概率论与数理统计
习题册与综合训练

习题解答

苏州科技大学 数理学院 概率统计课程组

习题 1-1　样本空间与随机事件

1．选择题

（1）设 A,B,C 为三个事件，则"A,B,C 中至少有两个发生"这一事件可表示为（ A ）

（A）$AB\bigcup AC\bigcup BC$　（B）$A\bigcup B\bigcup C$　（C）$AB\overline{C}\bigcup A\overline{B}C\bigcup\overline{A}BC$　（D）$\overline{A}\bigcup\overline{B}\bigcup\overline{C}$

（2）$A,\ B,\ C$ 是任意事件，在下列各式中，不成立的是（ B ）

（A）$(A-B)\cup B=A\cup B$　　（B）$(A\cup B)-A=B$

（C）$(A\cup B)-AB=A\overline{B}\cup\overline{A}B$　　（D）$(A\cup B)\overline{C}=(A-C)\cup(B-C)$

（3）设三个元件的寿命分别为 T_1,T_2,T_3，并联成一个系统，则只要有一个元件正常工作则系统能正常工作，事件"系统的寿命超过 t"可表示为（ D ）

(A) $\{T_1+T_2+T_3>t\}$　(B) $\{T_1T_2T_3>t\}$　(C) $\{\min\{T_1,T_2,T_3\}>t\}$　(D) $\{\max\{T_1,T_2,T_3\}>t\}$

2．用集合的形式表示下列随机试验的样本空间 Ω 与随机事件 A：对目标进行射击，击中后便停止射击，观察射击的次数；事件 A 表示"射击次数不超过 5 次"。

解：$\Omega=\{1,2,3,\cdots\}$；$A=\{1,2,3,4,5\}$。

3. 设 A,B,C 是随机试验 E 的三个事件，用 A,B,C 运算关系表示下列事件：

（1）仅 A 发生 $\underline{A\overline{B}\overline{C}}$；

（2）A,B,C 中至少有两个不发生 $\underline{\overline{A}\overline{B}\bigcup\overline{A}\overline{C}\bigcup\overline{B}\overline{C}}$；

（3）A,B,C 中不多于两个发生 $\underline{\overline{A}\bigcup\overline{B}\bigcup\overline{C}}$ 或 $\underline{\overline{ABC}}$；

（4）A,B,C 中恰有两个发生 $\underline{AB\overline{C}\bigcup A\overline{B}C\bigcup\overline{A}BC}$；

（5）A,B,C 中至多有一个发生 $\underline{\overline{A}\overline{B}\bigcup\overline{A}\overline{C}\bigcup\overline{B}\overline{C}}$ 或 $\underline{\overline{A}\overline{B}\overline{C}\bigcup A\overline{B}\overline{C}\bigcup\overline{A}B\overline{C}\bigcup\overline{A}\overline{B}C}$。

4．设某工人连续生产了 4 个零件，A_i 表示他生产的第 i 个零件是正品（$i=1,2,3,4$），试用 A_i 表示下列各事件：

（1）只有一个是次品；$\underline{\overline{A_1}A_2A_3A_4\cup A_1\overline{A_2}A_3A_4\cup A_1A_2\overline{A_3}A_4\cup A_1A_2A_3\overline{A_4}}$

（2）至多有三个不是次品；$\underline{\overline{A_1}\cup\overline{A_2}\cup\overline{A_3}\cup\overline{A_4}}$。

习题 1-2　随机事件的概率及计算

1．填空题

（1）已知 $A \subset B$，$P(A)=0.4$，$P(B)=0.6$，则　$P(\overline{A})=$ __0.6__，$P(AB)=$ __0.4__，

$P(A\overline{B})=$ __0__，$P(\overline{A}\,\overline{B})=$ __0.4__。

（2）设事件 A,B 仅有一个发生的概率为 0.3，且 $P(A)+P(B)=0.5$，则 A,B 至少有一个不发生的概率为__0.9__。

（3）盒子中有 10 个球，其中 3 个红球，接连不放回抽取五次，第一次抽到红球的概率__0.3__，第三次抽到红球的概率__0.3__。

2．选择题

（1）如果 A 与 B 互不相容，则（C ）

(A) $\overline{A}\overline{B}=\varnothing$　　(B) $\overline{A}=B$　　(C) $\overline{A}\cup\overline{B}=\Omega$　　(D) $A\cup B=\Omega$

（2）如果 $P(AB)=0$，则（ C ）

(A) A 与 B 互不相容　　(B) $\overline{A}$ 与 $\overline{B}$ 互不相容

(C) $P(A-B)=P(A)$　　(D) $P(A-B)=P(A)-P(B)$

(3) 两个事件 A 与 B 是对立事件的充要条件是（ C ）

（A） $P(AB)=P(A)P(B)$　　（B）$P(AB)=0$且$P(A\cup B)=1$

（C） $AB=\varnothing$且$A\cup B=\Omega$　　（D）$AB=\varnothing$

3．一批晶体管共 40 只，其中 3 只是坏的，今从中（不放回地）任取 5 只，求

（1）5 只全是好的的概率；

（2）5 只中有两只坏的的概率；

（3）5 只中至多有一只坏的的概率。

解：设 A_i 表示 5 只中有 i 只坏的晶体管 $i=0,1,\cdots,5$，则

（1）$P(A_0)=\dfrac{C_{37}^5}{C_{40}^5}$=0.6624　，（2）$P(A_2)=\dfrac{C_{37}^3C_3^2}{C_{40}^5}$=0.0354

（3）$P(A_0\cup A_1)=\dfrac{C_{37}^4C_3^1+C_{37}^5}{C_{40}^5}$=0.963 （不放回抽取与一次抽取 5 只结论一样）

4. （1）从 1，2，…，10 共十个数字中有放回地抽取，每次取一个，先后取出 5 个数字 ，求所得 5 个数字全不相同的概率；

（2）房间里有四个人，求至少两个人的生日在同一个月的概率（假设生日在每个月的概率相同）。

解：（1）设 A=“所得 5 个数字全不相同”，则

$$P(A)=\frac{P_{10}^{5}}{10^{5}}=\frac{10*9*8*7*6}{10^{5}}=\frac{30240}{100000}=0.3024$$

(2) 设 $B=$“至少有两个人的生日在同一个月”，则

$$P(B)=\frac{C_4^2C_{12}^1P_{11}^2+C_4^2C_{12}^2+C_4^3P_{12}^2+C_{12}^1}{12^4}=\frac{41}{96};$$

或 $$P(B)=1-P(\bar{B})=1-\frac{P_{12}^4}{12^4}=\frac{41}{96}。$$

5. 设 A、B、C 为三个随机事件，且 $P(A)=P(B)=P(C)=\frac{1}{4}$，$P(AB)=P(BC)=0$，$P(AC)=\frac{1}{6}$，求 A、B、C 至少有一个发生的概率。

解：因为 $ABC\subset AB$，所以 $P(ABC)\le P(AB)=0$

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

$$=\frac{1}{4}\times 3-\frac{1}{6}=\frac{7}{12}$$

习题 1-3　条件概率

1．选择题：

（1）设 A，B 为两个相互对立事件，且 $P(A)>0$，$P(B)>0$，则(C)。

（A）$P(B|A)>0$　（B）$P(A|B)=P(A)$　（C）$P(A|B)=0$　（D）$P(AB)=P(A)P(B)$

（2）一种零件的加工由两道独立工序组成，第一道工序的废品率为 p，第二道工序的废品率为 q，则该零件加工的成品率为（ C ）

（A）$1-p-q$　（B）$1-pq$　（C）$1-p-q+pq$　（D）$(1-p)+(1-q)$

（3）设 $P(A)=0.8$，$P(B)=0.7$，$P(A|B)=0.8$，则下列结论正确的是（ C ）。

（A）$B\supset A$；　（B）$P(A\cup B)=P(A)+P(B)$；

（C）事件 A 与事件 B 相互独立；　（D）事件 A 与事件 B 对立。

（4）设 $0<P(A)<1$，$0<P(B)<1$，$P(A|B)+P(\overline{A}|\overline{B})=1$，则（ D ）。

（A）事件 A 与 B 互不相容；　（B）事件 A 与 B 对立；

（C）事件 A 与 B 不相互独立；　（D）事件 A 与 B 相互独立。

2．填空题:

(1) 已知 $P(A)=0.5$，$P(A\bigcup B)=0.6$，若 A与B 互不相容，则 $P(B)=$ <u>0.1</u>；若 A与B 相互独立，则 $P(B)=$ <u>0.2</u>

(2) 一射手对同一目标独立地进行四次射击，若至少命中一次的概率为 80/81，则该射手的命中率 $p=\frac{2}{3}$。

(3) 已知 $P(A)=0.5$，$P(B)=0.6$，$P(B|A)=0.8$，$P(\overline{A}\,\overline{B})=$ <u>0.3</u>。

(4) 一批产品共有 10 个正品和 2 个次品，任意抽取两次，每次抽一个，抽出后不再放回，则第二次抽出的是次品的概率为 $1/6$。(1993 年考研题)

3．为防止意外，在矿内同时安装了两种报警系统 A 与 B，每种报警系统都使用时，对系统 A 其有效的概率是 0.92，对系统 B 其有效的概率为 0.93，在 A 失效的条件下，B 有效的概率为 0.85.求：（1）发生意外时，这两种报警系统至少有一个有效的概率；（2）B 失灵的条件下，A 有效的概率。

解：设 $A=$“报警系统 A 有效”，$B=$“报警系统 B 有效”

则　（1）$P(A\bigcup B)=1-P(\overline{A}\overline{B})=1-P(\overline{A})P(\overline{B}|\overline{A})=1-0.08\times0.15=0.988$

（2）因为：$P(AB)=P(A)+P(B)-P(A\bigcup B)=0.92+0.93-0.988=0.862$

$$P(A|\overline{B})=\frac{P(A\overline{B})}{P(\overline{B})}=\frac{P(A)-P(AB)}{1-P(B)}=\frac{0.058}{0.07}=0.829$$

。

4. 市场上有甲、乙、丙三家厂家生产的同一品牌的产品，已知三家工厂的市场占有率分别为：$\frac{1}{4}$，$\frac{1}{4}$，$\frac{1}{2}$，三家工厂的次品率分别为：0.002，0.01，0.005。求：

（1）市场上该品牌产品的次品率；

（2）若某人购买的该品牌产品为次品，问是甲厂生产的概率为多大。

解：设 A_1 =“甲厂生产”，A_2 =“乙厂生产”，A_3 =“丙厂生产”，B =“产品为次品”. 由题意

$P(A_1)=\frac{1}{4}$，$P(A_2)=\frac{1}{4}$，$P(A_3)=\frac{1}{2}$，$P(B|A_1)=0.002$，$P(B|A_2)=0.01$，$P(B|A_3)=0.005$.

则 （1） $P(B)=\sum_{i=1}^{3}P(A_i)P(B|A_i)$

$=0.25\times0.002+0.25\times0.01+0.5\times0.005=0.0055$

（2） $P(A_1|B)=\frac{P(A_1B)}{P(B)}=\frac{P(A_1)P(B|A_1)}{P(B)}=\frac{0.25\times0.002}{0.0055}=\frac{1}{11}$ =0.0909

5．已知一批产品中 96%是合格品，检查产品时，一个合格品被误认为是次品的概率是 0.02，一个次品被误认为是合格品的概率是 0.05，求在检查后认为是合格品的产品的确是合格品的概率。

解：设 A =“任取一产品，经检查是合格品”， B =“任取一产品确是合格品”，

则 $A=BA\cup\overline{B}A$

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

$$=0.96\times0.98+0.04\times0.05=0.9428,$$

所求概率为 $P(B|A)=\frac{P(B)P(A|B)}{P(A)}=\frac{0.96\times0.98}{0.9428}=0.998$.

6. 将编码分别为 A 和 B 的信息传递出去，当接收站收到时，A 被误收作 B 的概率为 0.02，B 被误收作 A 的概率为 0.01，信息 A 与信息 B 传递的频繁程度为 3 : 2。若接收站收到的信息是 A，问原发信息是 A 的概率是多少？

解：设 A =“发出的信息为 A”， B =“发出的信息为 B”， C =“收到信息为 A”

由题意，$P(A)=\frac{3}{5}$，$P(B)=\frac{2}{5}$，$P(C|A)=0.98$，$P(C|B)=0.01$，由贝叶斯公式，

$$P(A|C)=\frac{P(AC)}{P(C)}=\frac{P(A)P(C|A)}{P(A)P(C|A)+P(B)P(C|B)}=\frac{0.6\times0.98}{0.6\times0.98+0.4\times0.01}=0.9932$$

7．玻璃杯成箱出售，每箱 20 只，假设各箱含 0，1，2 只残次品的概率分别为 0.8，0.1，0.1，一顾客欲购一箱玻璃杯，售货员随意取一箱，顾客开箱随意地察看四只，若无残次品，则买下该箱，否则退回. 试求：

（1）顾客买下该箱的概率 α ；

（2）在顾客买下的一箱中，确无残次品的概率 β .

解　设 $A=$“顾客买下该箱”， $B_i=$“箱中恰有 i 件残次品”， $i=0,1,2$ ，

（1） $\alpha=P(A)=P(B_0)P(A|B_0)+P(B_1)P(A|B_1)+P(B_2)P(A|B_2)$

$$=0.8+0.1\times\frac{C_{19}^4}{C_{20}^4}+0.1\times\frac{C_{18}^4}{C_{20}^4}\approx 0.94$$ ；

（2） $\beta=P(B_0|A)=\dfrac{P(AB_0)}{P(A)}=\dfrac{0.8}{0.94}\approx 0.85$ 。

习题 2-1 随机变量及其分布函数

1．判断题：试说明下列函数是否可作为某随机变量的分布函数.

