四. 随机事件的关系及其运算

复习

随机事件的三种关系：

包含： $A \subset B$ → A发生必导致B发生

互斥：$AB = \Phi$ → A与B不同时发生

对立：$A \bigcup B = S, A \bigcap B = \Phi$ → A, B中有且仅有一个发生

随机事件的三种运算：

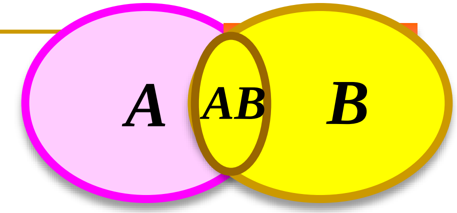
和事件 $A \bigcup B$ 发生 $\Leftrightarrow$ A, B至少有一个发生 $AB \neq \Phi$

$\Leftrightarrow$ A, B有且仅有一个发生 $AB = \Phi$

积事件 AB 发生 $\Leftrightarrow$ A 与 B 同时发生

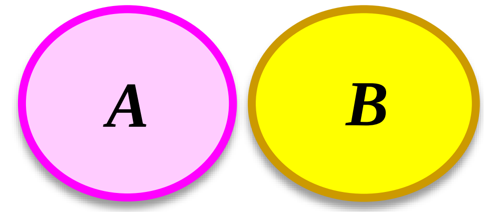
差事件 $A - B$ 发生 $\Leftrightarrow$ A 发生而 B 不发生

随机事件的关系及其运算



随机事件的三种关系：

包含： $A \subset B$ → A发生必导致B发生



互斥：$AB = \Phi$ → A与B不同时发生

对立：$A \cup B = S, A \cap B = \Phi$ → A, B中有且仅有一个发生

随机事件的三种运算：

和事件 $A \cup B$ 发生 $\Leftrightarrow$ A, B至少有一个发生　　$AB \neq \Phi$

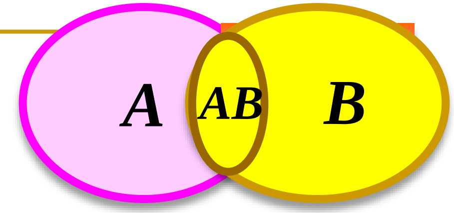
$\Leftrightarrow$ A, B有且仅有一个发生　　$AB = \Phi$

积事件 AB 发生 $\Leftrightarrow$ A 与 B 同时发生

差事件 $A - B$ 发生 $\Leftrightarrow$ A 发生而 B 不发生　$A - B = A\bar{B}$

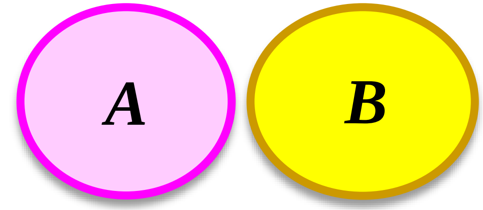
德. 摩根律： $\overline{A \cup B} = \bar{A} \cdot \bar{B}$　　$\overline{AB} = \bar{A} \cup \bar{B}$

随机事件的关系及其运算



随机事件的三种关系：

包含：$A \subset B$ → A发生必导致B发生



互斥：$AB = \Phi$ → A与B不同时发生

对立：$A \cup B = S, A \cap B = \Phi$ → A, B中有且仅有一个发生

随机事件的三种运算：

和事件 $A \cup B$ 发生 $\Leftrightarrow$ A, B至少有一个发生 $AB \neq \Phi$

$\Leftrightarrow$ A, B有且仅有一个发生 $AB = \Phi$

积事件 AB 发生 $\Leftrightarrow$ A 与 B 同时发生

差事件 $A - B$ 发生 $\Leftrightarrow$ A 发生而 B 不发生 $A - B = A\bar{B}$

德. 摩根律： $\overline{A \cup B} = \bar{A} \cdot \bar{B}$ $\overline{AB} = \bar{A} \cup \bar{B}$

概率的公理化定义

复习

设E是随机试验，S是它的样本空间，若对于E的每一个事件A都赋予一个实数 $P(A)$，它满足以下三个条件：

(1) 非负性： 对于每一事件A有： $0 \le P(A) \le 1$

(2) 规范性： $P(S)=1$

(3) 可列可加性：设 $A_1, A_2, \cdots A_k \cdots$ 是两两互斥的事件，

则 $P(A_1 \cup A_2 \cup \cdots \cup A_k \cdots) = P(A_1) + P(A_2) + \cdots + P(A_k) + \cdots$

有且仅有一个发生

则称 $P(A)$ 为事件A发生的概率。

小结

三. 概率的性质

性质1 $P(\Phi)=0$

性质2 若 A_1, A_2 互斥，$P(A_1 \cup A_2)=P(A_1)+P(A_2)$

性质4 一般情况，$P(A_1 \cup A_2)=P(A_1)+P(A_2)-P(A_1A_2)$

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

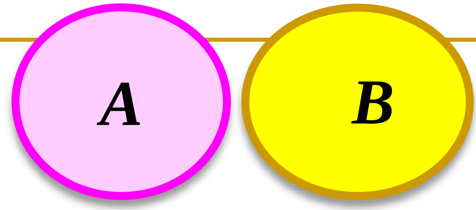
$$P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3)$$

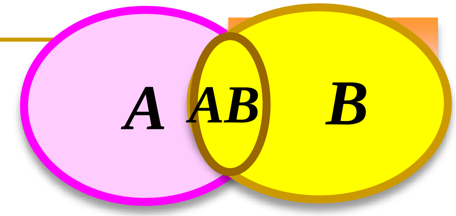
性质3 若 $A \subset B$，则有 $P(B-A)=P(B)-P(A)$

若 $A \not\subset B$，则有 $P(B-A)=P(B)-P(AB)=P(B\bar{A})$

性质5 $P(\bar{A})=1-P(A)$

概率的性质





性质1 $P(\Phi)=0$ $\quad P(A)=0 \not\Rightarrow A=\Phi$

性质2 $P(A\cup B)=P(A)+P(B)$ $\qquad AB=\Phi$

性质4 $P(A\cup B)=P(A)+P(B)-P(AB)$ $\qquad AB\neq\Phi$

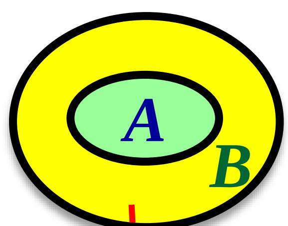
$$P(A_1\cup A_2\cup A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3)$$

性质3 $P(B-A)=P(B)-P(A)$ $\qquad A\subset B$

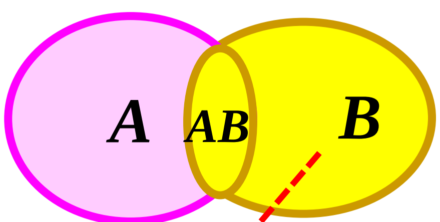
$P(B-A)=P(B\overline{A})=P(B)-P(AB)$ $\qquad A\not\subset B$

性质5 $P(\overline{A})=1-P(A)$

条件概率 $P(A|B)=\dfrac{P(AB)}{P(B)}=\dfrac{k}{m}$



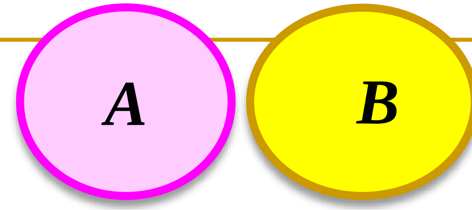
$B-A$

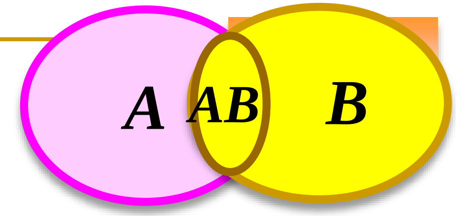


$B-A=B-AB$

$=B\overline{A}$

概率的性质





性质1 $P(\Phi)=0$ $P(A)=0 \not\Rightarrow A=\Phi$

性质2 $P(A\cup B)=P(A)+P(B)$ $AB=\Phi$

性质4 $P(A\cup B)=P(A)+P(B)-P(AB)$ $AB\neq\Phi$

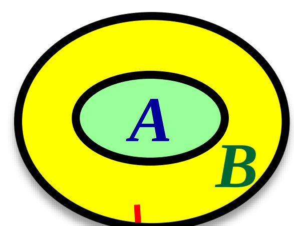
$$P(A_1\cup A_2\cup A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3)$$

性质3 $P(B-A)=P(B)-P(A)$ $A\subset B$

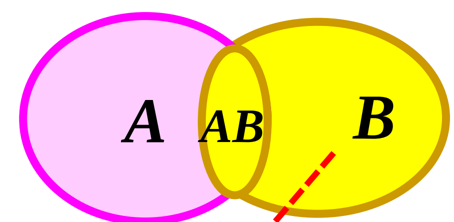
$P(B-A)=P(B\bar{A})=P(B)-P(AB)$ $A\not\subset B$

