第1次课主要内容

随机试验 $\xrightarrow{\text{定义}}$ 样本空间 $\xrightarrow{\text{子集}}$ 随机事件

- 基本事件
- 复杂事件
- 必然事件
- 不可能事件

$\longrightarrow$ 事件间的关系

- 子事件
- 和事件
- 积事件
- 互斥事件
- 对立事件
- 完备事件组事件

$\longrightarrow$ 事件的运算

- 交换律
- 结合律
- 分配律
- 对偶律
- 吸收律

都，都不，不都；至少，至多，恰好

第2次课主要内容

1.概率的定义方法

描述性定义　主观定义

统计定义　古典定义　几何定义

公理化定义

(1) 非负性 $P(A)\geq 0$;

(2) 规范性 $P(\Omega)=1$;

(3) 可列可加性 $P\left(\sum_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} P(A_i)$

2.概率的性质

性质1. $P(\varnothing)=0$.

性质2. (有限可加性)$P(\sum_{i=1}^{n} A_i)=\sum_{i=1}^{n} P(A_i)$.

性质3. $P(A)=1-P(\bar{A})$.

性质4. 若$A\subset B$，则$P(B-A)=P(B)-P(A)$.

$P(B-A)=P(B-AB)=P(B)-P(AB)$. (减法公式).

$P(B\bar{A})=P(B)-P(AB)$.

性质5. (单调性)若$A\subset B$,则$P(A)\le P(B)$.

性质6. (加法公式)$P(A\cup B)=P(A)+P(B)-P(AB)$.

第3次课主要内容

1.条件概率

$$P(B|A)=\frac{P(AB)}{P(A)}\quad (P(A)>0).$$

什么时候用？

怎么用？ (1)用定义$P(B|A)=\dfrac{P(AB)}{P(A)}, P(A)>0.$

(2)缩减样本空间.

乘法公式

$$P(AB)=P(A)P(B|A),\ \ (P(A)>0)$$

$$P(AB)=P(B)P(A|B),\ \ (P(B)>0).$$

$$P(AB)=P(A)+P(B)-P(A\cup B)$$

$$P(AB)=P(A)-P(A\bar{B})$$

全概率公式

$$P(B)=\sum_{i=1}^{n}P(A_iB)=\sum_{i=1}^{n}P(A_i)P(B|A_i),\ \left(B\subset\sum_{i=1}^{n}A_i\right).$$

贝叶斯公式

$$P(A_k|B)=\frac{P(A_kB)}{P(B)}=\frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)},\ \left(B\subset\sum_{i=1}^{n}A_i\right)$$

第4次课主要内容

1.两事件的独立性

A,B 独立 $\Leftrightarrow$ $P(AB)=P(A)P(B)$; “没概率,没独立”

$\overset{P(A)>0}{\Leftrightarrow}$ $P(B|A)=P(B)$; A,B独立指的是一事件的概率

$\overset{P(B)>0}{\Leftrightarrow}$ $P(A|B)=P(A)$; 与另一事件发生与否没有关系.

$\Leftrightarrow$ $A,\bar{B}$独立 $\Leftrightarrow$ $\bar{A},B$独立 $\Leftrightarrow$ $\bar{A},\bar{B}$独立

独立时,*bar*不*bar*没关系.

2.多个事件的独立性

$$A_1,A_2,A_3 \text{ 两两独立} \Leftrightarrow \left\{\begin{array}{l} P(A_1A_2)=P(A_1)P(A_2) \\ P(A_1A_3)=P(A_1)P(A_3) \\ P(A_2A_3)=P(A_2)P(A_3) \end{array}\right.$$

$$\left.\begin{array}{l} P(A_1A_2)=P(A_1)P(A_2) \\ P(A_1A_3)=P(A_1)P(A_3) \\ P(A_2A_3)=P(A_2)P(A_3) \\ P(A_1A_2A_3)=P(A_1)P(A_2)PA_3) \end{array}\right\} \Leftrightarrow A_1,A_2,A_3 \text{ 相互独立}$$

性质1 若$A_1,A_2,\cdots,A_n$相互独立, 则将其中一部分改为对立事件, 它们仍相互独立.

独立时,*bar*不*bar*没关系.

性质2 若$A_1,A_2,\cdots,A_n$相互独立, 则将他们任意分成两组(或多组), 组内施行“并,交,差,补”所得结果也独立.