（1）$F_1(x)=\begin{cases}0, & x<0,\\ \sin x, & 0\le x<\dfrac{\pi}{2},\\ 1, & x\ge\dfrac{\pi}{2}.\end{cases}$ （ 是 ）；（2）$F_2(x)=\begin{cases}0, & x<0,\\ \dfrac{\ln(1+x)}{1+x}, & x\ge 0.\end{cases}$ （ 否 ）。

2. 选择题：设 X 的分布函数为 $F(x)$（x 为任意实数），则下面选项中不一定正确的是（ B ）。

（A） $F(x)=P(X\le x)$　　（B）$F(-x)=1-F(x)$

（C） $P(a<X\le b)=F(b)-F(a)$　　（D）$P(X>x)=1-F(x)$

习题 2-2 离散型随机变量

1． 填空题

(1) 设随机变量 X 的分布律为:$P\{X=k\}=\dfrac{a}{N}$，$k=1,2,\cdots$，N，试确定 $a=$ __1__。

(2) 一批产品共 100 个，其中有 10 个次品，从中有放回取 5 次，每次任取一个，以 X 表示任意取出的产品中的次品数，则 X 的分布为 $\underline{B(5,0.1)}$。

(3) 某射手对一目标进行射击，直至击中为止，如果每次射击命中率都是 p，以 X

表示射击的次数，则 X 的分布列为 $\underline{P(X=k)=(1-p)^{k-1}p,\ \text{k}=1,2,\cdots}$。

2. 设盒中的 3 个零件里有一个次品 ，从中不放回依次取出进行检查，直到查到次品时为止，用 X 表示检查次数，试求：（1）X 的分布列；（2）求 X 的分布函数 $F(x)$。

解：设 X 表示检查次数，

（1）则 X 的分布列为：

X	1	2	3
p_k	1/3	$\dfrac{2}{3}\times\dfrac{1}{2}$	$\dfrac{2}{3}\times\dfrac{1}{2}\times 1$

（2）X 的分布函数 $F(x)$ 为：

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

3. 设某城市在一周内发生交通事故的次数服从参数为 0.3 的泊松分布,试问

(1) 在一周内恰好发生 2 次交通事故的概率是多少?

(2) 在一周内至少发生 1 次交通事故的概率是多少?

解：设一周内发生交通事故的次数为 X，则 $X\sim P(0.3)$，

（1） $P(X=2)=\frac{0.3^2}{2!}e^{-0.3}\approx 0.0333$ 。

（2） $P(X\ge 1)=1-P(X=0)=1-\frac{0.3^0}{0!}e^{-0.3}=1-e^{-0.3}=0.259$ 。

4．某人购买某种彩票，若已知中奖的概率为 0.001，现购买 2000 张彩票，试求：（1） 此人中奖的概率；（2）至少有 3 张彩票中奖的概率。

解：设中奖的彩票数为 X，则 $X\sim B(2000,0.001)$.

（1）$P(X\ge 1)=1-P(X=0)=1-(0.999)^{2000}\approx 0.8648$.

（2）$P(X\ge 3)=1-P(X=0)-P(X=1)-P(X=2)$

$=1-(C_{2000}^{0}0.001^{0}\times 0.999^{2000}+C_{2000}^{1}0.001^{1}\times 0.999^{1999}+C_{2000}^{2}0.001^{2}\times 0.999^{1998})=0.3233$.

习题 2-3 连续型随机变量

1. 设连续型随机变量 X 的密度函数为

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

试求：（1）常数 a 的值；（2）随机变量 X 的分布函数；（3）$P(\frac{1}{2}<X<\frac{3}{2})$。

解：（1）由于 $1=\int_{-\infty}^{+\infty}f(x)dx=\int_0^1 ax^2dx+\int_1^2(2-x)dx=\frac{a}{3}+\frac{1}{2}$. 故 $a=\frac{3}{2}$.

（2）当 $x<0$ 时，$F(x)=0$；

当 $0\le x\le 1$ 时，$F(x)=\int_0^x\frac{3}{2}t^2dt=\frac{1}{2}x^3$；

当 $1<x\le 2$ 时，$F(x)=\int_0^1\frac{3}{2}t^2dt+\int_1^x(2-t)dt=2x-\frac{1}{2}x^2-1$；

当 $x>2$ 时，$F(x)=1$.

故，

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

（3） $P(\frac{1}{2}<X<\frac{3}{2})=\int_{1/2}^1\frac{3}{2}x^2dx+\int_1^{3/2}(2-x)dx=\frac{13}{16}$.

2. 设连续型随机变量 X 的分布函数为 $F(x)=\begin{cases} A(1-e^{-x}), & x\ge 0 \\ 0, & x<0 \end{cases}$,

试求：（1）系数 A；（2）X 的密度函数；（3）$P(1<X<3)$。

解 （1）因 $F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$，所以 $\lim\limits_{x\to+\infty}A(1-e^{-x})=A=1$

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

（3）$P(1<X<3)=F(3)-F(1)=(1-e^{-3})-(1-e^{-1})=e^{-1}-e^{-3}$。

3. 设随机变量 X 的密度函数为 $f(x)=\begin{cases} Ax+1 & , \quad 0\le x\le 2 \\ 0 & , \quad x<0或x>2 \end{cases}$，

求：（1）A 的值；（2）$P(1.5<X<2.5)$；（3）X 的分布函数 F（x）。

解：(1) $1=\int_0^2 (Ax+1)dx=2A+2, 得A=-\frac{1}{2}$

(2) $P(1.5<X<2.5)=\int_{1.5}^{2}(-\frac{1}{2}x+1)dx=0.0625$

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

4. 设连续型随机变量 X 的分布函数为 $F(x)=\begin{cases} c, & x<0, \\ a+be^{-\frac{x^3}{2}}, & x\ge 0. \end{cases}$

试求：（1）常数 a,b,c 的值；（2）随机变量 X 的密度函数 $f(x)$；（3）$P(\sqrt[3]{2}<X<2)$。

解：(1) 由 $F(-\infty)=0$ 知 $c=0$；由 $F(+\infty)=1$ 知 $a=1$；由 $F(x)$ 在 $x=0$ 点连续知

$$F(0)=\lim_{x\to 0^{-0}}F(x),$$

即 $a+b=0$，故 $b=-1$。即：$F(x)=\begin{cases} 0, & x<0, \\ 1-e^{-\frac{x^3}{2}}, & x\ge 0. \end{cases}$

(2) 在 $F(x)$ 导数存在的处有 $f(x)=F'(x)$，所以

$$f(x)=\begin{cases} 0, & x\le 0, \\ \frac{3}{2}x^2e^{-\frac{x^2}{2}}, & x>0. \end{cases}$$

（3）$P(\sqrt[3]{2}<X<2)=F(2)-F(\sqrt[3]{2})=\left(1-e^{-4}\right)-\left(1-e^{-1}\right)=e^{-1}-e^{-4}$。

5. 设 K 在区间(0，5)内服从均匀分布, 求方程 $4x^2+4Kx+K+2=0$ 有实根的概率。

解：所求的概率为：

$$P(16K^2-16\left(K+2\right)\ge 0)=P\left(K\ge 2或K\le -1\right)$$

$$=P\left(K\ge 2\right)+P\left(K\le -1\right)=\int_2^5\frac{1}{5}dx+0=\frac{3}{5}。$$

6. .设随机变量$X \sim N(3,2^2)$，求

（1）$P(2 < X \le 5)$，$P(-4 < X \le 10)$，$P(|X| > 2)$；

（2）确定c，使得$P(X > c) = P(X \le c)$；

（3）若d满足$P(X > d) \ge 0.9$，问d至多为多少？

解：（1）$P(2 < X < 5) = \Phi(1) - \Phi(-0.5) = 0.8413 + 0.6915 - 1 = 0.5328$，

$$\begin{aligned}&P(-4 < X \le 10) = \Phi(3.5) - \Phi(-3.5) = 2\Phi(3.5) - 1\\&= 2 \times 0.9997674 - 1 = 0.99953\end{aligned},$$

$$\begin{aligned}&P(|X| > 2) = 1 - P(|X| \le 2) = 1 - [\Phi(-0.5) - \Phi(-2.5)] = 1 - [\Phi(2.5) - \Phi(0.5)]\\&= 1 - (0.9938 - 0.6915) = 1 - 0.3023 = 0.6977\end{aligned}$$

（2）求C使$P(X \le C) = P(X > C)$，则$C = 3$ 。

（3）若d满足$P(X > d) \ge 0.9$，则$P(X \le d) \le 0.1$，即$\Phi(\frac{d-3}{2}) \le 0.1$

即 $\Phi(-\frac{d-3}{2}) \ge 0.9$，查表有 $-\frac{d-3}{2} = 1.29$，解出 d 至多有$d = 0.42$。

7.某种型号的电子管寿命X (以小时计)具有以下概率密度

$$f(x) = \begin{cases} \dfrac{1000}{x^2}, & x > 1000 \\ 0, & 其他 \end{cases},$$

现有一大批此种管子（设各电子管损坏与否相互独立），任取 5 只，问其中至少有 2 只寿命大于 1500 小时的概率是多少？

解：电子管寿命大于 1500 小时的概率为：$P(X > 1500) = \int_{1500}^{+\infty} \frac{1000}{x^2} dx = -\frac{1000}{x}\Big|_{1500}^{+\infty} = \frac{2}{3}$ 。

设Y表示寿命大于 1500 小时的只数，则$Y \sim B(5, \frac{2}{3})$，于是所求概率为

$$P(Y \ge 2) = 1 - P(Y = 0) - P(Y = 1) = 1 - C_5^0 \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^5 - C_5^1 \frac{2}{3}\left(\frac{1}{3}\right)^4 = 1 - \frac{11}{3^5} = 0.9547 \text{ 。}$$

8. 已知某电子元件在电源电压不超过 200V、200~240V、240V 以上损坏的概率分别是：0.1，0.001，0.2，而电源电压X服从正态分布 N（220，225）（单位：V），求：

（1） 求电源电压X在区间$(-\infty, 200]$、$(200, 240]$、$(240, +\infty)$的概率；

（2）若该电子元件损坏，问电源电压在 200~240 之间的概率是多少。

解：（1）$$\begin{aligned}&P(X \le 200) = \Phi\left(\frac{200-220}{15}\right) = \Phi(-1.33) = 1 - \Phi(1.33)\\&= 1 - 0.90824 = 0.09176\end{aligned}$$

$$P(200 < X \le 240) = \Phi\left(\frac{240-220}{15}\right) - \Phi\left(\frac{200-220}{15}\right) = \Phi(1.33) - \Phi(-1.33)$$

$$= 2\Phi(1.33) - 1 = 0.81648$$

$$P(X \ge 240) = 1 - \Phi\left(\frac{240-220}{15}\right) = 1 - \Phi(1.33) = 1 - 0.90824 = 0.09176$$

（2）设事件 A_1 , A_2 , A_3 分别表示电源电压 X 落在区间（-∞，200]、（200,240]、（240，+∞），事件 B 表示电子元件损坏，则由贝叶斯公式

$$P(A_2|B) = \frac{P(A_2B)}{P(B)} = \frac{P(A_2)P(B|A_2)}{\sum_{i=1}^{3} P(A_i)P(B|A_i)}$$

$$= \frac{0.81648 * 0.001}{0.09176 * 0.1 + 0.81648 * 0.001 + 0.09176 * 0.2} = 0.0288$$

习题 2-4　二维随机变量及其分布

1．设随机变量 $Z \sim U(-2,2)$，随机变量

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

试求 (X,Y) 的联合分布列。

解 由 $Z \sim U(-2,2)$ 知其密度函数为

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

$$P(X=-1,Y=-1)=P(Z\le -1,Z\le 1)=P(Z\le -1)=\int_{-2}^{-1}\frac{1}{4}dz=\frac{1}{4}；$$

$$P(X=-1,Y=1)=P(Z\le -1,Z>1)=0 ；$$

$$P(X=1,Y=-1)=P(Z>-1,Z\le 1)=P(-1<Z\le 1)=\int_{-1}^{1}\frac{1}{4}dz=\frac{1}{2}；$$

$$P(X=1,Y=1)=P(Z>-1,Z>1)=P(Z>1)=\int_{1}^{2}\frac{1}{4}dz=\frac{1}{4}.$$

故，(X,Y) 的联合分布列为

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

2．完成下列表格并求 $P(X^2+Y^2=2)$

X \ Y	—1	0	1	$p_X(x_i)$
0	0.1	0.1	0.2	0.4
1	0.2	0.2	0.2	0.6
$p_Y(y_j)$	0.3	0.3	0.4	1

解：$P(X^2+Y^2=2)=P(X=1,Y=-1)+P(X=1,Y=1)=0.2+0.2=0.4$ 。

3.选择题：

（1）若 $f(x,y)$ 是二维连续型随机变量 (X,Y) 的联合密度函数，则（ C ）

A. $0 \le f(x) \le 1$　　B. $\lim\limits_{\substack{x\to+\infty \\ y\to+\infty}} f(x,y)=1$

C. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy=1$　　D. $f(x,y)$ 是连续函数

（2）若 $f(x,y)$ 是二维连续型随机变量 (X,Y) 的联合密度函数，则 X 的边缘密度函数为（ B ）

A. $\int_{-\infty}^{+\infty} f(x,y)dx$　B. $\int_{-\infty}^{+\infty} f(x,y)dy$　C. $\int_{-\infty}^{x} f(x,y)dy$　D. $\int_{-\infty}^{y} f(x,y)dx$

4. 设 (X,Y) 服从 $G=\{(x,y)\mid 0\le x\le 2, 0\le y\le 1\}$ 上的均匀分布，求：

（1）(X,Y) 的联合概率密度函数；（2）$P\{Y<X^2\}$；（3）X 和 Y 的边缘密度函数。

解：解：（1）由 (X,Y) 服从 G 上的均匀分布知，(X,Y) 的联合密度为：

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

（2）$P\left(Y<X^2\right)=\int_0^1\left(\int_{\sqrt{y}}^2 \frac{1}{2}dx\right)dy=\frac{2}{3}$。

（3）先求 X 的边缘密度：$f_X\left(x\right)=\int_{-\infty}^{+\infty} f\left(x,y\right)dy$。

当 $x<0$或$x>2$ 时，$f_X\left(x\right)=0$；当 $0\le x\le 2$ 时，$f_X\left(x\right)=\int_0^1 \frac{1}{2}dy=\frac{1}{2}$。

再求 Y 的边缘密度函数：$f_Y\left(y\right)=\int_{-\infty}^{+\infty} f\left(x,y\right)dx$

当 $y<0$或$y>1$ 时，$f_Y\left(y\right)=0$；当 $0\le y\le 1$ 时，$f_Y\left(y\right)=\int_0^2 \frac{1}{2}dx=1$。

5．设二维连续型随机变量 (X,Y) 的联合密度函数为

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

试求：(1)常数 k 的值；(2) X 与 Y 的边缘密度函数 $f_X(x)$ 及 $f_Y(y)$；(3) $P(X+Y<1)$ 及 $P(X<\frac{1}{2})$.