性质5 $P(\bar{A})=1-P(A)$

条件概率 $P(A\mid B)=\dfrac{P(AB)}{P(B)}=\dfrac{k}{m}$



$B-A$



$B-A=B-AB$

$=B\bar{A}$

第一章小结

概率的计算

1）统计定义： $f_n(A) \xrightarrow[n\to\infty]{}$ 稳定值 $= P(A)$

2）概率的性质：1~5 $P(A-B)$ $P(A\cup B)$ $P(\bar{A})$

3）等可能概型： $P(A)=\dfrac{m}{n}$

4）条件概率： $P(B|A)=\dfrac{k}{m}=\dfrac{P(AB)}{P(A)}$

独立

5）乘法定理： $P(AB)=P(A)P(B|A)=P(A)P(B)$

$A=AB_1\cup AB_2$ 互斥

★ 6）全概率公式： $P(A)=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)$

7）贝叶斯公式： $P(B_1|A)=\dfrac{P(B_1)P(A|B_1)}{P(A)}$

小结

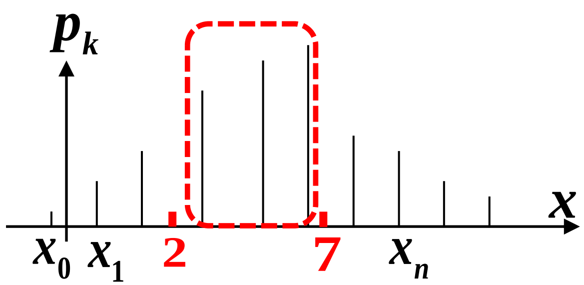
随机变量

	离散型随机变量	概率模型适用范围
重要分布	1） (0-1)分布 $P(X=k)=p^k(1-p)^{1-k}$ $k=0,1$	检查产品的质量(正品与次品) 有奖债券是否中奖(中与不中) 对婴儿性别进行登记(男与女)
	2）$B(n,p)$ ★ $P(X=k)=C_n^k p^k(1-p)^{n-k}$ $k=0,1,2,\cdots,n$	20台独立工作的相同设备中同一时刻发生故障的台数 4个婴儿中男孩的个数 3个产品中的次品数
	3）$P(\lambda)$ $k=0,1,2,\cdots$	120急救台收到的呼叫次数 机场降落的飞机数

上述问题都具有相同的概率模型，但其中参数随问题不同而不同。可以根据抽样得到的大量随机数据，利用数理统计中的统计推断方法-Ch7参数估计得到这些参数的估计值。

小结	随机变量
	离散型随机变量
其它概率分布（书上未列出）	4）几何分布 $G(p)$ $P(X=k)=p(1-p)^{k-1} \quad k=1,2,\cdots$
	5）对数分布 $L(p)$ $P(X=k)=-\dfrac{(1-p)^k}{(\ln p)k} \quad k=1,2,\cdots$
	6）超几何分布 $H(n,M,N)$ $P(X=k)=\dfrac{C_{N-M}^{n-k}C_M^k}{C_N^n}$ $\quad N,M,n$ 为整数 $\quad 0\le M,n\le N$

复习



随机变量

$$P(A)=P\{2<X\le 7\}$$

离散型

连续型

离散非离散

第2节 分布律

第4节 概率密度

第3节 分布函数

(0-1)分布

$$P(X=k)=p^k(1-p)^{1-k}$$

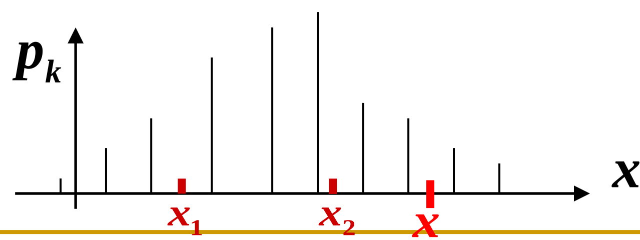
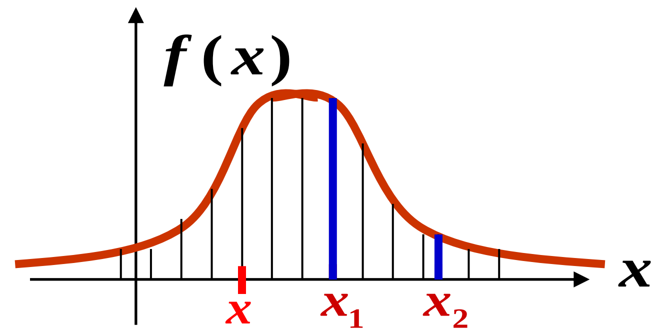
$B(n,p)$

$$P(X=k)=C_n^k p^k(1-p)^{n-k}$$

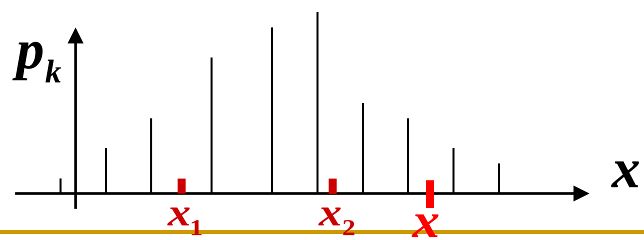
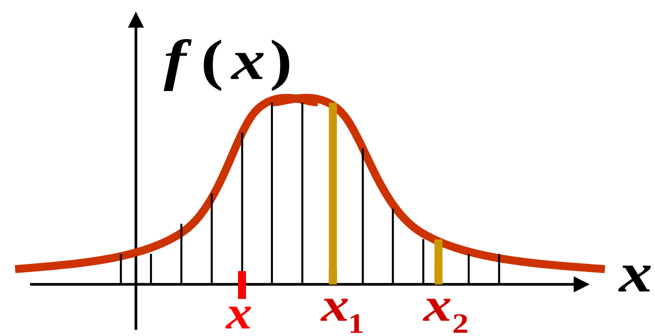
$P(\lambda)$

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$$

小结

<table>
<tr><th colspan="3">随机变量</th></tr>
<tr><th></th><th>离散型随机变量</th><th>连续型随机变量</th></tr>
<tr><td>分布函数
$F(x)=P(X\le x)$
x左侧区间上的概率和
不直观</td><td>$$F(x)=\sum_{x_k\le x} p_k$$
右连续</td><td>$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$
连续</td></tr>
<tr><td>概率分布
概率1分布
情况,直观</td><td>分布律: $\sum p_k=1$
$$\begin{array}{c|cccc} X & x_1 & x_2 & \cdots & x_k \\ \hline p_k & p_1 & p_2 & \cdots & p_k \end{array}$$

</td><td>概率密度: $\int_{-\infty}^{+\infty} f(t)\mathrm{d}t=1$

</td></tr>
<tr><td>概率计算</td><td>$$P(x_1<X\le x_2)=\sum_{x_1<x_k\le x_2} p_k$$
$$=F(x_2)-F(x_1)$$</td><td>$$P(x_1<X\le x_2)=\int_{x_1}^{x_2} f(t)\mathrm{d}t$$
$$=F(x_2)-F(x_1)$$</td></tr>
</table>

复习

	随机变量	
	离散型随机变量	连续型随机变量
分布函数 $F(x)=P(X\le x)$ x左侧区间上的概率和 不直观	$F(x)=\sum\limits_{x_k\le x} p_k$ 右连续	$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$ 连续
概率分布 概率1分布情况,直观	分布律: $\sum p_k=1$ $\begin{array}{c\|cccc} X & x_1 & x_2 & \cdots & x_k \\ \hline p_k & p_1 & p_2 & \cdots & p_k \end{array}$ 	概率密度: $\int_{-\infty}^{+\infty} f(t)\mathrm{d}t=1$ 
概率计算	$P(x_1<X\le x_2)=\sum\limits_{x_1<x_k\le x_2} p_k$ $=F(x_2)-F(x_1)$	$P(x_1<X\le x_2)=\int_{x_1}^{x_2} f(t)\mathrm{d}t$ $=F(x_2)-F(x_1)$

小结	随机变量	
	离散型随机变量	连续型随机变量
重要分布	1）(0-1)分布 $P(X=k)=p^k(1-p)^{1-k}$ $k=0,1$	1）$U(a,b)$ ★ $f(x)=\begin{cases}1/(b-a), & a<x<b\\ 0, & \text{其它}\end{cases}$
	2）$B(n,p)$ ★ $P(X=k)=C_n^k p^k(1-p)^{n-k}$ $k=0,1,\cdots n$	2）$E(\theta)$ ★ $f(x)=\begin{cases}1/\theta\, e^{-x/\theta}, & x>0\\ 0, & \text{其它}\end{cases}$
	3）$P(\lambda)$ $P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2\cdots$	3）$N(\mu,\sigma^2)$ ★ $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty<x<\infty$