3.独立重复试验

试验独立　　n重独立试验　伯努利试验　n重伯努利试验

n重伯努力试验中，A恰好发生k次的概率记为$P_n(k)$，则

$$P_n(k) = C_n^k p^k (1-p)^{n-k},\quad k = 0,1,2,\cdots,n.$$

多重伯努力试验中，A首次发生在第k次的概率记为$G(k)$，则

$$G(k) = p(1-p)^{k-1},\quad k = 1,2,3,\cdots$$

多重伯努力试验中，A第r次发生在第k次的概率记为$G_r(k)$，则

$$G_r(k) = C_{k-1}^{r-1} p^r (1-p)^{k-r},\quad k = r, r+1, r+2, \cdots$$

概率统计第一条主线

随机试验 ® 样本点 样本空间 ® 随机事件 关系 运算 ¾¾¾¾¾® 古典概型 几何概型 ® 概率 公理定义 性质 **五大公式** **条件概率** **事件独立性** ® 伯努利 **二项** 几何 负二项

第5次课主要内容

1.随机变量、分布函数

$$\{X \le x\} = \{\omega \mid X(\omega) \le x, \omega \in \Omega\},$$

$$F(x) = P\{X \le x\}, \quad x \in R$$

充要条件

2.有关X的各种事件的概率

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

$$P\{X < x\} = F(x - 0)$$

第6次课主要内容

1.离散型随机变量的分布律

$p_i = P\{X=a_i\}$， $i = 1, 2, \ldots$

X	a_1	a_2	$\ldots$	a_n	$\ldots$
P	p_1	p_2	$\ldots$	p_n	$\ldots$

性质：(1)非负性 $p_i \geq 0, i = 1,2,3,\cdots$； (2)正则性 $\sum_{i=1}^{+\infty} p_i = 1.$

与$F(x)$的关系

$$F(x) = P\{X \leq x\} = \sum_{a_i \leq x} p_i, \quad x \in R$$

$$p_i = P\{X = a_i\} = P\{X \leq a_i\} - P\{X < a_i\} = F(a_i) - F(a_i - 0)$$

2.常见的离散型随机变量的分布律

0--1分布 $X \sim B(1,p)$. 一次伯努利试验

$$P\{X=k\}=p^k(1-p)^{1-k},\ k=0,1,\ 0<p<1.$$

X	0	1
P	$1-p$	p

二项分布 $X \sim B(n,p)$. n重伯努利成功次数X

$$P\{X=k\}=C_n^k p^k(1-p)^{n-k},\ k=0,1,\cdots,n,\ 0<p<1.$$

几何分布 $X \sim G(p)$. 多重伯努利首次成功时试验次数X

$$P\{X=k\}=p(1-p)^{k-1},\ \ k=1,2,\cdots,\ \ 0<p<1.$$

泊松分布 $X \sim P(\lambda)$ 某时间段内某事件发生的次数X

$$P\{X=k\}=e^{-\lambda}\frac{\lambda^k}{k!},\ \ k=0,1,\cdots,\quad \lambda>0.$$

第7课主要内容

1.连续型随机变量的密度

$$f(x) \underset{\text{不唯一}}{\overset{\text{唯一}}{\underset{\leftarrow}{\rightarrow}}} F(x)$$

$X,\ f(x) \geq 0$,满足$F(x)=\int_{-\infty}^{x} f(t)dt,\ x\in R$,

车续型$r.v.$　密度.　连续型分布函数.几何

$(1)\ f(x)\geq 0,\ x\in R;$

$(2)\int_{-\infty}^{+\infty} f(x)dx=1;$

密度函数的充要条件

$(3)P\{a<X\leq b\}=\int_{a}^{b} f(x)dx;$ 哪求概率,哪求积分.

$(4)F(x)$是x的连续函数;$P\{X=a\}=0$. $f(x)$不一定连续

(5)在$f(x)$的可微点x处有$F'(x)=f(x)$.

2.常见的连续型随机变量的密度

(1)均匀分布 $X \sim U[a,b]$ 特殊几何概型

$$f(x)=\begin{cases}1/(b-a), & x\in[a,b]\\ 0, & \text{其它}\end{cases}\qquad F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\le x<b\\ 1, & b\le x\end{cases}$$

①不积分，量长度

(2)指数分布 $X \sim \underset{}{\overset{\Gamma(1,\lambda)}{E(\lambda)}}$ 寿命 某事件两次发生等待的时间

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

①无记忆性 $P\{X>s+t\mid X>s\}=P\{X>t\},\ \forall s,t\ge 0.$

②$P\{X>t\}=\int_t^{+\infty}\lambda e^{-\lambda x}dx=e^{-\lambda t},\ \forall t\ge 0.$