解　（1）$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy=\int_0^1\int_0^1 k(x+y)dxdy=k$.

（2）$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy$.

当 $x \le 0$ 或 $x \ge 1$ 时，$f_X(x) = 0$；

当 $0 < x < 1$ 时，$f_X(x) = \int_0^1 (x+y)dy = x + \frac{1}{2}$.

故，

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

类似地，

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

（3）$P(X+Y<1) = \int_0^1 (\int_0^{1-x} (x+y)dy)dx = \frac{1}{3}$；

$P(X < \frac{1}{2}) = \int_{-\infty}^{1/2} f_X(x)dx = \int_0^{1/2} (x + \frac{1}{2})dx = \frac{3}{8}$

或 $P(X < \frac{1}{2}) = \int_0^{1/2} (\int_0^1 (x+y)dy)dx = \frac{3}{8}$.

6．设二维随机变量 (X,Y) 的联合密度函数为：

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

求：（1）常数 c；（2）$P\{X+Y \le 1\}$；（3）X 和 Y 的边缘密度函数。

解：（1）由 $1 = \int_0^1 (\int_0^2 (x^2 + cxy)dy)dx = \frac{2}{3} + c$，得 $c = \frac{1}{3}$

（2）$P(X+Y \le 1) = \int_0^1 (\int_0^{1-x} (x^2 + \frac{1}{3}xy)dy)dx = \frac{7}{72}$

（3）X 的边缘密度函数：$f_X\left(x\right) = \int_{-\infty}^{+\infty} f\left(x,y\right)dy$，　当 $x<0$ 或 $x>1$ 时，$f_X\left(x\right) = 0$；

当 $0 \le x \le 1$ 时，$f_X\left(x\right) = \int_0^2 \left(x^2 + \frac{1}{3}xy\right)dy = 2x^2 + \frac{2}{3}x$ 。

Y 的边缘密度函数：$f_Y\left(y\right) = \int_{-\infty}^{+\infty} f\left(x,y\right)dx$，当 $y<0$ 或 $y>2$ 时，$f_Y\left(y\right) = 0$；

当 $0 \le y \le 2$ 时，$f_Y\left(y\right) = \int_0^1 \left(x^2 + \frac{1}{3}xy\right)dx = \frac{1}{3} + \frac{1}{6}y$ 。

习题 2-5　条件分布及随机变量的独立性

1．设二维离散型随机变量(X,Y)只取 $(0,0),(-1,1),(-1,2)$ 及 $(2,0)$ 四对值，相应概率依次为 $\frac{1}{12},\frac{1}{6},\frac{1}{3}$和$\frac{5}{12}$，试判断随机变量 X 与 Y 是否相互独立。

解：由已知(X,Y)的联合分布律为

$X \backslash Y$	0	1	2	$p_{i.}$
-1	0	$1/6$	$1/3$	$1/2$
0	$1/12$	0	0	$1/12$
2	$5/12$	0	0	$5/12$
$p_{.j}$	$1/2$	$1/6$	$1/3$	1

由于 $P(X=0)=\frac{1}{12},\quad P\left(Y=0\right)=\frac{1}{12}+\frac{5}{12}=\frac{1}{2},$

　　而 $P\left(X=0,\quad Y=0\right)=\frac{1}{12}\neq P\left(X=0\right)P\left(Y=0\right)=\frac{1}{24}$

所以，X 与 Y 不独立。

2. 设随机变量 X 与 Y 相互独立，试完成下表：

$X \backslash Y$	y_1	y_2	y_3	$p_{i.}$
x_1	1/24	1/8	1/12	1/4
x_2	1/8	3/8	1/4	3/4
$p_{.j}$	1/6	1/2	1/3	1

3．设二维连续型随机变量(X,Y)的联合密度函数为

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

求常数 c，并判断 X 与 Y 是否相互独立。

解：$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f\left(x,y\right)dxdy=\int_0^1\left(\int_0^1 cxy^2dy\right)dx=\frac{c}{6},$ 从而，$c=6$。

X 的边缘密度：$f_X\left(x\right)=\int_{-\infty}^{+\infty}f\left(x,y\right)dy$ 。

当$x\le 0或x\ge 1$时，$f_X\left(x\right)=0$；

当$0<x<1$时，$f_X\left(x\right)=\int_0^1 6xy^2dy=2x$ 。

Y 的边缘密度函数：$f_Y\left(y\right)=\int_{-\infty}^{+\infty}f\left(x,y\right)dx$ 。

当$y\le 0或y\ge 1$时，$f_Y\left(y\right)=0$；

当$0<y<1$时，$f_Y\left(y\right)=\int_0^1 6xy^2dx=3y^2$ 。

由于对任x，y，有$f\left(x,y\right)=f_X\left(x\right)f_Y\left(y\right)$。所以，$X$ 与Y 相互独立。

4*. 设 A 、B 为两个随机事件，已知 $P(A)=\frac{1}{4}$，$P(B|A)=\frac{1}{2}$，$P(A|B)=\frac{1}{4}$. 设随机变量 $X=\begin{cases}1 & A发生\\0 & A不发生\end{cases}$，$Y=\begin{cases}1 & B发生\\0 & B不发生\end{cases}$，问随机变量$X$与$Y$是否相互独立？

解：因为 $P(A)=\frac{1}{4}$，$P(B|A)=\frac{1}{2}$，$P(A|B)=\frac{1}{4}$.， 所以 $P(\bar{A})=\frac{3}{4}$，$P(\overline{B}|A)=\frac{1}{2}$，$P(\overline{A}|B)=\frac{3}{4}$.

$$P(X=0,Y=0)=P(\overline{A}\overline{B})=1-P(A\cup B)=1-P(A)-P(B)+P(AB)$$

$$=1-P(A)-P(B)+P(A)P(B|A)=1-\frac{1}{4}-P(B)+\frac{1}{4}\times\frac{1}{2}$$

$$P(X=0,Y=1)=P(\overline{A}B)=P(B)P(\overline{A}|B)=P(B)\times\frac{3}{4}$$

$$P(X=1,Y=0)=P(A\overline{B})=P(A)P(\overline{B}|A)=\frac{1}{4}\times\frac{1}{2}$$

$$P(X=1,Y=1)=P(AB)=P(A)P(B|A)=\frac{1}{4}\times\frac{1}{2}$$

由规范性，有 $\frac{7}{8}-P(B)+\frac{3}{4}P(B)+\frac{1}{8}+\frac{1}{8}=1$，解出，$P(B)=\frac{1}{2}$

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

经检验边缘分布列乘积等于联合分布列，故 X 与 Y 相互独立。

习题 2-6　随机变量函数的分布

1．设随机变量 X 的分布列为

X	-2	-1	0	1
p_k	1/6	1/3	1/6	1/3

试求：(1)$Y=2X-1$，(2) $Z=X^2$ 的分布列。

解：

$Y=2X-1$	-5	-3	-1	1
p_k	1/6	1/3	1/6	1/3

$Z=X^2$	0	1	4
p_k	1/6	2/3	1/6

2．设随机变量 $X\sim U(0,1)$，试求 $Y=e^X$ 的密度函数。

解：由 $X\sim U(0,1)$ 知其密度函数为 $f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & 其他.\end{cases}$

设 $F_Y(y)$ 为 Y 的分布函数，则 $F_Y(y)=P(Y\le y)=P(e^X\le y)=P(X\le \ln y)=F_X(\ln y)$，

从而 Y 的密度函数　$f_Y(y)=(F_Y(y))'=(F_X(\ln y))'=F_X'(\ln y)(\ln y)'=\dfrac{1}{y}f_X(\ln y)$

$$=\begin{cases}\dfrac{1}{y}\cdot 1, & 0<\ln y<1\\ 0, & 其它\end{cases}=\begin{cases}\dfrac{1}{y}, & 1<y<e\\ 0, & 其它\end{cases}。$$

3．设连续型随机变量 X 的密度函数为 $f(x)=\begin{cases}2(1-x), & 0\le x\le 1\\ 0, & 其它\end{cases}$，求函数 $Y=2X+3$ 的密度函数 $f_Y(y)$。

解：解法一：$F_Y(y)=P(Y\le y)=P(2X+3\le y)=P(X\le \dfrac{y-3}{2})=F_X(\dfrac{y-3}{2})$

$$=\begin{cases}0 & \dfrac{y-3}{2}<0\\ \int_0^{\frac{y-3}{2}}2(1-x)dx & 0\le \dfrac{y-3}{2}<1\\ 1 & \dfrac{y-3}{2}\ge 1\end{cases}$$

所以：$f_Y(y)=F_Y'(y)=\begin{cases}\frac{1}{2}(5-y) & 3\le y<5\\ 0 & 其他\end{cases}$。

或 $f_Y(y)=F_Y'(y)=(F_X(\frac{y-3}{2}))'=F_X'(\frac{y-3}{2})(\frac{y-3}{2})'$

$$=\begin{cases}2(1-\frac{y-3}{2}), & 0\le\frac{y-3}{2}\le 1\\ 0, & 其他\end{cases}=\begin{cases}\frac{1}{2}(5-y) & 3\le y<5\\ 0 & 其他\end{cases}$$

解法二：$\mathrm{y}=2\mathrm{x}+3$ 的反函数为：$x=\frac{y-3}{2}$，其导数：$x'=\frac{1}{2}$

代入公式：$f_Y(y)=\begin{cases}2*\frac{1}{2}(1-\frac{y-3}{2}) & 3\le y<5\\ 0 & 其他\end{cases}=\begin{cases}\frac{1}{2}(5-y) & 3\le y<5\\ 0 & 其他\end{cases}$。

4. 选择题：

（1）设两个相互独立的随机变量 X 和 Y 分别服从正态分布 $N(0,1)$ 和 $N(1,1)$。则（ B ）。(1999年考研题)

（A）$P(X+Y\le 0)=\frac{1}{2}$　　（B）$P(X+Y\le 1)=\frac{1}{2}$

（C）$P(X-Y\le 0)=\frac{1}{2}$　　（D）$P(X-Y\le 1)=\frac{1}{2}$

（2）设 $X\sim N(1,3)$，$Y\sim N(1,1)$，且 X 和 Y 相互独立，则下列说法正确的是（ A ）

（A）$X-Y\sim N(0,4)$　　（B）$X-Y\sim N(1,4)$

（C）$X-Y\sim N(2,4)$　　（D）$X-Y\sim N(0,2)$

习题 3-1 数学期望

1．填空题

（1）设二维随机变量 $(X,Y)\sim N(10,\ 2,\ 1,\ 1,\ 0)$，则 $E(-2XY+Y+5)=$ **-33**。

（2）设随机变量 $X\sim P(2)$，$Y\sim U(0,6)$，若 $Z=2X-3Y-3$，则 $E(Z)=$ **-8**。

（3）设 ξ,η 是两个相互独立且均服从正态分布 $\boldsymbol{N}\left(0,\left(\frac{1}{\sqrt{2}}\right)^2\right)$ 的随机变量，则随机变量 $|\xi-\eta|$ 的数学期望 $E\left(|\xi-\eta|\right)=\sqrt{\frac{2}{\pi}}$。（1996 年考研题）

2．设 X 的分布列为：

X	-1	0	$\frac{1}{2}$	1	2
P	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{4}$

求（1）$E(X)$；（2）$E(-X+1)$；（3）$E(X^2)$。

解：(1) $E(X)=(-1)\times\frac{1}{3}+\frac{1}{2}\times\frac{1}{6}+1\times\frac{1}{12}+2\times\frac{1}{4}=\frac{1}{3}$，

(2)$E(-X+1)=-E(X)+1=\frac{2}{3}$，

(3)$E(X^2)=(-1)^2\times\frac{1}{3}+(\frac{1}{2})^2\times\frac{1}{6}+1^2\times\frac{1}{12}+2^2\times\frac{1}{4}=\frac{35}{24}$。

3．设连续型随机变量 X 的密度函数为

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

求（1）$E(X)$，（2）$E\,|\,X-EX\,|$。

解：$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_0^1 x\bullet xdx+\int_1^2 x(2-x)dx=1$，

$E(|X-EX|)=\int_{-\infty}^{+\infty}|x-1|f(x)dx=\int_0^1(1-x)xdx+\int_1^2(x-1)(2-x)dx=\frac{1}{3}$。

4．设二维离散型随机变量 (X,Y) 的联合分布列为

X \ Y	0	1
0	0.3	0.4
1	0.2	0.1

求：（1）$E(X),E(Y)$；（2）$E(X-2Y),E(3XY)$。

解：（1）

X	0	1
p_k	0.7	0.3

$E(X)=0.3$

Y	0	1
p_k	0.5	0.5

$E(Y)=0.5$

(2) $E(X-2Y)=E(X)-2E(Y)=0.3-2*0.5=-0.7$，

$E(3XY)=3E(XY)=3(1\times1\times0.1)=0.3$。

5. 设(X,Y)服从在区域A上的均匀分布，其中A由x轴、y轴及直线$x+y+1=0$所围成，

求（1）$E(X)$；（2）$E(-3X+2Y)$；（3）$E(XY)$。

解：由题意知(X,Y)的联合密度为：

$$f(x,y)=\begin{cases}2 & (x,y)\in A\\ 0 & \text{其他}\end{cases}$$

（1）$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)dxdy=\int_{-1}^{0}(\int_{-x-1}^{0}2xdy)dx=-\dfrac{1}{3}$。