小结

正态分布

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$

	$N(\mu,\sigma^2)$	$N(0,1)$
概率密度	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty<x<\infty$	$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ $-\infty<x<\infty$
分布函数	$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t$ $-\infty<x<\infty$	$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\mathrm{d}t$ 查表 $-\infty<x<\infty$
概率计算	$F(x)=P(X\le x)=P(\frac{X-\mu}{\sigma}\le\frac{x-\mu}{\sigma})=\Phi(\frac{x-\mu}{\sigma})$ $P(a\le X\le b)=P(\frac{a-\mu}{\sigma}\le\frac{X-\mu}{\sigma}\le\frac{b-\mu}{\sigma})=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})$	

小结

	随机变量	
	连续型随机变量	其它概率分布
重要分布	1）$U(a,b)$ $$f(x)=\begin{cases}1/(b-a), & a<x<b\\ 0, & 其它\end{cases}$$	4）瑞利分布 $R(\mu)$ 5）贝塔分布 $\beta(p,q)$ 6）伽马分布 $\Gamma(\alpha,\beta)$ √
	2）$E(\theta)$ $$f(x)=\begin{cases}1/\theta\, e^{-x/\theta}, & x>0\\ 0, & 其它\end{cases}$$	7）χ^2 分布 $\chi^2(n)$ √ 8）t 分布 $t(n)$ √ 9）F 分布 $F(n_1,n_2)$ √
	3）$N(\mu,\sigma^2)$ $$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty<x<\infty$$	10）威布尔分布 $W(m,\alpha)$ 11）柯西分布 $C(\mu,\alpha)$

三. 连续型随机变量的函数的分布

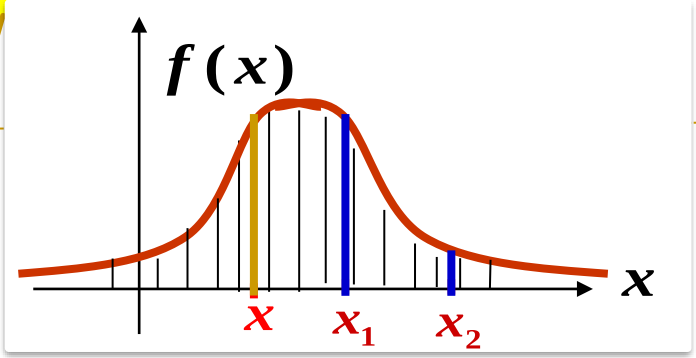
问题：已知 $Y = g(X)$ 和 $f_X(x)$，求 $f_Y(y)$

分布函数法：(1) 求 $F_Y(y)$

(2) $F_Y'(y) = f_Y(y)$

难点

复习



分布函数：$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$

概率计算：$P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f(t)\mathrm{d}t = F(x_2) - F(x_1)$

概率密度性质：$F'(x) = f(x)$

高数求导：$Z(x) = \int_{p(x)}^{g(x)} z(t)\mathrm{d}t \Rightarrow Z'(x) = z[g(x)]\,g'(x) - z[p(x)]\,p'(x)$

复习

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$

	正态分布	
	$N(\mu,\sigma^2)$	
概率密度	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty<x<\infty$	
分布函数	$F(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t$ $-\infty<x<\infty$	
概率计算	$P(a\le X\le b)=\int_{x_1}^{x_2} f(t)\mathrm{d}t=F(x_2)-F(x_1)$	

复习

$$F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

正态分布

	$N(\mu,\sigma^2)$	$N(0,1)$
概率密度	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty<x<\infty$	$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ $-\infty<x<\infty$
分布函数	$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t$ $-\infty<x<\infty$	$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\mathrm{d}t$ $-\infty<x<\infty$ 查表
概率计算	$F(x)=P(X\le x)=P(\frac{X-\mu}{\sigma}\le\frac{x-\mu}{\sigma})=\Phi(\frac{x-\mu}{\sigma})$ $P(a\le X\le b)=P(\frac{a-\mu}{\sigma}\le\frac{X-\mu}{\sigma}\le\frac{b-\mu}{\sigma})=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})$	

第二章小结

计算 $P(x_1 < X \le x_2)$ 的方法

随机变量

离散型：分布律 分布函数 $\longrightarrow P(x_1 < X \le x_2)$

$$P(x_1 < X \le x_2) = \sum_{x_1 < x_k \le x_2} p_k = F(x_2) - F(x_1)$$

非离散型：

连续型：概率密度 分布函数 $\longrightarrow P(x_1 < X \le x_2)$

$$P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f(t)\mathrm{d}t = F(x_2) - F(x_1)$$

非连续型：分布函数 $\longrightarrow P(x_1 < X \le x_2)$

$$P(x_1 < X \le x_2) = F(x_2) - F(x_1)$$

分布律	概率密度
$X \sim (0,1)$	$X \sim U(a,b)$
$X \sim B(n,p)$	$X \sim E(\theta)$
$X \sim P(\lambda)$	$X \sim N(\mu,\sigma^2)$

小结

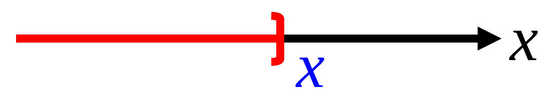
	随机变量	
	离散型随机变量	连续型随机变量
重要分布	1）(0-1)分布 $P(X=k)=p^k(1-p)^{1-k}$ $k=0,1$	1）$U(a,b)$ ★ $f(x)=\begin{cases}1/(b-a), & a<x<b\\ 0, & \text{其它}\end{cases}$
	2）$B(n,p)$ ★ $P(X=k)=C_n^k p^k(1-p)^{n-k}$ $k=0,1,\cdots n$	2）$E(\theta)$ ★ $f(x)=\begin{cases}1/\theta\, e^{-x/\theta}, & x>0\\ 0, & \text{其它}\end{cases}$
	3）$P(\lambda)$ $P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2\cdots$	3）$N(\mu,\sigma^2)$ ★ $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
函数的分布 $Y=g(X)$	X的分布律 → Y的分布律	★★ $f_X(x)\rightarrow f_Y(y)=F_Y'(y)$

两章关系	第二章 一维 X	第一节 二维(X,Y)	第二节 边缘分布	第三节 相依(条件分布)	第四节 独立
分布函数	$F(x)$	$F(x,y)$	$F_X(x)$ $F_Y(y)$	$P(A\|B)=\dfrac{P(AB)}{P(B)}$	$P(AB)=P(A)P(B)$
离散型分布律	$P\{X=x_k\}$ $=p_k$	$P\{X=x_i,Y=y_j\}$ $=p_{ij}$	$P\{X=x_i\}$ $P\{Y=y_i\}$	$P\{Y=y_j\|X=x_i\}$ $P\{X=x_i\|Y=y_j\}$	$P\{X=x_i,Y=y_j\}$ $=P\{X=x_i\}\cdot P\{Y=y_j\}$
连续型概率密度	$f(x)$	$f(x,y)$	$f_X(x)$ ★ $f_Y(y)$	$f_Y(y\|X=x)$ $f_X(x\|Y=y)$	$f(x,y)=$ $f_X(x)\cdot f_Y(y)$
算概率	$P(x_1<X\le x_2)=\int_{x_1}^{x_2}f(x)\mathrm{d}x$ $=\sum\limits_{x_1<x_k\le x_2}p_k$		$P\{(X,Y)\in G\}=\iint\limits_G f(x,y)\mathrm{d}x\mathrm{d}y$ ★ $=\sum\limits_{(x_i,y_j)\in G}p_{ij}$		
函数分布	$Y=g(X)$ $f_X(x)\rightarrow f_Y(y)=F_Y'(y)$		$Z=g(X,Y)$ 第五节 $f(x,y)\rightarrow f_Z(z)=F_Z'(z)$ 分布函数法 ★ ★		

