前三章练习题

1. **某人独立重复向一目标射击，每次命中目标的概率都是p（1>p>0）,则此人第四次射击恰好是第二次命中目标的概率是**

说明前三次命中一次和第四次命中

$$C_3^1 p(1-p)^2 p = 3p^2(1-p)^2$$

某人向同一目标独立重复射击，设他每次射击命中目标的概率为 $p\,(0<p<1)$，则此人第 5 次射击命中且恰好是第 3 次命中的概率是____________________。

$$C_4^2 p^3(1-p)^2$$

1.设 A，B 为两个随机事件，$P(A)=0.7$，$P(A-B)=0.3$，则 $P\left(\overline{AB}\right)=$ ________

2. 设随机变量 X 的概率分布为 $P\{X=k\}=\frac{c}{n}\cdot k$，$k=1, 2, \cdots, n$，则常数 $c=$ ________

3. 设随机变量 X 服从正态分布 $N(1, 1)$，已知 $z_{0.025}=1.96$，则常数 $P\{X\le 2.96\}=$ ________

A 0.95 B 0.975 C 0.025 D 0.005

4. 设 $X\sim N(0,\sigma_1^2)$，$Y\sim N(-1,\sigma_2^2)$，且 X 与 Y 相互独立，则 $P(X-Y>1)=$ ________。

(A) $\frac{1}{8}$ (B) $\frac{1}{4}$ (C) $\frac{1}{2}$ (D) 与 σ_1^2，σ_2^2 有关

5. 设随机变量 X,Y 相互独立，$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，则 $X-Y+1$ 服从的分布为________

A $N(\mu_1-\mu_2+1, \sigma_1^2+\sigma_2^2)$ B $N(\mu_1-\mu_2+1, \sigma_1^2-\sigma_2^2)$

C $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ D $N(\mu_1-\mu_2+1, \sigma_1^2-\sigma_2^2+1)$

1. 0.6 2. $\frac{2}{n+1}$ 3. B 4. C 5. A

6. 设随机变量 X 服从指数分布，其概率密度函数是 $f(x)$. 则下列能成为 X 的概率密度函数的是____________

A $g_1(x)=\begin{cases} f(x), & x>0 \\ f(-|x|), & x\le 0 \end{cases}$　　B $g_2(x)=\begin{cases} f(x), & x>0 \\ f(-x), & x\le 0 \end{cases}$

C $g_3(x)=\begin{cases} 0.5f(x), & x>0 \\ 0.5f(-|x|), & x\le 0 \end{cases}$　　D $g_4(x)=\begin{cases} 0.5f(x), & x>0 \\ 0.5f(-x), & x\le 0 \end{cases}$

$$f(x)=\begin{cases} \lambda e^{-\lambda x}, & x>0, \quad (\lambda>0) \\ 0, & \text{其它} \end{cases}$$

7. 设随机变量 X 的概率密度函数是 $\varphi(x)$，且有 $\varphi(-x)=\varphi(x)$，$F(x)$ 是 X 的分布函数，则对任意的实数 a，有______________。

（A）$F(-a)=\frac{1}{2}-\int_a^{+\infty}\varphi(x)dx$　　（B）$F(-a)=\frac{1}{2}-\int_{-\infty}^{a}\varphi(x)dx$

（C）$F(-a)=1-\int_0^{a}\varphi(x)dx$　　（D）$F(-a)=1-\int_{-\infty}^{a}\varphi(x)dx$

6.A　7.D

8. 设二维随机变量(X,Y)服从二维正态分布$N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,0)$，则以下说法错误的是________。

（A） X与Y相互独立　　（B） X与Y不相关

（C） $D(XY)=D(X)D(Y)$　　（D） $E(XY)=E(X)E(Y)$

9. 设随机变量X与Y相互独立，其分布函数分别为$F_1(x)$, $F_2(y)$，则$\max(X,Y)$的分布函数

$F_{\max}(z)=P(\max(X,Y)\le z)=$______________。

(A) $F_1(x)F_2(y)$　　(B) $F_1(z)F_2(z)$　　(C) $\max\{F_1(z),F_2(z)\}$　　(D) $1-\max\{F_1(z),F_2(z)\}$

10. 设随机变量X与Y相互独立，且都服从区间$[0,1]$上的均匀分布，则$P(X^2+Y^2\le 1)=$______。

(A) $\frac{1}{4}$　　(B) $\frac{1}{2}$　　(C) $\frac{\pi}{8}$　　(D) $\frac{\pi}{4}$

11. 设随机变量X的分布函数$F(x)=\begin{cases}0, & x<0\\ 1/2, & 0\le x<1\\ 1-e^{-x}, & x\ge 1\end{cases}$，则$P\{X=1\}=$____________。

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

8.C 9.B 10.D 11.C

12. 设随机变量$X \sim N(\mu,\sigma^2)(\sigma>0)$,

且二次方程$y^2+4y+X=0$无实根的概率为$\frac{1}{2}$,则$\mu=$ 4

解:

二次方程无实根$\Leftrightarrow 16-4X<0 \Leftrightarrow X>4$

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

$$=1-P\left(\frac{X-\mu}{\sigma}\le\frac{4-\mu}{\sigma}\right)=1-\Phi\left(\frac{4-\mu}{\sigma}\right)$$

$$\Phi\left(\frac{4-\mu}{\sigma}\right)=1-\frac{1}{2}=\frac{1}{2}$$

$$\frac{4-\mu}{\sigma}=0 \rightarrow \mu=4$$

13. 设X_1和X_2是任意两个相互独立的连续型随机变量，它们的概率密度分别为$f_1(x)$和$f_2(x)$，分布函数分别为$F_1(x)$和$F_2(x)$，则有：

(A) $f_1(x)+f_2(x)$必为某一随机变量的概率密度。

(B) $f_1(x)\cdot f_2(x)$必为某一随机变量的概率密度。

(C) $F_1(x)+F_2(x)$必为某一随机变量的分布函数。

(D) $F_1(x)\cdot F_2(x)$必为某一随机变量的分布函数。

解： $X=\max\{X_1,X_2\}$

$$F_X(x)=P(X\le x)=P(X_1\le x,X_2\le x)$$

$$=P(X_1\le x)\cdot P(X_2\le x)=F_1(x)\cdot F_2(x)$$

例1.设两箱内装有同种零件，第一箱装50件，有10件一等品。

第二箱装30件，有18件一等品。

先从两箱中任挑一箱，在从此箱中先后不放回地任取两个零件，

求：(1) 先取出的零件是一等品的概率p

(2) 在先取出的零件是一等品的条件下，后取出的仍是一等品的条件概率q

解：

(1) 设 $A_i=\{\text{第}i\text{次取出一等品}\},i=1,2$

$B_i=\{\text{挑到第}i\text{箱}\},i=1,2$ $\quad P(B_1)=P(B_2)=1/2$

$$p=P(A_1)=P(B_1)P(A_1|B_1)+P(B_2)P(A_1|B_2)$$

$$=\frac{1}{2}\times\frac{1}{5}+\frac{1}{2}\times\frac{3}{5}=\frac{2}{5}$$

$$P(A_1|B_1)=\frac{10}{50}=\frac{1}{5} \qquad P(A_1|B_2)=\frac{18}{30}=\frac{3}{5}$$

例1.设两箱内装有同种零件，第一箱装50件，有10件一等品。

第二箱装30件，有18件一等品。

先从两箱中任挑一箱，在从此箱中先后不放回地任取两个零件，

求：(1) 先取出的零件是一等品的概率p

(2) 在先取出的零件是一等品的条件下，后取出的仍是一等品的条件概率q

解：

$$(2)\ q = P(A_2|A_1) = \frac{P(A_1A_2)}{P(A_1)} \qquad P(B_1) = P(B_2) = 1/2$$

$$= \frac{1}{P(A_1)}[P(B_1)P(A_1A_2|B_1) + P(B_2)P(A_1A_2|B_2)]$$

$$= \frac{1}{2/5}[\frac{1}{2}\times\frac{10\times 9}{50\times 49} + \frac{1}{2}\times\frac{18\times 17}{30\times 29}] \approx 0.48557$$

例2、 设随机变量X在区间[2,5]上服从均匀分布，现在对X进行三次独立观测，试求至少有两次观测值大于3的概率。

解：X的密度函数为 $f(x)=\begin{cases} 1/3, & 2\le x\le 5 \\ 0, & 其他 \end{cases}$

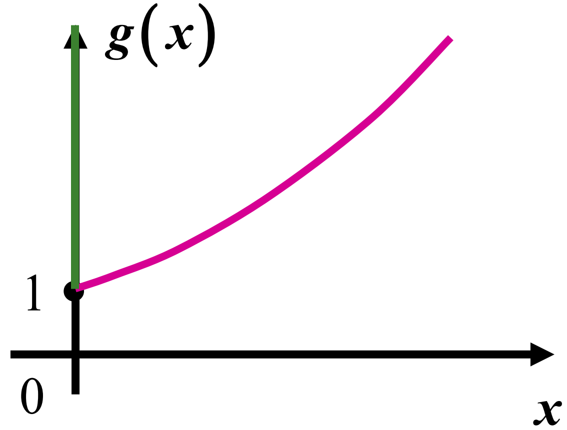
$$p=P\{X>3\}=\int_3^{+\infty} f(x)dx=\int_3^5 \frac{1}{3}dx+\int_5^{+\infty} 0dx=\frac{2}{3}$$

$Y=$ “三次独立观测中观测值大于3的次数”，

$$Y\sim \mathrm{B}(3,p)$$

$$P\{Y\ge 2\}=C_3^2(\frac{2}{3})^2\frac{1}{3}+C_3^3(\frac{2}{3})^3=\frac{4}{9}+\frac{8}{27}=\frac{20}{27}$$

例3-a 设 $f_X(x)=\begin{cases} e^{-x}, & x>0, \\ 0, & x\le 0. \end{cases}$ $Y=e^X$ 求 $f_Y(y)$