(3)正态分布 $X \sim N(\mu, \sigma^2)$ 众多微小因素影响的量

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

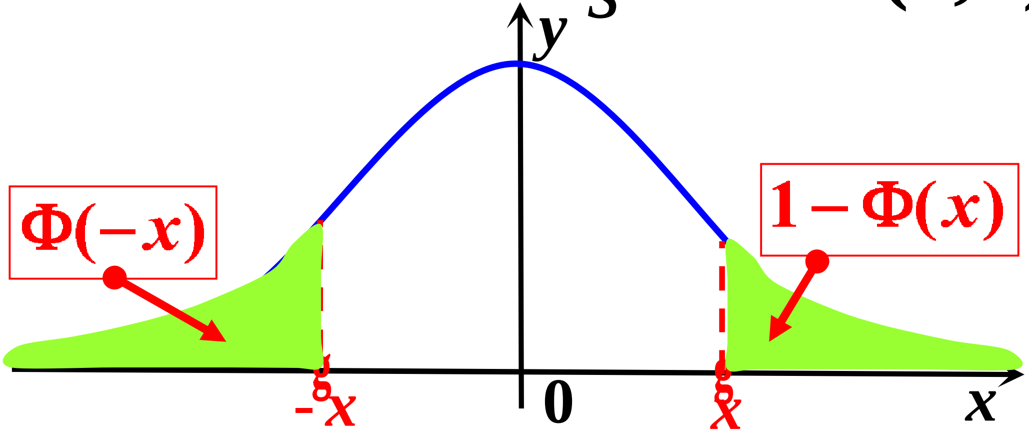
$F(x)$无法用初等函数表示

标准正态分布 $X \sim N(0,1)$ $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $\Phi(x)$

①$f(x)$关于$x=\mu$对称；

②$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$, $X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$.

③$\Phi(-x) = 1 - \Phi(x)$;



$X \sim N(0,1)$, $P\{|X| \le x\} = 2\Phi(x) - 1$.

第8次课主要内容随机变量函数的分布

已知$r.v.X$的$F_X(x)$ 求$r.v.Y=g(X)$的$F_Y(x)$. ($g(x)$连续)

离散型 列表法 分析法

连续型 定义法 已知$f_X(x),F_X(x)$,求$Y=f(X)$的分布.

$$F_Y(x)=P\{Y\le x\}=P\{f(X)\le x\}=P\{X\le \ ?\ \}=F_X(\ ?\ )$$

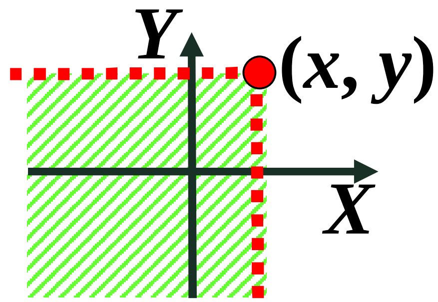
$$f_{aX+b}(x)=\frac{1}{|a|}f_X\left(\frac{x-b}{a}\right),\ x\in R.$$

$$f_{X^2}(x)=\begin{cases}\frac{1}{2\sqrt{x}}\left(f_X(\sqrt{x})+f_X(-\sqrt{x})\right), & x>0\\ 0, & x\le 0\end{cases}$$

公式法 $$f_Y(y)=\begin{cases}|h'(y)|\,f_X(h(y)), & a<y<b\\ 0, & \text{其它}\end{cases}$$

第9次课主要内容

1.二维$r.v.$及其分布函数



二维$r.v.$ (X,Y)

联合分布函数 $F(x,y)=P\{X\le x, Y\le y\}$

$$P\{x_1<X\le x_2, y_1<Y\le y_2\}$$

$$=F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1)$$

$F(x,y)$的性质

①有界性$F(-\infty,y)=F(x,-\infty)=F(-\infty,-\infty)=0,$

$$F(x,+\infty)=F_X(x),\ F(+\infty,y)=F_Y(y),\ F(+\infty,+\infty)=1.$$

②单调性 ③右连续性

2.离散型二维$r.v.$及其分布律

离散型二维$r.v.$ (X,Y)的可能取值有限或可列

联合分布律 $P(X=x_i, Y=y_j)=p_{ij}$， $i, j=1, 2, \ldots,$

联合分布律的性质

(1) 非负性 p_{ij} ³ 0, $i, j = 1, 2, \ldots$

(2) 正则性 SS $p_{ij} = 1.$

3.连续型二维*r.v.*及密度

(X, Y)， $F(x, y)$，若非负可积有 $f(x, y)$，使

$$F(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dvdu$$ 几何意义

则称 (X, Y) 为二维连续型*r.v.*； $f(x, y)$ 为联合密度函数。

(X,Y)的密度的性质

① $f(x,y) \geq 0, (x,y) \in R^2$; ② $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$;

③在(x, y)处连续， $f(x,y) = \dfrac{\partial^2 F(x,y)}{\partial x \partial y}$;

④$P\{(X,Y) \in G\} = \iint\limits_{G} f(x,y)dxdy$;

第10次课主要内容

1. $(X, Y) \sim U(D)$.

不积分，算面积

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

2. n维随机变量:$(X_1, X_2,\ldots,X_n)$

(联合)分布函数： $F(x_1, \ldots, x_n) = P\{ X_1 \le x_1, \ldots, X_n \le x_n\}$

联合分布律：$P\{X_1 = a_{1,k_1}, \cdots, X_n = a_{n,k_n}\}$.