3. 二维随机变量的分布函数

复习：一维随机变量的分布函数

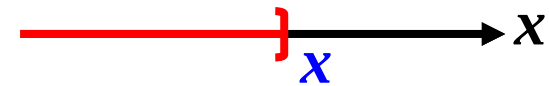
设 X 是一个一维随机变量，



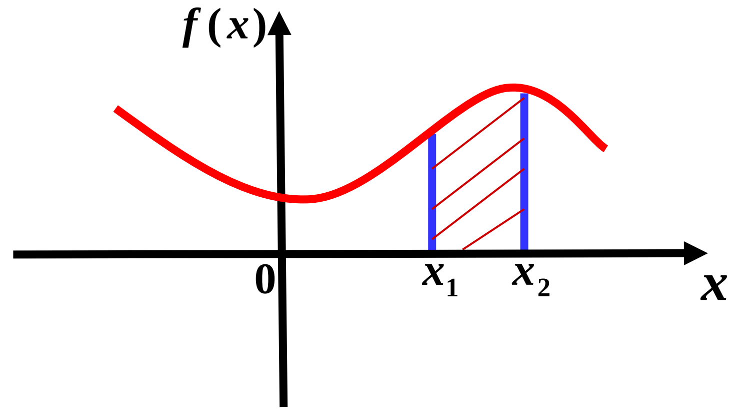
$$F(x) = P(X \le x)$$

复习：一维连续型随机变量

X 为连续型随机变量：



$$F(x)=P(X\le x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$$



性质1　$f(x)\ge 0$

性质2　$\int_{-\infty}^{\infty} f(x)\mathrm{d}x=1$

性质3　$P(x_1<X\le x_2)=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)\mathrm{d}x$

性质4　$F'(x)=f(x)$

复习

一维离散型随机变量的分布律

X 的分布律：

X	0	1	2	3	4	5
P_k	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{2}{9}$	$\frac{1}{9}$

➢分布律需满足：$(1)\ p_k \geq 0,\quad (2)\sum_{k=1}^{\infty} p_k = 1$

两章关系	第二章 一维 X	第一节 二维(X,Y)	第二节 边缘分布	第三节 相依(条件分布)	第四节 独立
分布函数	$F(x)$	$F(x,y)$			$P(AB)=P(A)P(B)$
离散型分布律	$P\{X=x_k\}$ $=p_k$	$P\{X=x_i,Y=y_j\}$ $=p_{ij}$	$P\{X=x_i\}$ $P\{Y=y_i\}$		$P\{X=x_i,Y=y_j\}$ $=P\{X=x_i\}\cdot P\{Y=y_j\}$
连续型概率密度	$f(x)$	$f(x,y)$	$f_X(x)$ ★ $f_Y(y)$		$f(x,y)=$ $f_X(x)\cdot f_Y(y)$
算概率	$P(x_1<X\le x_2)=\int_{x_1}^{x_2} f(x)\mathrm{d}x$ $=\sum\limits_{x_1<x_k\le x_2} p_k$		$P\{(X,Y)\in G\}=\iint\limits_G f(x,y)\mathrm{d}x\mathrm{d}y$ ★ $=\sum\limits_{(x_i,y_j)\in G} p_{ij}$		
函数分布	$Y=g(X)$ $f_X(x)\rightarrow f_Y(y)=F_Y'(y)$				

二维随机变量的分布函数

复习

设 (X,Y) 是二维随机变量，对任意的实数 x, y，

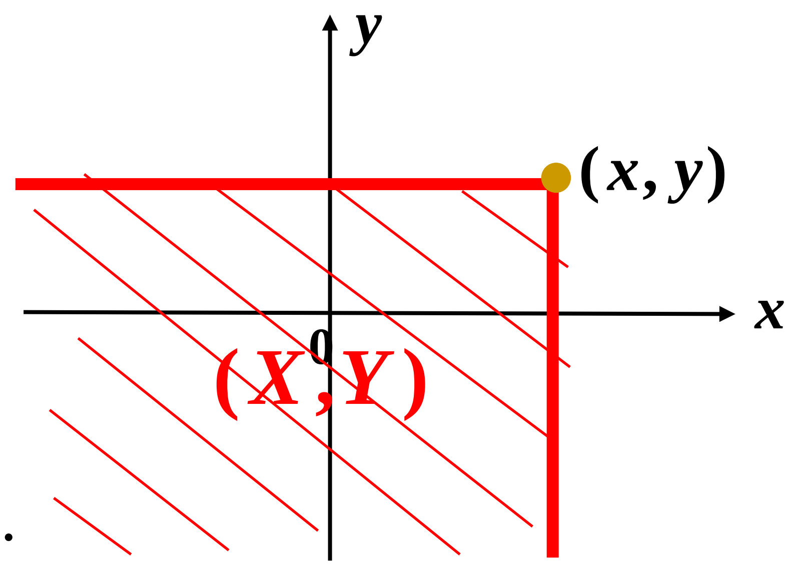
二元函数 $F(x,y)= P(X \le x, Y \le y)$ (积事件) 称为二维随机变量 (X,Y) 的分布函数或 X与Y 的联合分布函数。

注:

- $F(x, y)$ 几何意义：

  将(X, Y)看成是随机点的坐标，则

  $F(x, y)$ = 随机点 (X,Y)落在以 (x, y) 为顶点的左下方无穷矩形内的概率.



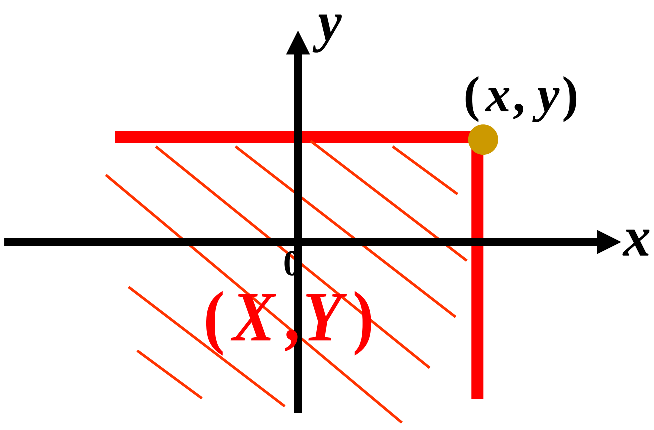
复习

二维连续型随机变量的概率密度

定义 对于二维随机变量(X,Y) 的分布函数$F(x, y)$，若存在非负函数 $f(x, y)$, 对任意的x, y有：

$$F(x,y)=P\{X\le x,Y\le y\}=\int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\mathrm{d}u\mathrm{d}v$$

则称 (X,Y) 是二维连续型随机变量，$f(x, y)$为 (X,Y) 的联合概率密度.

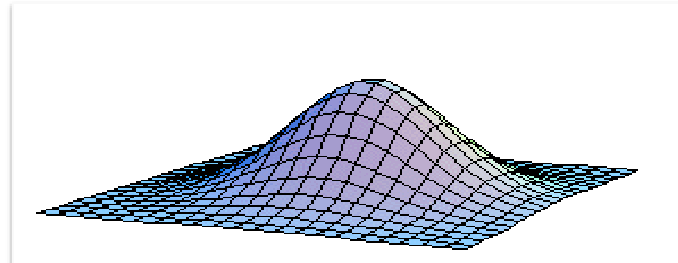


概率密度的性质

复习

$$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v$$

性质1 $f(x,y)\geq 0$



概率密度曲面下面的体积是1

性质2 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}x\,\mathrm{d}y=1$

性质3 $\dfrac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y)$

性质4 设 G 是 XOY 平面上的一个区域，则点(X,Y)落在G内的概率为:

$$\bigstar\ P\{(X,Y)\in G\}=\iint_G f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\iint_D f(x,y)\mathrm{d}x\mathrm{d}y$$

➢ D是积分区域G和概率密度取值非零区域的交集

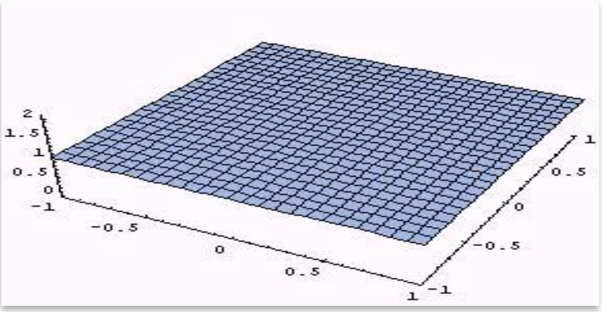
复习

1^0 若 $f(x,y)=\begin{cases}\dfrac{1}{(b_1-a_1)(b_2-a_2)} & \boxed{\begin{array}{l} a_1\le x\le b_1 \\ a_2\le y\le b_2\end{array}} \\ 0 & \text{其它}\end{cases}$

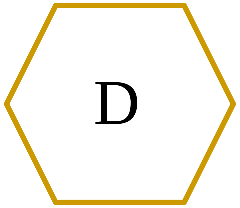
则称 (X,Y)在矩形区域上服从均匀分布。

$X \sim U(a,b)$

$$f(x)=\begin{cases}\dfrac{1}{b-a} & a<x<b \\ 0 & \text{其它}\end{cases}$$



1^0 若 $f(x,y)=\begin{cases}\dfrac{1}{A} & (x,y)\in D \\ 0 & \text{其它}\end{cases}$



则称 **(*X*,*Y*)** 平面区域***D***上服从均匀分布，***A***是***D***的面积。

$X \sim U(a,b)$

$$f(x)=\begin{cases}\dfrac{1}{b-a} & a<x<b \\ 0 & \text{其它}\end{cases}$$

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

(*X*,*Y*) 只能在平面有限区域上服从均匀分布。

复习

2^0 若 $f(x,y)=\begin{cases}\lambda^2 e^{-\lambda(x+y)} & x>0, y>0\\ 0 & \text{其它}\end{cases}$

则称 (X,Y) 服从参数为 λ 的指数分布.