（2） $E(-3X+2Y)=-3E(X)+2E(Y)=1+2E(Y)$

$$=1+2\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)dxdy$$

$$=1+2\int_{-1}^{0}(\int_{-1-y}^{0}2ydx)dy=\frac{1}{3}$$。

（3） $E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyf(x,y)dxdy=\int_{-1}^{0}(\int_{-1-x}^{0}xy\bullet 2dy)dx=\dfrac{1}{12}$。

习题 3-2 方差

1. 填空题

（1）从学校乘汽车到火车站途中有 8 个交通路口，假设在每个路口遇到红绿灯是独立的，且遇到红绿灯的概率均为 0.4。令 X 表示途中遇到红灯的次数，则 $X \sim \underline{B(8,0.4)}$， $E(X)=\underline{\ 3.2\ }$，

$D(X)=\underline{1.92\ \ }$。

（2）设随机变量 X_1， X_2， X_3 相互独立，其中 $X_1 \sim U(0,6)$， $X_2 \sim N(0,4)$， X_3 服从参数为 3 的泊松分布，记 $Y = X_1 - 2X_2 + 3X_3$，则 $D(Y)=\underline{\quad 46 \quad}$。

（3）设二维随机变量 $(X,Y) \sim N(1,\ 2,\ 1,\ 1,\ 0)$，则 $D(-2X+Y+5)=\underline{\ 5\ }$，$Z=-2X+Y$ 分布为 $\underline{N(0,5)}$。

2. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases} 0, & x<-1 \\ \dfrac{2}{\pi}\arctan x+\dfrac{1}{2}, & -1\le x<1 \\ 1, & x\ge 1 \end{cases},$$

求（1） X 的密度函数；（2） $E(X), D(X)$。

解：(1)由 $f(x)=F'(x)$ 知 $f(x)=\begin{cases} \dfrac{2}{\pi}\dfrac{1}{1+x^2} & -1<x<1 \\ 0 & 其他 \end{cases}$

（2）$E(X)=\int_{-\infty}^{+\infty} x f(x)dx=\int_{-1}^{1}\frac{2}{\pi}\frac{x}{1+x^2}dx=0$，$E(X^2)=\int_{-\infty}^{+\infty} x^2 f(x)dx=\int_{-1}^{1}\frac{2}{\pi}\frac{x^2}{1+x^2}dx=\frac{4}{\pi}-1$，

$D(X)=E(X^2)-(E(X))^2=\dfrac{4}{\pi}-1$。

3．设随机变量 $X \sim P(\lambda)$ 且 $E[(X-1)(X-2)]=1$，随机变量 $Y \sim B(8,\frac{1}{2})$ 且 X 与 Y 相互独立，试求 $E(X-3Y-4)$ 及 $D(X-3Y-4)$。

解：由 $X \sim P(\lambda)$ 知 $E(X)=\lambda$， $D(X)=\lambda$. 所以， $E(X^2)=D(X)+(E(X))^2=\lambda+\lambda^2$. 又

$1=E[(X-1)(X-2)]=E(X^2)-3E(X)+2=\lambda^2-2\lambda+2$，故 $\lambda=1$. 所以，$E(X)=1$，$D(X)=1$.

由于 $Y \sim B(8,\frac{1}{2})$，故 $E(Y)=4$， $D(Y)=2$。

所以， $E(X-3Y-4)=E(X)-3E(Y)-4=-15$。

由于 X 与 Y 相互独立，故 $D(X-3Y-4)=D(X)+9D(Y)=19$ 。

4．设 (X,Y) 的联合概率密度为 $f(x,y)=\begin{cases}12y^2, & 0\le y\le x\le 1\\ 0, & 其他\end{cases}$,试求 $D(X)$ 及 $D(Y)$ 。

解：
$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)dxdy=\int_0^1(\int_0^x x\cdot 12y^2dy)dx=\frac{4}{5},$$
$$E(X^2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x^2f(x,y)dxdy=\int_0^1(\int_0^x x^2\cdot 12y^2dy)dx=\frac{2}{3},$$
$$D(X)=E(X^2)-(E(X))^2=\frac{2}{75},$$
$$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)dxdy=\int_0^1(\int_0^x y\cdot 12y^2dy)dx=\frac{3}{5},$$
$$E(Y^2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}y^2f(x,y)dxdy=\int_0^1(\int_0^x y^2\cdot 12y^2dy)dx=\frac{2}{5},$$
$$D(Y)=E(Y^2)-(E(Y))^2=\frac{2}{5}-(\frac{3}{5})^2=\frac{1}{25}。$$

5． Warren Dinner 公司对 9 个不同的项目进行了投资，假设不同投资项目带来的收益是独立的，每个投资项目的收益服从正态分布，且均值为 500 个单位，标准差为 100 个单位。

（1） Warren Dinner 公司总的投资收益的平均值是多少？

（2） 总收益在 4000～5200 之间的概率是多少？

解:设 X 为公司总的投资收益，$X_i\ (i=1,2,\cdots 9)$ 为第 i 个投资项目的收益，则由题意，$X_1,X_2,\cdots X_9$ 相互独立且均服从均值为 500 个单位，标准差为 100 个单位的正态分布，且 $X=X_1+X_2+\cdots+X_9$，

（1）$E(X)=E(X_1+X_2+\cdots+X_9)=9\times E(X_1)=9\times 500=4500$（单位）

并且有：$D(X)=D(X_1+X_2+\cdots+X_9)=9\times D(X_1)=9\times 10000=90000$

$\sqrt{D(X)}=300$（单位），所以有 $X\sim N(4500,300^2)$ 。

（2）$P(4000<X<5200)=P(\frac{4000-4500}{300}<\frac{X-4500}{300}<\frac{5200-4500}{300})$

$=\Phi(\frac{7}{3})-\Phi(-\frac{5}{3})=\Phi(2.33)+\Phi(1.67)-1=0.99+0.9525-1=0.9425$ 。

习题 3-3 协方差与相关系数

1．填空题

（1）设随机变量 $X\sim P(2)$，$Y\sim U(0,6)$ 且 $\rho_{XY}=\dfrac{1}{\sqrt{6}}$，若 $Z=2X-3Y-3$，则 $D(Z)=$__23___。

（2）设 (X,Y) 服从二维正态分布，则 $\mathrm{Cov}(X,Y)=0$ 是 X 与 Y 相互独立的__充要__条件。

（3）设 (X,Y) 服从二维正态分布 $N(0,1,1,4,0.5)$，则 $E(2X^2-XY+3)=$__4__。

2. 选择题

（1）设 X 与 Y 的相关系数 $\rho_{XY}=0$，则必有__C__。

(A) X 与 Y 相互独立；　　(B) X 与 Y 不一定相关；

(C) X 与 Y 必不相关；　　(D) X 与 Y 必相关

（2）设随机变量 X 与 Y 的期望和方差存在，且 $D(X-Y)=D(X)+D(Y),$，则下列说法哪个是不正确的__B__ 。

(A) $D(X+Y)=D(X)+D(Y)$；　　(B) X 与 Y 独立；

(C) X 与 Y 不相关；　　(D) $E(XY)=E(X)\cdot E(Y)$

3. 已知二维离散型随机变量 (X,Y) 的联合分布列为

$X \backslash Y$	-1	0	1
-1	1/8	1/8	1/8
0	1/8	0	1/8
1	1/8	1/8	1/8

（1）求协方差 $\mathrm{Cov}(X,Y)$ 及相关系数 ρ_{XY}；（2）判断 X 与 Y 是否相互独立？是否不相关？

解：X 及 Y 的边缘分布列为：

X	-1	0	1
p_k	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{3}{8}$

Y	-1	0	1
p_k	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{3}{8}$

(1)$E(X)=0$，$E(Y)=0,\ E(XY)=1\times\frac{1}{8}-1\times\frac{1}{8}-1\times\frac{1}{8}+1\times\frac{1}{8}=0$ 。

故$Cov(X,Y)=E(XY)-E(X)E(Y)=0$。所以，$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=0$。

(2)由于$P(X=-1,Y=-1)=\dfrac{1}{8}\neq P(X=-1)P(Y=-1)=\dfrac{9}{64}$ 所以X与Y不独立。但$\rho_{XY}=0$，故X与Y不相关。

4．设二维连续型随机变量(X,Y)的联合密度函数为$f(x,y)=\begin{cases}1, & 0<x<1, 0<y<2x,\\ 0, & 其他.\end{cases}$

试求：(1) X与Y的边缘概率密度函数$f_X(x)$和$f_Y(y)$，并判断X与Y是否相互独立；(2)求X与Y的协方差$Cov(X,Y)$，判断X与Y是否相关。（3）求$P(X<\dfrac{1}{2})$。

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

当$x\le 0$或$x\ge 1$时，$f_X(x)=0$；当$0<x<1$时，$f_X(x)=\int_0^{2x}1dy=2x$。

$f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx$.

当$y\le 0$或$y\ge 2$时，$f_Y(y)=0$；当$0<y<2$时，$f_Y(y)=\int_{y/2}^{1}1dx=1-\dfrac{y}{2}$。

由于在$(x,y)\in\{0<x<1, 0<y<2x\}$区域内，

$f(x,y)\neq f_X(x)\cdot f_Y(y)$，

且区域$\{0<x<1, 0<y<2x\}$的面积不为0，所以，X与Y不相互独立。

（2）$E(X)=\int_0^1 x\cdot 2xdx=\dfrac{2}{3}, E(Y)=\int_0^2 y(1-\dfrac{y}{2})dy=\dfrac{2}{3}$

$E(XY)=\int_0^1[\int_0^{2x}xydydx=\dfrac{1}{2}$，

所以$Cov(X,Y)=E(XY)-E(X)E(Y)=\dfrac{1}{2}-\dfrac{2}{3}\cdot\dfrac{2}{3}=\dfrac{1}{18}\neq 0$，故$X$与$Y$相关。

（3）$P(X<\dfrac{1}{2})=\int_0^{0.5}dx\int_0^{2x}dy=\dfrac{1}{4}$

习题 4　大数定律与中心极限定理

1. 已知正常成人男性血液中每毫升含白细胞数的平均值是 7300 个，均方差是 700，利用切比雪夫不等式估计每毫升血液中白细胞数在 5200~9400 之间的概率。

解 以 X 表示每毫升含白细胞数，由题设

$$E(X)=7300, D(X)=700^2$$

而概率

$$P(5200<X<9400)=P(-2100<X-7300<2100)$$
$$=P(|X-7300|<2100)$$

在切比雪夫不等式

中，取 $\varepsilon=2100$，此时 $1-\dfrac{D(X)}{\varepsilon^2}=1-700^2/2100^2=8/9$，知

$$P(|X-7300|<2100)\geq 8/9=0.8889。$$

2. 某保险公司多年的资料表明，在索赔户中被盗索赔户占 20%。随机抽查 100 个索赔户，令 X 表示其中因被盗而向保险公司索赔的户数，求 $P(14\leq X\leq 30)$。

解：$X\sim B(100,0.2)$，$E(X)=np=100\times 0.2=20, D(X)=np(1-p)=100\times 0.2\times 0.8=16$

$$P(14\leq X\leq 30)=P(\frac{14-20}{\sqrt{16}}\frac{X-np}{\sqrt{np(1-p)}}\leq\frac{30-20}{\sqrt{16}})$$

$$\approx\Phi(2.5)-\Phi(-1.5)=0.99379+0.93319-1=0.92698。$$

3. 设随机变量 $X_1,X_2,\cdots,X_{48}$ 相互独立且都在[0，1]上服从均匀分布．令 $\overline{X}=\dfrac{1}{48}\sum_{i=1}^{48}X_i$，试用中心极限定理计算 $P(\left|\overline{X}-\dfrac{1}{2}\right|<0.04)$ 的值。

解：由中心极限定理，

$$P(\left|\overline{X}-\frac{1}{2}\right|<0.04)=P(\frac{\left|\overline{X}-1\right|}{\frac{1}{\sqrt{12\times 48}}}\leq\frac{0.04}{\frac{1}{\sqrt{12\times 48}}})=2\Phi(0.96)-1=0.663$$

4. 在人寿保险公司里有 10000 个同一年龄的人参加人寿保险。在这一年中，这些人的死亡率为 0.6%，参加保险的人在一年的头一天交付保险费 12 元，死亡时，家属可以从保险公司领取 1000 元。求

(1)保险公司一年中获利不少于 40000 元的概率；

(2)保险公司亏本的概率。

解： 设 X 表示一年中 **10000** 个同龄参保人中死亡的人数，则 $X \sim B(10000, 0.006)$，由题意，保险公司的收益为 $10000 \times 12 = 120000$ 元，支出为 **1000** X 。由中心极限定理

（**1**）保险公司一年中获利不少于 **40000** 元的概率为

$$P(120000 - 1000X > 40000) = P(X < 80)$$

$$= P\left(\frac{X - np}{\sqrt{np(1-p)}} < \frac{80 - 60}{\sqrt{59.64}}\right) \approx \Phi(2.59) = 0.9952$$

（**2**）保险公司亏本的概率为

$$P(1000X > 120000) = P(X > 120)$$

$$= 1 - P\left(\frac{X - np}{\sqrt{np(1-p)}} < \frac{120 - 60}{\sqrt{59.64}}\right) \approx 1 - \Phi(7.77) = 0$$

可见保险公司一般不会亏本。

5．某车间有同型号机床 200 部，每部开动的概率为 0.7，假定各机床开关是独立的，开动时每部要消耗电能 15 个单位。问电厂最少要供应这个车间多少电能，才能以 95%的概率保证不致因供电不足而影响生产。

解 设 X 表示同时开动机床的台数，则 $X \sim B(200,\ 0.7)$

$$E(X) = np = 200 \times 0.7 = 140,\ D(X) = np(1-p) = 200 \times 0.7 \times 0.3 = 42$$

又设需供电 **Q** 单位，而实际用电为 **15** X 单位，题意要求 $P(15X \le Q) \ge 0.95$，由中心极限定理

$$P(15X \le Q) = P\left(X \le \frac{Q}{15}\right) = P\left(\frac{X - np}{\sqrt{np(1-p)}} \le \frac{Q/15 - 140}{\sqrt{42}}\right) \approx \Phi\left(\frac{Q/15 - 140}{\sqrt{42}}\right)$$

