第一章　概率论的基本概念

习题 1—1　随机事件

1.设 A, B, C 表示三个事件，试将下列事件用 A, B, C 表示出来：

（1）A, C 都发生，B 不发生；　【 $A\bar{B}C$，　$AC-B$ 】

（2）三个事件中至少有一个发生；　【 $A\cup B\cup C$ 】

（3）三个事件中至少有两个．　【 $AB\cup AC\cup BC$，　$AB\bar{C}+A\bar{B}C+\bar{A}BC+ABC$ 】

2．设某人对一目标接连进行三次射击，设 $A_i=\{$第 i 次命中$\}$（$i=1,2,3$）；$B_j=\{$射击恰好命中 j 次$\}$（$j=0,1,2,3$）；$C_k=\{$三次射击至少命中 k 次$\}$（$k=0,1,2,3$）.

（1）通过 A_1, A_2, A_3 表示 B_2；　【 $B_2=A_1A_2\bar{A}_3\cup A_1\bar{A}_2A_3\cup\bar{A}_1A_2A_3$ 】

（2）通过 B_1, B_2, B_3 表示 C_2 ．　【 $C_2=B_2\cup B_3$ 】

3. 设 A, B, C 为三个事件，指出下列各等式成立的条件.

（1）$ABC=A$；　【 $A\subset BC$ 】

（2）$A\cup B\cup C=A$；　【 $B\cup C\subset A$ 】

（3）$A\cup B=AB$；　【 $A=B$ 】

（4）$(A\cup B)-A=B$ 。　【 $AB=\phi$ 】

习题 1—2　概　率

1．设 $P(A)=P(B)=P(C)=\dfrac{1}{4}$，$P(AB)=P(AC)=P(BC)=\dfrac{1}{8}$，$P(ABC)=\dfrac{1}{16}$，求下列事件的概率：

（1）$P(A\cup B\cup C)$；　（2）．$P(\bar{A}\bar{B}\bar{C})$

解 （1）$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)=\dfrac{3}{4}-\dfrac{3}{8}+\dfrac{1}{16}=\dfrac{7}{16}$

（2）$P(\bar{A}\bar{B}\bar{C})=P\left(\overline{A\cup B\cup C}\right)=1-P(A\cup B\cup C)=\dfrac{9}{16}$.

2．从 **5** 双不同尺码的鞋子中任取 **4** 只，求至少有 **2** 只配成一双的概率.

解　$p=\dfrac{C_5^1C_4^2C_2^1C_2^1+C_5^2}{C_{10}^4}=\dfrac{13}{21}$，　或 $p=1-\dfrac{C_5^4C_2^1C_2^1C_2^1C_2^1}{C_{10}^4}=\dfrac{13}{21}$ ．

3．从$[0,1]$中随机地取两个数，求下列事件的概率：（**1**）两数之和小于$\dfrac{5}{4}$；（**2**）两数之积大于$\dfrac{1}{4}$；（**3**）以上两个条件均满足**.**

解 （1）设 A：两数之和小于$\dfrac{5}{4}$， 则有 $P(A)=\dfrac{1-\dfrac{1}{2}\times\dfrac{3}{4}\times\dfrac{3}{4}}{1}=\dfrac{23}{32}$.

（2）设 B：两数之积大于$\dfrac{1}{4}$，则有 $P(B)=\dfrac{\int_{\frac{1}{4}}^{1}(1-\dfrac{1}{4x})dx}{1}=\dfrac{3}{4}-\dfrac{1}{2}\ln 2$.

（3） $P(AB)=\dfrac{\int_{\frac{1}{4}}^{1}(\dfrac{5}{4}-x-\dfrac{1}{4x})dx}{1}=\dfrac{3}{4}-\dfrac{1}{2}\ln 2-\dfrac{1}{2}\times\dfrac{3}{4}\times\dfrac{3}{4}=\dfrac{15}{32}-\dfrac{1}{2}\ln 2$.

4．旅行社 **100** 人中有 **43** 人会讲英语，**35** 人会讲日语，**32** 人会讲日语和英语，**9** 人会讲法语、英语和日语，且每人至少会讲英、日、法三种语言中的一种，在其中任意挑选一人，求此人会讲英语和日语，但不会讲法语的概率.

解 设 A：会讲英语，B：会讲日语，C：会讲法语.

则有：$P(AB\bar{C})=P(AB)-P(ABC)=\dfrac{32}{100}-\dfrac{9}{100}=0.23$.

习题 **1-3** 条件概率

1．根据对电路停电情况的研究，得到电路停电原因的一下经验数据：**5%**是由于变电器损坏；**80%**是由于电路线损坏；**1%**是由于两者同时损坏**.** 试求下列各种停电事件发生的概率。（**1**）在已知变电器损坏的条件下，电路线损坏；（**2**）变电器损坏但电路线完好；（**3**）在已知电路线没损坏的条件下，变电器损坏**.**

解 A：变电器损坏，，B：电路线损坏，则 $P(A)=0.05,\quad P(B)=0.8,\quad P(AB)=0.01$

（**1**） $P(B|A)=\dfrac{P(AB)}{P(A)}=\dfrac{0.01}{0.05}=0.05,\quad P(B)=0.8,\quad P(AB)=0.2$；

（**2**） $P(A\bar{B})=P(A)-P(AB)=0.05-0.01=0.04$

（**3**） $P(A|\bar{B})=\dfrac{P(A\bar{B})}{P(\bar{B})}=\dfrac{P(A)-P(AB)}{1-P(B)}=\dfrac{0.05-0.01}{1-0.8}=0.2$.

2．一批灯泡共 **100** 只，次品率为 **10%**，不放回的抽取 **3** 次，每次取一只，问第 **3** 次才取到合格品的概率是多少？

解 记 A_i：第 i 次取到合格品，$(i=1,\ 2,\ 3)$ 所求概率即为：

$$P(\bar{A}_1\bar{A}_2A_3)=P(\bar{A}_1)P(\bar{A}_2\mid\bar{A}_1)P(A_3\mid\bar{A}_1\bar{A}_2)=\frac{10}{100}\times\frac{9}{99}\times\frac{90}{98}=\frac{9}{1078}.$$

3．玻璃杯成箱的出售，每箱 **20** 只，假设各箱含 **0** 个，**1** 个，**2** 个次品的概率相应的为 $0.8, 0.1, 0.1$，一顾客欲买一箱玻璃杯，售货员随意地抽取一箱，顾客开箱后随意地查看 **4** 只，若无次品则买下这箱玻璃杯，否则退回，试求：（**1**）顾客买下该箱玻璃杯的概率；（**2**）若一个顾客买下了一箱玻璃杯，在顾客买下的这箱玻璃杯中确实无次品的概率。

解 （1）记 A：顾客买下该箱玻璃杯， B_k：该箱含有 k 只次品，$k=0,\ 1,\ 2$．则有

$$P(A)=\sum_{k=0}^{2}P(B_k)P(A\mid B_k)=0.8\times 1+0.1\times\frac{C_{19}^4}{C_{20}^4}+0.1\times\frac{C_{18}^4}{C_{20}^4}=0.8+0.1\times\frac{4}{5}+0.1\times\frac{12}{19}=\frac{448}{475}=0.94$$

（2）$P(B_0\mid A)=\dfrac{P(AB_0)}{P(A)}=\dfrac{P(B_0)P(A\mid B_0)}{P(A)}=\dfrac{0.8}{0.94}=\dfrac{95}{112}=0.85$.

习题 1—4　独立性

1．设 A, B 为两个事件，且 $P(A)=0.8, P(B)=0.6, P(A-B)=0.32$，问 A 与 B 是否相互独立，为什么？

解 因为 $P(A-B)=P(A)-P(AB)\Rightarrow P(AB)=P(A)-P(A-B)=0.8-0.32=0.48=P(A)P(B)$ ，

所以 A 与 B 独立.

2．某举重运动员在一次试举中能举起某一重量的概率为 p，如果他最多只能试举 **3** 次，且前面的试举情况对后面没有影响，求他能举起这个重量的概率。

解 记 A：能举起这个重量， B_k：他第 k 次能举起某一重量（k=1, 2, 3），则 $P(B_k)=p$ （k=1, 2, 3）则有

$$P(A)=P(B_1\cup\bar{B}_1B_2\cup\bar{B}_1\bar{B}_2B_3)=P(B_1)+P(\bar{B}_1B_2)+P(\bar{B}_1\bar{B}_2B_3)=p+p(1-p)+p(1-p)^2$$

$$=p^3-3p^2+3p .$$

3．一实习生用一台机器接连独立地制造 **3** 个同种零件，第 i 个零件是不合格的概率为 $p_i=\dfrac{1}{i+1}\ (i=1,2,3)$，求：（**1**）他制造`的三个零件中前两个为合格品，而第三个不是合格品的概率，（**2**）三个零件中至少有一个是合格品的概率。

解 记 A_k：第 k 零件为合格品（k=1, 2, 3），则 $P(\bar{A}_1)=\dfrac{1}{2},\ P(\bar{A}_2)=\dfrac{1}{3},\ P(\bar{A}_3)=\dfrac{1}{4}$，

（1）所求即为：$P(A_1A_2\bar{A}_3)=P(A_1)P(A_2)P(\bar{A}_3)=(1-\dfrac{1}{2})\times(1-\dfrac{1}{3})\times\dfrac{1}{4}=\dfrac{1}{12}$；

（2）所求即为：$P(A_1\cup A_2\cup A_3)=1-P(\bar{A}_1\bar{A}_2\bar{A}_3)=1-\dfrac{1}{2}\times\dfrac{1}{3}\times\dfrac{1}{4}=\dfrac{23}{24}$.