⇧

$X \sim E(\lambda)$

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

复习

3^0 若 $f(x,y)$

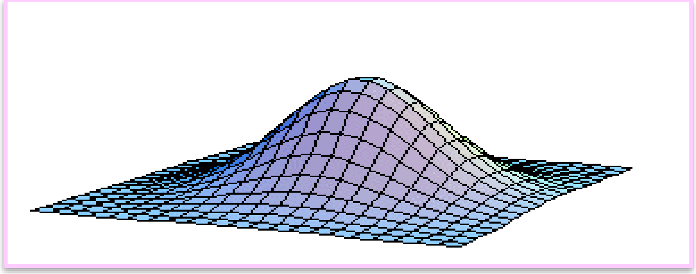
$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1{}^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2{}^2}\right]}$$

其中: $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 为5个常数,

则称 (X,Y)服从参数为 $\mu_1,\mu_2,\sigma_1,\sigma_2,\rho$ 的 正态分布.

$$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$$

$X\sim N(\mu,\sigma^2)$ $\quad f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



正态曲面下面的体积是1

复习

离散型随机变量的联合分布律和边缘分布律

二维随机变量 (X,Y) 的联合分布律：

$X \backslash Y$	y_0	y_1	$\cdots$	y_j	$\cdots$	$P(X=x_i)$
x_0	p_{00}	p_{01}	$\cdots$	p_{0j}	$\cdots$	$P_{0\bullet}$
x_1	p_{10}	p_{11}	$\cdots$	p_{1j}	$\cdots$	$P_{1\bullet}$
$\vdots$						$\vdots$
x_i	p_{i0}	p_{i1}	$\cdots$	p_{ij}	$\cdots$	$P_{i\bullet}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$P(Y=y_j)$	$P_{\bullet 0}$	$P_{\bullet 1}$	$\cdots$	$P_{\bullet j}$	$\cdots$	1

$P_{\bullet j} = P(Y = y_j) = \sum_{i=0}^{+\infty} p_{ij}$ 称为Y 的边缘分布律

$P_{i\bullet} = P(X = x_i) = \sum_{j=0}^{+\infty} p_{ij}$ 称为 X 的边缘分布律

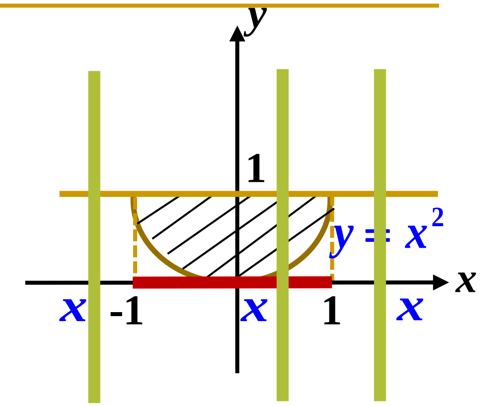
第三章小结　二维随机变量(X,Y)

	离散型	连续型
(X,Y) 整体	联合分布律 $P(X=x_i,Y=y_j)=p_{ij}$	联合概率密度 $f(x,y)$
(X,Y) 个体	边缘分布律 $P(X=x_i)=\sum_{j=1}^{\infty}p_{ij}=p_{i\cdot}$	边缘概率密度 ★ $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$
概率计算	$P\{(X,Y)\in G\}=\sum_{(x_i,y_j)\in G}p_{ij}$	★$P\{(X,Y)\in G\}=\iint_G f(x,y)\mathrm{d}x\mathrm{d}y$
X与Y 独立	$P(X=x_i,Y=y_j)$ $=P(X=x_i)P(Y=y_j)$	$f(x,y)=f_X(x)f_Y(y)$

复习

设二维随机变量 (X,Y) 的概率密度为

$$f(x,y)=\begin{cases}21/4\,x^2y, & \boxed{x^2\le y\le 1}\\ 0, & \text{其它}\end{cases}$$



求：边缘概率密度 $f_X(x)$，$f_Y(y)$.

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

当 $-1<x<1$ 时，$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y=\int_{x^2}^{1}\frac{21}{4}x^2y\mathrm{d}y=\frac{21}{4}x^2\int_{x^2}^{1}y\mathrm{d}y$

$$=\frac{21}{8}x^2y^2\Big|_{x^2}^{1}=\frac{21}{8}x^2(1-x^4)$$

当 $x\le -1$ 或 $x\ge 1$ 时，$f(x,y)=0$，$f_X(x)=0$

$$\therefore\quad f_X(x)=\begin{cases}\frac{21}{8}x^2(1-x^4), & -1<x<1\\ 0, & \text{其它}\end{cases}$$

1. 投影
2. 积分（分区间）
3. 画线（确定上下限）

小结 第三章中计算难点

1） $P\{(X,Y)\in G\}=\iint\limits_{G} f(x,y)\mathrm{d}x\mathrm{d}x = \iint\limits_{D\text{交集}} f(x,y)\mathrm{d}x\mathrm{d}x$

2） $f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y$ 分区间

X,Y 独立

$Z=X+Y$

3） $f_Z(z)=\int_{-\infty}^{+\infty} f(z-y,y)\mathrm{d}y=\int_{-\infty}^{+\infty} \underline{f_X(z-y)f_Y(y)}\mathrm{d}y$ 非零区域 分区间

4） $Z=g(X,Y)$ $f(x,y) \longrightarrow f_Z(z)=?$ $f_Z(z)=F_Z'(z)$

$$F_Z(z)=P(Z\le z)=P\{g(X,Y)\le z\}=\iint\limits_{g(x,y)\le z} f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\iint\limits_{D(z)\ \text{交集}} f(x,y)\mathrm{d}x\mathrm{d}y$$

D是积分区域 $g(x,y)\le z$ 与 $f(x,y)$ 取值非零区域的交集

小结

随机变量的数学期望

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

小结

随机变量的数学期望

	离散型随机变量	连续型随机变量
X	$E(X)=\sum_{k=1}^{\infty} x_k p_k$	$E(X)=\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x$
$Y=g(X)$ g连续	$E(Y)=E[g(X)]$ $=\sum_{k=1}^{\infty} g(x_k)p_k$	$E(Y)=E[g(X)]$ $=\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$
$Z=g(X,Y)$ g连续	$E(Z)=E[g(X,Y)]$ $=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} g(x_i,y_j)p_{ij}$	$E(Z)=E[g(X,Y)]$ $=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$
$E(X)$性质	$E(c)=c$ $E(cX)=cE(X)$ $E(X+Y)=E(X)+E(Y)$ X,Y 独立, $E(XY)=E(X)E(Y)$	

小结

三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	
	二项分布	$X \sim B(n, p)$	np	
	泊松分布	$X \sim P(\lambda)$	λ	
连续型	均匀分布	$X \sim U(a,b)$	$(a+b)/2$	
	指数分布	$X \sim E(\theta)$	θ	
	正态分布	$X \sim N(\mu,\sigma^2)$	μ	

小结

随机变量的数字特征

$E(X)$性质	$E(c)=c$ $E(cX)=cE(X)$ $E(X+Y)=E(X)+E(Y)$ X,Y独立 $E(XY)=E(X)E(Y)$
$D(X)$性质	$D(c)=0$ $D(cX)=c^2D(X)$ $D(X)=E(X^2)-[E(X)]^2$ X,Y独立，$D(X+Y)=D(X)+D(Y)$ $D(X)=0 \Leftrightarrow P(X=c)=1$

复习

随机变量的数学期望与方差

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

复习

数学期望与方差的性质

$E(X)$性质	$E(c)=c$　$E(cX)=cE(X)$　$E(X+Y)=E(X)+E(Y)$ X,Y 独立，$E(XY)=E(X)E(Y)$
$D(X)$性质	$D(c)=0$　$D(cX)=c^2D(X)$　$D(X)=0 \Leftrightarrow P(X=c)=1$ X,Y 独立，$D(X+Y)=D(X)+D(Y)$

复习

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	
	二项分布	$X \sim B(n, p)$	np	
	泊松分布	$X \sim P(\lambda)$	λ	
连续型	均匀分布	$X \sim U(a,b)$	$(a+b)/2$	
	指数分布	$X \sim E(\theta)$	θ	
	正态分布	$X \sim N(\mu,\sigma^2)$	μ	

