目　　录

第一章 预备知识

§1.1 概率

§1.2 一元随机变量及其分布

§1.3 多元随机变量及其分布

§1.4 数字特征

§1.5 极限定理

§1.1 概率

一、定义：将概率视为测度，且满足：

1° $P(A) \geq 0$

2° $P(S) = 1$

3° $A_1, A_2, \ldots, A_k, \ldots, A_i A_j = \varnothing$ （i ≠ j），

$$\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

称 P(A) 为事件 A 的概率。

二、概率的性质

$1^\circ\ P(\varnothing)=0$

$2^\circ\ \ A_1, A_2, \ldots, A_n,\ A_iA_j=\varnothing,\ i\neq j,$

$$\Rightarrow P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)$$

$3^\circ\ P(A)=1-P(\bar{A})$

4° 若$A \subset B$，则有 $P(B-A)=P(B)-P(A)$

$\Rightarrow P(B) \geq P(A)$, 于是有 $P(A) \leq P(S)=1$

5° 概率的加法公式：$P(A \cup B)=P(A)+P(B)-P(AB)$

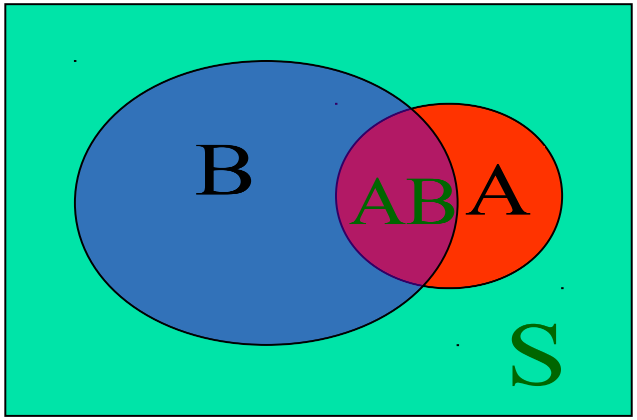
$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

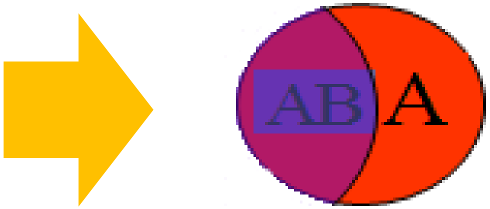
$$- P(AB) - P(AC) - P(BC) + P(ABC)$$

三、条件概率

1. 定义

$$P(B|A)=\frac{P(AB)}{P(A)} \qquad P(A)\neq 0$$





注： P(B|A) 应具有概率的所有性质，如

$$P(B\cup C\mid A)=P(B\mid A)+P(C\mid A)-P(BC\mid A)$$

2. 乘法公式

$$P(AB) = P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$$

$$P(ABC) = P(A)P(B \mid A)P(C \mid AB)$$

$$P(A_1A_2\cdots A_n)$$

$$=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)$$

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)\cdots P(A_n\mid A_1\cdots A_{n-1})$$

? $P(A_1A_2\cdots A_n|B)$

$$= P(A_1|B)P(A_2|BA_1)P(A_3|BA_1A_2)$$

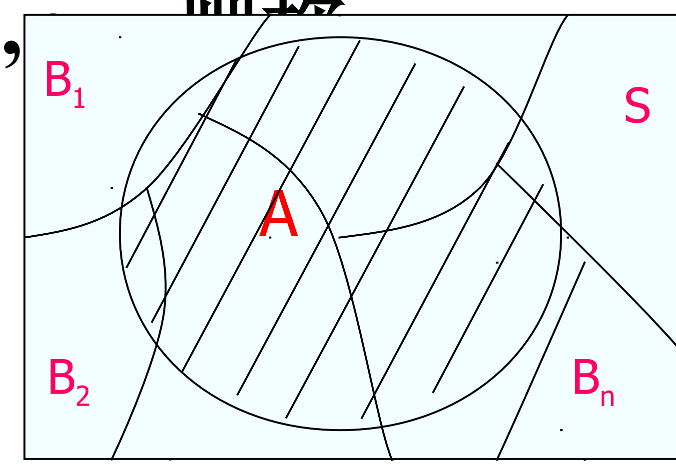
$$\cdots\cdots P(A_n|BA_1A_2\cdots A_{n-1})$$

