	7	8	9
次数	74	92	83	79	80	73	77	75	76	91

试在显著性水平下为0.05下, 检验每个数字出现的概率是否相同.

解：这是一个分布拟合优度检验, 总体共10类, 若记出现数字i的概率为p_i,

则要检验的假设为$H_0 : p_0 = p_1 = \cdots = p_9 = 0.1$. 这是$k = 10$, 检验拒绝域为$X^2 \geq X^2_{1-\alpha}(9)$.

若取$\alpha = 0.05$, 则查表$X^2_{0.95}(9) = 16.9190$.

检验统计量为$X^2 = \frac{(74-80)^2}{80} + \frac{(92-80)^2}{80} + \cdots + \frac{(91-80)^2}{80} = 50125$.

由于$X^2=5.125$未落入拒绝域, 故不拒绝原假设.

在显著性水平0.05下可以认为数字出现概率相同.

此处检验p值$p=PrX^2(9)\geq 50125=0.8233$.

8.33 设$A_i=\{x:\frac{i-1}{4}\leq x<\frac{i}{4}\}, i=1,2,3, A_4=\{x:\frac{3}{4}\leq x\leq 1\}$, 现取80个观测值，其中落入$A_i(i=1,2,3,4)$的频数分别为6，18，20，36.问在显著性水平$\alpha=0.01$下，可否认为总体的分布函数为

$$F_0(x)=\begin{cases}0 & x<0,\\ x^2 & 0\leq x<1,\\ 1 & x\geq 1\end{cases}$$

解：$H_0:F(x)=F_0(x), H_1:F(x)\neq F_0(x)$, 取$a_1=\frac{1}{4}, a_2=\frac{1}{2}, a_3=\frac{3}{4}, a_4=1$, 将$[0,1]$分成4个互不相交的区间$[0,\frac{1}{4}),[\frac{1}{4},\frac{1}{2}),[\frac{1}{2},\frac{3}{4}),[\frac{3}{4},1]$.

$$p_1=F(\frac{1}{4})-F(0)=\frac{1}{16}, p_2=F(\frac{1}{2})-F(\frac{1}{4})=\frac{3}{16}, p_3=F(\frac{3}{4})-F(\frac{1}{2})=\frac{5}{16}, p_4=F(1)-F(\frac{3}{4})=\frac{7}{16},$$

i	$[a_{i-1},a_i)$	f_i	np_i	$\frac{(f_i-np_i)^2}{np_i}$
1	$[0,\frac{1}{4})$	6	5	1/5
2	$[\frac{1}{4},\frac{1}{2})$	18	15	3/5
3	$[\frac{1}{2},\frac{3}{4})$	20	25	1
4	$[\frac{3}{4},1]$,	36	35	1/35

$$\chi^2=\sum_{i=1}^{4}\frac{(f_i-np_i)^2}{np_i}=1.83, \chi^2_{m-k-1}(\alpha)=\chi^2_{4-0-1}(0.01)=11.344, \chi^2<\chi^2_3(0.01),$$

$\therefore$ 不拒绝H_0, 认为总体分布函数为$F_0(x)$.

8.34 某赌徒被指责使用一颗灌过铅的骰子，但他辩称此骰子没有问题。一组记录保存了最近的60次掷该骰子所得的数据，结果如下表所示：

点数	1	2	3	4	5	6
观测次数	4	6	17	16	8	9

试问此骰子是否有问题（$\alpha = 0.05$）？

解：X表示掷骰子出现的点数，$H_0 : P\{X = i\} = \frac{1}{6}, i = 1, 2, 3, ..., 6, H_1$:至少存在一个i, 使$P\{X = i\} \neq \frac{1}{6}$,

X	p_i	np_i	f_i	$\frac{(f_i-np_i)^2}{np_i}$
1	$\frac{1}{6}$	10	4	3.6
2	$\frac{1}{6}$	10	6	3.6
3	$\frac{1}{6}$	10	17	4.9
4	$\frac{1}{6}$	10	16	3.6
5	$\frac{1}{6}$	10	8	0.4
6	$\frac{1}{6}$	10	9	0.1

$\chi^2 = \sum_{i=1}^{6} \frac{(f_i-np_i)^2}{np_i} = 14.2, \chi^2_{m-k-1}(\alpha) = \chi^2_5(0.05) = 11.07, \chi^2 > \chi^2_5(0.05)$, 所以拒绝$H_0$, 认为此骰子有问题.