由题意要求 $\Phi\left(\frac{Q/15 - 140}{\sqrt{42}}\right) \ge 0.95$

而由 $\Phi(1.645) = 0.95$ 得 $\frac{Q/15 - 140}{\sqrt{42}} \ge 1.645$

解得 $Q \ge 2259.91$，取 $Q = 2260$，所以应供电能 **2260** 个单位才能满足要求。

习题 5-1　习题 5-2　统计量和抽样分布

1．填空题

（1）设 $X \sim N(\mu,\sigma^2)$，$\overline{X}$，S^2 分别为样本均值与样本方差，n 为样本容量，则

$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim$ __t(n-1)__，　$\dfrac{(n-1)S^2}{\sigma^2} \sim$ __$\chi^2(n-1)$__ 。

（2）设随机变量 X 与 Y 相互独立且 $X \sim N(\mu,\sigma^2)$，$Y \sim \chi^2(n)$，则 $Z=\dfrac{X-\mu}{\sqrt{Y}\sigma}\sqrt{n} \sim$ __$t(n)$__ 。

(3) 设总体 X 服从正态分布 $N(0,1)$，而 $X_1,X_2,\cdots,X_{15}$ 是来自总体 X 的简单随机样本，则随机变量 $Y=\dfrac{X_1^2+\cdots+X_{10}^2}{2(X_{11}^2+\cdots+X_{15}^2)} \sim$ __$F(10,5)$__。

(4) 设 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$，且 X，Y 相互独立，则 $F=\dfrac{Y/n_2}{X/n_1} \sim$ __$F(n_2,n_1)$__。

2. 选择题

(1) $F_{0.95}(7,9)=$（　D　）

（A）$F_{0.95}(9,7)$　　（B）$\dfrac{1}{F_{0.95}(9,7)}$　　（C）$\dfrac{1}{F_{0.05}(7,9)}$　　（D）$\dfrac{1}{F_{0.05}(9,7)}$

(2) 设总体 $X \sim N(\mu,\sigma^2)$，其中 μ 已知，σ^2 未知，X_1,X_2,X_3 是从中抽取的简单随机样本，下列各项中不是统计量的是（ A　）

（A）$\dfrac{1}{\sigma^2}(X_1^2+X_2^2+X_3^2)$　　（B）$X_1+3\mu$

（C）$\max(X_1,X_2,X_3)$　　（D）$\dfrac{1}{3}(X_1+X_2+X_3)$

(3) 设随机变量 $X \sim t(n)\ (n>1), Y=\dfrac{1}{X^2}$，则（　C　）

(A)　$Y \sim \chi^2(n)$　(B) $Y \sim \chi^2(n-1)$　(C)　$Y \sim F(n,1)$　(D) $Y \sim F(1,n)$

3．设某种电灯泡的寿命 X 服从指数分布 $E(\lambda)$，从中抽取 100 只灯泡，求这一简单随机样本 $X_1,X_2,\cdots,X_{100}$ 的联合概率密度函数。

解：$f(x_1,x_2,\cdots,x_{100})=\prod\limits_{i=1}^{100} f(x_i)=\lambda^{100}e^{-\lambda\sum\limits_{i=1}^{100}x_i}$，　　其中 $x_i>0, i=1,2,\cdots,100$

4.抽取 10 只辽宁绒山羊产绒量（单位：g）：450,450,500,500,500,550,550,600,600,650,试计算其样本均值、样本方差和标准差。

解：样本均值 $\overline{x}=\dfrac{1}{40}\sum_{i=1}^{40}x_i=535$

样本方差 $s^2=\dfrac{1}{39}\sum_{i=1}^{40}(x_i-\overline{x})^2\approx 4472.222$

样本标准差 $s=\sqrt{\dfrac{1}{39}\sum_{i=1}^{40}(x_i-\overline{x})^2}\approx 66.875$

5.设 $X_1,X_2,\cdots,X_5$ 是独立且服从相同分布 $N(0,1)$ 的随机变量，

(1)试给出常数 c，使得 $c\cdot(X_1{}^2+X_2{}^2)$ 服从 χ^2 分布，并指出它的自由度；

(2)试给出常数 d，使得 $d\cdot\dfrac{X_1+X_2}{\sqrt{X_3{}^2+X_4{}^2+X_5{}^2}}$ 服从 t 分布，并指出它的自由度。

解：（1）因为 $X_1{}^2+X_2{}^2\sim\chi^2(2)$，所以 $c=1$，自由度为 2。

（2）因为 $\dfrac{\dfrac{X_1+X_2}{\sqrt{2}}}{\sqrt{\dfrac{X_3{}^2+X_4{}^2+X_5{}^2}{3}}}\sim t(3)$，所以 $d=\dfrac{\sqrt{3}}{\sqrt{2}}$，自由度为 3.

6*.设 $X_1,X_2,\cdots,X_n(n>2)$ 为来自总体 N(0,1)的简单随机样本，$\overline{X}$ 为样本均值，记

$Y_i=X_i-\overline{X},i=1,2,\cdots,n.$

求：（1） Y_i 的方差 $D(Y_i),i=1,2,\cdots,n$；

(2) Y_1 与 Y_n 的协方差 $Cov(Y_1,Y_n)$.

解：（I）$D(Y_i)=D\left(X_i-\overline{X}\right)=D(X_i)+D(\overline{X})-2Cov(X_i,\overline{X})=\sigma^2+\dfrac{\sigma^2}{n}-2Cov(X_i,\dfrac{X_i}{n})$

$$=\sigma^2+\frac{\sigma^2}{n}-2\cdot\frac{\sigma^2}{n}=\frac{(n-1)\sigma^2}{n}$$

（II）$Cov(Y_1,Y_n)=Cov(X_1-\overline{X},X_n-\overline{X})=-Cov(X_1,\overline{X})-Cov(X_n,\overline{X})+Cov(\overline{X},\overline{X})$

$$=-\frac{\sigma^2}{n}-\frac{\sigma^2}{n}+\frac{\sigma^2}{n}=-\frac{\sigma^2}{n}$$

习题 5-3　正态总体统计量的抽样分布

1. 填空题

(1) 设 $X_1,\cdots,X_7$ 为总体 $X\sim N(0,0.5^2)$ 的一个样本，则 $P(\sum_{i=1}^{7}X_i^2>4)=$ <u>0.025</u> .

(2) 设总体 $X\sim N(\mu,\sigma^2)$， $X_1,X_2,\cdots,X_n$ 为来自总体 X 的样本，$\overline{X}$ 为样本均值，则 $D(\overline{X})=$ $\underline{\dfrac{\sigma^2}{n}}$.

2.选择题

(1) 假设总体 $X\sim N(1,\ 2^2)$， $X_1,X_2,\cdots,X_{100}$ 是来自总体 X 的一个样本，$\overline{X}$ 为其样本均值，且 $\overline{X}\sim N(\mu,\ \sigma^2)$，则下列成立的是（ D ）

（A） μ=1， σ=0.04　　（B） μ=100， σ=0.2

（C） μ=0.01， σ=0.04　　（D） μ=1， σ=0.2

(2) 设 $X_1,X_2,\cdots,X_{100}$ 为来自总体 $X\sim N(\mu,4^2)$ 的一个样本，而 $Y_1,Y_2,\cdots,Y_{100}$ 为来自总体 $Y\sim N(\mu,3^2)$ 的一个样本，且两个样本独立，以 $\overline{X},\overline{Y}$ 分别表示这两个样本的样本均值，则 $\overline{X}-\overline{Y}$ 所服从的分布是（　B　）

（A） $N\left(0,\dfrac{7}{100}\right)$　　（B） $N\left(0,\dfrac{1}{4}\right)$　　（C） $N(0,7)$　　（D） $N(0,25)$

3．从正态总体 $N(3.4,6^2)$ 中抽取容量为 n 的样本，如果要求其样本均值位于区间（1.4,5.4）内的概率不小于 0.95，问样本容量 n 至少应取多大？

解：由题意

$$P(1.4<\overline{X}<5.4)=P\left(\frac{1.4-3.4}{6/\sqrt{n}}<\frac{\overline{X}-3.4}{6/\sqrt{n}}<\frac{5.4-3.4}{6/\sqrt{n}}\right)=\Phi\left(\frac{\sqrt{n}}{3}\right)-\Phi\left(-\frac{\sqrt{n}}{3}\right)$$

$$=2\Phi\left(\frac{\sqrt{n}}{3}\right)-1\geq 0.95，$$

即 $\Phi\left(\dfrac{\sqrt{n}}{3}\right)\geq 0.975$，查表得，$\dfrac{\sqrt{n}}{3}\geq 1.96$，所以 $n\geq 34.5744$，样本容量 n 至少应取 35.

4. 从正态总体 $N(\mu,0.5^2)$ 中抽取容量为 10 的样本 $X_1,X_2,\cdots,X_{10}$，

（1）已知 $\mu=0$，求 $\sum_{i=1}^{10}X_i{}^2\geq 4$ 的概率；（2）未知 μ，求 $\sum_{i=1}^{10}\left(X_i-\overline{X}\right)^2<3.67$ 的概率.

解：（1）当 $\mu=0$ 时，因为 $X_i \sim N(0,0.5^2)$，则 $\sum_{i=1}^{10}\left(\frac{X_i-0}{0.5}\right)^2 \sim \chi^2(10)$，

所以 $P\left(\sum_{i=1}^{10} X_i^{\,2} \geq 4\right)=P\left(\sum_{i=1}^{10}\left(\frac{X_i-0}{0.5}\right)^2 \geq 16\right)=P(\chi^2 \geq 16)$，

查附表 4 得上述概率为 0.1。

（2）当 μ 为未知时，因为 $X_i \sim N(\mu,0.5^2)$，则 $\sum_{i=1}^{10}\left(\frac{X_i-\bar{X}}{0.5}\right)^2 \sim \chi^2(9)$，

所以 $P\left(\sum_{i=1}^{10}\left(X_i-\bar{X}\right)^2<3.67\right)=P\left(\sum_{i=1}^{10}\left(\frac{X_i-\bar{X}}{0.5}\right)^2<14.68\right)=P\left(\chi^2<14.68\right)$，

查附表 4 得 $P\left(\chi^2 \geq 14.68\right)=0.1$，故上述概率为 0.9.

5. 设总体 $X \sim N(50,6^2)$，总体 $Y \sim N(46,4^2)$，从总体 X 中抽取容量为 10 的样本，其样本方差计为 S_1^2；从总体 Y 中抽取容量为 8 的样本，其样本方差记为 S_2^2，求下列概率：

（1）$P(0<\bar{X}-\bar{Y}<8)$；　（2）$P\left(\frac{S_1^2}{S_2^2}<8.28\right)$

解：（1）因为 $U=\frac{(\bar{X}-\bar{Y})-(50-46)}{\sqrt{\frac{6^2}{10}+\frac{4^2}{8}}}=\frac{\bar{X}-\bar{Y}-4}{\sqrt{5.6}} \sim N(0,1)$

则 $P\left(0<\bar{X}-\bar{Y}<8\right)=P\left(\frac{0-4}{\sqrt{5.6}}<\frac{\bar{X}-\bar{Y}-4}{\sqrt{5.6}}<\frac{8-4}{\sqrt{5.6}}\right)=\Phi(1.69)-\Phi(-1.69)$

$=2\Phi(1.69)-1=0.909$

（2）因为 $F=\frac{S_1^2/6^2}{S_2^2/4^2}=\frac{4S_1^{\,2}}{9S_2^{\,2}} \sim F(9,7)$

则 $P\left(\frac{S_1^{\,2}}{S_2^{\,2}}<8.28\right)=P\left(\frac{4S_1^{\,2}}{9S_2^{\,2}}<3.68\right)=P\left(F<3.68\right)$

查附表 6 得 $F_{0.05}(9,7)=3.68$，即 $P\left(F \geq F_{0.05}(9,7)\right)=P\left(F \geq 3.68\right)=0.05$

由此得所求的概率 $P\left(\frac{S_1^{\,2}}{S_2^{\,2}}<8.28\right)=P\left(F<3.68\right)=1-P\left(F \geq 3.68\right)=0.95$

6*.设总体 $X \sim N(\mu,\sigma^2)(\sigma>0)$，从该总体中抽取样本 $X_1,X_2,\cdots,X_{2n}(n\ge 2)$，其样本的均值 $\overline{X}=\frac{1}{2n}\sum_{i=1}^{2n}X_i$，求统计量 $Y=\sum_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2$ 的数学期望 $E(Y)$。

解：$E(Y)=E\left(\sum_{i=1}^{n}(X_i+X_{n+i}-2\overline{X})^2\right)=\sum_{i=1}^{n}E(X_i+X_{n+i}-2\overline{X})^2$

$$=\sum_{i=1}^{n}E(X_i-\overline{X}+X_{n+i}-\overline{X})^2=\sum_{i=1}^{n}\left[E(X_i-\overline{X})^2+E(X_{n+i}-\overline{X})^2+2E(X_i-\overline{X})(X_{n+i}-\overline{X})\right]$$

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

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

习题 6-1 点估计

1. 选择题

（1）设 $X_1, X_2, \cdots, X_n$ 是取自总体 X 的一个简单随机样本，则 $E(X^2)$ 的矩估计是（ D ）

（A）$S_1^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ （B）$S_2^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ （C）$S_1^2+\overline{X}^2$ （D）$S_2^2+\overline{X}^2$

（2）设总体 X 服从正态分布 $N\left(\mu,\sigma^2\right), X_1, X_2, \cdots, X_n$ 是来自 X 的样本，则 σ^2 的最大似然估计为（ A ）

（A）$\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$ （B）$\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$ （C）$\frac{1}{n}\sum_{i=1}^{n}X_i^2$ （D）$\bar{X}^2$

（3）在（2）题条件下，σ^2 的无偏估计量是（ B ）

（A）$\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$ （B）$\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$ （C）$\frac{1}{n}\sum_{i=1}^{n}X_i^2$ （D）$\bar{X}^2$