3. 全概率公式、贝叶斯公式

定理：设试验 E 的样本空间为 S， A 为 E 的事件。$B_1,B_2,\ldots,B_n$ 为 S 的一个划分，$P(B_i)>0$，$i=1,2,\ldots$，则称

$$P(A)=\sum_{j=1}^{n}P(B_j)\cdot P(A\,|\,B_j)$$

为**全概率公式**



更进一步，设试验 E 的样本空间为 S， A 和 C 为 E 的事件。 $B_1, B_2, \ldots, B_n$ 为 S 的一个划分， $P(B_iC)>0$ $i=1, 2, \ldots, n$； $P(C)>0$，则称：

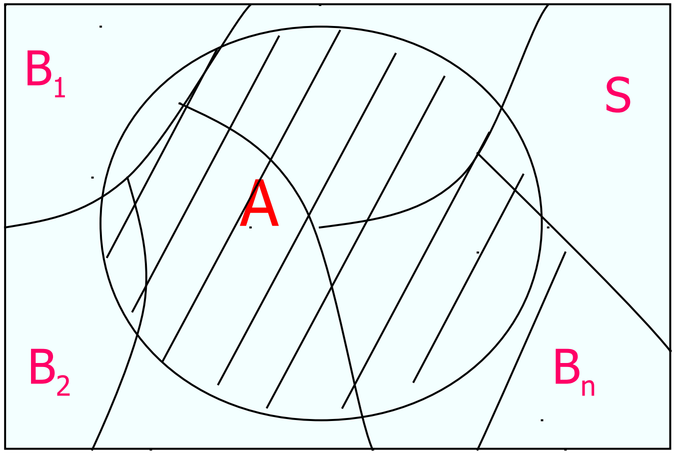
$$P(A|C)=\sum_{j=1}^{n}P(AB_j|C)=\sum_{j=1}^{n}P(B_j|C)\cdot P(A|B_jC)$$

为**条件概率的全概率公式**。

定理：接上面全概率公式的条件，P(A)>0，则

$$P(B_i \mid A)=\frac{P(B_i)P(A \mid B_i)}{\sum_{j=1}^{n} P(B_j)P(A \mid B_j)}$$

称此式为 **Bayes 公式**。



4. 事件的独立性

- 定义：设A， B为两随机事件，如果P(AB)=P(A)*P(B)，则称A， B 相互独立.

- 定义：设A， B， C为三个随机事件，如果

$$P(AB \mid C)=P(A \mid C)*P(B \mid C)$$

则称在事件C发生的条件下A， B 相互独立.

设$A_1, A_2, \cdots, A_n$为n个随机事件，若对$2 \le k \le n$,

均有：$P\left(A_{i_1} A_{i_2} \cdots A_{i_k}\right) = \prod_{j=1}^{k} P\left(A_{i_j}\right)$