第二章　随机变量及其分布

习题 2—1　随机变量及其分布函数

1． 已知随机变量 X 的分布函数为

$$F(x)=\begin{cases} a+be^{-\frac{x^2}{2}}, & x>0, \\ 0, & x\le 0. \end{cases}$$

求系数 a,b 的值.

解　由 $\lim\limits_{x\to+\infty}F(x)=1$ 及 $\lim\limits_{x\to 0^+}F(x)=F(0)$（处处右连续）得 $a=1,b=-1$

2． 下列函数中可以作为某个随机变量的分布函数的是（　）

(A) $F(x)=\dfrac{1}{1+x^2}$，$-\infty<x<+\infty$　　(B) $F(x)=\dfrac{1}{2}+\dfrac{1}{\pi}\arctan x$，$-\infty<x<+\infty$

(C) $F(x)=\begin{cases} \dfrac{1}{2}(1-e^{-x}), & x>0, \\ 0, & x\le 0. \end{cases}$　　(D) $f(x)=\int_{-\infty}^{x}f(t)dt$，其中$\int_{-\infty}^{+\infty}f(x)dx=1$

解　因为 $F(+\infty)=\lim\limits_{x\to+\infty}F(x)=\lim\limits_{x\to+\infty}P\{X\le x\}=P\{X<+\infty\}=P\{\Omega\}=1$

否（A）（C），而（D）中未有 $f(x)\ge 0$的条件. 正确选项（B）

习题 2—2　离散型随机变量及其分布

1． 已知袋中编号分别为 **1**，**2**，**3**，**4**，**5** 的五只球，现从中任意抽取三只，以 X 表示取出的三只球中最小编号，求 X 的分布律和分布函数，并画出分布函数的图形.

解　$P(X=1)=\dfrac{C_1^1C_4^2}{C_5^3}=\dfrac{6}{10}$，　$P(X=2)=\dfrac{C_1^1C_3^2}{C_5^3}=\dfrac{3}{10}$，　$P(X=3)=\dfrac{C_1^1C_2^2}{C_5^3}=\dfrac{1}{10}$

则 X 的分布律为

X	1	2	3
p_k	$\frac{6}{10}$	$\frac{3}{10}$	$\frac{1}{10}$

故 X 的分布函数为 $F(x)=P\{X\le x\}=\begin{cases} 0, & x<1, \\ 0.6, & 1\le x<2, \\ 0.9, & 2\le x<3, \\ 1, & x\ge 3. \end{cases}$　图形略.

2．已知实验室有同类设备 **4** 台，每台设备一年里需要维修的概率为 **0.25**，求一年里（**1**）需要维修的设备台数 X 的分布律；（**2**）没有设备需要维修的概率；（**3**）至少有两台设备需要维修的概率**.**

解　（**1**）$X \sim B(4,0.25)$，其分布律为 $P(X=k)=C_4^k\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{4-k}$，$k=0,1,2,3,4$；

（**2**）$P(X=0)=C_4^0\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^4=\frac{81}{256}\approx 0.316$；

（**3**）$P(X\ge 2)=1-P(X<2)=1-P(X=0)-P(X=1)=1-\frac{81}{256}-C_4^1\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^3=\frac{67}{256}\approx 0.262$

3．一批产品共有 **10** 件，其中 **7** 件正品，**3** 件次品，每次随机地抽取一件产品，分别在下列情况下，求直到取出正品为止所需抽取的次数 X 的分布律。（**1**）采取无放回抽样；（**2**）采取有放回抽样**.**

解　（1）无放回抽样时　设 A_k：第次取到正品，$k=1,2,3,4$，则有

$$P(X=1)=P(A_1)=\frac{7}{10};\quad P(X=2)=P(\overline{A_1}A_2)=P(\overline{A_1})P(A_2|\overline{A_1})=\frac{3}{10}\cdot\frac{7}{9}=\frac{7}{30};$$

$$P(X=3)=P(\overline{A_1}\,\overline{A_2}A_3)=P(\overline{A_1})P(\overline{A_2}|\overline{A_1})P(A_3|\overline{A_1}\,\overline{A_2})=\frac{3}{10}\cdot\frac{2}{9}\cdot\frac{7}{8}=\frac{7}{120};$$

$$P(X=4)=P(\overline{A_1}\,\overline{A_2}\,\overline{A_3}A_4)=P(\overline{A_1})P(\overline{A_2}|\overline{A_1})P(\overline{A_3}|\overline{A_1}\,\overline{A_2})P(A_4|\overline{A_1}\,\overline{A_2}\,\overline{A_3})=\frac{3}{10}\cdot\frac{2}{9}\cdot\frac{1}{8}\cdot 1=\frac{1}{120};$$

X	1	2	3	4
p_k	$\frac{7}{10}$	$\frac{7}{30}$	$\frac{7}{120}$	$\frac{1}{120}$

(2)　有放回抽样时　$\{X=k\}$表示前 $k-1$ 次取到的均为次品，而第 k 次取到的才是正品. 故

$$\boldsymbol{P\{X=k\}=P(\overline{A_1}\,\overline{A_2}\cdots\overline{A}_{k-1}A_k)=P(\overline{A_1})P(\overline{A_2})\cdots P(\overline{A}_{k-1})P(A_k)}=\frac{3}{10}\cdot\frac{3}{10}\cdot\cdots\cdot\frac{3}{10}\cdot\frac{7}{10}=\frac{7}{10}\left(\frac{3}{10}\right)^{k-1}\quad \boldsymbol{k}=1,2,\cdots$$

习题 2—3　连续型随机变量及其分布

1.　设随机变量 X 的概率密度为

$$\boldsymbol{f(x)}=\begin{cases}\boldsymbol{cx^3}, & 0\le \boldsymbol{x}\le 1,\\ 0, & \text{其他}.\end{cases}$$

求（**1**）常数 C **.**　（**2**）X 的分布函数 $F(x)$；　**(3)**　$P\{-1\le X\le \frac{1}{2}\}$

解　（**1**）由 $\int_{-\infty}^{+\infty}f(x)dx=1\Rightarrow\int_0^1 cx^3dx=1\Rightarrow\frac{c}{4}=1\Rightarrow c=4$

（2）$F(x)=\begin{cases} 0 & x<0 \\ \int_0^x 4x^3dx, & 0\le x<1 \\ 1 & x\ge 1 \end{cases}=\begin{cases} 0 & x<0 \\ x^4, & 0\le x<1, \\ 1 & x\ge 1 \end{cases}$

（3） $P\{-1\le X\le \frac{1}{2}\}=F(\frac{1}{2})-F(-1^+)=(\frac{1}{2})^4=\frac{1}{16}$，　或

$$P\{-1\le X\le \frac{1}{2}\}=\int_{-1}^{\frac{1}{2}} f(x)dx=\int_0^{\frac{1}{2}} 4x^3dx=\frac{1}{16}$$

2. 设连续型随机变量 X 的分布函数为

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

求（1） X 的概率密度 $f(x)$.　（2） $P\{X<2\}$, $P\{0<X\le 3\}$

解 （1） X 的概率密度 $f(x)=F'(x)=\begin{cases} \frac{1}{x}, & 1<x<e \\ 0 & 其他 \end{cases}$

（2） $P\{X<2\}=F(2)=\ln 2$;　　$P\{0<X\le 3\}=F(3)-F(0)=1$

3. 设某年级学生的数学考试成绩（百分制）服从正态分布 $X\sim N(\mu,\sigma^2)$，平均成绩为 **72** 分.
(1) 若 $\sigma=10$，且规定 **90** 分以上为“优秀”，则“优秀”考生占总学生数的百分之几？
(2) 若 σ 未知，但已知 **96** 分以上的占考生总数的 **2.3%**，试求考生的数学成绩在 **60** 分至 **84** 分之间的概率.

解 （1）设 X 为考生的数学成绩，由题意 $X\sim N(72,10^2)$，所以

$P\{X>90\}=1-\Phi(\frac{9-72}{10})=1-\Phi(1.8)=0.0359=3.6\%$，即“优秀”考生占总学生数的百分之 3.6.