$f(x_1,\cdots,x_n)$为$(X_1,\cdots,X_n)$的联合密度函数 $\Leftrightarrow$

$$F(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f(t_1,t_2,\cdots,t_n)dt_1dt_2\cdots dt_n.$$

3.边缘分布 $(X,Y), F(x,y),$

边缘分布函数 $F_X(x) = P\{X \le x\} = F(x,+\infty);$

$$F_Y(y) = P\{Y \le y\} = F(+\infty, y).$$

边缘分布律： $P\{X = a_i\} = \sum_{j=1}^{\infty} p_{ij} = p_{i\cdot};$

$$P\{Y = b_j\} = \sum_{i=1}^{\infty} p_{ij} = p_{\cdot j};$$

边缘密度: $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy;$

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

4随机变量的独立性

X与Y相互**独立** $\Leftrightarrow F(x,y)=F_X(x)F_Y(y), \quad \forall x,y\in R;$

$\Leftrightarrow p_{ij}=p_{i\cdot}p_{\cdot j}, \quad i,j=1,2,\cdots;$

$\Leftrightarrow f(x,y)=f_X(x)f_Y(y), \quad \forall x,y\in R$

X与Y独立，则$g(X)$与$h(Y)$也独立.

多维类似.

联合密度 $\overset{独立}{\Longleftarrow}$ $\Rightarrow$ $\not\Leftarrow$ 边缘密度

第11次课主要内容

1.离散型$r.v.$的条件分布

在$Y=b_j$条件下, $r.v.X$的**条件分布律**和**条件分布函数**为

$$P\{X=a_i|Y=b_j\}=\frac{P\{X=a_i,Y=b_j\}}{P\{Y=b_j\}}=\frac{p_{ij}}{p_{\cdot j}},\quad i=1,2,\ldots$$

$$F_{X|Y}(x|b_j)=P\{X\le x|Y=b_j\}=\sum_{a_i\le x}P\{X=a_i|Y=b_j\},x\in R$$

在$X=a_i$条件下,$r.v.Y$的**条件分布律**和**条件分布函数**为

$$P\{Y=b_j|X=a_i\}=\frac{P\{X=a_i,Y=b_j\}}{P\{X=a_i\}}=\frac{p_{ij}}{p_{i\cdot}},\quad j=1,2,\ldots$$

$$F_{Y|X}(y|a_i)=P\{Y\le y|X=a_i\}=\sum_{b_j\le y}P\{Y=b_j|X=a_i\},y\in R$$

2.连续型*r.v.*的条件分布

在$Y = y$下X的**条件分布函数**：$F_{X|Y}(x|y) = \int_{-\infty}^{x} \frac{f(u,y)}{f_Y(y)}du$

条件密度：$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$， $x \in R$, $f_Y(y) > 0$.

在$X = x$下Y的**条件分布函数**：$F_{Y|X}(y|x) = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)}dv$

条件密度：$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$， $y \in R$, $f_X(x) > 0$.

注： 联合密度、边缘密度和条件密度三者知二求一.

第12次课主要内容

二维离散型*r.v.*函数的分布 列表法,分析法

二维连续型*r.v.*函数的分布

1.极值函数的分布

X, Y独立, $Z_1 = \max\{X, Y\}$和$Z_2 = \min\{X, Y\}$

$$F_{Z_1}(z) = F_X(z)F_Y(z);$$

$$f_{Z_1}(z) = F'_{Z_1}(z) = f_X(z)F_Y(z) + F_X(z)f_Y(z)$$

$$F_{Z_2}(z) = 1 - (1 - F_X(z))(1 - F_Y(z))$$

$$f_{Z_2}(z) = F'_{Z_2}(z) = f_X(z)(1 - F_Y(z)) + (1 - F_X(z))f_Y(z)$$

1.极值函数的分布(推广)

若$X_1, \cdots, X_n$独立同分布

记$Z_{(n)} = \max(X_1, \cdots, X_n)$, $Z_{(1)} = \min(X_1, \cdots, X_n)$,

$$F_{Z_{(n)}}(z) = \left[F(z)\right]^n,$$

$$f_{Z_{(n)}}(z) = F'_{Z_{(n)}}(z) = n[F(z)]^{n-1} f(z);$$

$$F_{Z_{(1)}}(z) = 1 - \left[1 - F(x)\right]^n$$

$$f_{Z_{(1)}}(z) = F'_{Z_{(1)}}(z) = n[1 - F(z)]^{n-1} f(z).$$

2.和函数$X+Y$的分布

已知(X,Y)的$f(x,y)$,求$Z=X+Y$的密度.

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)dx=\int_{-\infty}^{+\infty}f(z-y,y)dy,\ z\in R.$$

卷积公式

若X,Y相互独立,

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

同理可求$Z=aX+bY(a\neq 0,b\neq 0)$的密度.

$$f_Z(z)=\frac{1}{|b|}\int_{-\infty}^{+\infty}f\left(x,\frac{z-ax}{b}\right)dx=\frac{1}{|a|}\int_{-\infty}^{+\infty}f\left(\frac{z-by}{a},y\right)dy,\ z\in R$$

第13次课主要内容

1.公式法求$Z=X+Y$的密度的步骤

1.背公式($f(x,y)$中只含x就对x积分)

$$f_Z(z)=\int_{-\infty}^{+\infty}f(x,z-x)\,dx=\int_{-\infty}^{+\infty}f(z-y,y)\,dy,\ z\in R.$$