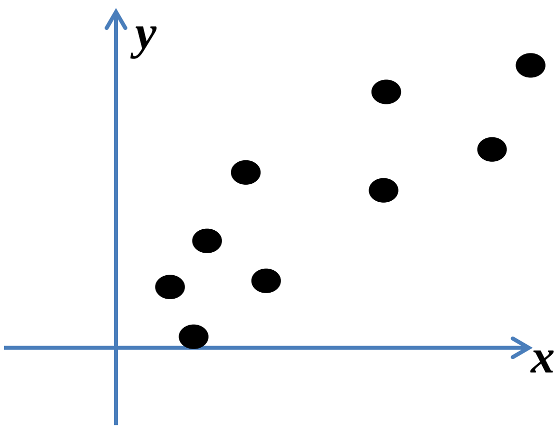
小结

六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	$X \sim E(\theta)$	θ	θ^2
	正态分布	$X \sim N(\mu,\sigma^2)$	μ	σ^2

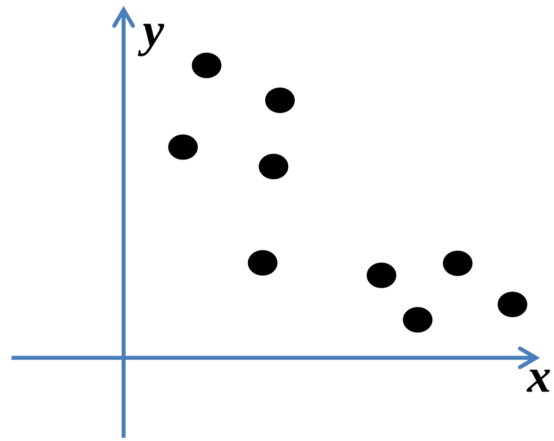
小结:

研究两个随机变量X, Y之间的关系:



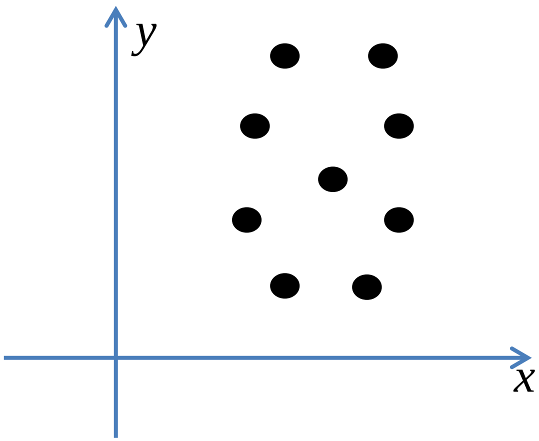
X与Y正相关时,

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})>0$$



X与Y负相关时,

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})<0$$



X与Y不相关时,

$$\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})=0$$

结论: 可以用$\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})$的取值情况来描述$X$与$Y$的相互关系.

小结

随机变量的数字特征

$E(X)$性质	$E(c)=c \quad E(cX)=cE(X) \quad E(X+Y)=E(X)+E(Y)$ X,Y独立 $E(XY)=E(X)E(Y)$
$D(X)$性质	$D(X)=E(X^2)-[E(X)]^2 \quad D(c)=0 \quad D(cX)=c^2D(X)$ X,Y独立, $D(X+Y)=D(X)+D(Y) \quad D(X)=0 \Leftrightarrow P(X=c)=1$
协方差性质	$Cov(X,Y)=E\{[X-E(X)][(Y-E(Y)]\}$ $=E(XY)-E(X)E(Y)$ $D(X+Y)=D(X)+D(Y)+2\operatorname{cov}(X,Y)\overset{独立}{=}D(X)+D(Y)$ (1) $Cov(X,Y)=Cov(Y,X)$ (2) $Cov(X,c)=0$ (3) $Cov(aX,bY)=ab\,Cov(X,Y)$ (4) $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

小结

随机变量的数学期望与方差

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

小结

六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

小结

随机变量的数字特征

$E(X)$性质	$E(c)=c \quad E(cX)=cE(X) \quad E(X+Y)=E(X)+E(Y)$ 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, $D(X+Y)=D(X)+D(Y) \quad D(X)=0 \Longleftrightarrow P(X=c)=1$
协方差	$Cov(X,Y)=E\{[X-E(X)][(Y-E(Y)]\}$ $=E(XY)-E(X)E(Y)$ $D(X+Y)=D(X)+D(Y)+2\mathrm{cov}(X,Y)\overset{\text{独立}}{=}D(X)+D(Y)$
相关系数	$\rho_{XY}=\dfrac{Cov(X,Y)}{\sqrt{D(X)}\cdot\sqrt{D(Y)}}$ (1) $\lvert\rho_{XY}\rvert\le 1$ (2) $\lvert\rho_{XY}\rvert=1 \Leftrightarrow$ 存在常数 a,b, 使得: $P(Y=aX+b)=1$

小结

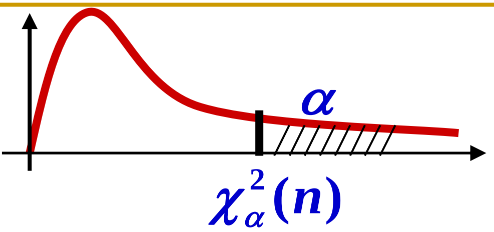
常用统计量

设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的一个样本,

样本均值: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$E(X)$	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} E(X)$ Ch5 大数定律
样本方差: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$	$D(X)$	
样本 k 阶矩: $A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$	$E(X^k)$	$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{P} E(X^k)$ Ch5 大数定律

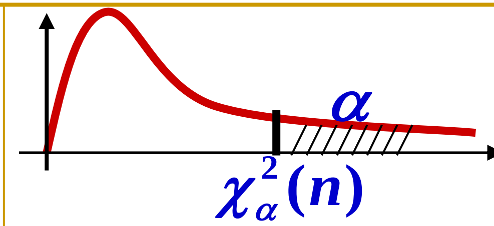
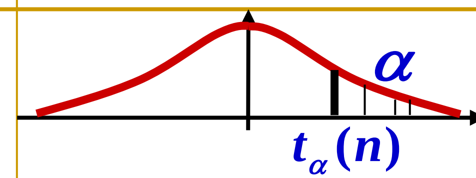
小结

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 ★$\chi^2=\sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$	

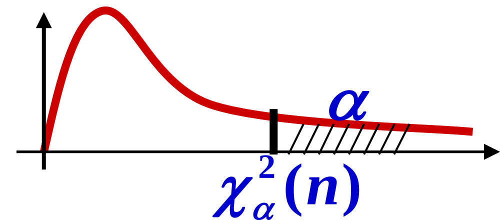
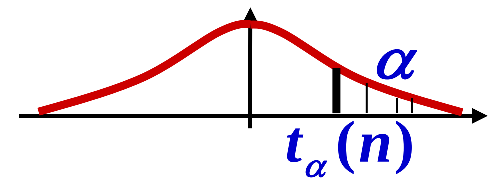
小结

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 ★$\chi^2=\sum_{i=1}^{n}X_i^2 \sim \chi^2(n)$	
t 分布	$X\sim N(0,1),\ Y\sim\chi^2(n),$ 独立 ★$t=\dfrac{X}{\sqrt{Y/n}}\sim t(n)$	 $t_\alpha(n)=-t_{1-\alpha}(n)$

小结

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 ★$\chi^2=\sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$	
t 分布	$X \sim N(0,1),\ Y \sim \chi^2(n)$, 独立 ★$t=\dfrac{X}{\sqrt{Y/n}} \sim t(n)$	
$X \sim N(\mu,\sigma^2)$ $X_1,X_2,\cdots,X_n$ $\bar{X},S^2$ $\bar{X}=\dfrac{1}{n}\sum_{i=1}^{n} X_i$	★*Th1* $\bar{X} \sim N(\mu,\sigma^2/n)$, 独立 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$	*Th2* ★ $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

小结

常用统计量及抽样分布

$X \sim N(\mu,\sigma^2)$

$X_1, X_2, \cdots, X_n$

	统计量	概率分布	
χ^2统计量	$\chi^2 = \sum_{i=1}^{n} X_i^2$	$\sim \chi^2(n)$	$X_i \sim N(0,1)\ i=1,2,\cdots,n$ 独立
t 统计量	$t = \dfrac{X}{\sqrt{Y/n}}$	$\sim t(n)$	$X \sim N(0,1),\ Y \sim \chi^2(n)$, 独立
样本均值	$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$	$\sim N(\mu, \dfrac{\sigma^2}{n})$	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
样本方差	$\dfrac{(n-1)S^2}{\sigma^2}$	$\sim \chi^2(n-1)$	$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\sim t(n-1)$	