解 $Y=e^X$ $x>0$时，$e^x>1$，

易知 $\forall y\le 1,\quad f_Y(y)=0$

$$\forall y>1, F_Y(y)=P\{Y\le y\}=P\{e^X\le y\}$$

$$=P\{X\le \ln y\}=F_X(\ln y)$$

求导得 $f_Y(y)=f_X(\ln y)\cdot\dfrac{1}{y}=\dfrac{1}{y^2}$

$$\therefore f_Y(y)=\begin{cases} \dfrac{1}{y^2}, & y>1, \\ 0, & y\le 1. \end{cases}$$

例3-b、设随机变量X在区间（1，2）上服从均匀分布，试求随机变量 $Y=e^{2X}$ 的概率密度 $f(y)$

解　X的概率密度 $f(x)=\begin{cases}1, & 1<x<2\\ 0, & \text{其他}\end{cases}$　$Y=e^{2X}\in(e^2,e^4)$

分布函数法：

$y\le e^2$ 和 $y\ge e^4$ 时，$f_Y(y)=0$

$e^2<y<e^4$ 时 $F(y)=P\{Y\le y\}=P\{e^{2X}\le y\}=P\{X\le\frac{1}{2}\ln y\}$

$$=\int_{-\infty}^{\frac{1}{2}\ln y}f(x)dx=\int_{1}^{\frac{1}{2}\ln y}1dx=\frac{1}{2}\ln y-1$$ 求导数

$$f(y)=F'(y)=\begin{cases}\frac{1}{2y}, & e^2<y<e^4\\ 0, & \text{其他}\end{cases}$$

例3-c. 设随机变量X的概率密度函数为$f(x)=ax^2, 0\le x\le 2$

其中a是常数。随机变量$Y=\sqrt{X}$

问：（1）试确定常数a

（2）求Y的概率密度函数。

1)由于$\int_{-\infty}^{+\infty} f(x)dx=1$

所以$\int_0^2 ax^2dx=1$ 解得：$a=\dfrac{3}{8}$

2) 首先 $0\le Y\le\sqrt{2}$

对 $y\in\left[0,\sqrt{2}\right]$，有 $F_Y(y)=P\{Y\le y\}$

$$=P\left\{\sqrt{X}\le y\right\}=P\left\{X\le y^2\right\}=\int_0^{y^2}\frac{3}{8}x^2dx=\frac{y^6}{8}$$

所以，概率密度函数

$$f_Y(y)=\left(F_Y(y)\right)_y'=\frac{3}{4}y^5, 0\le y\le\sqrt{2}$$

例4-a 设(X, Y)的密度为 $f(x,y)=\begin{cases} Ax, & 0<y<x<1, \\ 0, & 其他 \end{cases}$ G

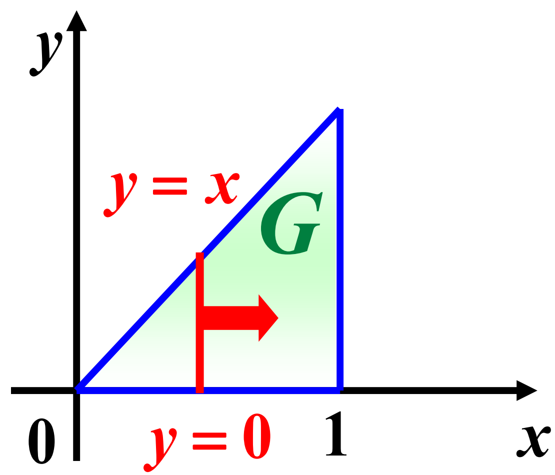
求： $i)\ A$

$ii)\ P\{X+Y<1\}$. $iii)\ f_X(x), f_Y(y)$

解 $i)\ \because \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy$

$$=\iint_G Axdxdy = \int_0^1 dx\int_0^x Axdy$$

$$=\frac{A}{3}=1 \longrightarrow A=3$$



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

例4-a $f(x,y)=\begin{cases}3x, & 0<y<x<1,\\ 0, & \text{其他}\end{cases}$ $\quad P\{X+Y<1\}$

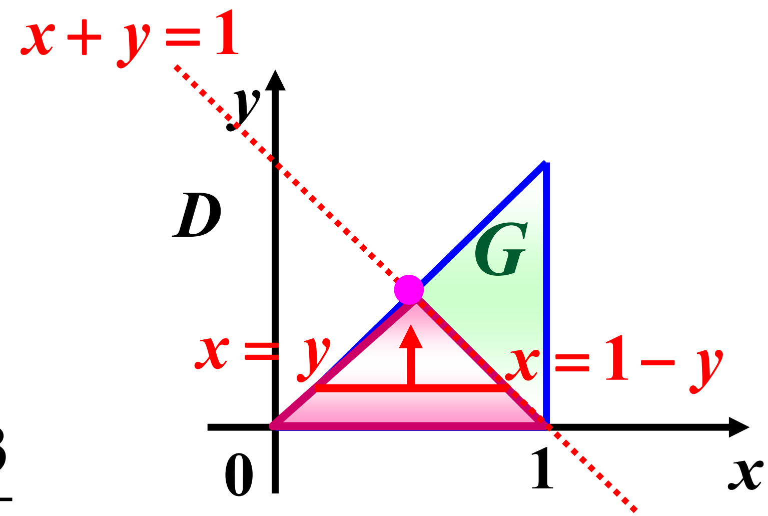
解 $ii)\ \{X+Y<1\}=\{(X,Y)\in D\},\ D=\{(x,y)\mid x+y<1\}$

$P\{X+Y<1\}$

$$=\iint_{GD} f(x,y)\,dxdy$$

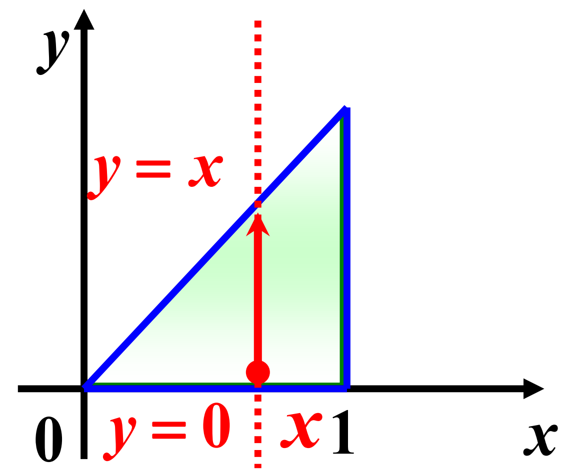
$$=\iint_{GD} 3x\,dxdy$$

$$=\int_0^{1/2} dy\int_y^{1-y} 3x\,dx \ =\frac{3}{8}$$



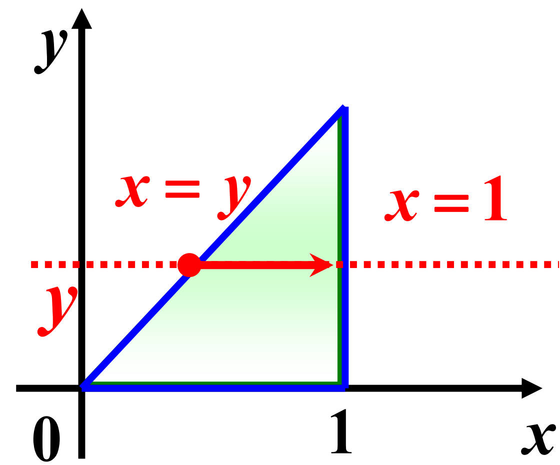
例4-a $f(x,y)=\begin{cases}3x, & 0<y<x<1,\\ 0, & 其他\end{cases}$

求：$iii)\ f_X(x), f_Y(y)$



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

$$=\begin{cases}\int_0^x 3xdy = 3x^2, & 0<x<1,\\ 0, & 其他\end{cases}$$



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

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

例5. 设二维随机变量 (X,Y) 的联合概率密度为 $f(x,y)=\begin{cases} xe^{-(x+y)}, & x>0, y>0 \\ 0, & \text{其他} \end{cases}$，

(1) 求边缘概率密度 $f_X(x), f_Y(y)$；　(2) 判断 X 与 Y 是否相互独立，并说明原因；

(3) 计算 $E(2X+Y)$；　(4) 求条件概率密度 $f_{X|Y}(x|y)$。

解 (1) $x\le 0$ 时，$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=0$；

$x>0$ 时，$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^{+\infty} xe^{-(x+y)}dy=xe^{-x}$

因此，$f_X(x)=\begin{cases} xe^{-x}, & x>0 \\ 0, & \text{其他} \end{cases}$

$y\le 0$ 时，$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=0$；

$y>0$ 时，$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^{+\infty} xe^{-(x+y)}dx=e^{-y}$

因此，$f_Y(y)=\begin{cases} e^{-y}, & y>0 \\ 0, & \text{其他} \end{cases}$

(2) 由于 $f(x,y)=f_X(x)f_Y(y)$，故而 X 与 Y 相互独立。

例5. 设二维随机变量(X,Y)的联合概率密度为$f(x,y)=\begin{cases} xe^{-(x+y)}, & x>0,y>0 \\ 0, & 其他 \end{cases}$，

(1) 求边缘概率密度$f_X(x), f_Y(y)$； (2) 判断X与Y是否相互独立，并说明原因；

(3) 计算$E(2X+Y)$； (4) 求条件概率密度$f_{X|Y}(x|y)$。

解 (3) 由（1）可得，$E(X)=\int_{-\infty}^{+\infty} xf_X(x)dx=\int_0^{+\infty} x^2e^{-x}dx=2$

$$E(Y)=\int_{-\infty}^{+\infty} yf_Y(y)dy=\int_0^{+\infty} ye^{-y}dy=1$$

$$E(2X+Y)=2E(X)+E(Y)=5$$

或

$$E(2X+Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(2x+y)f(x,y)dxdy$$

$$=\int_0^{+\infty}\int_0^{+\infty}2x^2e^{-x}e^{-y}dxdy+\int_0^{+\infty}\int_0^{+\infty}xe^{-x}\cdot ye^{-y}dxdy=4+1=5$$

例5. 设二维随机变量(X,Y)的联合概率密度为$f(x,y)=\begin{cases} xe^{-(x+y)}, & x>0,y>0 \\ 0, & \text{其他} \end{cases}$，

(1) 求边缘概率密度$f_X(x), f_Y(y)$； (2) 判断X与Y是否相互独立，并说明原因；

(3) 计算$E(2X+Y)$； (4) 求条件概率密度$f_{X|Y}(x|y)$。

解 (4) $y>0$时，$f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}=\begin{cases} xe^{-x}, & x>0 \\ 0, & \text{其他} \end{cases}$

类似例6 设二维随机变量的概率密度函数为

$$f(x,y)=\begin{cases} x^2+\dfrac{xy}{3}, & 0\le x\le 1,\ 0\le y\le 2 \\ 0\ , & \text{其 它} \end{cases}$$

试求：（1）边缘概率密度函数$f_X(x)$和$f_Y(y)$，并判断X和Y的独立性；

（2）条件概率密度函数$f_{X|Y}(x|1)$； （3）概率$P\{X+Y>1\}$.