则称$A_1, A_2, \cdots, A_n$相互独立

§1.2 一元随机变量及其分布

一、定义：取值至多可数的随机变量为离散型的随机变量。概率分布（分布律）为

X	x_1	x_2	$\cdots$	x_i	$\cdots$
P	p_1	p_2	$\cdots$	p_i	$\cdots$

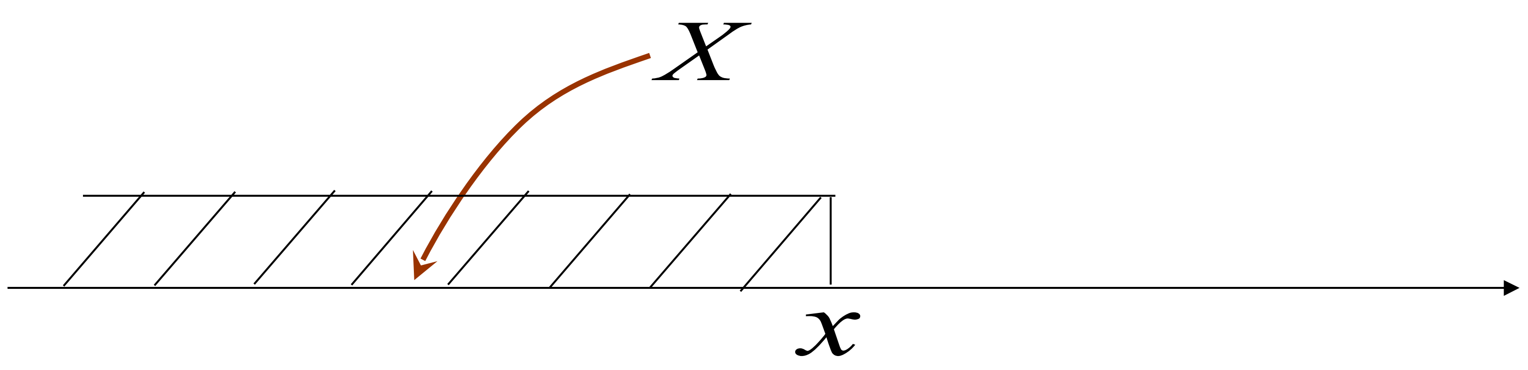
$$p_i \geq 0, \sum_{i=1}^{\infty} p_i = 1$$

二、分布函数

定义：随机变量X,对任意实数x,称函数

$F(x)=P(X\le x)$ 为 X 的概率分布函数，简称分布函数。

$F(x)$的几何意义：



$F(x)$的性质：

1) $0 \le F(x) \le 1$

2) $F(x)$单调不减，且$F(-\infty)=0$，$F(+\infty)=1$

3) $F(x)$右连续,即$F(x+0)=F(x)$.