（2）依题意有 $X\sim N(\mu,\sigma^2)$，且 $\mu=72$. 但 σ^2 未知. 故

$P\{X>96\}=1-P\{X\le 96\}=1-\Phi(\frac{96-72}{\sigma})=1-\Phi(\frac{24}{\sigma})=0.023$，$\Phi(\frac{24}{\sigma})=1-0.023=0.977.$

查表得 $\frac{24}{\sigma}\approx 2.0\Rightarrow\sigma=12$. 即 $X\sim N(72,12^2)$. 则

$$P\{60\le X\le 84\}=P\{\left|\frac{X-72}{12}\right|\le 1\}=2\Phi(1)-1=0.6826.$$

4. 设随机变量 X 的概率密度为

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

以 Y 表示对 X 的三次独立重复观察中事件 $\{X\le\frac{1}{2}\}$ 出现的次数，求 $P\{Y=2\}$.

解 由于 $p=P\{X\le\frac{1}{2}\}=\int_0^{\frac{1}{2}}2xdx=\frac{1}{4}$，故 $Y\sim B(3,\frac{1}{4})$. 于是 $P\{Y=2\}=C_3^2(\frac{1}{4})^2(\frac{3}{4})=\frac{9}{64}$.

习题 2—4（随机变量函数的分布）

1. 设离散型随机变量 X 的分布律为：

X	-2	-1	0	1	3
p_k	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{15}$	a

试求：（1）确定常数 a；（2）$Y=X^2+2$ 的分布律。

解 （1）由 $\sum_i p_i=1\Rightarrow a=\frac{11}{30}$；（2）

Y	2	3	6	11
p_k	$\frac{1}{5}$	$\frac{7}{30}$	$\frac{1}{5}$	$\frac{11}{30}$

其中

$$P(Y=3)=P(X=-1)+P(X=1)=\frac{1}{6}+\frac{1}{15}=\frac{7}{30}$$

2. 设随机变量 $X\sim N(0,1)$，求 $Y=e^X$ 的概率密度函数.

解 $F_Y(y)=P\{Y\le y\}=P\{e^X\le y\}=\begin{cases}P\{X\le\ln y\}, & y>0\\ 0, & y\le 0\end{cases}=\begin{cases}\int_{-\infty}^{\ln y}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, & y>0\\ 0, & y\le 0\end{cases}$

所以 $f_Y(y)=F_Y{}'(y)=\begin{cases}\frac{1}{\sqrt{2\pi}y}e^{-\frac{\ln^2 y}{2}}, & y>0\\ 0 & y\le 0\end{cases}$.

第三章　多维随机变量及其分布

习题 3—1　二维随机变量及其分布

1．设一袋中有四个球，它们依次标有数字$1, 2, 2, 3$．从此袋中任取一球后不放回袋中，再从袋中任取一球，以分别X、Y记第一、二次取得的球上标有的数字，求：（**1**）(X,Y)的联合分布律，（**2**）$P\{X+Y\geq 4\}$的值。

解　（1）

Y \ X	1	2	3
1	0	$1/6$	$1/12$
2	$1/6$	$1/6$	$1/6$
3	$1/12$	$1/6$	0

其中

$$P(X=2,\ Y=1)=\frac{2}{4}\cdot\frac{1}{3}=\frac{1}{6}．\ P(X=1,\ Y=3)=\frac{1}{4}\cdot\frac{1}{3}=\frac{1}{12}，\ P(X=1,\ Y=1)=\frac{1}{4}\cdot 0=0．$$

（2）$P\{X+Y\geq 4\}=1-P\{X+Y<4\}=1-P\{X=1,y=1\}+P\{X=1,y=2\}+P\{X=2,y=1\}=\dfrac{2}{3}$

2．设二维随机变量(X,Y)在区域$D=\{(x,y)|0<x<2,\ -1<y<2)\}$上服从均匀分布，试求（**1**）$P\{X\leq Y\}$，　（**2**）$P\{X+Y>1\}$．

解　X的概率密度为：$f(x,y)=\begin{cases}\dfrac{1}{6} & (x,y)\in D\\ 0 & 其它\end{cases}$，　则

（1）$P\{X\leq Y\}=\iint\limits_{x\leq y}f(x,y)dxdy=\int_0^2 dy\int_0^y\frac{1}{6}dx=\frac{1}{3}$；

(2）$P\{X+Y>1\}=\int_0^2 dx\int_{1-x}^2\frac{1}{6}dy=\frac{2}{3}$．

3．设二维随机变量的联合概率密度函数为：

$$f(x,y)=\begin{cases}Ce^{-2(x+y)} & 0<x<+\infty,\ 0<y<+\infty \\ 0 & \text{其它}\end{cases}$$

求：（**1**）常数C的值；

（**2**）$P\{(X,Y)\in D\}$的值，其中$D=\{(x,y)\mid x>0,\ \ y>0,\ \ x+y\le 1\}$；

(3)）随机变量X与Y至少有一个小于 **2** 的概率。

解　（1）因为$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1\Rightarrow C\int_0^{+\infty}dx\int_0^{+\infty}e^{-2(x+y)}dy=1\Rightarrow\dfrac{C}{4}=1$，所以$C=4$；

（2）$P\{(X,Y)\in D\}=\iint\limits_D f(x,y)dxdy=\int_0^1 dx\int_0^{1-x}4e^{-2(x+y)}dy=1-3e^{-2}$.

（3）$P(\{X\le 2\}\bigcup\{Y\le 2\})=1-P\{X\ge 2,\ \ Y\ge 2\}=1-\int_2^{+\infty}dx\int_2^{+\infty}4e^{-2(x+y)}dy=1-e^{-8}$.

习题 3—2　边缘分布

1．一射手进行射击，每次击中目标的概率为0.7，射击进行到击中目标两次为止。设X表示第一次击中目标所进行的射击次数，以Y表示总共射击次数。试求：（**1**）(X,Y)的联合分布律；（**2**）(X,Y)关于X与Y的边缘分布律**.**

解　（1）$\{X=m,\ Y=n\}$的含义：第n次射击时恰好第二次击中目标，前$n-1$射击中仅有一次击中目标，

故
$$P\{X=m,\ Y=n\}=\begin{cases}0.7^2\times 0.3^{n-2} & n=2,3,\cdots,\ m=1,2,\cdots n-1 \\ 0 & \text{其它}\end{cases},$$

（2）$\{X=m\}$的含义：第m次射击时恰好第一次击中目标，前$m-1$射击均未击中目标，故

$$P\{X=m\}=0.7\times 0.3^{m-1},\ \ m=1,2,\cdots$$

$\{Y=n\}$的含义：第n次射击时恰好第二次击中目标，前$n-1$射击中有一次击中目标，$n-2$次射击未中.
故

$$P\{Y=n\}=(n-1)0.7^2\times 0.3^{n-2},\ \ n=2,3,\cdots.$$

2. 设二维连续型随机变量的联合概率密度函数为：

$$f(x,y)=\begin{cases} A(x^2+xy), & 0\le x\le 1, 0\le y\le 2, \\ 0, & \text{其它}. \end{cases}$$

试求：(**1**) 常数 A 的值；

(**2**) (X,Y) 的联合分布函数；

(**3**) (X,Y) 关于 X 与 Y 的边缘概率密度函数和边缘分布函数。

解 (1) 因为 $A\int_0^1 dx\int_0^2 (x^2+xy)dy=1 \Rightarrow A\times\frac{5}{3}=1\Rightarrow A=\frac{3}{5}$；

(2) $$F(x,y)=\begin{cases} 0 & x\le 0 \ \text{或} y \ \le 0 \\ \int_0^x dx\int_0^y \frac{3}{5}(x^2+xy)dy=\frac{3}{5}(\frac{x^3}{3}y+\frac{x^2y^2}{4}) & 0<x<1,\ 0<y<2 \\ \int_0^x dx\int_0^2 \frac{3}{5}(x^2+xy)dy=\frac{2}{5}x^3+\frac{3x^2}{5} & 0<x<1,\ y\ge 2 \\ \int_0^1 dx\int_0^y \frac{3}{5}(x^2+xy)dy=\frac{y}{5}+\frac{3y^2}{20} & x\ge 1,\ 0<y<2 \\ 1 & x\ge 1,\ y\ge 2 \end{cases}$$

(3) $$F_X(x)=\begin{cases} 0 & x<0 \\ \frac{2}{5}x^3+\frac{3x^2}{5} & 0\le x<1 \\ 1 & x\ge 1 \end{cases}, \quad f_X(x)=\begin{cases} \frac{6}{5}x^2+\frac{6x}{5} & 0<x<1 \\ 0 & \text{其它} \end{cases},$$

$$F_Y(y)=\begin{cases} 0 & y\le 0 \\ \frac{y}{5}+\frac{3y^2}{20} & 0<y<2 \\ 1 & y\ge 2 \end{cases} \quad f_Y(y)=\begin{cases} \frac{1}{5}+\frac{3y}{10} & 0<x<2 \\ 0 & \text{其它} \end{cases}.$$

习题 3—3 随机变量的独立性

1.设二维随机变量 (X,Y) 的联合密度函数为

$$f(x,y)=\begin{cases} \frac{1}{x^2y} & 1\le x<+\infty,\ 1\le y\le e \\ 0, & \text{其它} \end{cases},$$

判断 X 与 Y 是否相互独立。

解 因为 $f_X(x)=\int_1^e \frac{1}{x^2y}dy=\begin{cases} \frac{1}{x^2} & 1\le x<+\infty \\ 0 & \text{其他} \end{cases}$， $f_Y(y)=\int_1^{+\infty}\frac{1}{x^2y}dx=\begin{cases} \frac{1}{y}, & 1\le y\le e \\ 0, & \text{其它} \end{cases}$,

所以 $f(x,y)=f_X(x)f_Y(y)$，因而 X与Y 是独立.