例6 设二维随机变量的概率密度函数为

$$f(x,y)=\begin{cases} x^2+\dfrac{xy}{3}, 0\le x\le 1,\ 0\le y\le 2 \\ 0\ , \qquad 其\ 它 \end{cases}$$

试求：（1）边缘概率密度函数 $f_X(x)$ 和 $f_Y(y)$，并判断 X 和 Y 的独立性；

（2）条件概率密度函数 $f_{X|Y}(x|1)$；　　（3）概率 $P\{X+Y>1\}$.

（1）$0\le x\le 1$ 时，$f_X(x)=\int_0^2\left(x^2+\dfrac{xy}{3}\right)dy=2x^2+\dfrac{2x}{3}$，其余为零；

$0\le y\le 2$ 时，$f_Y(y)=\int_0^1\left(x^2+\dfrac{xy}{3}\right)dx=\dfrac{1}{3}+\dfrac{y}{6}$，其余为零；

显然 $f(x,y)\ne f_X(x)f_Y(y)$，所以 X 和 Y 的不独立；

（2）$f_{X|Y}(x|1)=\dfrac{f(x,1)}{f_Y(1)}=\begin{cases} 2x^2+\dfrac{2x}{3},\ 0\le x\le 1 \\ 0\qquad ,\ 其\ 它 \end{cases}$

（3）$P\{X+Y>1\}=\int_0^1 dx\int_{1-x}^2\left(x^2+\dfrac{xy}{3}\right)dy=\int_0^1\left(\dfrac{5}{6}x^3+\dfrac{4x^2}{3}+\dfrac{x}{2}\right)dx=\dfrac{65}{72}$

卷积公式法　　X, Y 相互独立，连续型

例7-a. 设 X,Y 相互独立具有相同的概率密度：

$$f(x)=\begin{cases}1 & 0\le x\le 1\\ 0 & \text{其　它}\end{cases}\qquad f(y)=\begin{cases}1 & 0\le y\le 1\\ 0 & \text{其　它}\end{cases}$$

求：$Z=X+Y$ 的概率密度

法一：由题意知

因 X,Y 相互独立，由卷积公式：

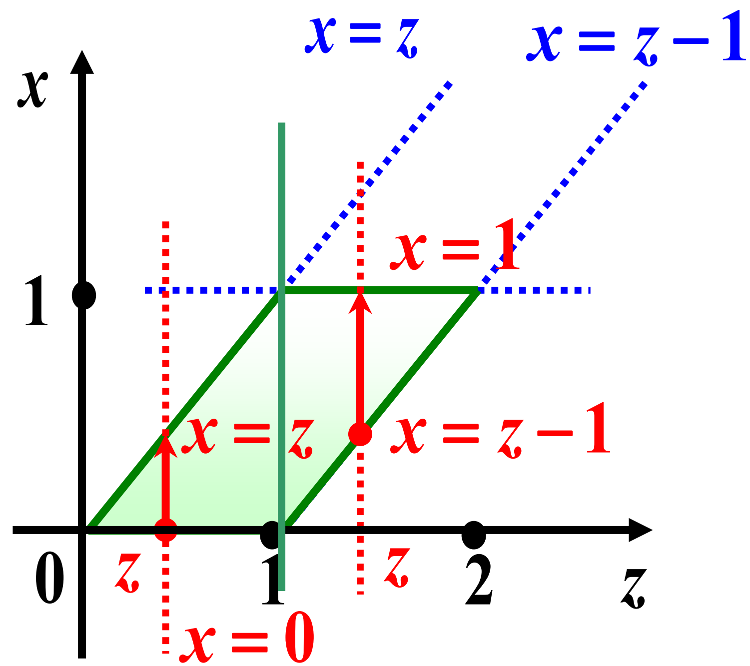
$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\,dx=\int_{-\infty}^{+\infty} f(x)f(z-x)dx$$

确定积分限，找出使被积函数不为 0 的区域：$0\le z\le 2$

$$\begin{cases}0\le x\le 1\\ 0\le z-x\le 1\end{cases}\qquad\begin{cases}0\le x\le 1\\ -1\le x-z\le 0\end{cases}$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x)\,dx$$

区域 $\begin{cases} 0 \le x \le 1 \\ -1 \le x - z \le 0 \end{cases}$ 如图示:



于是得:

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x) f(z-x)dx$$

$$= \begin{cases} \int_0^z dx = z & 0 \le z < 1 \\ \int_{z-1}^1 dx = 2 - z & 1 \le z < 2 \\ 0 & \text{其它} \end{cases}$$

$$f(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{其　它} \end{cases}$$

例7-b 设 X,Y 相互独立，概率密度如下：　求：$f_{X+Y}(z)$

$$f_X(x)=\begin{cases} e^{-x}, & x>0, \\ 0, & x\le 0. \end{cases} \qquad f_Y(y)=\begin{cases} 2e^{-2y}, & y>0, \\ 0, & y\le 0. \end{cases}$$

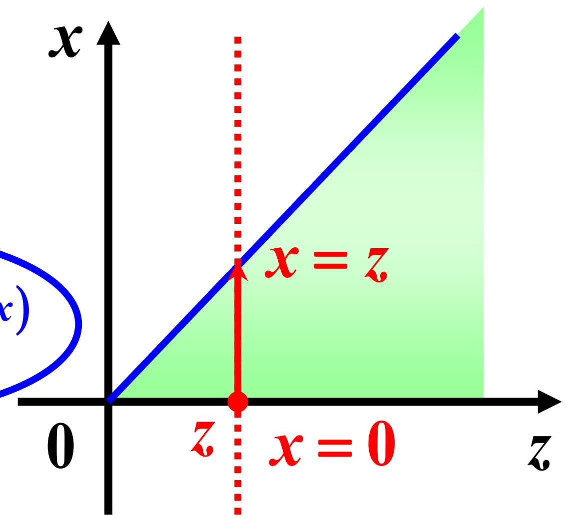
解　因 X,Y 相互独立，由卷积公式：

$$f_{X+Y}(z)=\int_{-\infty}^{+\infty} f_X(x)\,f_Y(z-x)\,dx$$

先找使被积函数不为 0 的区域：$z>0$,

$$\begin{cases} x>0 \\ z-x>0 \end{cases} \quad 即 \quad \begin{cases} x>0 \\ x<z \end{cases}$$

$$e^{-x}\cdot 2e^{-2(z-x)}$$



$z>0$:

$$f_{X+Y}(z)=\int_0^z 2e^{-2z+x}\,dx = 2e^{-2z}\left(e^{z}-1\right)$$

$$f_{X+Y}(z)=\begin{cases} 2e^{-2z}\left(e^{z}-1\right), & z>0, \\ 0, & z\le 0. \end{cases}$$

1、卷积公式法 X, Y 相互独立，连续型

2、分布函数法 连续型

例7−a 设 X,Y 相互独立具有相同的概率密度：

$$f(x)=\begin{cases}1 & 0\le x\le 1\\ 0 & \text{其　它}\end{cases}$$ 求：$Z=X+Y$ 的概率密度