2.画出非零密度区域,确定z的分界点,

(注意坐标轴的变化,对x积分就是画x,z轴)

(分界点就是非零密度区域顶点对应在z轴上的点)

3.分区间积分. (对谁积分,画平行谁的直线,定积分上下限)

类似可做

$$f_{Z=aX+BY}(z)=\frac{1}{|b|}\int_{-\infty}^{+\infty}f\left(x,\frac{z-ax}{b}\right)dx=\frac{1}{|a|}\int_{-\infty}^{+\infty}f\left(\frac{z-by}{a},y\right)dy,\ z\in R$$

2.定义法求$Z=g(X+Y)$的密度

①$F_Z(z)=P\{g(X,Y)\le z\}$

先画出$f(x,y)$的非负区域,再画满足$g(x,y)\le z$的积分区域,

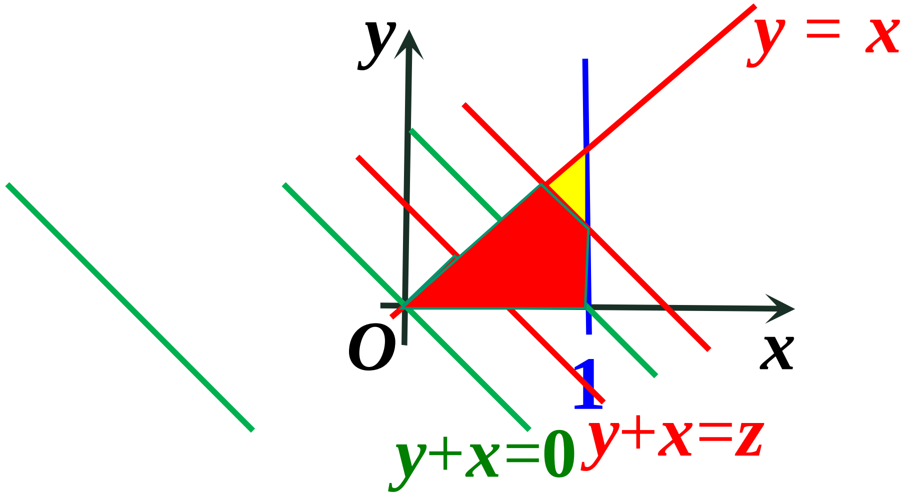
不同的z积分区域与非负密度区域的交集不同,

突变处就是z的分界点.

分区间化二重积分为二次积分

$$\iint\limits_{x+y\le z} f(x,y)dxdy$$

②$f_Z(z)=F'_Z(z)$.



3.二维正态分布

$\left(X_1, X_2\right) \sim N\left(m_1, m_2; s_1^2, s_2^2; r\right)$ 二维正态分布

则$X_1 \sim N\left(m_1, s_1^2\right)$, $X_2 \sim N\left(m_2, s_2^2\right)$.

X_1, X_2独立 Û $r = 0$;

当b_1, b_2 ¹ 0时,线性组合$Z_3 = b_1 X_1 + b_2 X_2 + b$服从正态分布.

Z_1, Z_2独立且都服从正态分布,则(Z_1, Z_2)服从二维正态分布.

第14次课主要内容

1.数学期望的定义

数学期望 = 可能的取值与其概率乘积之和

$$EX=\sum_{k=1}^{\infty}x_k p_k \qquad EX=\int_{-\infty}^{\infty}x f(x)dx$$

$$E[g(X)]=\begin{cases}\displaystyle\sum_{i=1}^{\infty}g(a_i)p_i, & X\text{离散型},\\ \displaystyle\int_{-\infty}^{\infty}g(x)f(x)dx, & X\text{连续型}.\end{cases}$$

$$E[g(X,Y)]=\begin{cases}\displaystyle\sum_{i=1}^{+\infty}\sum_{j=1}^{+\infty}g(a_i,b_j)p_{ij}, & X\text{离散型},\\ \displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)dxdy, & X\text{连续型}.\end{cases}$$

2.常见分布的数学期望

X	$B(n,p)$	$P(\lambda)$	$G(p)$	$U[a,b]$	$\Gamma(1,\lambda)$	$N(\mu,\sigma^2)$
EX	np	λ	$1/p$	$(a+b)/2$	$1/\lambda$	μ

3.数学期望的性质

1.设c是常数,则有$Ec=c$;

2.设X是$r.v.$,a,b是常数,则$E(aX+b)=aEX+b$;

3.设X,Y是两个$r.v.$,则$E(aX+bY)=aEX+bEY$;

4.若X,Y独立,则$E(XY)=EX\cdot EY$.

第15次课主要内容

1.方差的定义

$\boxed{DX = E(X - EX)^2}$ $\boxed{DX = E(X^2) - (EX)^2}$

$\sqrt{DX}$为X的**标准差**或**均方差**,记作s .