3.设随机随机变量Y的密度函数 $f_Y(y)=\begin{cases}e^{-y}, & y>0\\ 0, & y\le 0\end{cases}$，定义随机变量$X_1,X_2$为 $X_k=\begin{cases}2, & Y\le k\\ 3, & Y>k\end{cases}$

$(k=1,\ 2)$， 求X_1和X_2的联合分布，并判断X_1与X_2是否相互独立.

解 （1）

X_2 \ X_1	2	3	$P(X_2=\ )$
2	$1-e^{-1}$	$e^{-1}-e^{-2}$	$1-e^{-2}$
3	0	e^{-2}	e^{-2}
$P(X_1=\ )$	$1-e^{-1}$	e^{-1}	1

其中

$P\{X_1=2,X_2=2\}=P\{Y\le 1\}=\int_0^1 e^{-y}dy=1-e^{-1}$，$P\{X_1=2,X_2=3\}=P(\{Y\le 1\}\cap\{Y>2\})=0$，

$P\{X_1=3,X_2=2\}=P(\{Y>1\}\cap\{Y\le 2\})=\int_1^2 e^{-y}dy=e^{-1}-e^{-2}$，

$P\{X_1=3,X_2=3\}=P(\{Y>2\})=\int_2^{+\infty} e^{-y}dy=e^{-2}$；

（2）因为$P\{X_1=2,X_2=3\}\ne P\{X_1=2\}P\{X_2=3\}$，所以$X_1$与$X_2$是不独立的.

3．设X与Y是相互独立的随机变量，X在$(0,1)$服从均匀分布，Y的概率密度函数为

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

（**1**）求X与Y的联合概率密度;

（**2**）设含有a的二次方程为$a^2+2Xa+Y^2=0$，试求该方程有实根的概率.

解 （1）因为X与Y是相互独立的随机变量，所以其联合概率密度函数为

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}\frac{1}{2}e^{-\frac{y}{2}} & 0<x<1,\quad y>0\\ 0 & \text{其它}\end{cases};$$

（2）设A：该二次方程有实根.

则有$P(A)=P(\Delta\ge 0)=P\{4X^2-4Y^2\ge 0\}=P\{Y^2\le X^2\}=\int_0^1 dx\int_0^x \frac{1}{2}e^{-\frac{y}{2}}dy=\int_0^1(1-e^{-\frac{x}{2}})dx=2e^{-\frac{1}{2}}-1$.

习题 3—4　条件分布

1. 设二维随机变量的联合概率密度函数为

$$f(x,y)=\begin{cases}\frac{1}{2}\sin(x+y), & 0<x<\frac{\pi}{2}, 0<y<\frac{\pi}{2}\\ 0, & \text{其它}\end{cases}$$

试求 $f_{X|Y}(x|y)$ 和 $f_{Y|X}(y|x)$ $(0<x<\frac{\pi}{2}, 0<y<\frac{\pi}{2})$.

解　$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy=\int_0^{\frac{\pi}{2}}\frac{1}{2}\sin(x+y)dy=\begin{cases}\frac{1}{2}(\cos x+\sin x) & 0<x<\frac{\pi}{2}\\ 0 & \text{其它}\end{cases}$

$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx=\int_0^{\frac{\pi}{2}}\frac{1}{2}\sin(x+y)dx=\begin{cases}\frac{1}{2}(\cos y+\sin y) & 0<y<\frac{\pi}{2}\\ 0 & \text{其它}\end{cases}$

则

$$f_{X|Y}(x|y)=\begin{cases}\frac{\sin(x+y)}{\cos y+\sin y} & 0<x<\frac{\pi}{2},\ 0<y<\frac{\pi}{2}\\ 0 & \text{其它}\end{cases},$$

$$f_{Y|X}(y|x)=\begin{cases}\frac{\sin(x+y)}{\cos x+\sin x} & 0<x<\frac{\pi}{2},\ 0<y<\frac{\pi}{2}\\ 0 & \text{其它}\end{cases}$$

2．在 **10** 件产品中有 **2** 件一级品，**7** 件二级品和 **1** 件次品，从 **10** 件产品中无放回抽取 **3** 件，用 X 表示其中的一级品，用 Y 示其中的二级品数，求（**1**）$X=0$ 在的条件下 Y 的条件分布；（**2**）在 $Y=2$ 的条件下 X 的条件分布。

解

Y \ X	0	1	2	$P(Y=\ )$
0	0	0	$1/120$	$1/120$
1	0	$14/120$	$7/120$	$21/120$
2	$21/120$	$42/120$	0	$63/120$
3	$35/120$	0	0	$35/120$
$P(X=\ )$	$56/120$	$56/120$	$8/120$	1

其中 $P\{X=1,\ Y=1\}=\frac{C_2^1C_7^1C_1^1}{C_{10}^3}=\frac{14}{120}$，$P\{X=1,\ Y=2\}=\frac{C_2^1C_7^2}{C_{10}^3}=\frac{42}{120}$.

（1）$X=0$在的条件下Y的条件分布

$$P\{Y=2\mid X=0\}=\frac{21/120}{56/120}=\frac{3}{8}，\text{或}\ P\{Y=2\mid X=0\}=\frac{C_7^2C_1^1}{C_8^3}=\frac{3}{8},$$

$Y=2\mid X=0$ 的含义：已知取出的三件中无一级品，即在剩余的 8 件中取三件，其中有两件二级品和一件次品.

$$P\{Y=3\mid X=0\}=\frac{35/120}{56/120}=\frac{5}{8}，\text{或}\ P\{Y=3\mid X=0\}=\frac{C_7^3}{C_8^3}=\frac{5}{8}；$$

$Y=3\mid X=0$ 的含义：已知取出的三件中无一级品，即在剩余的 8 件中取三件，其中有三件二级品.

（2）$P\{X=0\mid Y=2\}=\dfrac{21/120}{63/120}=\dfrac{1}{3}$，　　$P\{X=1\mid Y=2\}=\dfrac{42/120}{63/120}=\dfrac{2}{3}$.

或 $P\{X=0\mid Y=2\}=\dfrac{C_1^1}{C_3^1}=\dfrac{1}{3}$，　　$P\{X=1\mid Y=2\}=\dfrac{C_2^1}{C_3^1}=\dfrac{2}{3}$.

习题 3—5　二维随机变量函数的分布

1.设随机变量X与Y相互独立，且它们的分布率分别为

X	-1	-2
p	$\frac{1}{4}$	$\frac{3}{4}$

Y	1	2
p	$\frac{2}{5}$	$\frac{3}{5}$

求（**1**）$U=2X+Y$ 的分布律；　　（**2**）$V=X^2+Y^2$的分布律。

解（1）

U	-3	-2	-1	0
p_k	$\frac{3}{10}$	$\frac{9}{20}$	$\frac{1}{10}$	$\frac{3}{20}$

其中

$$P(U=-3)=P(X=-2,\ Y=1)=P(X=-2)\cdot P(Y=1)=\frac{3}{4}\cdot\frac{2}{5}=\frac{3}{10}；$$

（2）

V	2	5	8
p_k	$\frac{1}{10}$	$\frac{9}{20}$	$\frac{9}{20}$

其中

$$P(V=5)=P(X=-1,\ Y=2)+P(X=-2,\ Y=1)=\frac{1}{4}\cdot\frac{3}{5}+\frac{3}{4}\cdot\frac{2}{5}=\frac{9}{20}.$$

2. 设 X,Y 是相互独立的随机变量，它们分别服从参数为 λ_1,λ_2 泊松分布，证明 $Z=X+Y$ 服从参数为 $\lambda_1+\lambda_2$ 的泊松分布。

证　对任意的非负整数 k，有

$$P\{Z=k\}=P\{X+Y=k\}=\sum_{i=0}^{k}P\{X=i,Y=k-i\}=\sum_{i=0}^{k}P\{X=i\}P\{Y=k-i\}$$

$$=\sum_{i=0}^{k}\frac{\lambda_1{}^i}{i!}e^{-\lambda_1}\times\frac{\lambda_2{}^{k-i}}{(k-i)!}e^{-\lambda_2}=\frac{e^{-(\lambda_1+\lambda_2)}}{k!}\sum_{i=0}^{k}\frac{k!}{i!(k-i)!}\lambda_1{}^i\lambda_2{}^{k-i}=\frac{e^{-(\lambda_1+\lambda_2)}}{k!}\sum_{i=0}^{k}C_k^i\lambda_1{}^i\lambda_2{}^{k-i}$$

$$=\frac{e^{-(\lambda_1+\lambda_2)}}{k!}(\lambda_1+\lambda_2)^k=\frac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)},$$

即 $P\{X+Y=k\}=\dfrac{(\lambda_1+\lambda_2)^k}{k!}e^{-(\lambda_1+\lambda_2)}$， $k=0,1,2,\cdots$，所以 $X+Y\sim P(\lambda_1+\lambda_2)$.