三、连续型随机变量

定义：对于随机变量 X 的分布函数 $F(x)$，若存在非负的函数 $f(x)$，使对于任意实数 x，有：

$$\boxed{F(x)=\int_{-\infty}^{x} f(t)dt}$$

则称 X 为连续型随机变量，

其中$f(x)$称为X的概率密度函数，简称概率密度。

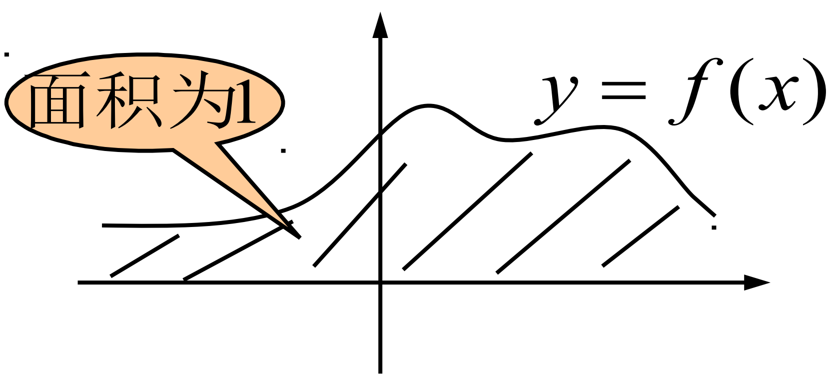
$f(x)$的性质：

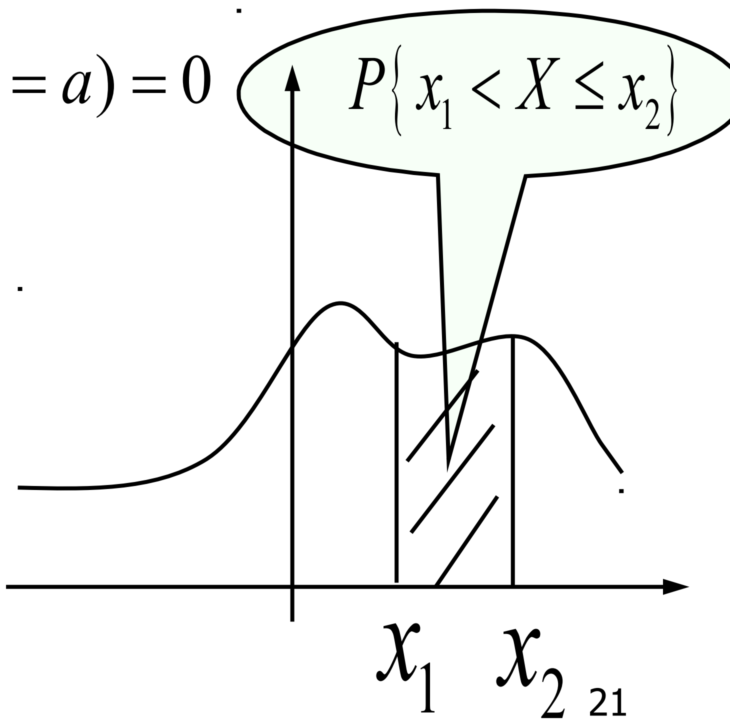
1) $f(x) \geq 0$

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

3) 对于任意的实数x_1，$x_2(x_2 > x_1)$

$$P\{x_1 < X \leq x_2\} = \int_{x_1}^{x_2} f(t)\,dt \Rightarrow P(X = a) = 0$$





4) 在$f(x)$连续点x, $F'(x) = f(x)$

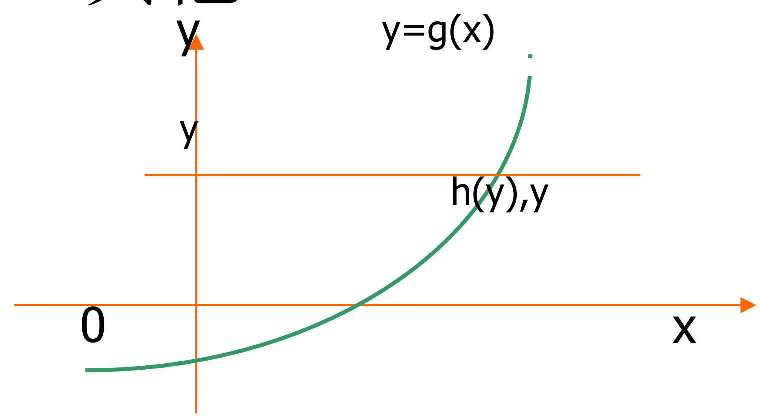
即在$f(x)$的连续点

$$f(x) = F'(x) = \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{P(x < X \le x + \Delta x)}{\Delta x}$$

四、随机变量函数的分布

定理：设$X \sim f_X(x), -\infty < x < +\infty$，$g'(x) > 0$ (或$g'(x) < 0$)。$Y = g(X)$, 则Y具有概率密度为：

$$f_Y(y) = \begin{cases} f_X(h(y)) \cdot |h'(y)|, & \alpha < y < \beta \\ 0, & \text{其他} \end{cases}$$



其中$\alpha = g(-\infty),\ \beta = g(+\infty)$,

$(+\infty)$, {当$g'(x) < 0$时$\alpha = g(+\infty),\ \beta = g(-\infty)$}

推论：设$X \sim f_X(x)$, $\{x \mid f(x)>0\}=(a,b)$,

当$a<x<b$时$g'(x)>0$ (或$g'(x)<0$)。

$Y=g(X)$, 则Y具有概率密度为：

$$f_Y(y)=\begin{cases} f_X(h(y))\cdot\left|h'(y)\right|, & \alpha<y<\beta \\ 0, & \text{其他} \end{cases}$$