2.常见分布的期望和方差

X	$B(n,p)$	$P(l)$	$G(p)$	$U[a,b]$	$G(1,l)$	$N(m,s^2)$
EX	np	l	$1/p$	$(a+b)/2$	$1/l$	m
DX	npq	l	q/p^2	$(b-a)^2/12$	$1/l^2$	s^2

3.方差的性质

1.设c是常数,则有$Dc = 0$;

2.a,b是$\forall$常数,则$D(aX + b) = a^2 DX$;

3.$D(X \pm Y) = DX + DY \pm 2E[(X - EX)(Y - EY)]$;

若X,Y独立,则$D(X \pm Y) = DX + DY$;

4.$\forall x \in R$,由$DX = E(X - EX)^2 \le E(X - x)^2$,

且等号成立 $\Leftrightarrow x = EX$;

5.(切比雪夫不等式)设DX存在,则$\forall \varepsilon \in R^+$,有

$$P\{|X - EX| \ge \varepsilon\} \le DX / \varepsilon^2$$

$$P\{|X - EX| < \varepsilon\} \ge 1 - DX / \varepsilon^2$$

6.$DX = 0 \Leftrightarrow P\{X = EX\} = 1$.

第16次课主要内容

1.协方差和相关系数的概念

协方差

$$\mathrm{cov}(X,Y)=E\left[(X-EX)(Y-EY)\right]$$

$$\mathrm{cov}(X,Y)=E(XY)-EX\cdot EY$$

相关系数

$$r(X,Y)=\mathrm{cov}(X^*,Y^*)$$

$$r(X,Y)=\frac{\mathrm{cov}(X,Y)}{\sqrt{DX\cdot DY}}$$

$$X^*=\frac{X-EX}{\sqrt{DX}},$$

$$Y^*=\frac{Y-EY}{\sqrt{DY}},$$

当$r(X,Y)=\pm1$,称X与Y正(负)相关;

当$r(X,Y)=0$,称X与Y不(线性)相关.

2.协方差和相关系数的性质

(1)设c为常数,则$\operatorname{cov}(c, X) = 0$;

(2)$\operatorname{cov}(X, X) = DX$; (3)$\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$;

(4)设a,b为常数,则$\operatorname{cov}(aX, bY) = ab\operatorname{cov}(X, Y)$;

(5)$\operatorname{cov}(X+Y, Z) = \operatorname{cov}(X, Z) + \operatorname{cov}(Y, Z)$;

(6)$D(X \pm Y) = DX + DY \pm 2\operatorname{cov}(X, Y)$;

(1)$\left|\rho(X,Y)\right| \le 1$;(2)$\rho(X,Y) = \pm 1 \Leftrightarrow P\left\{\dfrac{X-EX}{\sqrt{DX}} = \pm\dfrac{Y-EY}{\sqrt{DY}}\right\} = 1$

相互独立 $\underset{\not\leftarrow}{\rightarrow}$ 不相关 $\Leftrightarrow \rho(X,Y) = 0 \Leftrightarrow \operatorname{cov}(X,Y) = 0$

$\Leftrightarrow E(XY) = EX \cdot EY$

$\Leftrightarrow D(X \pm Y) = DX + DY.$

第17次课主要内容

1.矩

X的k阶原点矩： $E(X^k)$

X的k阶中心矩： $E(X-EX)^k$

X与Y的$k+l$阶混合原点矩： $E(X^kY^l)$

X与Y的$k+l$阶混合中心矩： $E\left[(X-EX)^k(Y-EY)^l\right]$

2.随机变量序列收敛

$\{X_n\}$依概率收敛于X：$\forall \varepsilon \in R$有 $\lim\limits_{n\to+\infty} P\{|X_n - X| < \varepsilon\} = 1$

$$X_n \xrightarrow[n\to+\infty]{P} X.$$

$\{X_n\}$依分布收敛于X：$\forall x \in R$有 $\lim\limits_{x\to+\infty} F_n(x) = F(x)$

$$X_n \xrightarrow[n\to+\infty]{L} X.$$

3.大数定律

条件: $\{X_i\}$独立, EX_i, DX_i存在, DX_i有界

或者$\{X_i\}$独立且同分布, EX_i, DX_i存在,

结论: $\sum_{i=1}^{n} X_i \xrightarrow[n\to+\infty]{P} E\left(\sum_{i=1}^{n} X_i\right)$; $\bar{X}_n \xrightarrow[n\to+\infty]{P} E\bar{X}_n$.

4.中心极限定理

条件: $\{X_i\}$独立且同分布, EX_i, DX_i存在,

结论: $\sum_{i=1}^{n} X_i$ 依分布收敛于$N\left(E(\sum_{i=1}^{n} X_i), D(\sum_{i=1}^{n} X_i)\right)$;

$\bar{X}_n$ 依分布收敛于$N\left(E\bar{X}_n, D\bar{X}_n\right)$.