复习

常用统计量及抽样分布

$X \sim N(\mu,\sigma^2)$ $E(X)=\mu$

$X_1, X_2, \cdots, X_n$ $D(X)=\sigma^2$ 未知 S^2

	统计量	概率分布	
χ^2统计量	$\chi^2=\sum_{i=1}^{n} X_i^2$	$\sim \chi^2(n)$	$X_i \sim N(0,1)$ $i=1,2,\cdots,n$ 独立
t 统计量	$t=\dfrac{X}{\sqrt{Y/n}}$	$\sim t(n)$	$X \sim N(0,1),\ Y \sim \chi^2(n)$, 独立
样本均值	$\bar{X}=\dfrac{1}{n}\sum_{i=1}^{n} X_i$	$\sim N(\mu,\dfrac{\sigma^2}{n})$	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
样本方差	$\dfrac{(n-1)S^2}{\sigma^2}$	$\sim \chi^2(n-1)$	$S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$
	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\sim t(n-1)$	$\dfrac{(n-1)S^2}{\sigma^2}=\sum_{i=1}^{n}(\dfrac{X_i-\bar{X}}{\sigma})^2$ μ

小结

常用统计量及抽样分布

$X \sim N(\mu,\sigma^2)$

$X_1, X_2, \cdots, X_n$

	统计量	概率分布	
χ^2统计量	$\chi^2 = \sum_{i=1}^{n} X_i^2$	$\sim \chi^2(n)$	$X_i \sim N(0,1)\ i=1,2,\cdots,n$ 独立
t统计量	$t = \dfrac{X}{\sqrt{Y/n}}$	$\sim t(n)$ √	$X \sim N(0,1),\ Y \sim \chi^2(n)$, 独立
样本均值	$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$\sim N(\mu, \frac{\sigma^2}{n})$ √	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
样本方差	$\dfrac{(n-1)S^2}{\sigma^2}$	$\sim \chi^2(n-1)$ √	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\sim t(n-1)$	

小结

常用统计量及抽样分布 $X \sim N(\mu, \sigma^2)$ $X_1, X_2, \cdots, X_n$

	统计量	概率分布	
χ^2统计量	$\chi^2 = \sum_{i=1}^{n} X_i^2$	$\sim \chi^2(n)$ √	$X_i \sim N(0,1)$ $i = 1,2,\cdots,n$ 独立
t 统计量	$t = \dfrac{X}{\sqrt{Y/n}}$	$\sim t(n)$ √	$X \sim N(0,1),\ Y \sim \chi^2(n),$ 独立
样本均值	$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$	$\sim N(\mu, \dfrac{\sigma^2}{n})$	$\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
样本方差	$\dfrac{(n-1)S^2}{\sigma^2}$	$\sim \chi^2(n-1)$	$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
	$\dfrac{\bar{X} - \mu}{S/\sqrt{n}}$	$\sim t(n-1)$	

小结

常用统计量及抽样分布 $X \sim N(\mu,\sigma^2)$ $X_1, X_2, \cdots, X_n$

	统计量	概率分布	
χ^2统计量	$\chi^2 = \sum_{i=1}^{n} X_i^2$	$\sim \chi^2(n)$ √	$X_i \sim N(0,1)\ i=1,2,\cdots,n$ 独立
t 统计量	$t = \dfrac{X}{\sqrt{Y/n}}$	$\sim t(n)$	$X \sim N(0,1),\ Y \sim \chi^2(n),$ 独立
样本均值	$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$	$\sim N(\mu, \dfrac{\sigma^2}{n})$ √	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
样本方差	$\dfrac{(n-1)S^2}{\sigma^2}$	$\sim \chi^2(n-1)$ √	$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\sim t(n-1)$	

小结

常用统计量及抽样分布　$X \sim N(\mu, \sigma^2)$　$X_1, X_2, \cdots, X_n$

	统计量	概率分布	
χ^2统计量	$\chi^2 = \sum_{i=1}^{n} X_i^2$	$\sim \chi^2(n)$	$X_i \sim N(0,1)$ $i = 1,2,\cdots,n$ 独立
t 统计量	$t = \dfrac{X}{\sqrt{Y/n}}$	$\sim t(n)$	$X \sim N(0,1),\ Y \sim \chi^2(n)$, 独立
样本均值	$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$	$\sim N(\mu, \dfrac{\sigma^2}{n})$	$\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
样本方差	$\dfrac{(n-1)S^2}{\sigma^2}$	$\sim \chi^2(n-1)$	$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ √
	$\dfrac{\bar{X} - \mu}{S/\sqrt{n}}$	$\sim t(n-1)$ √	

参数估计问题

复习

设总体 $X \sim F(x,\theta)$, θ 是未知参数,

对未知参数 θ 进行估计 ---- 参数估计

$$E(X)=\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x$$

方法:

$EX_i = EX$　　$DX_i = DX$

1) 抽样: $X_1, X_2, \cdots, X_n$ (独立且与X同分布)

$x_1, x_2, \cdots, x_n$

不含任何未知量

2) 构造统计量: $g(X_1, X_2, \cdots, X_n)$ --- θ 的估计量

$g(x_1, x_2, \cdots, x_n)$ --- θ 的估计值

$\hat{\theta}$

参数估计
- 点估计: 给出 θ 的估计值或近似值 $\hat{\theta}$
- 区间估计: 给出 θ 的估计值或近似值 $\hat{\theta}$ 的误差范围与可信程度。

参数估计问题

复习

设总体 $X \sim F(x,\theta)$, θ 是未知参数,

对未知参数 θ进行估计 ---- 参数估计

方法:

1) 抽样: $X_1, X_2, \cdots, X_n$ (独立且与X同分布)

$x_1, x_2, \cdots, x_n$

2) 构造统计量: $g(X_1, X_2, \cdots, X_n)$ --- θ的估计量

$g(x_1, x_2, \cdots, x_n)$ --- θ的估计值

（二者均记为 $\hat{\theta}$）

参数估计
- 点估计:
  - 矩估计法: θ 的矩估计量 § 7.1
  - 极大似然估计法: θ 的极大似然估计量 § 7.1
- 区间估计: $\hat{\theta}$ 的误差范围与可信程度 § 7.4, § 7.5

➢ 需要讨论估计量的评价标准 § 7.3

小结

统计量的无偏性

设$X_1, X_2, \ldots, X_n$ 是总体 X 的一个样本

样本均值：$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$	$E(X)$	$E(\bar{X})=E(X)$
样本方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$	$D(X)$	$E(S^2)=D(X)$
样本 k 阶矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$	$E(X^k)$	$E(A_k)=E(X^k)$

结论：$\bar{X}, S^2, A_k$ 分别是 $E(X), D(X), E(X^k)$ 的无偏估计量.

参数估计问题

复习

设总体 $X \sim F(x,\theta), \theta$ 是未知参数，

对未知参数 θ 进行估计 ---- 参数估计

方法： $EX_i = EX$ $DX_i = DX$

1) 抽样： $X_1, X_2, \cdots, X_n$ (独立且与 X 同分布)

$x_1, x_2, \cdots, x_n$

2) 构造统计量： $g(X_1, X_2, \cdots, X_n)$ --- θ 的估计量

$g(x_1, x_2, \cdots, x_n)$ --- θ 的估计值

$\hat{\theta}$

参数估计

点估计: 给出 θ 的估计值或近似值 $\hat{\theta}$

区间估计: 给出 θ 的估计值或近似值 $\hat{\theta}$ 的误差范围及可信程度。

参数估计问题

复习

设总体 $X \sim F(x,\theta)$, θ 是未知参数,

对未知参数 θ进行估计 ---- 参数估计

方法:

1) 抽样: $X_1, X_2, \cdots, X_n$ (独立且与X同分布)

$x_1, x_2, \cdots, x_n$

2) 构造统计量: $g(X_1, X_2, \cdots, X_n)$ --- θ的估计量

$g(x_1, x_2, \cdots, x_n)$ --- θ的估计值

$\hat{\theta}$

参数估计
- 点估计:
  - 矩估计法: θ 的矩估计量
  - 极大似然估计法: θ 的极大似然估计量
- 区间估计: $\hat{\theta}$ 的误差范围与可信程度

➢ 需要讨论估计量的评价标准: 1. 无偏性 2. 有效性

复习

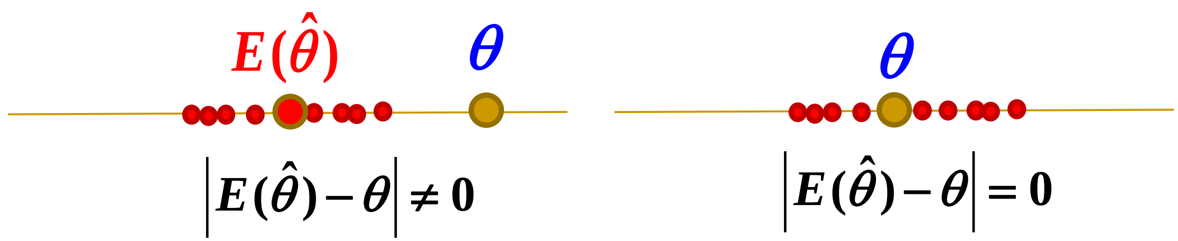
一. 无偏性

设未知参数 θ 的估计量和估计值分别为：

$$\hat{\theta}=\hat{\theta}\,(X_1,X_2\cdots,X_n)\qquad \hat{\theta}=\hat{\theta}\,(x_1,x_2\cdots,x_n)$$

估计量是随机变量，对于不同的样本值会得到不同的估计值。

希望估计值在未知参数真值 θ 附近摆动，即它的期望值等于未知参数的真值 $E(\hat{\theta})=\theta$ 。 $\left|E(\hat{\theta})-\theta\right|$ 系统误差



定义：设 $\hat{\theta}$ 是 θ 的估计量，若 $E(\hat{\theta})$ 存在且 $E(\hat{\theta})=\theta$

则称 $\hat{\theta}$ 是 θ 的无偏估计量.