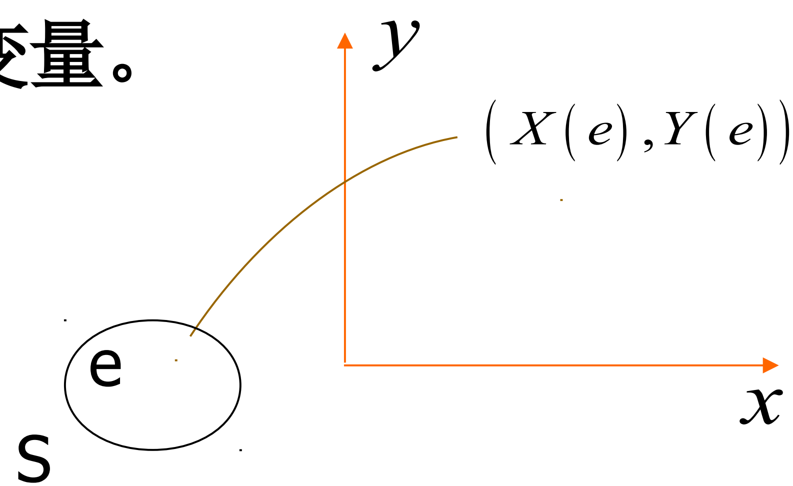
其中$\alpha=\min(g(a),g(b))$, $\beta=\max(g(a),g(b))$,

$$h(y)=x \Leftrightarrow y=g(x)$$

§1.3　多元随机变量及其分布

- 一、定义：假设 E 是一个随机试验，样本空间 S={e}；设 X=X(e) 和 Y=Y(e) 是定义在 S 上的随机变量

由它们构成的向量 (X, Y) 叫做二元随机变量或二维随机变量。



二、定义：若二元随机变量（X, Y）全部可能取到的不同值是有限对或可列无限对，则称（X, Y）是离散型随机变量。

1、联合概率分布律

设(X,Y)所有可能取值为

$(x_i, y_i),\ i,j=1,2,\cdots$

二元离散型随机变量 (X, Y)

的联合概率分布律可以用如

右表格，或如下公式表示

X \ Y	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\vdots$	$\cdots$		$\cdots$		$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\cdots$		$\cdots$		$\cdots$

$$P\left(X=x_i, Y=y_j\right)=p_{ij},\ i,j=1,2,\cdots$$

2. 分布律的性质

$$1^\circ \quad p_{ij} \geq 0, \quad i, j = 1, 2, \cdots$$

$$2^\circ \quad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_{ij} = 1$$

3、边际（边缘）分布律

$$P(X=x_i)=P(X=x_i,\ Y\le+\infty)=\sum_{j=1}^{\infty}p_{ij}\overset{\text{记为}}{=\!=}p_{i\bullet}\quad i=1,2,\cdots$$

$$P(Y=y_j)=P(X\le+\infty,\ Y=y_j)=\sum_{i=1}^{\infty}p_{ij}\overset{\text{记为}}{=\!=}p_{\bullet j}\quad j=1,2,\cdots$$

X \ Y	y_1	y_2	$\cdots$	y_j	$\cdots$	$P(X=x_i)$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$p_{2\cdot}$
$\vdots$	$\cdots$		$\cdots$		$\cdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\vdots$	$\cdots$		$\cdots$		$\cdots$	$\vdots$
$P(Y=y_j)$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$	1

4、条件分布律

定义：设 (X, Y) 是二维离散型随机变量，对于固定的

，y_j 若$P(Y=y_j)>0$，则称

$$P(X=x_i \mid Y=y_j)=\frac{P(X=x_i,\ Y=y_j)}{P(Y=y_j)}=\frac{P_{ij}}{P_{\bullet j}} \quad i=1,2\cdots$$

为在$Y=y_j$条件下，随机变量X的**条件分布律**；

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	$P(X = x_i)$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$p_{2\cdot}$
$\vdots$	$\cdots$		$\cdots$		$\cdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\vdots$	$\cdots$		$\cdots$		$\cdots$	$\vdots$
$P(Y = y_j)$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$	1

同样，对于固定的 x_i 若 $P(X=x_i)>0$，则称：

$$P(Y=y_j \mid X=x_i)=\frac{P(X=x_i,\ Y=y_j)}{P(X=x_i)}=\frac{p_{ij}}{p_{i\bullet}}\quad j=1,2\cdots$$