概率统计第二条主线

一维二维随机变量 离散® 分布律 连续 ® 密度 分布 ® 数字特征 ® 大数中心

非负性

归一性

独立性

联合

边缘

条件

函数的密度

常见分布

EX

DX

cov

r

第18次课主要内容

1.数理统计的基本概念

总体：研究对象全体 ® 数据 ® 某$r.v$的分布 X

样本：从总体中抽的部分个体 $X_1, X_2, \mathrm{L}, X_n$

样本容量：样本中个体的数量　　　**简单样本**

样本值：　样本的一次观测值 $x_1, x_2, \mathrm{L}, x_n$　**样本空间**

样本分布的计算 (注意样本是独立同分布的)

总体分布函数的猜测　经验分布函数

总体密度函数的猜测　　直方图

2.样本矩统计量及性质

(1)样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ $\quad \bar{X}=M_1 \quad S^2=\frac{n}{n-1}M_2^*$

(2)样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\bar{X}^2\right)$

样本标准方差 $S=\sqrt{S^2}$ $\quad M_2^*=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}^2$

(3)样本k阶原点矩 $M_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k;k=1,2,\cdots$

(4)样本k阶中心矩 $M_k^*=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k;k=1,2,\cdots$

$E\bar{X}=EX,\ D\bar{X}=\frac{1}{n}DX,\ EM_2^*=\frac{n-1}{n}DX,\ ES^2=DX.$

3.样本顺序统计量

顺序统计量：$X_{(1)}, X_{(2)}, \cdots, X_{(n)}$

最小顺序统计量：$X_{(1)}$

最小顺序统计量：$X_{(n)}$

$\left.\right\rangle$ 分布函数与密度的计算

样本极差：$R = X_{(n)} - X_{(1)}$

样本中位数：$\tilde{X} = \begin{cases} X_{(\frac{n+1}{2})}, & \text{当}n\text{为奇数时} \\ \dfrac{1}{2}\left(X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}\right), & \text{当}n\text{为偶数时} \end{cases}$

第19次课主要内容

1.三大抽样分布

1)卡方分布 $\chi^2 \sim \chi^2(n) \Rightarrow E\chi^2 = n, D\chi^2 = 2n$ 线性性质

独立的标准正态平方和 $X \sim N(0,1) \Rightarrow X^2 \sim \chi^2(1)$

2) t分布 密度对称 $n>1$时$ET=0$; $n>2$时$DT = n/(n-2)$.

独立、标准正态除以卡方与自由度比值开方

3) F分布 $F \sim F(m,n) \Rightarrow \dfrac{1}{F} \sim F(n,m)$; $T \sim t(n) \Rightarrow T^2 \sim F(1,n)$.

独立、两卡方与自由度比值的商

若$P\{X \pounds v_p\} = p$,则称v_p为X的(下侧) **p分位数**.

$$u_p = -u_{1-p} \qquad t_p(n) = -t_{1-p}(n) \qquad F_p(m,n) = \frac{1}{F_{1-p}(n,m)}$$

2.抽样分布定理单正态总体

若$X \sim N(\mu, \sigma^2)$, $X_1, X_2, \cdots, X_n$为总体X的样本,则

$$(1)\ \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right),$$

$$(2)\ \boxed{\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)}; \qquad (3)\ \bar{X}\text{与}S^2\text{独立}.$$

$$(4)\ \boxed{T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)}, \qquad (5)\ ES^2 = \sigma^2,\ DS^2 = \frac{2\sigma^4}{n-1}.$$

2.抽样分布定理双正态总体(了解)

若$X \sim N(\mu_1, \sigma_1^2), X_1, X_2, \cdots, X_m$为总体$X$的样本,

$Y \sim N(\mu_2, \sigma_2^2), Y_1, Y_2, \cdots, Y_n$为总体$Y$的样本,则

$$(1) F = \frac{S_X^2 / S_Y^2}{\sigma_1^2 / \sigma_2^2} \sim F(m-1, n-1);$$

$$(2)当\sigma_1^2 = \sigma_2^2 = \sigma^2时, T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t(m+n-2)$$

$$其中 S_w^2 = \frac{(m-1)S_X^2 + (n-1)S_Y^2}{m+n-2}.$$

第20次课主要内容

1.矩估计　样本矩等于总体矩(近似)。

$$(1)EX^k=\int_{-\infty}^{+\infty}x^k f(x,\theta)dx,\ \text{末时参数的个数}=m,$$

$$(2)\frac{1}{n}\sum_{i=1}^{n}x_i^k=EX^k=\int_{-\infty}^{+\infty}x^k f(x,\theta)dx,\quad k=1,2,L,m$$

(3) 求解以上m个方程组得到其解.

末知参数的个数＝2时

(1)计算EX,DX，(2) $\begin{cases}\bar{x}=EX\\ m_2^*=DX\end{cases}$　(3)解此方程组.