3. 设随机变量 (X,Y) 的联合密度函数为

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

试求 $Z=X-Y$ 的概率密度函数.

解　$F_Z(z)=P\{Z\le z\}=P\{X-Y\le z\}=\iint\limits_{x-y\le z}f(x,y)dxdy$,

$$=\begin{cases}0 & z<0\\ 1 & z\ge1\\ 1-\int_z^1dx\int_0^{x-z}3xdy & 0\le z<1\end{cases}=\begin{cases}0 & z<0\\ 1 & z\ge1\\ \dfrac{3z}{2}-\dfrac{z^3}{2} & 0\le z<1\end{cases},$$

所以 $f_Z(z)=F_Z'(z)=\begin{cases}\dfrac{3}{2}(1-z^2) & 0<z<1\\ 0 & \text{其它}\end{cases}$.

4. 设连续型随机变量 (X,Y) 的联合概率密度函数为

$$f(x,y)=\begin{cases}x+y & 0\le x\le1,\ 0\le y\le\\ 0 & \text{其它}\end{cases}$$

求（**1**）$U=\max(X,Y)$ 的分布函数和概率密度函数；

（**2**）$V=\min(X,Y)$ 的分布函数和概率密度函数。

解 （1） $F_U(u)=P\{\max(X,Y)\le u\}=P\{X\le u,Y\le u\}=\iint_D f(x,y)dxdy$，

$$=\begin{cases}0 & u<0\\ \int_0^u dx\int_0^u (x+y)dx & 0\le u<1\\ 1 & u\ge 1\end{cases}=\begin{cases}0 & u<0\\ u^3 & 0\le u<1\\ 1 & u\ge 1\end{cases}，$$ 所以 $f_U(u)=F_U'(u)=\begin{cases}3u^2 & 0<u<1\\ 0 & 其它\end{cases}$；

(2) $F_V(v)=P\{\min(X,Y)\le v\}=1-P(\min(X,Y)>v)=1-P(X>v,Y>v)$

$$=1-\iint_D f(x,y)dxdy=\begin{cases}0 & v\le 0\\ 1-\int_u^1 dx\int_u^1 (x+y)dx, & 0<v<1\\ 1 & v\ge 1\end{cases}=\begin{cases}0 & v\le 0\\ v+v^2-v^3 & 0<v<1\\ 1 & v\ge 1\end{cases}$$

所以 $f_V(v)=F_V{}'(v)=\begin{cases}1+2v-3v^2 & 0<v<1\\ 0 & 其它\end{cases}$

第四章　数字特征

习题 4—1　数学期望

1．将 n 只球随机地放到 m 个盒子中，每个盒子可装任意多个球，每个球以相同的概率落入每个盒子中，求有球的盒子数 X 的数学期望。

解 设 $X_k=\begin{cases}1 & 第k盒子中有球\\ 0 & 第k盒子是空的\end{cases}$，$k=1,2,\cdots,m$，则

X_k	0	1
$P(X_k=\ )$	$(\frac{m-1}{m})^n$	$1-(\frac{m-1}{m})^n$

所以 $E(X_k)=1-(\frac{m-1}{m})^n$.

设 X 表示有球的盒子数，则 $X=\sum_{k=1}^{m}X_k$，由期望的性质得 $E(X)=m(1-(\frac{m-1}{m})^n)=m-m(\frac{m-1}{m})^n$.

2．设 (X,Y) 的密度函数为 $f(x,y)=\begin{cases}4xye^{-(x^2+y^2)}, & x>0,y>0\\ 0, & 其它\end{cases}$，

求 （1）$E(X)$；（2）$E(Z)$，其中 $Z=\sqrt{X^2+Y^2}$.

解 （1）$E(X)=\int_{-\infty}^{+\infty}dx\int_{-\infty}^{+\infty}xf(x,y)dy=4\int_0^{+\infty}dx\int_0^{+\infty}x^2ye^{-(x^2+y^2)}dy=-\int_0^{+\infty}xde^{-x^2}\cdot\int_0^{+\infty}e^{-y^2}dy^2$

$$=\left(xe^{-x^2}\Big|_0^{+\infty}-\int_0^{+\infty}e^{-x^2}dx\right)\left(e^{-y^2}\Big|_0^{+\infty}\right)=\int_0^{+\infty}e^{-x^2}dx=\frac{\sqrt{\pi}}{2}$$

（2）$E(Z)=4\int_0^{+\infty}dx\int_0^{+\infty}xy\sqrt{x^2+y^2}e^{-(x^2+y^2)}dy=4\int_0^{\frac{\pi}{2}}\sin\theta\cos\theta d\theta\int_0^{+\infty}r^4e^{-r^2}dr$ （分部积分）

$$=-r^3e^{-r^2}\Big|_0^{+\infty}+3\int_0^{+\infty}r^2e^{-r^2}dr=\frac{3\sqrt{\pi}}{4}.$$

【注】概率积分 $\int_{-\infty}^{+\infty}e^{-x^2}dx=\sqrt{\pi}$

习题 4—2　方 差

1.设离散型随机变量 X 的分布律为

X	-2	0	2
p	0.4	0.3	0.3

求：（1）$D(X)$；（2）$D(-3X^2-5)$

解（1）

$E(X)=-2\times0.4+0\times0.3+2\times0.3=-0.2$，

$E(X^2)=(-2)^2\times0.4+0^2\times0.3+2^2\times0.3=2.8$，所以 $D(X)=E(X^2)-(EX)^2=2.76$；

（2）$E(X^4)=(-2)^4\times0.4+0^4\times0.3+2^4\times0.3=11.2$，

$D(X^2)=E(X^4)-(E(X^2))^2=11.2-2.8^2=3.36$，所以 $D(-3X^2-5)=9D(X^2)=9\times3.36=30.24$.

2.设连续型随机变量 X 的分布函数为：$F(x)=\begin{cases}0, & x<0\\ kx+b, & 0\le x\le\pi\\ 1, & x>\pi\end{cases}$，求：（1）常数 k,b 的值；（2）$D(X)$.

解 法 1 （1）因为 X 是连续型随机变量，所以它的分布函数应该是连续的，因而有 $F(0^-)=F(0)$，

由此可得 $b=0$；$F(\pi^+)=F(\pi)$，由此可得 $k\pi=1$，即 $k=\frac{1}{\pi}$，所以 $F(x)=\begin{cases}0 & x\le0\\ \frac{x}{\pi} & 0<x<\pi\\ 1 & x\ge\pi\end{cases}$，

即 $X\sim U(0,\pi)$；

（2）因为 $X\sim U(0,\pi)$，所以 $D(X)=\frac{\pi^2}{12}$.

法 2 （1）概率密度函数 $f(x)=F'(x)=\begin{cases}k, & 0<x<\pi\\ 0, & 其他\end{cases}$，

由$\int_{-\infty}^{+\infty} f(x)dx=1$及$\lim\limits_{x\to\pi^+}F(x)=F(\pi)$得$\begin{cases}\int_0^{\pi} kdx=1, \\ k\pi+b=1\end{cases}\Rightarrow\begin{cases}k=\dfrac{1}{\pi}, \\ b=0\end{cases}$，所以$f(x)=\begin{cases}\dfrac{1}{\pi}, & 0<x<\pi \\ 0, & 其他\end{cases}$，

（2）$E(X)=\int_0^{\pi} x\cdot\frac{1}{\pi}dx=\frac{\pi}{2}$，$E(X^2)=\int_0^{\pi} x^2\cdot\frac{1}{\pi}dx=\frac{\pi^2}{3}$，所以$D(X)=E(X^2)-(EX)^2=\frac{\pi^2}{12}$.