8.35 某建筑工地每天发生事故数现场记录如下：

一天发生的事故数	0	1	2	3	4	5	≥ 6	合计
天数	102	59	30	8	0	1	0	200

试在显著性水平$\alpha = 0.05$下检验这批数据是否服从泊松分布？

解：设每天发生的事故数为X, $H_0 : P\{X = i\} = \frac{\lambda^i}{i!}\mathrm{e}^{-\lambda}, H_1$: 至少存在一个$i$, 使得$P\{X = i\} \neq \frac{\lambda^i}{i!}\mathrm{e}^{-\lambda}$, $\hat{\lambda} = \bar{x} = \frac{1}{200}\sum_{i=1}^{6} f_i x_i = 0.74, p_i = \frac{0.74^i}{i!}\mathrm{e}^{-0.74}$

X	p_i	np_i	f_i	$\frac{(f_i-np_i)^2}{np_i}$
0	0.4771	95.42	102	0.4538
1	0.353	70.6	59	1.906
2	0.1306	26.12	30	0.576
3	0.0322	6.44	8	$\}$ 0.1653
4	0.006	1.2	0	
5	0.0009	0.18	1	
6	0.0002	0.04	0	

$\chi^2=\sum_{i=0}^{3}\frac{(f_i-np_i)^2}{np_i}=3.1011, \chi^2_{m-k-1}(\alpha)=\chi^2_2(0.05)=5.992, \chi^2<\chi^2_2(0.05)$, 所以不拒绝$H_0$, 即这批数据服从Poisson分布.

8.36 为考察高中生的性别与是否喜欢数学课程之间的关系，在某城市的某校高中生中随机抽取300名学生，得到如下列联表：

	喜欢数学课程	不喜欢数学课程	合计
男	37	85	122
女	35	143	178
合计	72	228	300

试在显著性水平0.05下判断高中生的性别与是否喜欢数学课程之间有无关系？

解：A表示性别，有两个属性，A_1表示男，A_2表示女，B表示爱好，有两个属性，B_1表示喜欢数学，B_2表示不喜欢数学，

$H_0: p_{ij}=p_{i\cdot}p_{\cdot j}, i=1,2, j=1,2,$　　　　H_1：至少有一对$(i,j), s.t. p_{ij}\neq p_{i\cdot}p_{\cdot j}$.

当H_0成立时，$n_{1\cdot}=122, n_{2\cdot}=178, n_{\cdot 1}=72, n_{\cdot 2}=228, n=300$,

$$\chi^2=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(nn_{ij}-n_{i\cdot}n_{\cdot j})^2}{nn_{i\cdot}n_{\cdot j}}=4.514.\chi_1^2(0.05)=3.843<4.514,$$

所以拒绝原假设，认为高中生性别与是否喜欢数学课有关系.

8.37 为研究慢性气管炎与吸烟量的关系，调查了813人，各类人数统计数字如下：

健康状况	吸烟量			合计
	0	$1\sim 5$	>5	
患病	126	245	49	420
健康	152	209	32	393
合计	278	454	81	813

是否可以认为吸烟量对慢性气管炎没有影响（取$\alpha=0.05$）？

解：A表示健康状况，有两个属性，A_1表示患病，A_2表示健康，B表示吸烟量，有3个水平，B_1, B_2, B_3分别为表8.13中的三种情况，

H_0：吸烟量对慢性气管炎没有影响，即$p_{ij}=p_{i\cdot}p_{\cdot j}, i=1,2, j=1,2,3, n_{1\cdot}=420, n_{2\cdot}=393, n_{\cdot 1}=278, n_{\cdot 2}=454, n_{\cdot 3}=81.$

$$\chi^2=\sum_{i=1}^{2}\sum_{j=1}^{3}\frac{(nn_{ij}-n_{i\cdot}n_{\cdot j})^2}{nn_{i\cdot}n_{\cdot j}}=9.0592, \chi_2^2(0.05)=5.992<9.0592,$$

所以拒绝H_0, 不能认为对慢性气管炎没有影响.

8.38 某调查机构连续三年对某城市的居民进行热点调查：对下列四个问题：(1)收入、(2)物价、(3)住房、(4)交通，要求被调查者选择其中之一作为最关心的问题，调查结果如下：

问题	收入	物价	住房	交通	和
1997年	155	232	87	50	524
1998年	134	201	100	75	510
1999年	176	114	165	61	516
和	465	547	352	186	1550

是否可以认为各年该城市居民对社会热点问题的看法保持不变（取$\alpha = 0.05$）？

解：A表示时间，A_1表示1997年，A_2表示1998年，A_3表示1999年，B表示关注的问题，有4个属性，B_1表示收入，B_2表示物价，B_3表示住房，B_4表示交通，

$$H_0 : p_{ij} = p_{i\cdot} p_{\cdot j}, i = 1, 2, 3, j = 1, 2, 3, 4$$

$n_{1\cdot} = 524, n_{2\cdot} = 510, n_{3\cdot} = 516, n_{\cdot 1} = 465, n_{\cdot 2} = 547, n_{\cdot 3} = 352, n_{\cdot 4} = 186$

$$\chi^2 = \sum_{i=1}^{3} \sum_{j=1}^{4} \frac{(nn_{ij} - n_{i\cdot} n_{\cdot j})^2}{nn_{i\cdot} n_{\cdot j}} = 81.2423, \chi_6^2(0.05) = 12.592 < 81.2423,$$

所以拒绝H_0, 不可以认为各年该城市居民对社会对社会热点问题的看法保持不变.