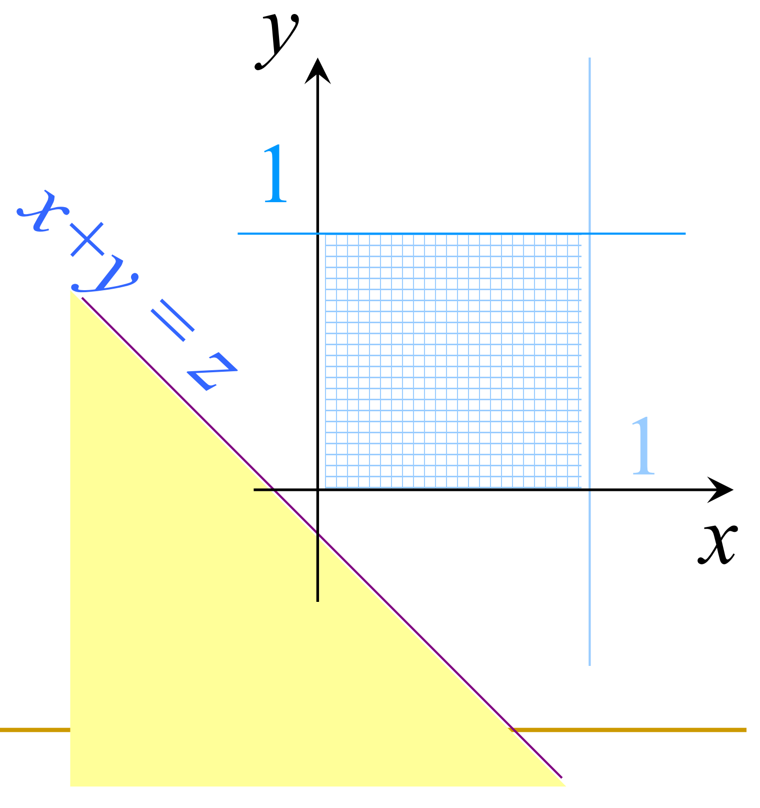
解法二 从分布函数出发

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

当 $z<0,\ F_Z(z)=0$

密度函数非零取值区间为：

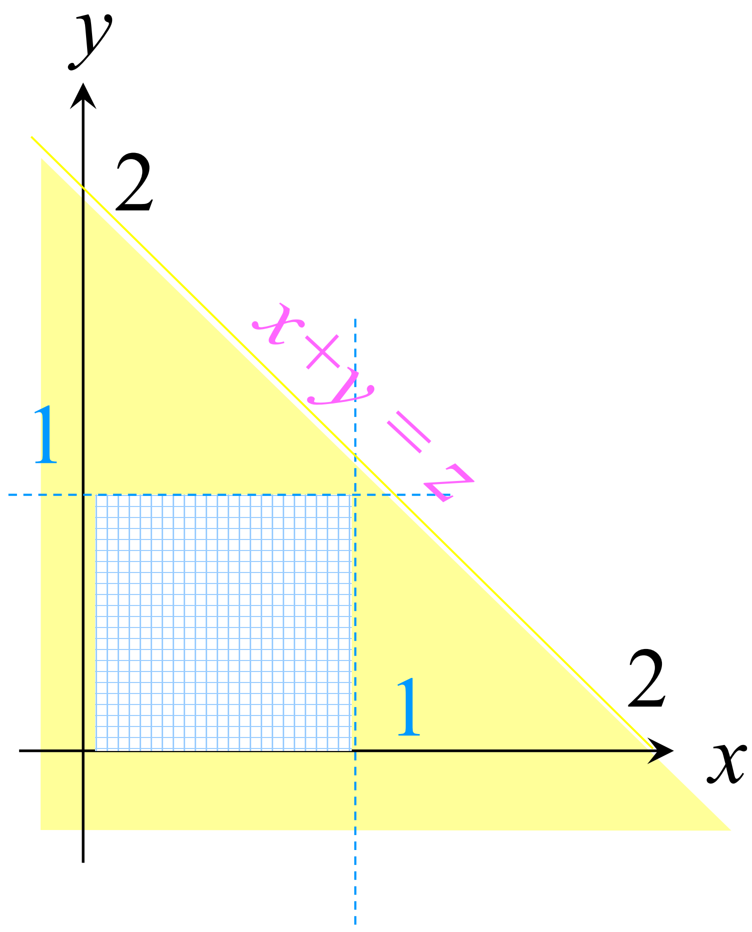
$0<z<2,$



当$2 \le z$时，

$F_Z(z) = 1 \quad f_Z(z) = 0$

$$f_Z(z) = \begin{cases} z, & 0 \le z < 1 \\ 2 - z, & 1 \le z < 2 \\ 0, & \text{其他} \end{cases}$$



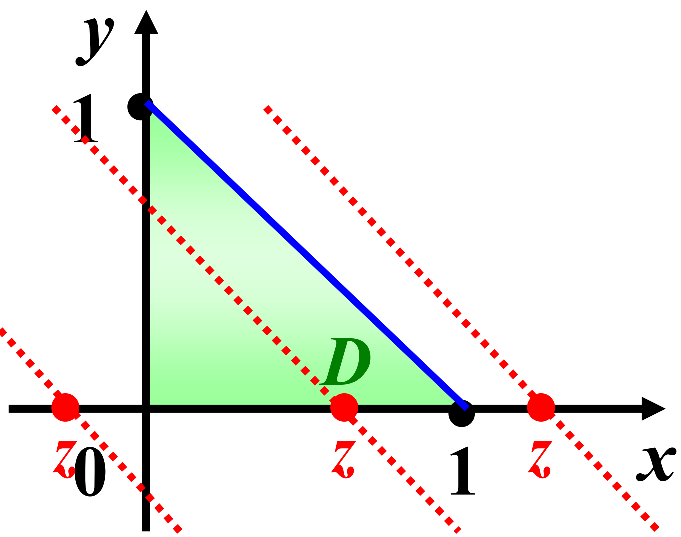
分布函数法 连续型 不知X, Y是否相互独立

不能用卷积公式！

例8 记 $D: x+y<1, x>0, y>0.$

(X, Y) 服从区域上D 的均匀分布 $\longrightarrow f_{X+Y}(z)$

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



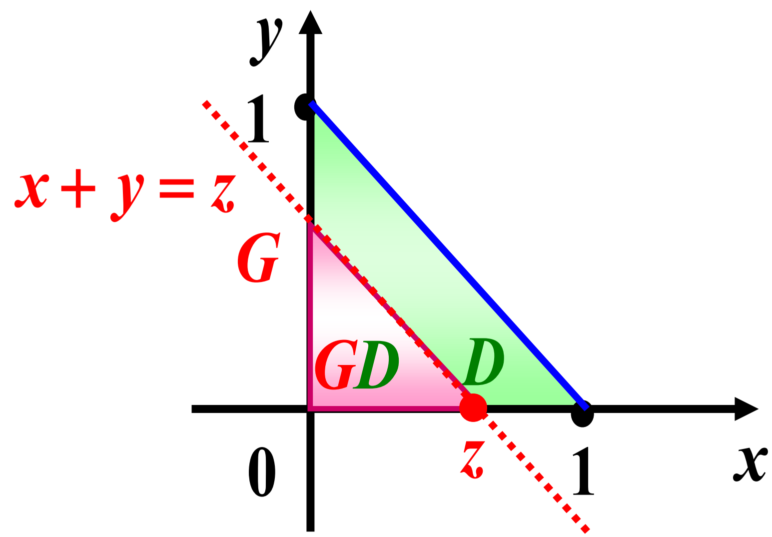
$$F_{X+Y}(z) = P\{X+Y\le z\} = \iint\limits_{x+y\le z} f(x,y)\,dxdy = \iint\limits_{\substack{x+y\le z\\ D}} 2\,dxdy$$

例8 记 $D: x+y<1, x>0, y>0.$

(X, Y) 服从区域上D 的均匀分布 $\longrightarrow f_{X+Y}(z)$

解 $F_{X+Y}(z)=P\{X+Y\le z\}$

$$=\iint\limits_{\substack{x+y\le z\\ D}} 2dxdy$$



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

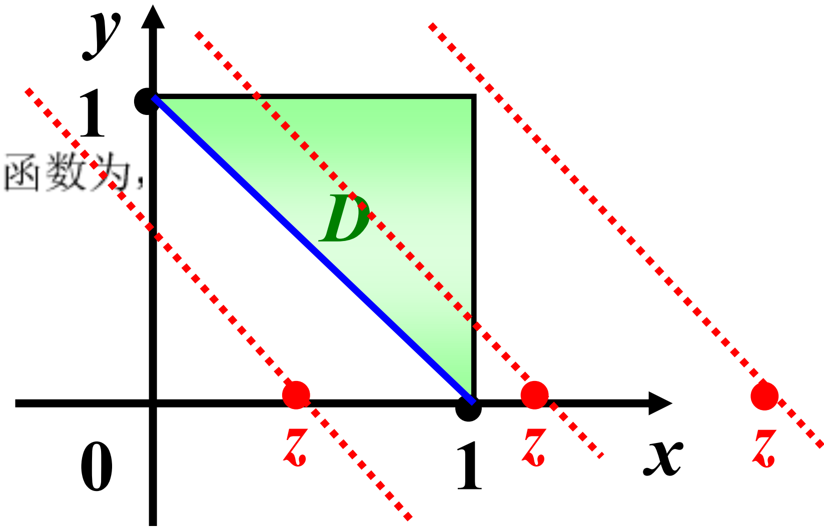
求导

$\Rightarrow$ $$f_{X+Y}(z)=\begin{cases} 2z, & 0<z<1,\\ 0, & \text{其他}\end{cases}$$

例9. 设二维随机变量(X, Y)在以$(0, 1), (1, 0), (1, 1)$为顶点的三角形区域D上服从均匀分布，试求：

（1）$Z = X + Y$ 的概率密度函数；

（2）$D(X+Y)$.



解（1）D的面积为0.5，则(X, Y)的联合概率密度函数为，

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

当$z<1$时，$F_Z(z)=0$

当$1\le z<2$时，$F_Z(z)=P(X+Y\le z)=\iint\limits_{1\le x+y\le 2} 2dxdy = 4z-z^2-3$

当$z\ge 2$时，$F_Z(z)=1$，

例9. 设二维随机变量(X,Y)在以$(0,1),(1,0),(1,1)$为顶点的三角形区域D上服从均匀分布，试求：

（1）$Z=X+Y$ 的概率密度函数；

（2）$D(X+Y)$.