习题 4—3 重要分布的期望和方差

1．设随机变量X与Y相互独立，且$X\sim N(2,1)$，$Y\sim N(-2,4)$，$Z=3X-2Y+4$，试求（**1**）$D(Z)$；（**2**）$P\{Z\le 9\}$的值**.**

解 （1）因为$X\sim N(2,1)$，$Y\sim N(-2,4)$，则有$EX=2,\ DX=1;\quad EY=-2,\ DX=4$，

所以 $D(Z)=9D(X)+4D(Y)=9+16=25$；

（2）$EZ=E(3X-2Y+4)=3EX-2EY+4=14$，$D(Z)=25$，由题设可知$Z\sim N(14,\ 5^2)$，

所以 $P\{Z\le 9\}=P\left\{\frac{Z-14}{5}\le\frac{9-14}{5}\right\}=\Phi(-1)=1-\Phi(1)=0.1587$.

习题 4—4 协方差、相关系数与矩

1．设随机变量(X,Y)服从区域$D=\{(x,y)|0<x<1,0<y<x\}$上的均匀分布，试求**: (1)** X与Y的协方差$\mathrm{cov}(X,Y)$； （**2**）相关系数ρ_{XY}．

解 (X,Y)的联合密度为$f(x,y)=\begin{cases}2 & (x,y)\in D \\ 0 & 其他\end{cases}$

（1）$E(X)=\int_0^1 dx\int_0^x 2xdy=\frac{2}{3}$，$E(Y)=\int_0^1 dx\int_0^x 2ydy=\frac{1}{3}$，$E(XY)=\int_0^1 dx\int_0^x 2xydy=\frac{1}{4}$，

所以 $\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{4}-\frac{2}{3}\times\frac{1}{3}=\frac{1}{36}$，

（2）$D(X)=E(X^2)-(EX)^2=\int_0^1 dx\int_0^x 2x^2dy-(\frac{2}{3})^2=\frac{1}{18}$，同理$D(Y)=\int_0^1 dx\int_0^x 2y^2dy-(\frac{1}{3})^2=\frac{1}{18}$

所以 $\rho_{XY}=\frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)D(Y)}}=\frac{1}{2}$.

2. 随机变量X的概率密度函数为：$f(x)=\frac{1}{2}e^{-|x|},\quad -\infty<x<+\infty$,试证明$X$与$|X|$不相关，但不独立**.**

证明 因为$E(X)=\frac{1}{2}\int_{-\infty}^{+\infty} xe^{-|x|}dx=0$，$E(X\cdot|X|)=\frac{1}{2}\int_{-\infty}^{+\infty} x|x|e^{-|x|}dx=0$，

所以 $\operatorname{cov}(X,|X|)=E(X\cdot|X|)-E(X)E(|X|)=0$，因而 X 与 $|X|$ 是不相关的；

又因为 $P\{X\le 1,|X|\le 1\}=P\{|X|\le 1\}\ne P\{X\le 1\}\cdot P\{|X|\le 1\}$，所以 X 与 $|X|$ 不是相互独立的.

（$P\{X\le 1\}=\int_{-\infty}^{1}\frac{1}{2}e^{-|x|}dx=\int_{-\infty}^{0}\frac{1}{2}e^{x}dx+\int_{0}^{1}\frac{1}{2}e^{-x}dx=1-\frac{1}{2}e^{-1|}$，$P\{|X|\le 1\}=2\int_{0}^{1}\frac{1}{2}e^{-x}dx=1-e^{-1}$）

3. 已知 $E(X)=1,\ E(Y)=2,\ E(Z)=-1$，$D(X)=1$，$D(Y)=2$，$D(Z)=3$，$\rho_{XY}=0$，$\rho_{XZ}=\frac{1}{2}$，$\rho_{YZ}=-\frac{1}{2}$，求：（**1**）$D(X+Y+Z)$，（**2**）$E[(X+Y+Z)^2]$**.**

解 （1）因为 $\operatorname{cov}(X,Y)=\rho_{XY}\sqrt{D(X)D(Y)}=0$，$\operatorname{cov}(X,Z)=\rho_{XZ}\sqrt{D(X)D(Z)}=\frac{\sqrt{3}}{2}$，

$\operatorname{cov}(Y,Z)=\rho_{YZ}\sqrt{D(Y)D(Z)}=-\frac{\sqrt{6}}{2}$， 所以

$D(X+Y+Z)=D(X)+D(Y)+D(Z)++2\operatorname{cov}(X,Y)+2\operatorname{cov}(X,Z)+2\operatorname{cov}(Y,Z)=6+\sqrt{3}-\sqrt{6}$；

（2）$E((X+Y+Z)^2)=D(X+Y+Z)+(E(X+Y+Z))^2=10+\sqrt{3}-\sqrt{6}$.

第五章 大数定律与中心极限定理

习题 5—1 中心极限定理

1. 一册 400 页的书中每一页的印刷错误个数服从参数为 $\lambda=0.2$ 的泊松分布，各页有多少个错误是相互独立的，求这册书的错误个数不多于 90 个的概率。

解 设 X_i 表示第 i 页中印刷错误的个数，则 $X_i\sim P(0.2)$，$i=1,2,\cdots,400$，且 $X_1,\ldots,X_{400}$ 相互独立，

记 X 表示这册书中印刷错误的个数，则 $X=\sum_{i=1}^{400}X_i$，由于 $E(X)=400\times 0.2=80,\quad D(X)=80$，

由中心极限定理知 $X\sim N(80,\ 80)$，所求概率为

$$P\{X\le 90\}=P\{\frac{X-80}{\sqrt{80}}\le\frac{90-80}{\sqrt{80}}\}\approx\Phi(\frac{\sqrt{5}}{2})\approx\Phi(1.118)=0.8683.$$

2. 某单位设置一电话总机，共有 200 架电话分机。设每个电话分机是否使用外线通话是相互独立的。设同一时刻每个分机有 5% 的概率要使用外线通话。问总机需要多少外线才能以不低于 90% 的概率保证每个分机要使用外线时可供使用**.**

解 设 X 表示某一时刻要同时使用外线电话的分机个数，

则 $X\sim B(200,\ 0.05)$，由中心极限定理知 $X\sim N(10,\ 9.5)$，若总机有 n 条外线，由题设则有

$$P\{X \le n\} \ge 0.9，而 P\{X \le n\} \approx \Phi(\frac{n-10}{\sqrt{9.5}})，$$

由此可得，$\Phi(\frac{n-10}{\sqrt{9.5}}) \ge 0.9$，因而有 $\frac{n-10}{\sqrt{9.5}} \ge 1.282$，即 $n \ge 10 + 1.282 \times \sqrt{9.5} = 13.95$，

所以应该取 $n = 14$，也即总机需要 14 条外时线才能以不低于 90% 的概率保证每个分机要使用外线时可供使用.

第六章 数理统计的基本概念

习题 6—1 样本与统计量

1．设总体 X 的期望 $EX = \mu$ 已知，方差 $DX = \sigma^2$ 未知，$(X_1, X_2, \cdots, X_n)$ 是总体的一个样本，试判别

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$，$\frac{1}{2}(X_1 + X_2) - \mu$，$\min(X_1, X_2, \cdots, X_n)$，$\frac{(n-1)S^2}{\sigma^2}$，$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ 之中哪些是统计量，哪些不是统计量，为什么？

解 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$，$\frac{1}{2}(X_1 + X_2) - \mu$，$\min(X_1, X_2, \cdots, X_n)$ 都是统计量，而 $\frac{(n-1)S^2}{\sigma^2}$，$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ 则不是统计量，因为含未知参数 σ .

2．设总体 X 的期望 $EX = \mu$，方差 $DX = \sigma^2$ 已知，$X_1, X_2, \cdots, X_n$ 是总体的一个样本，$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，

$S_n^{\ 2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$，求 $E(\bar{X}), D(\bar{X})$ 和 $E(S_n^{\ 2})$.

解 $E(\bar{X}) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X) = \mu$；

$$D(\bar{X}) = D(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2} D(\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} D(X) = \frac{\sigma^2}{n}，$$

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

习题 6—2 抽样分布

1．设总体 $X \sim N(0,1)$，$X_1, X_2, \cdots, X_n$ 是其一个样本，试求下列统计量的分布.