2. 设总体 X 具有分布列 ：

X	1	2	3
p	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 $\theta(0<\theta<1)$ 为未知参数，已知取得了样本值 $x_1=1$，$x_2=2$，$x_3=1$ 试求 θ 的矩估计值和最大似然估计值。

解： $\mu_1=E(X)=\theta^2+4\theta(1-\theta)+3(1-\theta)^2=3-2\theta$

由矩估计法 $\hat{\mu}_1=3-2\hat{\theta}=\bar{x}$ **得** $\hat{\theta}=\frac{3-\bar{x}}{2}$

因 $\bar{x}=\frac{1}{3}(1+2+1)=\frac{4}{3}$ **，所以 θ 的矩估计为** $\hat{\theta}=\frac{3-\bar{x}}{2}=\frac{3-4/3}{2}=\frac{5}{6}$；

似然函数 $L(x_1,x_2,x_3;\theta)=\theta^2\bullet 2\theta(1-\theta)\bullet\theta^2=2\theta^5-2\theta^6$

令导数 $L'(\theta)=10\theta^4-12\theta^5=0$ **（或先取对数，再令导数等于 0）**

得 θ 的最大似然估计为 $\hat{\theta}=\frac{10}{12}=\frac{5}{6}$。

3. 设总体 X 的概率密度函数为

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

$X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，求 θ 的矩估计和最大似然估计。

解：1) 由题意 $\mu_1 = E(X) = \int_0^1 x\theta x^{\theta-1} dx = \frac{\theta}{\theta+1} x^{\theta+1} \Big|_0^1 = \frac{\theta}{\theta+1}$

解之得：$\theta = \frac{\mu_1}{1-\mu_1}$，由矩估计法 $\hat{\mu}_1 = \overline{X}$，得 θ 的矩估计： $\hat{\theta} = \frac{\hat{\mu}_1}{1-\hat{\mu}_1} = \frac{\overline{X}}{1-\overline{X}}$。

2) 构造似然函数

$$L(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} \theta x_i^{\theta-1} = \theta^n (\prod_{i=1}^{n} x_i)^{\theta-1} \qquad 0 < x_i < 1, i = 1, 2, \cdots, n .$$

两边取对数得 $\ln L = n \ln \theta + (\theta - 1) \ln(\prod_{i=1}^{n} x_i)$

对 θ 求导并令其等于零,得似然方程 $\frac{d \ln L}{d\theta} = \frac{n}{\theta} + \ln(\prod_{i=1}^{n} x_i) = 0$，

解之得参数 θ 的最大似然估计值为 $\hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln x_i}$，

与它相应的最大似然估计量 $\hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \ln X_i}$。

4. 设总体 X 服从正态分布 $N(\mu, \sigma^2)$， X_1, X_2 是从此总体中抽取的一个样本．试验证下面三个估计量：

（1）$\hat{\mu}_1 = \frac{2}{3} X_1 + \frac{1}{3} X_2$，　（2）$\hat{\mu}_2 = \frac{1}{4} X_1 + \frac{3}{4} X_2$，　（3）$\hat{\mu}_3 = \frac{1}{2} X_1 + \frac{1}{2} X_2$

都是 μ 的无偏估计，并指出哪一个估计量最有效．

证:

$$E(\hat{\mu}_1) = E(\frac{2}{3} X_1 + \frac{1}{3} X_2) = \frac{2}{3} E(X_1) + \frac{1}{3} E(X_2) = \frac{2}{3}\mu + \frac{1}{3}\mu = \mu$$

$$E(\hat{\mu}_2) = E(\frac{1}{4} X_1 + \frac{3}{4} X_2) = \frac{1}{4} E(X_1) + \frac{3}{4} E(X_2) = \frac{1}{4}\mu + \frac{3}{4}\mu = \mu$$

$$E(\hat{\mu}_3) = E(\frac{1}{2} X_1 + \frac{1}{2} X_2) = \frac{1}{2} E(X_1) + \frac{1}{2} E(X_2) = \frac{1}{2}\mu + \frac{1}{2}\mu = \mu$$

因为：

$$D(\hat{\mu}_1) = D(\frac{2}{3} X_1 + \frac{1}{3} X_2) = (\frac{2}{3})^2 D(X_1) + (\frac{1}{3})^2 D(X_2)$$

$$= \frac{4}{9}\sigma^2 + \frac{1}{9}\sigma^2 = \frac{5}{9}\sigma^2$$

$$D(\hat{\mu}_2)=D(\frac{1}{4}X_1+\frac{3}{4}X_2)=\frac{5}{8}\sigma^2$$

$$D(\hat{\mu}_3)=D(\frac{1}{2}X_1+\frac{1}{2}X_2)=(\frac{1}{2})^2D(X_1)+(\frac{1}{2})^2D(X_2)$$
$$=\frac{1}{4}\sigma^2+\frac{1}{4}\sigma^2=\frac{1}{2}\sigma^2$$

所以 $\hat{\mu}_3$ 最有效。

5. 设 $X_1,X_2,\cdots,X_n$ 为总体 X 的一个简单随机样本，$E(X)=\mu,D(X)=\sigma^2$，若 $\hat{\theta}^2=C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$ 为 σ^2 的无偏估计，求 C 值。

解　由题意知：$E(X_i)=\mu,D(X_i)=\sigma^2$，$\therefore E(X_i^2)=D(X_i)+[E(X_i)]^2=\sigma^2+\mu^2$

而
$$E[C\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2]=C\sum_{i=1}^{n-1}[E(X_{i+1}^2)-2E(X_{i+1}X_i)+E(X_i^2)]$$
$$=C\sum_{i=1}^{n-1}[\sigma^2+\mu^2-2\mu^2+\sigma^2+\mu^2]=C\cdot 2(n-1)\sigma^2=\sigma^2$$

所以　$C=\frac{1}{2(n-1)}$.

习题 6-2 区间估计

1. 设有一组来自正态总体 $N(\mu,\sigma^2)$ 的样本观测值：

0.497，0.506，0.518，0.524，0.488，0.510，0.510，0.515，0.512，

⑴ 已知 $\sigma=0.01$，求 μ 的置信区间（设置信度为 0.95）；

⑵ σ^2 未知，求 μ 的置信区间（设置信度为 0.95）.

解：由样本算得 $\bar{x}=0.50889 \quad s=0.01088 \quad$ 而$1-\alpha=0.95$

（1）$\sigma=0.01$时，查得$u_{\alpha/2}=u_{0.025}=1.96$，μ 的置信区间为

$[\bar{x}\text{-}u_{0.025}\dfrac{\sigma_0}{\sqrt{n}},\bar{x}+u_{0.025}\dfrac{\sigma_0}{\sqrt{n}}]$ =$[0.50889\text{-}1.96\times\dfrac{0.01}{\sqrt{9}},0.50889+1.96\times\dfrac{0.01}{\sqrt{9}}]$ =[0.50204，0.5154]。

（2）σ^2 未知时，对 $\alpha=0.05$，查得 $t_{\alpha/2}(n-1)=t_{0.025}(8)=2.306$，μ 的置信区间为

$[\bar{x}\text{-}t_{0.025}(n-1)\dfrac{s}{\sqrt{n}},\bar{x}+t_{0.025}(n-1)\dfrac{s}{\sqrt{n}}]$ =$[0.50889\text{-}2.306\times\dfrac{0.01088}{\sqrt{9}},0.50889+2.306\times\dfrac{0.01088}{\sqrt{9}}]$

=[0. 5006，0.5172]。

2. 某厂生产一批金属材料，其抗弯强度服从正态分布，现从这批金属材料中抽取 11 个测试件，测得它们的抗弯强度为（单位：kg）：

42.5　42.7　43.0　42.3　43.4　44.5　44.0　43.8　44.1　43.9　43.7

试求抗弯强度标准差 σ 的置信度为 0.90 的置信区间。

解：金属抗弯强度 $X\sim N(\mu,\sigma^2)$**，由样本算得** $\bar{x}=43.445 \quad s=0.722$，

对于 $1-\alpha=0.90,\ \alpha=0.1,\ \chi^2_{\alpha/2}(n-1)=\chi^2_{0.05}(10)=18.307,\ \chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.95}(10)=3.94$

所以σ 的置信区间为 ：

$\left[\sqrt{\dfrac{n-1}{\chi^2_{0.05}(10)}}s,\sqrt{\dfrac{n-1}{\chi^2_{0.95}(10)}}s\right]=\left[\sqrt{\dfrac{10}{18.307}}\times0.722,\sqrt{\dfrac{10}{3.94}}\times0.722\right]$ **=[0.53，1.15]。**

3. 某厂利用两条自动化流水线灌装番茄酱，分别在两条流水线上抽取样本：$X_1,X_2,\cdots,X_{12}$ 及 $Y_1,Y_2,\cdots,Y_{17}$，算出 $\overline{X}=10.6(g),\overline{Y}=9.5(g),S_1^2=2.4,S_2^2=4.7$，假设这两条流水线上灌装的番茄酱的重量都服从正态分布，且相互独立，其均值分别为 μ_1,μ_2，（1）设两总体方差 $\sigma_1^2=\sigma_2^2$，求 $\mu_1-\mu_2$ 置信水平为95 %的置信区间；（2）求 σ_1^2/σ_2^2 的置信水平为95 %的置信区间。

解：　总体$X \sim N(\mu_1,\sigma_1^2)$，$Y \sim N(\mu_2,\sigma_2^2)$

（1）$\sigma_1^2=\sigma_2^2$未知，$\mu_1-\mu_2$的置信度为 0.95 的置信区间为

$$\left[\ \bar{x}-\bar{y}-t_{\alpha/2}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ \bar{x}-\bar{y}+t_{\alpha/2}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\ \right]$$

对于$1-\alpha=0.95$　　$\alpha=0.05$　　查表$t_{\alpha/2}(n_1+n_2-2)=t_{0.025}(27)=2.0518$

由样本算得$s_w=\sqrt{\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}=\sqrt{\dfrac{11\times2.4+16\times4.7}{27}}=1.94$，　而 $\bar{x}=10.6,\ \bar{y}=9.5$

故$\mu_1-\mu_2$的置信度为 0.95 的置信区间为

$[10.6-9.5-2.0518\times1.94\sqrt{\frac{1}{12}+\frac{1}{17}},10.6-9.5+2.0518\times1.94\sqrt{\frac{1}{12}+\frac{1}{17}}]$ = **[-0.401，2.601]**。

(2) μ_1，μ_2 未知

$$F=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$$

所以σ_1^2/σ_2^2的置信度为$1-\alpha$的置信区间为$\left[\ \dfrac{S_1^2/S_2^2}{F_{\alpha/2}(n_1-1,n_2-1)},\dfrac{S_1^2/S_2^2}{F_{1-\alpha/2}(n_1-1,n_2-1)}\ \right]$

对于$1-\alpha=0.95$　　$\alpha=0.05$

$$\text{查表}F_{0.025}(11,16)=2.94,F_{0.975}(11,16)=\frac{1}{F_{0.025}(16,11)}=\frac{1}{3.28}$$

又$s_1^2=2.4,\quad s_2^2=4.7$

故可得σ_1^2/σ_2^2的 0.95 的置信区间为：$\left[\ \dfrac{1}{2.94}\times\dfrac{2.4}{4.7},\ 3.28\times\dfrac{2.4}{4.7}\ \right]$=**[0.128，1.283]**。

习题 6-3 非正态总体均值的置信区间

1*.假定每次试验时，事件 A 发生的概率 p 未知。若在 60 次独立试验中，A 发生 15 次。求概率 p 的置信度为 0.95 的置信区间。

解：在大样本情形下，两点分布的样本均值近似服从正态分布，

即有：$\dfrac{\bar{x}-p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0,1)$

因此，查分位点 $u_{\frac{\alpha}{2}}$ 得概率 p 的置信区间的近似表达式为：

$\left[\ \bar{X}\text{-}u_{\alpha/2}\dfrac{\sqrt{\bar{X}(1-\bar{X})}}{\sqrt{n}}, \bar{X}+u_{\alpha/2}\dfrac{\sqrt{\bar{X}(1-\bar{X})}}{\sqrt{n}}\ \right]$，代入 $n=60, \bar{x}=0.25$,$u_{0.025}=1.96$

得到：[0.1404, 0.3596]

习题 6-4 单侧置信限

1*. 从汽车轮胎厂生产的某种轮胎中抽取 10 个样品进行磨损试验， 直至轮胎磨损到破坏为止，测得它们的行驶路程（Km）如下：

41250 41010 42650 38970 40200 42550 43500 40400 41870 39800

设汽车轮胎行驶路程服从正态分布 $N(\mu,\sigma^2)$，求：

（1） μ 的置信度为 95%的单侧置信下限；

（2） σ^2 的置信度为 95%的单侧置信上限。

解：汽车轮胎行驶路程 $X \sim N(\mu,\sigma^2)$。$\bar{x}=41220 \quad s=1424.84$

（1）方差 σ^2 未知，对于 $1-\alpha=0.95 \qquad \alpha=0.05 \qquad$ 查表 $t_{\alpha}(n-1)=t_{0.05}(9)=1.833$

所以参数 μ 的置信度为 0.95 的单侧置信下限为

$$\hat{\mu}_l=\bar{x}-t_{\alpha}(n-1)\frac{s}{\sqrt{n}}=41220-1.833\times\frac{1424.84}{\sqrt{10}}=40394.1 \quad 。$$

（2） μ 未知，对于 $1-\alpha=0.95 \qquad \alpha=0.05 \qquad$ 查表 $\chi^2_{1-\alpha}=\chi^2_{0.95}(9)=3.325$

所以参数 σ^2 的置信度为 0.95 的单侧置信上限为

$$\hat{\sigma}_u^2=\frac{(n-1)s^2}{\chi^2_{1-\alpha}(n-1)}=\frac{9}{3.325}\times1424.84^2=2344.18^2 \text{ 。}$$