解（1）D的面积为0.5，则(X,Y)的联合概率密度函数为，

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

$$F_Z(z)=\begin{cases}0, & z<1\\ 4z-z^2-3, & 1\le z<2\\ 1, & z\ge 2\end{cases}$$

当$z<1$时，$F_Z(z)=0$

当$1\le z<2$时，$F_Z(z)=P(X+Y\le z)=\iint\limits_{1\le x+y\le 2} 2dxdy=4z-z^2-3$

当$z\ge 2$时，$F_Z(z)=1$，

$$f_Z(z)=F_Z'(z)=\begin{cases}4-2z, & 1\le z<2\\ 0, & \text{其它}\end{cases}$$

类似例10

设随机变量 X 服从区间[0,2]上的均匀分布，Y 服从区间[0,1]上的均匀分布，两者相互独立，求 $Z=X+Y$ 的概率密度函数。

解法一 $f_X(x)=\begin{cases}\frac{1}{2}, & 0<x<2\\ 0, & 其他\end{cases}$，$f_Y(y)=\begin{cases}1, & 0<y<1\\ 0, & 其他\end{cases}$，

由于 X,Y 相互独立，利用卷积公式 $f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx$

$$=\begin{cases}0, & z\le 0 \text{ 或 } z\ge 3\\ \int_0^z \frac{1}{2}dx=\frac{z}{2}, & 0<z<1\\ \int_{z-1}^z \frac{1}{2}dx=\frac{1}{2}, & 1\le z<2\\ \int_{z-1}^2 \frac{1}{2}dx=\frac{3-z}{2}, & 2\le z<3\end{cases}$$

例10

设随机变量 X 服从区间[0,2]上的均匀分布，Y 服从区间[0,1]上的均匀分布，两者相互独立，求 $Z=X+Y$ 的概率密度函数。

解法二 分布函数法

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

$$F_Z(z)=P(Z\le z)=P(X+Y\le z)$$

当 $z\le 0$ 时，$F_Z(z)=0$；　　当 $z\ge 3$ 时，$F_Z(z)=1$

当 $0<z<1$ 时，$F_Z(z)=P(X+Y\le z)=\int_0^z dx\int_0^{z-x}\frac{1}{2}dy=\frac{1}{4}z^2$

当 $1\le z<2$ 时，$F_Z(z)=P(X+Y\le z)=\int_0^{z-1}dx\int_0^1\frac{1}{2}dy+\int_{z-1}^{z}dx\int_0^{z-x}\frac{1}{2}dy=\frac{z}{2}-\frac{1}{4}$

当 $2\le z<3$ 时，$F_Z(z)=P(X+Y\le z)=\int_0^{z-1}dx\int_0^1\frac{1}{2}dy+\int_{z-1}^{2}dx\int_0^{z-x}\frac{1}{2}dy=\frac{3}{2}z-\frac{z^2}{4}-\frac{5}{4}$

$$f_Z(z)=\begin{cases}0, & z\le 0 \text{ 或 } z\ge 3\\ \dfrac{z}{2}, & 0<z<1\\ \dfrac{1}{2}, & 1\le z<2\\ \dfrac{3-z}{2}, & 2\le z<3\end{cases}$$

第四章 随机变量的数字特征

随机变量的数字特征

	离散型随机变量	连续型随机变量
X	$E(X)=\sum_{k=1}^{\infty}x_k p_k$ ★	$E(X)=\int_{-\infty}^{+\infty}xf(x)dx$ ★
$D(X)=E[X-E(X)]^2$	$D(X)=\sum_{k=1}^{\infty}(x_k-E(X))^2 p_k$	$D(X)=\int_{-\infty}^{+\infty}(x-E(X))^2 f(x)dx$
$Y=g(X)$ g连续	$E(Y)=E[g(X)]=\sum_{k=1}^{\infty}g(x_k)p_k$	$E(Y)=E[g(X)]=\int_{-\infty}^{+\infty}g(x)f(x)dx$ ★
$Z=g(X,Y)$ g连续	$E(Z)=E[g(X,Y)]=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}g(x_i,y_j)p_{ij}$	$E(Z)=E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy$ ★

随机变量的数字特征

★ $E(X)$性质	$E(c)=c \quad E(cX)=cE(X) \qquad E(X+Y)=E(X)+E(Y)$ X,Y独立 $\quad E(XY)=E(X)E(Y)$
★ $D(X)$性质	$D(X)=E(X^2)-[E(X)]^2 \quad D(c)=0 \quad D(cX)=c^2D(X)$ X,Y独立,$D(X+Y)=D(X)+D(Y) \quad D(X)=0 \Leftrightarrow P(X=c)=1$
★ 协方差	$Cov(X,Y)=E\{[X-E(X)][(Y-E(Y)]\}$ $=E(XY)-E(X)E(Y)$ $D(X+Y)=D(X)+D(Y)+2\mathrm{cov}(X,Y) \overset{独立}{=} D(X)+D(Y)$
相关系数 X,Y相互独立 $\not\Uparrow \; \Downarrow$ $\rho_{X,Y}=0$	$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}$ (1). $\lvert\rho_{XY}\rvert\le 1$ (2). $\lvert\rho_{XY}\rvert=1 \Leftrightarrow$ 存在常数 a,b,使得:$P(Y=aX+b)=1$

切比雪夫不等式

$$P\{\lvert X-\mu\rvert\ge\varepsilon\}\le\frac{\sigma^2}{\varepsilon^2}$$

★几种常见分布的数学期望和方差

	概率分布		$E(X)$	$D(X)$
离散型	(0-1)分布	$X \sim B(1,p)$	p	pq
	二项分布	$X \sim B(n,p)$	np	npq
	泊松分布	$X \sim P(\lambda)$	λ	λ
连续型	均匀分布	$X \sim U(a,b)$	$(a+b)/2$	$(b-a)^2/12$
	指数分布	$X \sim Exp(\theta)$	θ	θ^2
	正态分布	$X \sim N(\mu,\sigma^2)$	μ	σ^2

1. 设$X_1,X_2,\cdots,X_n$独立同分布，且其方差为$\sigma^2>0$.

令$Y=\frac{1}{n}\sum_{i=1}^{n}X_i$ 则：(A)正确

$(A)\ \mathrm{cov}(X_1,Y)=\frac{\sigma^2}{n}$ $(B)\ \mathrm{cov}(X_1,Y)=\sigma^2$

$(C)\ D(X_1+Y)=\frac{(n+2)\sigma^2}{n}$ $(D)\ D(X_1-Y)=\frac{(n+1)\sigma^2}{n}$

解：

$$\mathrm{cov}(X_1,Y)=\mathrm{cov}(X_1,\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n}\mathrm{cov}(X_1,\sum_{i=1}^{n}X_i)$$

$$=\frac{1}{n}[\underbrace{\mathrm{cov}(X_1,X_1)}_{\sigma^2}+\underbrace{\mathrm{cov}(X_1,X_2)}_{0}+\cdots+\underbrace{\mathrm{cov}(X_1,X_n)}_{0}]=\frac{\sigma^2}{n}$$

2. 设A,B是随机事件，且$P(A)=\frac{1}{4}, P(B|A)=\frac{1}{3}, P(A|B)=\frac{1}{2}$,

令$X=\begin{cases}1, & A\text{发生}\\ 0, & A\text{不发生}\end{cases}$ $\quad Y=\begin{cases}1, & B\text{发生}\\ 0, & B\text{不发生}\end{cases}$

求：(1)(X,Y)的概率分布　(2) ρ_{XY}

解： $P(X=0,Y=0)=P(\overline{A}\,\overline{B})=P(\overline{A\cup B})$

$=1-P(A\cup B)$

$=1-[P(A)+P(B)-P(AB)]=\frac{2}{3}$

$$\frac{1}{3}=P(B|A)=\frac{P(AB)}{P(A)} \rightarrow P(AB)=\frac{1}{12},$$

$$\frac{1}{2}=P(A|B)=\frac{P(AB)}{P(B)} \rightarrow P(B)=\frac{1}{6},$$

2. 设A,B是随机事件，且$P(A)=\frac{1}{4}, P(B|A)=\frac{1}{3}, P(A|B)=\frac{1}{2}$,

$$令X=\begin{cases}1, & A发生\\ 0, & A不发生\end{cases} \qquad Y=\begin{cases}1, & B发生\\ 0, & B不发生\end{cases}$$

求：(1)(X,Y)的概率分布 $P(B)=\frac{1}{6}$, $P(AB)=\frac{1}{12}$,

$Y \backslash X$	0	1	
0	2/3	1/6	5/6
1	1/12	1/12	1/6
	3/4	1/4	1

$$P(X=0,Y=0)=P(\bar{A}\bar{B})=P(\overline{A\cup B})$$

$$=1-P(A\cup B)$$

$$=1-[P(A)+P(B)-P(AB)]=\frac{2}{3}$$

$$P(X=1,Y=0)=P(A\bar{B})$$

$$=P(A)-P(AB)=1/6$$

$$P(X=0,Y=1)=P(\bar{A}B)=P(B)-P(AB)=1/12$$

$$P(X=1,Y=1)=P(AB)=1/12$$

2. 设A,B是随机事件，且$P(A)=\dfrac{1}{4}, P(B|A)=\dfrac{1}{3}, P(A|B)=\dfrac{1}{2}$，

$$令X=\begin{cases}1, & A发生\\ 0, & A不发生\end{cases} \qquad Y=\begin{cases}1, & B发生\\ 0, & B不发生\end{cases}$$

求：(2) ρ_{XY}

Y \ X	0	1	
0	2/3	1/6	5/6
1	1/12	1/12	1/6
	3/4	1/4	1

$$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)$$
$$=1/12-1/4\cdot 1/6=1/24$$

$E(X)=1/4$, $E(Y)=1/6$,

$E(XY)=1/12$,

$E(X^2)=1/4$, $E(Y^2)=1/6$,

$D(X)=E(X^2)-E^2(X)=3/16$,

$D(Y)=5/36$,

$$\rho_{XY}=\frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}=\frac{1/24}{\sqrt{3}/4\cdot\sqrt{5}/6}=\frac{1}{\sqrt{15}}$$

3. 设随机变量X的概率密度为：$f(x)=\begin{cases}\dfrac{1}{2}\cos\dfrac{x}{2}, & 0\le x<\pi\\ 0, & \text{其它}\end{cases}$

对X独立重复地观察4次，用Y表示观察值大于$\dfrac{\pi}{3}$的次数，求：$E(Y^2)$　$E(Y^2)=D(Y)+E^2(Y)$

解： $P(X>\dfrac{\pi}{3})=\int_{\frac{\pi}{3}}^{\infty}f(x)dx=\int_{\frac{\pi}{3}}^{\pi}\dfrac{1}{2}\cos\dfrac{x}{2}dx=\dfrac{1}{2}$

Y：4次观察中$\{X>\dfrac{\pi}{3}\}$的次数，$Y\sim B(4,\dfrac{1}{2})$

$E(Y)=4\times\dfrac{1}{2}=2$　　$D(Y)=4\times\dfrac{1}{2}\times\dfrac{1}{2}=1$

$E(Y^2)=D(Y)+E^2(Y)=5$

4. 设 $X_1, X_2, \cdots, X_n$ 是来自总体 $N(\mu, \sigma^2)$ 的简单随机样本。记统计量 $T = \frac{1}{n}\sum_{i=1}^{n} X_i^2$，

则 $E(T) =$ ________________。　　　　$\mu^2 + \sigma^2$

5 设二维离散型随机变量 (X, Y) 的联合分布律为

X \ Y	0	1	2
0	1/4	0	1/4
1	0	1/3	0
2	1/12	0	1/12

(1) 求 $P(X = 2Y)$；　(2) 求 $Cov(X - Y, Y)$.