复习

统计量的无偏性

设 $X \sim F(x,\theta)$, $X_1, X_2, \ldots, X_n$ 是总体 X 的一个样本

样本均值：$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$	$E(X)$	$E(\bar{X}) = E(X)$
样本方差：$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$	$D(X)$	$E(S^2) = D(X)$
样本 k 阶矩：$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$	$E(X^k)$	$E(A_k) = E(X^k)$

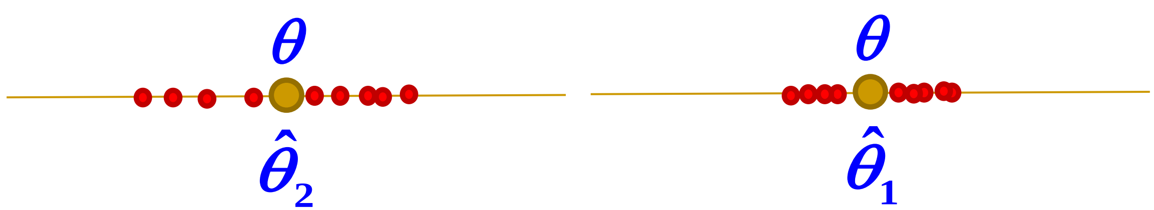
结论: $\bar{X}, S^2, A_k$ 分别是 $E(X), D(X), E(X^k)$ 的无偏估计量.

复习

二. 有效性

若 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是参数 θ 的无偏估计量，

则可通过比较 $\overset{D(\hat{\theta}_1)}{E(\hat{\theta}_1-\theta)^2}$ 和 $\overset{D(\hat{\theta}_2)}{E(\hat{\theta}_2-\theta)^2}$ 的大小来决定二者谁更优 。



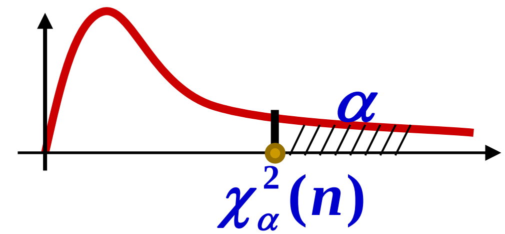
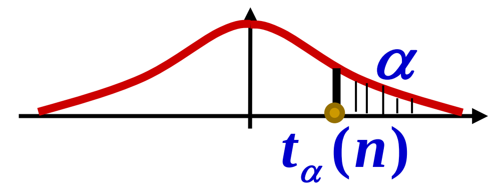
$E\hat{\theta}_2=\theta$ $\qquad$ $E\hat{\theta}_1=\theta$

定义：设 $\hat{\theta}_1=\hat{\theta}_1(X_1,X_2,\cdots,X_n)$ 与 $\hat{\theta}_2=\hat{\theta}_2(X_1,X_2,\cdots,X_n)$

都是 θ 的无偏估计量，且两个样本的容量相等。

若 $D(\hat{\theta}_1)\le D(\hat{\theta}_2)$，则称 $\hat{\theta}_1$ 较 $\hat{\theta}_2$ 有效。

常用统计量及抽样分布

χ^2 分布	$X_1, X_2, \cdots, X_n \sim N(0,1)$, 独立 $\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$	
t 分布	$X \sim N(0,1), Y \sim \chi^2(n)$, 独立 $t = \dfrac{X}{\sqrt{Y/n}} \sim t(n)$	
$EX = \mu, DX = \sigma^2$ $X \sim N(\mu, \sigma^2)$ $X_1, X_2, \cdots, X_n$ $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$	**Th1** $\bar{X} \sim N(\mu, \sigma^2/n)$, 独立 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$	**Th2** $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

小结

$X \sim N(\mu, \sigma^2)$，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	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$(\bar{X} \pm \dfrac{\sigma}{\sqrt{n}} z_{\alpha/2})$

$x_1, x_2, \cdots, x_{16}$

例1. 某实验室测量铝的比重 16 次，得平均值 $\bar{x} = 2.705$,

设总体 $X \sim N(\mu, 0.029^2)$ ，求：μ 的 95% 的置信区间.

解: 由已知：$\because\ 1-\alpha = 95\%$ $\quad \therefore\ \alpha = 5\%$,

$$\left(\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$$

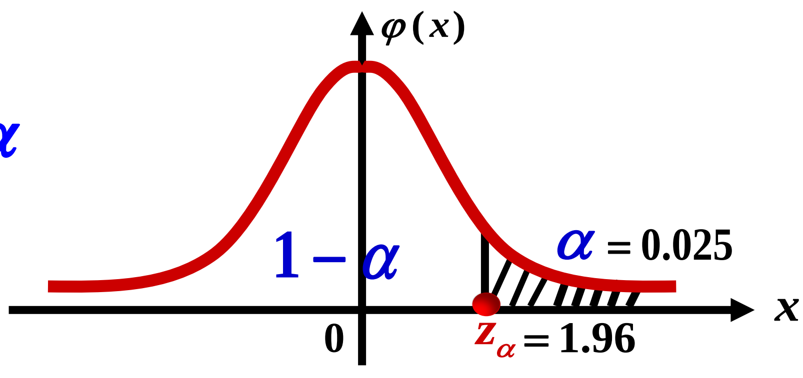
查正态分布表得：$z_{\alpha/2} = z_{0.05/2} = z_{0.025}$

$$\Phi(z_{0.025}) = 1 - 0.025 = 0.975 \qquad z_{0.025} = 1.96$$

复习

$$\Phi(z_\alpha) = \int_{-\infty}^{z_\alpha} \varphi(x)\mathrm{d}x = 1-\alpha$$

（附表2上）



小结

$X \sim N(\mu,\sigma^2)$, 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	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$(\bar{X} \pm \dfrac{\sigma}{\sqrt{n}} z_{\alpha/2})$
2)求 μ的置信区间 σ^2为未知	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$(\bar{X} \pm \dfrac{S}{\sqrt{n}} t_{\alpha/2}(n-1))$

小结

$X \sim N(\mu,\sigma^2)$，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	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$(\bar{X} \pm \dfrac{\sigma}{\sqrt{n}} z_{\alpha/2})$
2)求 μ 的置信区间 σ^2 为未知	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$(\bar{X} \pm \dfrac{S}{\sqrt{n}} t_{\alpha/2}(n-1))$
3)求 σ^2 的置信区间	$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$

小结

总体 $X \sim F(x,\theta)$，对 θ 进行估计　$X_1 X_2,\cdots,X_n$　$x_1 x_2,\cdots,x_n$

点估计	统计量 $\hat{\theta}=\hat{\theta}(X_1,X_2,\cdots,X_n)\to\theta$ 的估计量 ★1)矩估计法：　求解：$E(X^k)=A_k,\ k=1,2$ ★2)极大似然估计法：　求解：$L(\hat{\theta})=\max\limits_{\theta\in H}L(\theta)$
估计量的评选标准	★1)无偏性：$E(\hat{\theta})=\theta$ 2)有效性：$D(\hat{\theta}_1)<D(\hat{\theta}_2)$
★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

小结

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

小结

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

小结

$X \sim N(\mu,\sigma^2)$, 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	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$(\bar{X} \pm \dfrac{\sigma}{\sqrt{n}} z_{\alpha/2})$
2) 求 μ的置信区间 σ^2为未知	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$(\bar{X} \pm \dfrac{S}{\sqrt{n}} t_{\alpha/2}(n-1))$

复习

统计量的无偏性

设 $X_1, X_2, \ldots, X_n$ 是总体 X 的一个样本

样本均值：$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$	$E(X)$	$E\bar{X}=EX$
样本方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$	$D(X)$	$ES^2=DX$
样本 k 阶矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$	$E(X^k)$	$EA_k=E(X^k)$

结论：$\bar{X}, S^2, A_k$ 分别是 $E(X), D(X), E(X^k)$ 的无偏估计量.