(1) $\dfrac{X_1-X_2}{\sqrt{X_3^2+X_4^2}}$; (2) $\dfrac{\sqrt{n-1}X_1}{\sqrt{\sum\limits_{i=2}^{n}X_i^2}}$; (3) $\dfrac{(n-3)\sum\limits_{i=1}^{3}X_i^2}{3\sum\limits_{i=4}^{n}X_i^2}$

解 （1）根据简单随机样本的性质，可知 $X_1,X_2,\cdots,X_n$ 相互独立且都服从分布 $N(0,1)$，于是有

$$X_1-X_2\sim N(0,2)\Rightarrow\frac{X_1-X_2}{\sqrt{2}}\sim N(0,1),\quad X_3^2+X_4^2\sim\chi^2(2)$$，且 X_1-X_2 与 $X_3^2+X_4^2$ 相互独立，

所以有

$$\frac{(X_1-X_2)\big/\sqrt{2}}{\sqrt{(X_3^2+X_4^2)\big/2}}=\frac{X_1-X_2}{\sqrt{X_3^2+X_4^2}}\sim t(2)$$

（2）根据简单随机样本的性质，可知 $X_1,X_2,\cdots,X_n$ 相互独立且都服从分布 $N(0,1)$，于是有

$X_1\sim N(0,1),\ \sum\limits_{i=2}^{n}X_i^2\sim\chi^2(n-1)$，且 X_1 与 $\sum\limits_{i=2}^{n}X_i^2$ 相互独立，所以有

$$\frac{X_1}{\sqrt{\sum\limits_{i=2}^{n}X_i^2\Big/(n-1)}}=\frac{\sqrt{n-1}X_1}{\sqrt{\sum\limits_{i=2}^{n}X_i^2}}\sim t(n-1)$$

（3）根据简单随机样本的性质，可知 $X_1,X_2,\cdots,X_n$ 相互独立且都服从分布 $N(0,1)$，于是有

$\sum\limits_{i=1}^{3}X_i^2$ 与 $\sum\limits_{i=4}^{n}X_i^2$ 相互独立都服从 x^2 分布，自由度分别为 3 与 $n-3$，因此

$$\frac{\sum\limits_{i=1}^{3}X_i^2\Big/3}{\sum\limits_{i=4}^{n}X_i^2\Big/(n-3)}=\frac{(n-3)\sum\limits_{i=1}^{3}X_i^2}{3\sum\limits_{i=4}^{n}X_i^2}\sim F(3,n-3).$$

2. 设总体 $X\sim N(\mu,\sigma^2)$，$X_1,X_2,\cdots,X_n$ 是其一个样本， **(1)** 求 $P\left\{(\overline{X}-\mu)^2\le\dfrac{\sigma^2}{n}\right\}$**;** **(2)** 当 $n=6$ 时，求 $P\left\{(\overline{X}-\mu)^2\le\dfrac{4S^2}{n}\right\}$； **(3)** 当 n 很大时，求 $P\left\{(\overline{X}-\mu)^2\le\dfrac{4S^2}{n}\right\}$**.**

解 （**1**） 因为 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，所以 $P\left\{(\overline{X}-\mu)^2\le\dfrac{\sigma^2}{n}\right\}=P\left\{-1\le\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\le1\right\}=2\Phi(1)-1=0.6826$；

（**2**）因为 $\dfrac{\overline{X}-\mu}{S/\sqrt{6}}\sim t(5)$，所以

$$P\left\{(\overline{X}-\mu)^2 \le \frac{4S^2}{n}\right\} = P\left\{\left|\frac{\overline{X}-\mu}{S/\sqrt{6}}\right| \le 2\right\} = 1-P\left\{\left|\frac{\overline{X}-\mu}{S/\sqrt{6}}\right| > 2\right\} = 1-2P\left\{\frac{\overline{X}-\mu}{S/\sqrt{6}} > 2\right\} = 1-2\times 0.05 = 0.90$$

（**3**）当 n 很大时，$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$，所以 $P\left\{\left|\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\right| \le 2\right\} = 2\Phi(2)-1 = 2\times 0.9774-1 = 0.9544$.

第七章　参数估计

习题 7—1　点估计

1．设总体 X 的概率密度为 $f(x;\theta)=\dfrac{1}{2}e^{-|x-\theta|}, -\infty<x<+\infty$，$X_1,X_2,\cdots,X_n$ 为其样本，则未知参数 θ 的矩估计量.

解　$EX=\int_{-\infty}^{+\infty} x\cdot\dfrac{1}{2}e^{-|x-\theta|}dx=\int_{-\infty}^{\theta} x\cdot\dfrac{1}{2}e^{(x-\theta)}dx+\int_{\theta}^{+\infty} x\cdot\dfrac{1}{2}e^{-(x-\theta)}dx=\theta$，令 $\theta=\overline{X}$.

得 θ 的矩估计量为 $\theta_M=\overline{X}$.

2．设总体 X 的分布函数为 $F(x;\beta)=\begin{cases}1-\dfrac{1}{x^{\beta}}, & x>1,\\ 0, & x\le 1,\end{cases}$ 其中未知参数 $\beta>1$，$X_1,X_2,\cdots,X_n$ 为其样本，求 β 的矩估计量 $\hat{\beta}_M$ 和极大似然估计量 $\hat{\beta}_L$.

解　（1）X 的概率密度为

$$\boldsymbol{f}(\boldsymbol{x};\beta)=\boldsymbol{F}'(\boldsymbol{x};\beta)=\begin{cases}\dfrac{\beta}{\boldsymbol{x}^{\beta+1}}, & \boldsymbol{x}>1,\\ 0, & \boldsymbol{x}\le 1.\end{cases}$$

$EX=\int_{-\infty}^{+\infty} xf(x;\beta)dx=\int_{1}^{+\infty} x\bullet\dfrac{\beta}{x^{\beta+1}}=\dfrac{\beta}{\beta-1}\overset{\Delta}{=}\overline{X}\Rightarrow \beta=\dfrac{\overline{X}}{\overline{X}-1}$. 因此 β 的矩估计量为 $\beta_M=\dfrac{\overline{X}}{\overline{X}-1}$，

其中 $\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$.

（2）对于总体 X 的样本值 $x_1,x_2,\cdots,x_n$，似然函数为

$$L(\beta)=\begin{cases}\dfrac{\beta^n}{(x_1x_2\cdots x_n)^{\beta+1}}, & x_i>1,\\ 0, & \text{其他},\end{cases}\quad i=1,2,\cdots,n.$$

当 $x_i>1(i=1,2,\cdots,n)$ 时，$L(\beta)>0$，取对数得 $\ln L(\beta)=n\ln\beta-(\beta+1)\sum\limits_{i=1}^{n}\ln x_i$，对 β 求导数，得

$\dfrac{d\ln L(\beta)}{d\beta}=\dfrac{n}{\beta}-\sum\limits_{i=1}^{n}\ln x_i$，令 $\dfrac{d\ln L(\beta)}{d\beta}=0\Rightarrow\beta=\dfrac{n}{\sum\limits_{i=1}^{n}\ln x_i}$，

因此参数β的极大似然估计量为 $\beta_L=\frac{n}{\sum\limits_{i=1}^{n}\ln X_i}$.

3．设总体$X\sim P(\lambda)$，其中λ为未知参数．$(X_1,X_2,\cdots,X_n)$为来自总体X的样本，试求（1）λ的矩估计量λ_M（2）λ的极大似然估计量λ_L．

解 （1）由于只有一个未知参数λ，故只需建立一个方程．由$\overline{X}=EX=\lambda$，解得$\lambda_M=\overline{X}$．

（2）似然函数为

$$L(\lambda)=\prod_{i=1}^{n}(\frac{\lambda^{x_i}}{x_i\,!}e^{-\lambda})=\lambda^{\sum\limits_{i=1}^{n}x_i}(\prod_{i=1}^{n}x_i\,!)^{-1}e^{-n\lambda},$$

所以

$$\ln L(\lambda)=\sum_{i=1}^{n}x_i\ln\lambda-\sum_{i=1}^{n}\ln(x_i\,!)-n\lambda,\quad \frac{d\ln L(\lambda)}{d\lambda}=\sum_{i=1}^{n}x_i\cdot\frac{1}{\lambda}-n,\quad \text{令}\frac{d\ln L(\lambda)}{d\lambda}=0\text{，解得}$$

$\lambda_L=\frac{1}{n}\sum\limits_{i=1}^{n}x_i=\bar{x}$．则极大似然估计量为$\lambda_L=\frac{1}{n}\sum\limits_{i=1}^{n}X_i=\overline{X}$．

习题 7—2　估计量的评价标准

1．设$(X_1,X_2,\cdots,X_n)$为正态总体$N(\mu,\sigma^2)$的一个样本，试确定常数C，使$C\sum\limits_{i=1}^{n}|X_i-\mu|$为$\sigma$无偏估计.