$\frac{1}{4}$　　$Cov(X-Y,Y) = Cov(X,Y) - Cov(Y,Y) = EXY - EXEY - DY = -2/3$

第五章 极限定理

第一节 大数定律

第二节 中心极限定理

大数定律及中心极限定理

定理1 (贝努利)	$X_1, X_2, \cdots, X_n, \cdots$相互独立 $\sim (0-1)$分布(参数p)	$\dfrac{n_A}{n} = \dfrac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} p$
定理2 (切比雪夫)	$X_1, X_2, \cdots, X_n, \cdots$相互独立 $E(X_k) = \mu$ $D(X_k) = \sigma^2$	$\dfrac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu$
定理3 (辛钦)	$X_1, X_2, \cdots, X_n, \cdots$相互独立 $E(X_k) = \mu$ 同分布	$\dfrac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu$
定理1 (德莫弗)	$\eta_1, \eta_2 \cdots \eta_n \cdots$ 相互独立, $\eta_n \sim B(n, p)$	$\dfrac{\eta_n - np}{\sqrt{np(1-p)}} \overset{近似}{\sim} N(0,1)$
定理2 (林德)	$X_1, X_2, \cdots, X_n, \cdots$相互独立 同分布 $E(X_k) = \mu$ $D(X_k) = \sigma^2$	$\dfrac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \overset{近似}{\sim} N(0,1)$

1 设 $X_1,\cdots,X_n$ 是来自总体 X 的容量为 n 的简单随机样本，已知 $E(X)=\mu$，则对于任意的 $\epsilon>0$，有

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i-\mu\right|<\epsilon\right)=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

2 设 n_A 是 n 次独立试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率，则对于任意的

$\varepsilon>0$，$\lim_{n\to+\infty} P\left\{\left|\frac{n_A}{n}-p\right|\geq\varepsilon\right\}=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$。

1，0

3 在次品率为$\frac{1}{6}$的一批产品中，任意抽取 300 件产品，其中次品的数量记作 X。

（1）写出 X 的分布律，数学期望 EX 及方差 DX；

（2） k 取何值时，概率 $P\{X=k\}$ 最大？

（3）利用切比雪夫不等式估计次品数在 40 到 60 之间的概率；

（4）利用中心极限定理计算次品数在 40 到 60 之间的概率。已知 $\Phi(1.55)=0.939$，$\sqrt{\frac{5}{3}}=1.29$。

【解】（1） X 服从二项分布 $B\left(300,\frac{1}{6}\right)$，

$$P\{X=k\}=C_{300}^{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{300-k}=C_{300}^{k}\frac{5^{300-k}}{6^{300}},0\le k\le 300.$$

$$EX=300\times\frac{1}{6}=50,\quad DX=300\times\frac{1}{6}\times\frac{5}{6}=\frac{125}{3}。$$

2）当 $(n+1)p$ 不为整数时，在 $P\{X=k\}$ 在 $k=[(n+1)p]=[\frac{301}{6}]=50$ 达到最大。

3 在次品率为$\frac{1}{6}$的一批产品中，任意抽取 300 件产品，其中次品的数量记作 X。

（1）写出 X 的分布律，数学期望 EX 及方差 DX；

（2）k 取何值时，概率 $P\{X=k\}$ 最大？

（3）利用切比雪夫不等式估计次品数在 40 到 60 之间的概率；

（4）利用中心极限定理计算次品数在 40 到 60 之间的概率。已知 $\Phi(1.55)=0.939$，$\sqrt{\frac{5}{3}}=1.29$。

（3）由切比雪夫不等式有

$$P\{40 \le X \le 60\} = P\{-10 \le X-50 \le 10\} = P\{|X-50| \le 10\} \ge 1-\frac{125/3}{10^2} = \frac{7}{12}$$

（4）由中心极限定理，$\dfrac{X-np}{\sqrt{np(1-p)}} = \dfrac{X-50}{\sqrt{125/3}} \sim N(0,1)$

$$P\{40 \le X \le 60\} = P\left\{\frac{40-50}{\sqrt{125/3}} \le \frac{X-50}{\sqrt{125/3}} \le \frac{60-50}{\sqrt{125/3}}\right\}$$

$$= P\left\{-1.55 \le \frac{X-50}{\sqrt{125/3}} \le 1.55\right\} \approx 2\Phi(1.55)-1 = 0.878$$

第六章 数理统计的基本概念

第一节 总体与随机样本

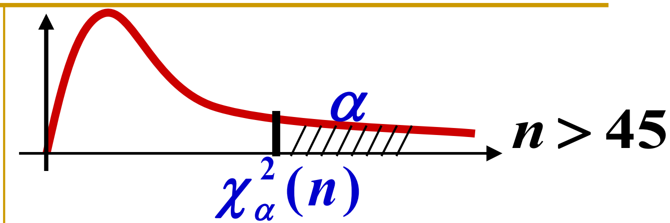
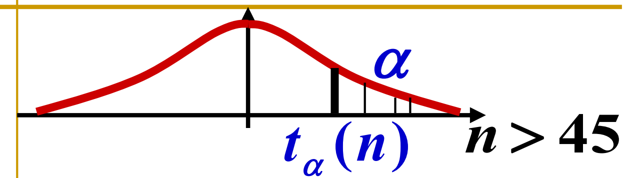
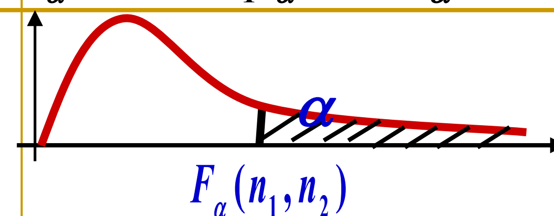
第二节 统计量及其分布

第三节 常用的重要统计量及其分布

几个常用的统计量

名称	统计量	观察值
样本均值	$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$	$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$
样本方差	$S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$	$s^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$
样本标准差	$S=\sqrt{S^2}$	$s=\sqrt{s^2}$
样本 k 阶(原点)矩	$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$	$a_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k$
样本k阶中心矩	$B_k=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^k$	$b_k=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^k$

常用统计量及抽样分布

χ^2分布	$X_i \sim N(0,1)\ i=1,2,\cdots,n$ 独立 ★$\chi^2=\sum_{i=1}^{n}X_i^2 \sim \chi^2(n)$ $E(\chi^2)=n \quad D(\chi^2)=2n$	 $\chi_\alpha^2(n)\approx 1/2(z_\alpha+\sqrt{2n-1})^2$
t分布	$X\sim N(0,1),\ Y\sim\chi^2(n)$，独立 ★$t=\dfrac{X}{\sqrt{Y/n}}\sim t(n)$	 $t_\alpha(n)=-t_{1-\alpha}(n),\ t_\alpha(n)\approx z_\alpha$
F分布	$U\sim\chi^2(n_1),V\sim\chi^2(n_2)$，独立 ★$F=\dfrac{U/n_1}{V/n_2}\sim F(n_1,n_2)$ $1/F\sim F(n_2,n_1)$	 $F_{1-\alpha}(n_1,n_2)=1/F_\alpha(n_2,n_1)$
$X\sim N(\mu,\sigma^2)$ $X_1,X_2,\cdots,X_n$	★$Th1\quad \bar{X}\sim N(\mu,\sigma^2/n)$, $(n-1)S^2/\sigma^2\sim\chi^2(n-1)$ 独立	$Th2$ ★ $\dfrac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$

第七章 参数估计

第一节 矩估计法

第二节 极大似然估计法

第三节 估计量的优良性

第四节 区间估计

第五节 正态总体均值与方差的区间估计

	总体$X \sim F(x,\theta)$, $\begin{matrix} X_1 X_2, \cdots, X_n \\ x_1 x_2, \cdots, x_n \end{matrix}$ 对θ进行估计
★ 点估计	统计量$\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n) \to \theta$估计量 1)矩估计法：求解：$\mu_i = A_i, i = 1,2,\cdots,k$ 2)极大似然估计法：求解：$L(\hat{\theta}) = \max\limits_{\theta \in H} L(\theta)$
★ 估计量的优良性	1)无偏性：$E(\hat{\theta}) = \theta$ 2)有效性：$D(\hat{\theta}_1) < D(\hat{\theta}_2)$ 3)相合性： $n \to \infty, \hat{\theta}_n \xrightarrow{p} \theta$
★区间估计	$P(\underline{\theta} < \theta < \overline{\theta}) = 1 - \alpha$ $(\underline{\theta}, \overline{\theta})$是置信度为$1-\alpha$的置信区间
	$X \sim N(\mu, \sigma^2)$, 对μ, σ^2进行区间估计 1)求μ的置信区间σ^2为已知 2)求μ的置信区间σ^2为未知 3)求σ^2的置信区间

$X \sim N(\mu,\sigma^2)$, 对μ,σ^2进行区间估计置信度$1-\alpha$

	统计量	置信区间
1)求μ的置信区间，σ^2为已知	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	★ $(\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2})$
2)求μ的置信区间，σ^2为未知	$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	★ $(\bar{X} \pm \frac{s}{\sqrt{n}} t_{\alpha/2}(n-1))$
3)求σ^2的置信区间	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	★ $(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)})$

例1 设总体 $X \sim \begin{pmatrix} 1 & 2 & 3 \\ 1-\theta^2 & \dfrac{\theta^2}{2} & \dfrac{\theta^2}{2} \end{pmatrix}$, $\theta\left(0<\theta<1\right)$未知.

现得一样本值1, 3, 2, 3, 求 θ 最大似然估计值.