2.极大似然估计

使样本发生概率最大的参数值作为参数的估计值

(1) 求似然函数 $L(\theta)=\prod_{i=1}^{n} f(x_i,\theta)$ 或 $\prod_{i=1}^{n} P\{X=x_i\}$

(2) $\dfrac{\partial}{\partial \theta_i}L(\theta)=0,\quad i=1,2,\cdots,k,$　　似然方程组

为了计算方便,似然方程组常换为

$\dfrac{\partial}{\partial \theta_i}\ln L(\theta)=0,\ i=1,2,\cdots,k,$　　对数似然方程组

(3)解(对数)似然方程组,得极大似然估计$\hat{\theta}_1,\cdots,\hat{\theta}_k$.

第21次课主要内容

1.估计的评价

无偏估计：$E\hat{\theta}=\theta$，

渐进无偏估计：$\lim\limits_{n\to+\infty}E\hat{\theta}_n=\theta$，　**系统误差**：$E\hat{\theta}-\theta$.

设$\hat{\theta}_1$和$\hat{\theta}_2$都是θ的无偏估计，若$D\hat{\theta}_1<D\hat{\theta}_2$，称$\hat{\theta}_1$比$\hat{\theta}_2$**有效**.

最优无偏估计：无偏估计中方差最小者.

$\hat{\theta}_n$为θ的**相合估计**：$\hat{\theta}_n \xrightarrow[n\to+\infty]{P} \theta$.

$\lim\limits_{n\to\infty}E\hat{\theta}_n=\theta$，$\lim\limits_{n\to\infty}D\hat{\theta}_n=0 \Rightarrow \hat{\theta}_n$是$\theta$的相合估计。

2.区间估计

置信区间，置信度，置信下限，置信上限.

(1)单正态总体μ的置信度$1-\alpha$的区间估计

若σ^2已知:

$$\left(\bar{X}-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X}+u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right).$$

若σ^2未知:

$$\left(\bar{X}-t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}},\ \bar{X}+t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)$$

(2)单正态总体σ^2的置信度$1-\alpha$的区间估计

$$\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)},\ \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right).$$

第22次课主要内容

1.假设检验的两类错误

$$\alpha = P\{犯第一类错误\} = P\left\{拒绝H_0 \middle| H_0成立\right\}$$

$$\beta = P\{犯第二类错误\} = P\left\{接受H_0 \middle| H_0不成立\right\}$$

2.假设检验的步骤

1）提出原假设H_0与备择假设H_1；

2）选出统计量，给出的拒绝域的形式K_0；

3）求出临界值，确定拒绝域K_0；

4）作出是否拒绝H_0的判断。

H_0	H_1	拒绝域	条件	临界值
$\mu=\mu_0$	$\mu \neq \mu_0$	$\{\lvert\bar{x}-\mu_0\rvert > c\}$	σ^2已知	$c = u_{1-\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}$
			σ^2未知	$c = t_{1-\frac{\alpha}{2}}(n-1) \dfrac{S}{\sqrt{n}}$
$\mu=\mu_0$ $\mu \leq \mu_0$	$\mu > \mu_0$	$\{\bar{x}-\mu_0 > c\}$	σ^2已知	$c = u_{1-\alpha} \dfrac{\sigma}{\sqrt{n}}$
			σ^2未知	$c = t_{1-\alpha}(n-1) \dfrac{S}{\sqrt{n}}$
$\mu=\mu_0$ $\mu \geq \mu_0$	$\mu < \mu_0$	$\{\bar{x}-\mu_0 < c\}$	σ^2已知	$c = u_{\alpha} \dfrac{\sigma}{\sqrt{n}}$
			σ^2未知	$c = t_{\alpha}(n-1) \dfrac{S}{\sqrt{n}}$

H_0	H_1	拒绝域	临界值
$\sigma^2=\sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\left\{\frac{s^2}{\sigma_0^2}>c_2\text{或}\frac{s^2}{\sigma_0^2}<c_1\right\}$	$c_1=\frac{1}{n-1}\chi^2_{\frac{\alpha}{2}}(n-1)$ $c_2=\frac{1}{n-1}\chi^2_{1-\frac{\alpha}{2}}(n-1)$
$\sigma^2=\sigma_0^2$ $\sigma^2 \leq \sigma_0^2$	$\sigma^2>\sigma_0^2$	$\left\{\frac{s^2}{\sigma_0^2}>c\right\}$	$c=\frac{1}{n-1}c^2_{1-a}(n-1)$
$\sigma^2=\sigma_0^2$ $\sigma^2 \geq \sigma_0^2$	$\sigma^2<\sigma_0^2$	$\left\{\frac{s^2}{\sigma_0^2}<c\right\}$	$c=\frac{1}{n-1}\chi^2_{\alpha}(n-1)$

概率统计第三条主线

总体 ® 样本 ® 统计量 ® 三大抽样分布 抽样分布定理 ® 参数估计 ® 假设检验

三层次 独立同分布

$\bar{X}$

S^2

M_2^*

$X_{(1)}$

$X_{(n)}$

1)卡方分布

2) t分布

3) F分布

$$\frac{(n-1)S^2}{s^2} \sim c^2(n-1);$$

$$\frac{\bar{X}-m}{S/\sqrt{n}} \sim t(n-1)$$

矩估计

极大似然

区间估计

单正态总体

m, s^2