解　由题意 $E\left[C\sum\limits_{i=1}^{n}|X_i-\mu|\right]=\sigma$，而$Y_i=\frac{X_i-\mu}{\sigma}\sim N(0,1)$，故有

$$E\left[C\sigma\sum_{i=1}^{n}\left|\frac{X_i-\mu}{\sigma}\right|\right]=\sigma\Rightarrow C\sigma E\left(\sum_{i=1}^{n}\left|\frac{X_i-\mu}{\sigma}\right|\right)=\sigma\text{，即}Cn\cdot E|Y_i|=1,$$

而

$$E|Y_i|=E|Y|=\int_{-\infty}^{+\infty}|y|\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}ye^{-\frac{y^2}{2}}dy=\sqrt{\frac{2}{\pi}},\quad \text{则有}\quad Cn\cdot\sqrt{\frac{2}{\pi}}=1\Rightarrow C=\sqrt{\frac{\pi}{2}}\cdot\frac{1}{n}.$$

2．设X_1,X_2为取自总体X的一个样本，$EX=\mu,DX=\sigma^2$均存在，C_1，C_2为常数，且$C_1+C_2=1$.证明：

（1） $\hat{X}=C_1X_1+C_2X_2$为$EX=\mu$的无偏估计；**（2）** 当$C_1=C_2=\frac{1}{2}$时，其方差$D(\hat{X})$最小.

解　(1) $E\hat{X}=E(C_1X_1+C_2X_2)=C_1EX_1+C_2EX_2=C_1EX+C_2EX=(C_1+C_2)EX=(C_1+C_2)\mu=\mu$，

所以当$C_1+C_2=1$时，$\hat{X}=C_1X_1+C_2X_2$为$EX=\mu$的无偏估计；

(2) $D(\hat{X})=D(C_1X_1+C_2X_2)=C_1^2DX_1+C_2^2DX_2=C_1^2DX+C_2^2DX=(C_1^2+C_2^2)DX$

$$=\left[C_1^2+(1-C_1^2)\right]\sigma^2=(2C_1^2-2C_1+1)\sigma^2=\left[2\left(C_1-\frac{1}{2}\right)^2+\frac{1}{2}\right]\sigma^2$$

当$C_1+C_2=1$且当$C_1=C_2=\dfrac{1}{2}$时，其方差$D(\hat{X})$最小. $D(\hat{X})=\dfrac{1}{2}\sigma^2$

习题 7—3　区间估计

1．某公司希望估计其职工实际探亲的平均天数μ，为此，通过抽取n个职工调查，并且希望其估计误差不超过 **2** 天，且置信度不低于 **0.9**，假定职工实际探亲天数$X\sim N(\mu,15^2)$，问至少应调查多少职工？

解　因为$X\sim N(\mu,15^2)$，所以$\overline{X}\sim N(\mu,\dfrac{15^2}{n})\Rightarrow\dfrac{\overline{X}-\mu}{15/\sqrt{n}}\sim N(0,1)$，

由题意知

$$P\left\{(\left|\overline{X}-\mu\right|\le 2)\right\}=P\left\{(\left|\frac{\overline{X}-\mu}{15/\sqrt{n}}\right|\le\frac{2}{15/\sqrt{n}})\right\}\ge 0.9,$$

即　$2\Phi(\dfrac{2}{15/\sqrt{n}})-1\ge 0.9\Rightarrow\dfrac{2}{15/\sqrt{n}}\ge 1.65\Rightarrow n=154$.　即至少应调查 154 名职工才能达到要求.

2．为比较两种品牌的小卡车的燃料经济效益，做一项实验. 取 **13** 辆 **A** 品牌车和 **10** 辆 **B** 品牌车，以 **90km/h** 的不变速度来使用，测得结果为：$\bar{x}_A=16$ **km/L**, $s_A=1.0$ **km/L**, $\bar{x}_B=11$ **km/L**, $s_B=0.8$ **km/L**, 假设每辆卡车每升油所能行驶的距离近似服从正态分布；设$X_A\sim N(\mu_1,\sigma_1{}^2)$，$X_B\sim N(\mu_2,\sigma_2{}^2)$.

(1)求标准差比$\dfrac{\sigma_1}{\sigma_2}$的置信度为**98%**的置信区间；

解　$n_1=13,\ n_2=10$，$s_1^2=1.0^2$，$s_2^2=0.8^2$，　$\dfrac{s_1^2}{s_2^2}=\dfrac{1.0}{0.64}=1.5625$，$1-\alpha=0.98$，　$\dfrac{\alpha}{2}=0.01$，

$$F_{0.01}(12,9)=5.11,\quad F_{0.99}(12,9)=\frac{1}{F_{0.01}(9,12)}=\frac{1}{4.39},$$

所以 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信系数为的98%置信区间为

$$(\frac{s_1^2}{s_2^2}\cdot\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,n_2-1)},\frac{s_1^2}{s_2^2}\cdot\frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)})=(\frac{1.5625}{5.11},1.5625\times4.39)=(0.3058,6.859).$$

则标准差比$\frac{\sigma_1}{\sigma_2}$的置信度为 98%的置信区间为

$$\left(\sqrt{0.3058},\sqrt{6.859}\right)=\left(0.553,2.620\right).$$

(2)若$\sigma_1^{\ 2}=\sigma_2^{\ 2}$，求期望值差$\mu_1-\mu_2$的置信度为**98%**的置信区间**.**

解 $n_1=13,\ n_2=10,\ n_1+n_2-2=21$，$1-\alpha=0.98,\ \frac{\alpha}{2}=0.01$，$t_{0.01}(21)=2.5177$，

又 $\bar{x}_A-\bar{x}_B=16-11=5$， $s_\omega=\sqrt{\frac{1}{21}(12\times1.0^2+9\times0.8^2)}=\sqrt{0.8457}=0.92$.

由实际抽样的随机性可知两样本相互独立，且两总体的方差相等，故$\mu_A-\mu_B$的置信系数为98%的置信区间为

$$\left((\bar{X}-\bar{Y})-t_{\frac{\alpha}{2}}(n_1+n_2-2)S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\quad (\bar{X}-\bar{Y})+t_{\frac{\alpha}{2}}(n_1+n_2-2)S_\omega\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

$$=(5-2.5177\times0.92\times\sqrt{\frac{1}{13}+\frac{1}{10}},\ 5+2.5177\times0.92\times\sqrt{\frac{1}{13}+\frac{1}{10}})=(4.026,5.974).$$

第八章　假设检验

习题 8—1　单个正态总体的假设检验

1.某大学大一女生平均身高为**162.5cm**，标准差为**6.9cm**，在α=0.02的显著水平下，若从现在的班上随机选出**50**名女生，其平均身高为**165.2cm**，试问是否有理由相信平均身高改变了？

解 设H_0: $\mu=\mathbf{162.5}$，H_1: $\mu\neq\mathbf{162.5}$，这是双边假设检验问题，

H_0为真时，σ^2已知，选取统计量$U=\frac{\bar{X}-162.5}{\sigma/\sqrt{n}}\sim N(0,1)$，

在$\alpha=0.05$下，拒绝域$I_c=\{U\,|\,|U|\geq U_{0.025}=1.96\}$. 将$\bar{x}=165.2$，$s=6.9$，$n=50$代入得

$U=\frac{165.2-162.5}{6.9/\sqrt{50}}=2.6\in I_c$，有理由相信平均身高改变了.

2.设某次考试的学生成绩服从正态分布，从中随机地抽取**36**位考生的成绩，算得平均成绩为**66.5**分，标准差为**15**分，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？并给出检验过程**.**

（附：$t_{0.025}(35)=2.0301$，$t_{0.05}(35)=1.6896$，$t_{0.025}(36)=2.0281$，$t_{0.05}(36)=1.6883$）

解 设H_0：$\mu=\mu_0=70$，H_1：$\mu\neq\mu_0=70$，这是双边假设检验问题，H_0为真时，σ^2未知，选取统计量

$T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$，$\alpha=0.05$下，拒绝域$I_c=\{t\mid |t|\geq t_{0.025}(35)=2.0301\}$. 将$\bar{x}=66.5$，$s=15$，$n=36$

代入得 $|t|=\left|\dfrac{66.5-70}{15/6}\right|=1.4<2.0301$，故接受$H_0$，可以认为这次考试全体考生的平均成绩为 70 分.

3.某品牌香烟的尼古丁含量服从正态分布，其标准差为**1.3mg**，若随机抽取此牌香烟**8**支，其标准差为**s=1.8**，在α=0.05显著性水平下，检验假设$H_0:\sigma=1.3$，$H_1:\sigma\neq1.3$

解 由题设知要检验的是假设$\sigma=1.3$？

H_0：$\sigma=1.3$，H_1：$\sigma\neq1.3$，取$\chi^2=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$为检验统计量，$\alpha=0.05$下，

H_0拒绝域为$I_c=\left\{\chi^2\geq\chi^2_{0.025}(7)\text{或}\chi^2\leq\chi^2_{0.975}(7)\right\}=\left\{\chi^2\geq16.013\text{或}\chi^2\leq1.690\right\}$，

将$\sigma=1.3,\ n=8,\ s=1.8$代入得$\chi^2=\dfrac{7\times1.8^2}{1.3^2}=13.420\notin I_c$，所以要接受$H_0$，即认为$\sigma=1.3$.