解 $P\left\{\left(X_1, X_2, X_3, X_4\right)=\left(1,3,2,3\right)\right\}$

$$=P\left\{X_1=1,\ X_2=3,\ X_3=2,\ X_4=3\right\}$$

$$=P\left\{X_1=1\right\}P\left\{X_2=3\right\}P\left\{X_3=2\right\}P\left\{X_4=3\right\}$$

$$=P\left\{X=1\right\}P\left\{X=3\right\}P\left\{X=2\right\}P\left\{X=3\right\}$$

$$=\frac{1}{8}\theta^6\left(1-\theta^2\right)\triangleq L(\theta)$$

例1 设总体 $X \sim \begin{pmatrix} 1 & 2 & 3 \\ 1-\theta^2 & \dfrac{\theta^2}{2} & \dfrac{\theta^2}{2} \end{pmatrix}$, $\theta(0<\theta<1)$未知.

现得一样本值1, 3, 2, 3, 求 θ 最大似然估计值.

解 $$L(\theta)=\frac{1}{8}\theta^6\left(1-\theta^2\right),\quad 0<\theta<1$$

$$\ln L(\theta)=\ln\frac{1}{8}+6\ln\theta+\ln\left(1-\theta^2\right)$$

令 $$\frac{d}{d\theta}\ln L(\theta)=\frac{6}{\theta}-\frac{2\theta}{1-\theta^2}=0$$

解出 $\hat{\theta}=\dfrac{\sqrt{3}}{2}$ 最大似然估计值

例2 设总体 $X \sim U(a,b)$，a, b未知，$X_1, X_2, \ldots, X_n$是来自X的样本，$x_1, x_2, \ldots, x_n$是样本值。求a, b的极大似然估计量 $\hat{a}, \hat{b}$

解:
$$f(x,a,b)=\begin{cases} 1/(b-a), & a \le x \le b \\ 0, & \text{其它} \end{cases}$$

$$L(a,b)=\prod_{i=1}^{n} f(x_i,a,b)=\frac{1}{(b-a)^n},$$

当似然函数不可微或方程组无解时，则应根据定义直接寻求能使 $L(\theta_1,\theta_2,\cdots,\theta_k)$ 达到最大值的解 $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_k$ 作为极大似然估计值(量)。

例2. $X \sim U(a,b)$, 求a, b的极大似然估计量 $\hat{a}, \hat{b}$

解:
$$f(x,a,b)=\begin{cases} 1/(b-a), & a \le x \le b \\ 0, & \text{其它} \end{cases}$$

$$L(a,b)=\frac{1}{(b-a)^n},$$

用谁估计a?用谁估计b?

$$x_{(1)}=\min\{x_1,x_2,\cdots,x_n\} \qquad x_{(n)}=\max\{x_1,x_2,\cdots,x_n\}$$

a,b的极大似然估计值：$\hat{a}=x_{(1)}=\min\limits_{1\le i\le n} x_i$, $\hat{b}=x_{(n)}=\max\limits_{1\le i\le n} x_i$

a,b的极大似然估计量：$\hat{a}=\min\limits_{1\le i\le n} X_i$, $\hat{b}=\max\limits_{1\le i\le n} X_i$

1. 设总体 X 的概率密度函数为 $f(x)=\begin{cases} e^{-(x-\theta)}, & x>\theta \\ 0, & x\le\theta \end{cases}$，其中 $\theta>0$ 为未知参数，$X_1,\cdots,X_n$ 是来自总体 X 的样本，$x_1,\cdots,x_n$ 是样本值，试求：

(1) θ 的矩估计量；　(2) θ 的极大似然估计量

【解】(1) $E(X)=\int_\theta^{+\infty} xe^{-(x-\theta)}dx=\theta+1$

令 $E(X)=\bar{X}$，

解出矩估计量为 $\hat{\boldsymbol{\theta}}=\bar{X}-1=\frac{1}{n}\sum_{i=1}^{n}X_i-1$

(2) 似然函数为 $L(\theta)=\begin{cases} \prod_{i=1}^{n} e^{-(x_i-\theta)}=e^{n(\theta-\bar{x})}, & \theta\le\min\{x_1,\cdots,x_n\} \\ 0 & \text{其他} \end{cases}$，

由于 $L(\theta)$ 关于 θ 单增，故而极大似然估计值为 $\hat{\boldsymbol{\theta}}=\min\{x_1,\cdots,x_n\}$。

从而极大似然估计量为 $\hat{\boldsymbol{\theta}}=\min\{X_1,\cdots,X_n\}$.

2. 设总体 X 的概率密度函数为 $f(x)=\begin{cases}\dfrac{2\alpha^2}{x^3}, & x>\alpha\\ 0, & x\le\alpha\end{cases}$，其中 $\alpha>0$ 为未知参数，$X_1,\cdots,X_n$ 是来自总体 X 的样本，$x_1,\cdots,x_n$ 是样本值，试求：

(1) α 的矩估计量；　　(2) α 的极大似然估计量。

【解】(1) $E(X)=\int_\alpha^{+\infty} x\cdot\dfrac{2\alpha^2}{x^3}dx=2\alpha^2\int_\alpha^{+\infty}\dfrac{1}{x^2}dx=2\alpha$

令 $E(X)=\bar{X}$，　解出矩估计量为 $\hat{\alpha}=\dfrac{\bar{X}}{2}=\dfrac{1}{2n}\sum_{i=1}^{n}X_i$

(2) 似然函数为 $L(\alpha)=\begin{cases}\prod_{i=1}^{n}\dfrac{2\alpha^2}{x_i^3}=\dfrac{2^n\alpha^{2n}}{(x_1x_2\cdots x_n)^3}, & \alpha\le\min\{x_1,\cdots,x_n\}\\ 0, & 其他\end{cases}$,

由于 $L(\alpha)$ 关于 α 单增，故极大似然估计值为 $\hat{\alpha}=\min\{x_1,\cdots,x_n\}$，

极大似然估计量为 $\hat{\alpha}=\min\{X_1,\cdots,X_n\}$

3. 总体 $X \sim U(\theta, 2\theta)$ **服从均匀分布,其中**$\theta > 0$**是未知参数**

$X_1, \ldots, X_n$是取自该总体的样本, $\overline{X}$ 为样本均值.

证明: $\hat{\theta} = \frac{2}{3}\overline{X}$ **是参数**θ **的无偏估计和相合（一致）估计.**

证明(1) $X \sim U(\theta, 2\theta) \Rightarrow E(X) = \frac{\theta + 2\theta}{2} = \frac{3}{2}\theta$

$E(\hat{\theta}) = E\left(\frac{2}{3}\overline{X}\right) = \frac{2}{3}E(\overline{X}) = \frac{2}{3} \cdot \frac{3\theta}{2} = \theta$是参数$\theta$ 的无偏估计

(2) $D(X) = \frac{(2\theta - \theta)^2}{12} = \frac{1}{12}\theta^2$

$$D(\hat{\theta}) = \frac{4}{9}D(\overline{X}) = \frac{4}{9n}D(X) = \frac{4\theta^2}{9 \times 12n} = \frac{\theta^2}{27n}$$

由切比雪夫不等式得 $P(|\hat{\theta} - \theta| < \varepsilon) \geq 1 - \frac{1}{\varepsilon^2}D(\hat{\theta})$

$\lim\limits_{n \to \infty} P(|\hat{\theta} - \theta| < \varepsilon) = 1$是参数 θ的相合估计

4. 确定某种溶液的溶剂含量，现任取4个样品，测得样本均值为 $\bar{x}=8.34$ $s=0.03$

现溶液中溶剂含量近似服从正态分布 $N(\mu,\sigma^2)$

求：μ 的置信度为 95% 的置信区间

解 σ^2未知 估计量：$\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

$$P\left\{\left|\frac{\bar{X}-\mu}{S/\sqrt{n}}\right|<t_{\frac{\alpha}{2}}(n-1)\right\}=1-\alpha$$

μ的置信区间为$\left(\bar{X}\pm\dfrac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$

$\because 1-\alpha=95\%$ 故 $\alpha=0.05,$

查 t 分布表得：$t_{\frac{\alpha}{2}}(n-1)=t_{0.025}(3)=3.1824$

从而 μ 的 95%的置信区间为：(8.2923，8.3877)

5. 设样本$X_1,\cdots,X_n$来自总体X，$E(X)=\mu$，$D(X)=\sigma^2$均未知，则下列正确的是________。

（A）$\frac{1}{n\text{-}1}\sum_{i=1}^{n}X_i$是$\mu$的无偏估计

（B）$\frac{1}{n}\sum_{i=1}^{n}X_i$是$\mu$的无偏估计

（C）$\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$是$\sigma^2$的无偏估计

（D）$\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})$是$\sigma$的无偏估计

B

第八章 假设检验

第一节 假设检验的基本概念

第二节 单个正态总体的假设检验

总体X, $X \sim N(\mu,\sigma^2)$, $X_1X_2,\cdots,X_n$　$x_1x_2,\cdots,x_n$

对μ,σ^2进行假设检验　　显著性水平α,

	原假设H_0	备择假设H_1	检验统计量	拒绝域
1) μ的检验 σ^2为已知	$\mu=\mu_0$	$\mu\neq\mu_0$	$U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$	$\left(-\infty,-z_{\alpha/2}\right)\cup\left(z_{\alpha/2},+\infty\right)$
	$\mu\leq\mu_0$	$\mu>\mu_0$		$\left(z_\alpha,+\infty\right)$
	$\mu\geq\mu_0$	$\mu<\mu_0$		$\left(-\infty,-z_\alpha\right)$
2) μ的检验 σ^2为未知	$\mu=\mu_0$	$\mu\neq\mu_0$	$t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$	$\left(-\infty,-t_{\alpha/2}(n-1)\right)\cup\left(t_{\alpha/2}(n-1),+\infty\right)$
	$\mu\leq\mu_0$	$\mu>\mu_0$		$\left(t_\alpha(n-1),+\infty\right)$
	$\mu\geq\mu_0$	$\mu<\mu_0$		$\left(-\infty,-t_\alpha(n-1)\right)$
3) σ^2的检验	$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$	$\left(0,\chi^2_{1-\alpha/2}(n-1)\right)\cup\left(\chi^2_{\alpha/2}(n-1),+\infty\right)$
	$\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\left(\chi^2_\alpha(n-1),\ +\infty\right)$
	$\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\left(0,\ \chi^2_{1-\alpha}(n-1)\right)$

例1 在正常情况下，袋装糖的重量（公斤）服从 $N(0.5,\ 0.015^2)$. 某天随机抽取9袋糖，算得平均重量为 $\bar{x}=0.511$, 问这天机器是否正常? $(\alpha=0.05)$

解　这天袋装糖的重量 $X \sim N(\mu,\ 0.015^2)$

假设：$H_0: \mu=\mu_0=0.5$　　$H_1: \mu \neq \mu_0$

取统计量

$$U=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

拒绝域：$\left|\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| > z_{\alpha/2}=1.96$

即 $(-\infty,-1.96)\cup(1.96,+\infty)$

又　$n=9,\ \bar{x}=0.511,\ \sigma=0.015,\ \mu_0=0.5$

算得 $\left|\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|=2.2\ >1.96$　故拒绝 H_0

即认为这天机器工作不正常.