习题 7-1 假设检验的基本概念

1. 填空题

（1）设显著性水平为α，当原假设H_0正确时，由于样本的随机性，作出了“拒绝接受假设”的决策，因而犯了错误，称为犯了<u>**第一类**</u>错误，犯该错误的概率<u>$\le\alpha$</u>。

（2）假设检验的统计思想是概率很小的事件在一次试验中可以认为基本上不会发生，该原理称为<u>小概率事件实际不发生原理</u>。

（3）假设检验的步骤为_**（1）**<u>**作出假设**</u>；**（2）**<u>**构造检验统计量及其分布**</u>；**（3）**<u>**确定拒接域**</u>；**（4）**<u>**作出拒接或接受原假设的判断**</u>。

在假设检验中，用α和β分别表示犯第一类错误和第二类错误的概率，则当样本容量一定时，下列结论正确的是（ **B** ）

(A) α减少β也减少　　(B) α与β其中一个减少时另一个往往会增大

(C) α增大β也增大　　(D) A 和 C 同时成立

习题 7-2 正态总体参数的假设检验

1. 填空题

（1）设$X_1,\ldots,X_n$为来自总体$X\sim N(\mu,\sigma^2)$的随机样本，且σ^2已知，要检验假设$H_0:\mu=\mu_0(\mu_0\text{为已知常数})$时，选用的统计量为<u>$U=\dfrac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}$</u>，当$H_0$成立时，该统计量服从<u>**N(0,1)**</u>分布。

（2*）某种导线，要求其电阻的标准差不得超过 0.005(欧姆)。今在生产的一批导线中取样品 9 根，测得 s=0.007(欧姆)，设总体为正态分布。在显著性水平$\alpha=0.05$下检验这批导线的标准差是否显著地偏大，应该选取<u>右单侧</u>检验方式，拒绝区域形式为<u>$\chi_2^2\ge\chi_\alpha^2(n-1)$</u>。

2. 选择题

设总体$X\sim N(\mu,\sigma^2)$，对数学期望μ进行假设检验，如果在显著水平$\alpha=0.05$下接受了$H_0:\mu=\mu_0(\mu_0\text{为已知常数})$，那么在显著水平$\alpha=0.01$下（ **A** ）

(A) 必接受H_0　　(B) 必拒绝H_0

(C) 可能接受也可能拒绝H_0　　(D) 不接受也不拒绝H_0

3. 已知某炼铁厂铁水含碳量服从正态分布 $N(4.550, 0.108^2)$，现观测了九炉铁水，其平均含碳量为 4.484，如果方差没有变化，可否认为现在生产的铁水平均含碳量仍为 4.550(α＝0.05)?

解：设 X 表示铁水含碳量，且 $X \sim N(\mu, 0.108^2)$

假设 $H_0: \mu = 4.550$； $H_1: \mu \neq 4.550$

当 H_0 成立时，统计量 $U = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

由样本算得 $\overline{x} = 4.484$，**所以** $|u| = \left|\dfrac{\overline{x} - 4.550}{0.108/\sqrt{9}}\right| = 1.833$

查表的 $u_{\alpha/2} = u_{0.025} = 1.96$， **因** $|u| = 1.833 < 1.96 = u_{0.025}$。

故接受 H_0，即不能否认现在生产的铁水平均含碳量仍为 4.550。

4*. 过去某工厂向 A 公司订购原材料，自订货日开始至交货日止，平均为 49.1 日，现改为向 B 公司订购原料，随机抽取向 B 公司订的 8 次货，交货天数为：46 38 40 39 52 35 48 44，问 B 公司交货天数是否较 A 公司为短(α＝0.05)?

解：检验的假设 $H_0: \mu \geq 49.1$； $H_1: \mu < 49.1$

使用统计量 $T = \dfrac{\overline{X} - 49.1}{S/\sqrt{n}} \sim t(n-1)$，

此处 $\alpha = 0.05$，**自由度** $n-1=7$，**查 t 分布临界值表得** $t_{0.01}(7) = 1.895$，

故 H_0 在检验水平α＝0.05 的拒接域为 $t < -1.895$

由样本值算得 $\overline{x} = 42.75,\ s^2 = 32.7832$， $s = 5.7257$。

$$t = \frac{42.75 - 49.1}{5.7257/\sqrt{8}} = -3.137 < -1.895 = t_{\alpha}(n-1),$$

所以应拒接 H_0，即可以认为 B 公司交货日期显著比 A 公司要短。

5. 用一台自动包装机包装葡萄糖，假定在正常情况下，糖的净重服从正态分布.根据长期资料表明，标准差为 15 克.现从某一班的产品中随机取出 9 袋，测得重量为：497 506 518 511 524 510 488 515 512. 问包装机标准差有无显著变化？(α＝0.05)

解 糖的袋重 $X \sim N(\mu, \sigma^2)$， μ **未知。**

假设： $H_0: \sigma = 15;\ H_1: \sigma \neq 15$

当假设H_0成立时，统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$

对$\alpha=0.05$，查χ^2分布上侧分位数表得

$$\lambda_1=\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0..975}(8)=2.18\text{，}\lambda_2=\chi^2_{\alpha/2}(n-1)=\chi^2_{0.025}(8)=17.535$$

由样本值得$\bar{x}=509$，$(n-1)s^2=118.75\times8=950$，$\chi_0^2=\dfrac{(n-1)s^2}{\sigma_0^2}=\dfrac{950}{15^2}=4.2$。

由于$2.18<\chi_0^2<17.535$。

故应接受H_0，即不能认为标准差有显著变化。

6*. 现测定某种溶液中的水份，它的 10 个测定值给出 s=0.037%，设测定值总体为正态分布，σ^2 为总体方差。试在水平$\alpha=0.05$下检验假设 H_0：$\sigma\geq$0.04%；H_1：σ<0.04%。

解 由题意要检验的假设是

$$H_0:\sigma\geq0.04\%;H_1:\sigma<0.04\%.$$

因为μ未知，已知$\sigma_0=0.04,n=10$，计算样本均方差得$s=0.037$

由此得统计量的观测值$\chi_2^2=\dfrac{(n-1)s^2}{\sigma_0^2}=\dfrac{9\times0.037^2}{0.04^2}=7.700625$

查卡方附表 **4** 得$\chi^2_{1-\alpha}(n-1)=\chi^2_{0.95}(9)=3.3251$

因为$\chi_2^2\geq\chi^2_{0.95}(9)$ ，未落入拒绝域，所以保留原假设H_0。

7. 设某次考试的考生成绩服从正态分布，从中随机抽取 36 名考生的成绩，算得平均成绩为 66.5 分，标准差为 15 分，问在显著性水平$\alpha=0.05$下，是否可以认为这次考生的平均成绩为 70 分。

解：检验假设$H_0:\mu=70;H_1:\mu\neq70$

计算样本均值及样本标准差得 $\bar{x}=66.5,s=15$，$n=36$

因为σ^2未知, 由式（**7.2.3**）算得统计量T的观测值

$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}=\frac{66.5-70}{15/\sqrt{36}}=-1.4$$

查附表 **5** 得

$$t_{\alpha/2}(n-1)=t_{0.025}(35)=2.0301$$

因为$|t|<t_{0.025}(35)$, 所以接受原假设H_0，即可以认为平均成绩为 70 分.

8.填空题：

（1） 设 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 为来自总体 $N(\mu_1,\sigma_1^2)$ 的样本均值，$\overline{Y}=\frac{1}{m}\sum_{i=1}^{m}Y_i$ 为来自总体 $N(\mu_2,\sigma_2^2)$ 的样本均值，两总体相互独立。如果 σ_1和σ_2 已知，要求检验的假设为 $H_0:\mu_1=\mu_2$ 时，所选用的统计量为 __ $U=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$ __ ；

当 H_0 成立时，该统计量服从 __N（0,1)__ 分布。如果 σ_1和σ_2 未知，但 $\sigma_1=\sigma_2$，要求检验的假设同样为 $H_0:\mu_1=\mu_2$ 时，选用的统计量为 _$T=\dfrac{\overline{X}-\overline{Y}}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$，_$S_w=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$___ ；

当 H_0 成立时，该统计量服从 __$t_\alpha(n_1+n_2-2)$__ 分布。

（2）设两个正态总体 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$ 相互独立，用它们各自的一个样本 $X_1,\ldots,X_{n_1}$；$Y_1,\ldots,Y_{n_2}$ 来检验假设：$H_0:\sigma_1^2=\sigma_2^2$，在 H_0 成立的情况下，统计量 $F=\dfrac{S_1^2}{S_2^2}$ 服从__$F(n_1-1,n_2-1)$ 分布。

9. 选择题

（1） 设正态总体 $X\sim N(\mu_1,\sigma^2)$，$Y\sim N(\mu_2,\sigma^2)$ (σ^2未知)，X、Y 相互独立，分别抽取样本 $X_1,\ldots,X_n$；$Y_1,\ldots,Y_n$ ，检验假设：$H_0:\mu_1=\mu_2$，应选用的统计量为（ **C** ）

(A) $\dfrac{\overline{X}-\overline{Y}}{\sqrt{S_1^2+S_2^2}}\sqrt{n-1}$ (B) $\dfrac{\overline{X}-\overline{Y}}{\sqrt{S_1^2+S_2^2}}\dfrac{1}{\sqrt{n-1}}$ (C) $\dfrac{\overline{X}-\overline{Y}}{\sqrt{S_1^2+S_2^2}}\sqrt{n}$ (D) $\dfrac{\overline{X}-\overline{Y}}{\sqrt{S_1^2+S_2^2}}\dfrac{1}{\sqrt{n}}$

(2)设 $X_1,\ldots,X_{n_1}$ 是来自正态总体 $X\sim N(\mu_1,\sigma_1^2)$ 的样本，$Y_1,\ldots,Y_{n_2}$ 是来自正态总体 $Y\sim N(\mu_2,\sigma_2^2)$，$X$、$Y$ 相互独立，则当（ **B** ）时有

$$\frac{\overline{X}-\overline{Y}}{\sqrt{\sum_{i=1}^{n_1}(X_i-\overline{X})^2+\sum_{j=1}^{n_2}(Y_j-\overline{Y})^2}}\sqrt{\frac{n_1n_2(n_1+n_2-2)}{n_1+n_2}}\sim t(n_1+n_2-2)$$

(A) $\mu_1=\mu_2$ (B) $\mu_1=\mu_2,\sigma_1^2=\sigma_2^2$

(C) $\mu_1=\mu_2,\sigma_1^2$和σ_2^2已知 (D) μ_1和μ_2已知,$\sigma_1^2=\sigma_2^2$

10. 设用甲、乙两种方法生产同一种药品，其成品得率的方差分别为$\sigma_1^2 = 0.46, \sigma_2^2 = 0.37$.现测得甲方法生产的药品得率的 25 个数据，得 $\overline{x} = 3.81$；乙方法生产的药品得率的 30 个数据，得 $\overline{y} = 3.56$(单位：g/L) .设药品得率服从正态分布.问甲、乙两种方法的药品平均得率是否有显著差异？$(\alpha = 0.05)$

解：甲、乙两种方法生产药品得率分别为$X \sim N(\mu_1, \sigma_1^2)$**及**$Y \sim N(\mu_2, \sigma_2^2)$**。**

由题意，需要检验的假设为$H_0: \mu_1 = \mu_2$; $H_1: \mu_1 \neq \mu_2$

当H_0**成立时，统计量**$U = \dfrac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$

对$\alpha = 0.05$**，查得** $u_{\alpha/2} = u_{0.025} = 1.96$

由样本值算得统计量的观察值$u = \dfrac{3.81 - 3.56}{\sqrt{\dfrac{0.46}{25} + \dfrac{0.37}{30}}} = 1.426$

由于$|u| = 1.426 < 1.96 = u_{0.025}$**，所以应接受**$H_0$**，可以认为甲、乙两种方法的药品平均得率没有显著差异。**

11. 下表分别给出两个文学家马克 •吐温的 8 篇小品文以及斯诺特格拉斯的 10 篇小品文中由 3 个字母组成的词的比例.

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总体，且两总体方差相等，两样本相互独立，问两个作家所写的小品文中包含由 3 个字母组成的词的比例是否有显著的差异（$\alpha = 0.05$）?

解　这是一个两总体的正态分布的检验问题，σ_1^2**及**σ_2^2**未知，这里**$\sigma_1^2 = \sigma_2^2$**.**

由题意，需要检验的假设为$H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

选取统计量$T = \dfrac{\overline{X} - \overline{Y}}{S_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$

计算得$\overline{x} = 0.232, \overline{y} = 0.2097$**，**　$s_w = \sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = 12.1 \times 10^{-3}$

统计量的观测值$t=\dfrac{0.232-0.2097}{12.1\times10^{-3}\times\sqrt{\dfrac{1}{8}+\dfrac{1}{10}}}=3.918$

$\alpha=0.05$，查表知$t_{\alpha/2}(n_1+n_2-2)=2.1199$

$|t|>t_{\alpha/2}(n_1+n_2-2)$，因而拒绝$H_0$，即有显著差异.

12. 为比较甲、乙两种安眠药的疗效，将 20 名患者分成两组，每组 10 人，如服药后延长的睡眠时间分别近似服从正态分布，其数据如表所示

	a	b	c	d	e	f	g	h	i	j
甲	1.9	0.8	1.1	0.1	−0.1	4.4	5.5	1.6	4.6	3.4
乙	0.7	−1.6	−0.2	−1.2	−0.1	3.4	3.7	0.8	0	2.0

问在显著水平$\alpha=0.05$下，两种安眠药的疗效有无显著差异?