例2 某编织物强力指标 X 的均值 $\mu_0 = 21$ 公斤。改进工艺后生产了一批编织物，今从中取 30 件，测得 $\bar{x} = 21.55$ 公斤。假设强力X指标服从正态分布 $N(\mu, \sigma^2)$，且已知 $\sigma = 1.2$ 公斤

问：在显著性水平 $\alpha = 0.01$ 下，新生产编织物比过去的编织物强力是否有提高？

解：提出假设：$H_0 : \mu \le 21 \quad \Leftrightarrow \quad H_1 : \mu > 21$

取统计量：$U = \dfrac{\bar{X} - 21}{\sigma/\sqrt{n}} \sim N(0,1)$

拒绝域：$(z_\alpha, +\infty) = (2.33, +\infty) \qquad z_{0.01} = 2.33$

计算观察值：$u = 2.51 > 2.33$

故拒绝H_0,认为新生产编织物比过去强力有提高。

3 某罐头厂生产的水果罐头重量和维生素 C 的含量长期以来分别服从正态分布 $N(\mu_1,0.4),N(\mu_2,\sigma_2^2)$，根据生产要求每个水果罐头的维生素 C 含量不能小于 4。现从该厂生产的一批产品中抽取 9 个罐头测得重量的样本方差为 0.64；维生素 C 含量平均为 3.4，方差为 0.81。

(1) 这批产品的重量的波动较以往是否有显著变化。(取显著性水平 $\alpha=0.1$)

(2) 是否可以认为这批产品的维生素 C 含量符合生产要求。(取显著性水平 $\alpha=0.1$)

(3) 求这批产品的平均维生素 C 含量的置信度为 0.9 的置信区间.

可能需要用到的数据： $\chi^2_{0.05}(8)=15.507$， $\chi^2_{0.95}(8)=2.733$， $t_{0.1}(8)=1.3968$， $t_{0.05}(8)=1.8595$，

【解】1） $H_0:\sigma_1^2=0.4=\sigma_0^2, H_1:\sigma_1^2\neq 0.4$

选择检验统计量为 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$

拒绝域 $C=(0,\chi^2_{0.95}(8))\cup(\chi^2_{0.05}(8),+\infty)=(0,2.733)\cup(15.507,+\infty)$

计算得 $\chi^2=\dfrac{8\times 0.64}{0.4}=12.8\notin C$，

因此可以认为这批产品的重量的波动较以往没有显著性变化.

3 某罐头厂生产的水果罐头重量和维生素 C 的含量长期以来分别服从正态分布 $N(\mu_1,0.4),N(\mu_2,\sigma_2^2)$，根据生产要求每个水果罐头的维生素 C 含量不能小于 4。现从该厂生产的一批产品中抽取 9 个罐头测得重量的样本方差为 0.64；维生素 C 含量平均为 3.4，方差为 0.81。

(1) 这批产品的重量的波动较以往是否有显著变化。(取显著性水平 $\alpha=0.1$)

(2) 是否可以认为这批产品的$\boxed{\text{维生素 C 含量}}$符合生产要求。(取显著性水平 $\alpha=0.1$)

(3) 求这批产品的平均维生素 C 含量的置信度为 0.9 的置信区间.

可能需要用到的数据： $\chi_{0.05}^2(8)=15.507$， $\chi_{0.95}^2(8)=2.733$，$t_{0.1}(8)=1.3968$，$t_{0.05}(8)=1.8595$，

2）$H_0:\mu_2\ge 4=\mu_0, H_1:\mu_2<4$

选择检验统计量为 $T=\dfrac{\bar{X}-\mu_0}{\dfrac{S}{\sqrt{n}}}$，

拒绝域为 $C=(-\infty,-t_{0.1}(8))=(-\infty,-1.3968)$

计算得 $t=\dfrac{\bar{x}-\mu_0}{\dfrac{s}{\sqrt{n}}}=\dfrac{3.4-4}{\dfrac{\sqrt{0.81}}{\sqrt{9}}}=-2\in C$，因此认为这批产品的维生素 C 含量不符合生产要求。

3 某罐头厂生产的水果罐头重量和维生素 C 的含量长期以来分别服从正态分布 $N(\mu_1, 0.4), N(\mu_2, \sigma_2^2)$，根据生产要求每个水果罐头的维生素 C 含量不能小于 4。现从该厂生产的一批产品中抽取 9 个罐头测得重量的样本方差为 0.64；维生素 C 含量平均为 3.4，方差为 0.81。

(1) 这批产品的重量的波动较以往是否有显著变化。(取显著性水平 $\alpha = 0.1$)

(2) 是否可以认为这批产品的维生素 C 含量符合生产要求。(取显著性水平 $\alpha = 0.1$)

(3) 求这批产品的平均维生素 C 含量的置信度为 0.9 的置信区间.

可能需要用到的数据： $\chi^2_{0.05}(8) = 15.507$， $\chi^2_{0.95}(8) = 2.733$，$t_{0.1}(8) = 1.3968$，$t_{0.05}(8) = 1.8595$，

(3) $\alpha = 0.1$， $n = 9, s^2 = 0.81, \bar{x} = 3.4$

σ_2^2 未知，选择统计量为 $T = \dfrac{\bar{X} - \mu_0}{\dfrac{S}{\sqrt{n}}} \sim t(\mathrm{n}-1)$

所求置信区间为 $(\bar{X} - \dfrac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1), \bar{X} + \dfrac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)) = (3.4 - \dfrac{\sqrt{0.81}}{\sqrt{9}} t_{0.05}(8), 3.4 + \dfrac{\sqrt{0.81}}{\sqrt{9}} t_{0.05}(8))$

$=(3.4 - 0.3 \times 1.8595, 3.4 + 0.3 \times 1.8595) \approx (2.84, 3.96)$

4 在某次实验中需要测量某物体的质量。一组测量结果如下（单位：g）

8.6　8.5　8.9　8.4　8.4　8.7　8.8　8.3　8.8。

均值和方差分别记作 μ 和 σ^2。

问题：（1）求均值 μ 的置信区间，置信度为 0.95；

（2）是否可以认为物体的质量是 $8.2g$？显著性水平 $\alpha = 0.05$；

（3）是否可以认为物体的质量 $\mu \le 8.2$？显著性水平 $\alpha = 0.05$。

已知数据：$z_{0.05} = 1.65$；$t_{0.05}(8) = 1.860$；$t_{0.05}(9) = 1.833$；$t_{0.05}(10) = 1.813$；

$z_{0.025} = 1.96$；$t_{0.025}(8) = 2.306$；$t_{0.025}(9) = 2.262$；$t_{0.025}(10) = 2.228$；$\sqrt{0.045} = 0.212$。

【解】构造 $t-$ 统计量 $t = \dfrac{\overline{X} - \mu}{S/3}$，则 $t \sim t(8)$。计算得到 $\overline{X} = 8.6$，$S^2 = 0.045$。

（1）置信度为 0.95，则 $t_{0.025}(8) = 2.306$，

置信区间为 $\left(\overline{X} \pm \dfrac{S}{3} t_{0.025}(8)\right) = \left(8.6 \pm \dfrac{0.212}{3} \times 2.306\right) = (8.437, 8.763)$。

4　在某次实验中需要测量某物体的质量。一组测量结果如下（单位：g）

8.6　8.5　8.9　8.4　8.4　8.7　8.8　8.3　8.8。

均值和方差分别记作 μ 和 σ^2。

问题：（1）求均值 μ 的置信区间，置信度为 0.95；

（2）是否可以认为物体的质量是 $8.2g$？显著性水平 $\alpha = 0.05$；

（3）是否可以认为物体的质量 $\mu \le 8.2$？显著性水平 $\alpha = 0.05$。

（2）作双边检验 $H_0: \mu = 8.2, H_1: \mu \ne 8.2$，拒绝域为 $|t| = \left|\dfrac{\overline{X} - \mu_0}{S/3}\right| > t_{0.025}(8) = 2.306$，

本题 $|t| = \left|\dfrac{8.6 - 8.2}{0.212/3}\right| = 5.66 > 2.306$，因此拒绝原假设 H_0，不能认为物体的质量是 $8.2g$。

（3）作单边检验 $H_0: \mu \le 8.2, H_1: \mu > 8.2$，　拒绝域为 $t = \dfrac{\overline{X} - \mu}{S/3} > t_{0.05}(8) = 1.860$，

本题 $t = \dfrac{8.6 - 8.2}{0.212/3} = 5.66 > 1.860$，因此拒绝原假设 H_0，认为物体的质量 $\mu > 8.2$。

集中答疑时间：

11月20日周三晚6：00-8：00 地点：理化楼308

11月22日周五下午3：10-5：00 地点：逸夫楼504

考试时间：

11月23日（第10周）周六
上午8：30-10：30

考的都会，蒙的都对

预祝同学们考一个满意的成绩！