解： 此题需先检验方差再检验期望，设甲组服药延长的睡眠时间$X\sim N(\mu_1,\sigma_1^2)$，乙组服药后延长的睡眠时间$Y\sim N(\mu_2,\sigma_2^2)$。

待检验的假设是：(1) H_0 ：$\sigma_1^2=\sigma_2^2$，(2) H_1 ：$\mu_1=\mu_2$

(1) 假设 $H_0:\sigma_1^2=\sigma_2^2$；$H_1:\sigma_1^2\neq\sigma_2^2$

选取统计量 $F=\dfrac{S_1^2}{S_2^2}$， 当H_0成立时，$F\sim F(n_1-1,n_2-1)$

由$n_1=n_2=10$，计算$\bar{x}=2.33,\ \bar{y}=0.75,\ s_1^2=4.009,\ s_2^2=3.20$， $s_w^2=3.605,\ s_w=1.9$。

从而 $F_0=\dfrac{4.009}{3.2}=\mathbf{1.25}$

对$\alpha=0.05$，查 F 临界值表，得$F_{0.025}(9,9)=4.03$，

由于 $\dfrac{1}{4.03}<1.25<4.03$. 故接受$H_0$。

(2) $H_0:\mu_1=\mu_2$；$H_1:\mu_1\neq\mu_2$。

由于$\sigma_1^2=\sigma_2^2$，选取统计量 $T=\dfrac{\overline{X}-\overline{Y}}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$。 当$H_0$成立时，$T\sim t(n_1+n_2-2)$。

查$\alpha=0.05$，自由度为 18 的 t 分布临界值表，得 $t_{0.05}(18)=2.101$.

而统计量T的观察值$T_0=\dfrac{2.33-0.75}{1.899\sqrt{\dfrac{1}{10}+\dfrac{1}{10}}}=1.86$。

由于$|T_0|=1.86<2.101=t_{0.05}(18)$，故接受$H_0$，即不能认为两种安眠药有显著差异。

13*.环保公司利用新旧两种方法合成新型纳米吸附剂，通过考察对染料红的吸附率模拟吸附效果，对于浓度为 10mg/L 染料红，两种方法的吸附结果（吸附百分率，%）检测如下：

新方法	90	86	85	84	85	87	91	88	86
旧方法	83	81	82	88	86	78	79	89	80

若测定数据均服从正态分布，

（1）那么能否认为新方法的吸附效果显著优于旧方法（显著性水平$\alpha = 0.05$）？

（2）若新方法显著优于旧方法，那么新方法的吸附效果较之旧方法提高了多少？（$1-\alpha = 0.95$）。

解：（1）首先检验方差是否相等：因$F = \dfrac{S_1^2}{S_2^2}$=0. 359431>0. 290858$= F_{0.95}(8,8)$，

故可以认为方差相等。

再检验新方法的吸附效果是否显著优于旧方法 ：

$H_0: \mu_{新} \le \mu_{旧}$， $H_1: \mu_{新} \rangle\ \mu_{旧}$

选用 $t = \dfrac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{(m-1)S_x^2 + (n-1)S_Y^2}\sqrt{\dfrac{m+n}{mn(m+n-2)}}} \sim t(m+n-2)$

计算：$t = 2.6049 > 1.7459 = t_{0.05}(16)$ **拒绝原假设。**

（2）由：$t = \dfrac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{(m-1)S_x^2 + (n-1)S_Y^2}\sqrt{\dfrac{m+n}{mn(m+n-2)}}} \sim t(m+n-2)$

查表 $t_{0.05}(16) = 1.7459$

得$\mu_1 - \mu_2$**的 95%的单侧置信区间是：**$(1.35864, +\infty)$

或者求双侧置信区间： 查双侧分位数：$t_{0.025}(16) = 2.1199$

得$\mu_1 - \mu_2$**的 95%的双侧置信区间是：**$(0.7447, 7.2553)$

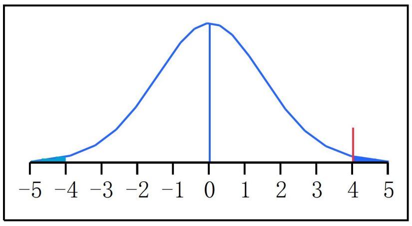
	新方法	旧方法
平均	86.88888889	82.88888889
方差	5.611111111	15.61111111
观测值	9	9
合并方差	10.61111111	
假设平均差	0	
df	16	
t Stat	2.604868579	
P(T<=t) 单尾	0.009575227	
t 单尾临界	1.745883669	
P(T<=t) 双尾	0.019150453	
t 双尾临界	2.119905285	

t 检验

新方法-旧方法

假定方差相等

差值	4.00000	t 比率	2.604869
差值标准误差	1.53559	自由度	16
差值置信上限	7.25530	概率>|t|	0.0192*
差值置信下限	0.74470	概率>t	0.0096*
置信	0.95	概率<t	0.9904



习题 8-1　单因素试验方差分析

1．填空题

（1） 在使用单因素试验方差分析方法时，需要满足的基本条件是 <u>正态性</u> 、 <u>方差齐性</u> 、<u>独立性</u> 。

（2）在单因素试验方差分析中，因素 A 有 k 个水平 $A_1, A_2, \cdots A_k$，检验的原假设 H_0 : <u>$H_0: \mu_1 = \mu_2 = \cdots = \mu_k$</u>，备择假设 H_0 : <u>$H_1: \mu_1, \mu_2, \cdots, \mu_k$ 不全相等</u>。

2．一批由同样原料织成的布，用五种不同的染整工艺处理，然后进行缩水试验，设每种工艺处理 4 块布样，测得缩水率的结果如下表：

布样号	缩　水　率				
	A_1	A_2	A_3	A_4	A_5
1	4.3	6.1	6.5	9.3	9.5
2	7.8	7.3	8.3	8.7	8.8
3	3.2	4.2	8.6	7.2	11.4
4	6.5	4.1	8.2	10.1	7.8

问不同的工艺对布的缩水率是否有显著的影响？$(\alpha = 0.01)$

解：这是单因素试验的方差分析，通过计算总平方和、因子平方和，可以求出误差平方和，将计算结果填入下面方差分析表：

方差分析

差异源	SS	df	MS	F	P-value	F crit
组间	55.537	4	13.88425	6.059022	0.004157	4.89321
组内	34.3725	15	2.2915			
总计	89.9095	19				

由检验统计量：

$$F_A = \frac{SS_A/(a-1)}{SS_e/(n-1)(a-1)} = \frac{MS_A}{MS_e} \sim F(4,15)$$，计算得 F=6.05922>4.89321

故在 0.01 的水平上可以认为不同的工艺对布的缩水率有显著的影响。

习题 8-2　双因素试验方差分析

1．填空题

（1）在双因素无重复试验方差分析中，因素 A 有 k 个水平 $A_1, A_2, \cdots A_k$，因素 B 有 l 个水平 $B_1, B_2, \cdots B_l$，每对水平组合下只做一次试验，试验数据为 x_{ij} $(i = 1,2,\cdots k;\ j = 1,2,\cdots l)$，则总平方

和分解式：$\underline{SS_T=\sum_{i=1}^{k}\sum_{j=1}^{l}(x_{ij}-\bar{x}_{..})^2 = SS_A+SS_B+SS_e}$，总自由度分解式：

$\underline{kl-1=(k-1)+(l-1)+(k-1)(l-1)}$。

2．在某种化工产品的生产过程中，选择 3 种浓度：A_1 =2%，A_2 =4%，A_3 =6%；4 种不同的温度：B_1=10℃，B_2 =24℃，B_3=38℃，B_4 =52℃．每种浓度和温度的组合都重复试验 2 次，得到产品的收率如下：

浓度	温度			
	B_1	B_2	B_3	B_4
A_1	10 14	11 11	9 13	10 12
A_2	7 9	8 10	7 11	6 10
A_3	5 11	13 14	12 13	10 14

试利用 Excel 的数据分析工具库分析不同的浓度、不同的温度以及不同浓度与不同温度的交互作用对产品的收率是否有影响？（$\alpha=0.05$）

解：这是双因素有重复试验的方差分析，通过计算总平方和、因子平方和，可以求出误差平方和，将计算结果填入下面方差分析表：

方差分析

差异源	SS	df	MS	F	P-value	F crit
样本	44.33333333	2	22.16666667	4.092308	0.044153	3.885294
列	11.5	3	3.833333333	0.707692	0.565693	3.490295
交互	27	6	4.5	0.830769	0.568369	2.99612
内部	65	12	5.416666667			
总计	147.8333333	23				

由检验统计量：$F_A=\dfrac{SS_A/(a-1)}{SS_e/ab(r-1)}=\dfrac{MS_A}{MS_e}\sim F(a-1,ab(r-1))$

计算得 $F_A=4.092308>3.885294$，故在 0.05 的水平上认为不同的浓度有显著影响。

$$F_B=\frac{SS_B/(b-1)}{SS_e/ab(r-1)}=\frac{MS_B}{MS_e}\sim F(b-1,ab(r-1))$$

计算得 $F_B=0.707692<3.490295$，故在 0.05 的水平上不认为不同的温度有显著影响。

$$F_{A\times B}=\frac{SS_{A\times B}/(a-1)(b-1)}{SS_e/ab(r-1)}=\frac{MS_{A\times B}}{MS_e}\sim F((a-1)(b-1),ab(r-1))$$

计算得 $F_{A\times B}=0.830769<2.99612$，故在 0.05 的水平上不认为不同的浓度与温度有交叉影响。

3. 某市空气质量检测站在该市不同时间点、不同地点对空气中的颗粒状物进行监测，数据如下：

因素 A（时间）	因素 B（地点）				
	B1	B2	B3	B4	B5
A1	76	67	81	56	51
A2	82	69	96	59	70
A3	68	59	67	54	42
A4	63	56	64	58	37

欲解决的问题是“不同时间、不同地点空气中的颗粒状物是否有区别？”，请回答以下问题：

（1） 写出检验的原假设和备择假设；

$$H_{01}:\alpha_1=\alpha_2=\cdots=\alpha_a=0\text{；}H_{02}:\beta_1=\beta_2=\cdots=\beta_b=0\ .$$

（2） 写出检验统计量及拒绝区域的形式；

$$F_A=\frac{SS_A/(a-1)}{SS_e/(a-1)(b-1)}=\frac{MS_A}{MS_e}\sim F(a-1,(a-1)(b-1))$$

$$F_B=\frac{SS_B/(b-1)}{SS_e/(a-1)(b-1)}=\frac{MS_b}{MS_e}\sim F(b-1,(a-1)(b-1))$$

（3） 完成 Excel 输出的表格：

差异源	SS	df	MS	F
因素 A	1182.95	3	394.3167	10.72241
因素 B	1947.5	4	486.875	13.23929
误差	441.3	12	36.775	
总计	3571.75	19		

（4） 根据输出以及拒绝域，给出你的判断。

拒绝域形式：对于给定的显著性水平 α，若 $F_A\ge F_\alpha(a-1,(a-1)(b-1))$，则拒绝 H_{01}，否则接受 H_{01}. 若 $F_B\ge F_\alpha(b-1,(a-1)(b-1))$，则拒绝 H_{02}，否则接受 H_{02}

可以查表：F(3,12)=3.4903, F(4.12)=3.2592 ,所以因素 A 和因素 B 均显著。

习题 8-3 一元线性回归

1．填空题:

(1) 一元线性回归方程 $y=\beta_0+\beta_1x$ 参数的最小二乘估计 $\hat\beta_1=$ $\underline{\hat\beta_1=\dfrac{l_{xy}}{l_{xx}}}$，$\hat\beta_0=$ $\underline{\hat\beta_0=\overline{y}-\hat\beta_1\overline{x}}$。

(2) 由样本 $(x_1,y_1),\cdots(x_n,y_n)$ 建立一元线性经验回归方程 $\hat y=\hat\beta_0+\hat\beta_1x$，当 $x=x_0$ 时，Y 的值 Y_0 的预测区间是 $\underline{[\hat y_0-\delta,\hat y_0+\delta]}$，其中 $\underline{\delta=t_{\alpha/2}(n-2)\hat\sigma\sqrt{1+\dfrac{1}{n}+\dfrac{(x_0-\overline{x})^2}{l_{xx}}}}$。

2．在钢线碳含量$x(\%)$对于电阻y(20 ℃时，微欧）效应的研究中，得到以下的数据

x	0.01	0.30	0.40	0.55	0.70	0.80	0.95
y	15	18	19	21	22．6	23.8	26

设对于给定的x，y为正态变量，且方差与x无关.

（1）求线性回归方程$\hat{y}=\hat{\beta}_0+\hat{\beta}_1 x$；

（2）检验回归方程的显著性；（$\alpha=0.05$）

（3）求β_1的置信区间（置信水平为 0.95）；

（4）求y在$x=0.50$处的置信水平为 0.95 的预测区间.

解：（1）由一元线性回归的最小二乘估计$\hat{\beta}_1=\dfrac{l_{xy}}{l_{xx}}$，$\hat{\beta}_0=\overline{y}-\hat{\beta}_1\overline{x}$，可以求出

$\hat{\beta}_1=11.63219, \hat{\beta}_0=14.60637$，故线性回归方程为：$\hat{y}=14.60637+11.63219x$

（2）回归的方差分析表为：

方差分析

	df	SS	MS	F	Significance F
回归分析	1	83.72851325	83.72851	1369.131	2.72E-07
残差	5	0.305772463	0.061154		
总计	6	84.03428571			

由于 F 检验统计量的值为 F=1369.131，检验 p 值为 2.72E-07，所以回归方程显著。

（3）β_1的置信区间为(10.82408, 12.4403)

(4) y在$x=0.50$处的点估计是：$\hat{y}_0=14.60637+11.63219*0.5=20.422465$

预测区间为，$[\hat{y}_0-\delta, \hat{y}_0+\delta]$，其中 $\delta=t_{\alpha/2}(n-2)\hat{\sigma}\sqrt{1+\dfrac{1}{n}+\dfrac{(x_0-\overline{x})^2}{l_{xx}}}$，计算得：

（19.74263616，21.10229）。

3．从某林场随机抽取了 100 株云杉，测量其胸径和树高数据，并按龄级分组得平均胸径D和平均树高H列表如下：

平均胸径D(cm)	15	20	25	30	35	40	45	50
平均树高H(m)	13.5	17.1	20.0	22.1	24.0	25.6	27.0	28.3

（1） 试用合适的函数建立H关于D的回归方程；

（2） 在平均胸径D为 28 cm 时，平均树高H的预测值.

解：（1）可以使用指数模型拟合：$\hat{H}=2.7799D^{0.6015}$；

（2）$\hat{H}(28)=20.6298(\mathrm{m})$