为在$X=x_i$条件下，随机变量Y的**条件分布律**。

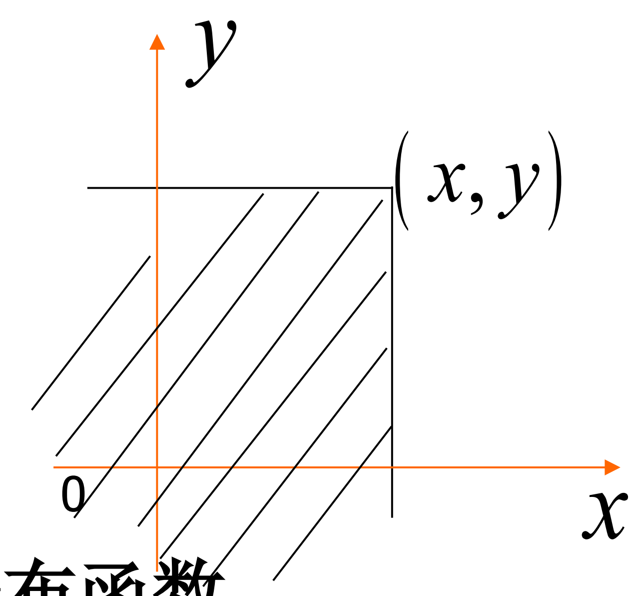
三、联合分布函数

定义：设 (X,Y) 是二元随机变量，对于任意实数

x,y， 二元函数

$$F(x,y)=P\{(X\le x)\cap(Y\le y)\}$$

$$\overset{\text{记成}}{=\!=} P(X\le x, Y\le y)$$



称为二元随机变量 (X, Y) 的联合分布函数。

1. 分布函数的性质

1° $F(x,y)$关于x,y单调不减，即：

$$x_1 < x_2 \Rightarrow F(x_1, y) \le F(x_2, y)$$

$$y_1 < y_2 \Rightarrow F(x, y_1) \le F(x, y_2)$$

2° $0 \le F(x,y) \le 1,\ F(+\infty,+\infty)=1$

对任意x,y

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

3° $F(x,y)$关于x,y右连续，即：

$$\lim_{\varepsilon\to 0^+} F(x+\varepsilon,y)=F(x,y)$$

$$\lim_{\varepsilon\to 0^+} F(x,y+\varepsilon)=F(x,y)$$

4° 若$x_1<x_2, y_1<y_2$

$$\Rightarrow F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\ge 0$$

➢边际分布函数

$$F_X(x) = F(x, +\infty)$$

$$F_Y(y) = F(+\infty, y)$$

四、二元连续型随机变量

定义：对于二元随机变量(X,Y)的分布函数$F(x,y)$，

如果存在非负函数$f(x,y)$，使对于任意x,y，

有$$\boxed{F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)dudv}$$

称(X,Y)为连续型的二维随机变量

称$f(x,y)$为二元随机变量(X,Y)的

(联合)概率密度

➢ 联合概率密度的性质

1. $f\left(x,y\right)\geq 0$

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

3. 设G是xoy平面上的区域，点$\left(X,Y\right)$落在G内的概率为：

$$P((X,Y)\in G)=\iint_G f(x,y)dxdy$$

4. 在$f(x,y)$的连续点(x, y)，有$\dfrac{\partial^2 F(x,y)}{\partial x\partial y}=f(x,y)$

➢边际（边缘）概率密度

X,Y 的边际概率密度为：

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

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

➢ 条件概率密度

若对于固定的y, $f_Y(y) > 0$,

则称$\dfrac{f(x,y)}{f_Y(y)}$为在$Y = y$的条件下，X的条件概率密度,

记为：$f_{X|Y}(x \mid y) = \dfrac{f(x,y)}{f_Y(y)}$

若对于固定的x, $f_X(x)>0$，且$f_X(x)$连续，

在$X=x$条件下，Y的条件概率密度为：

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}$$

五、随机变量的独立性

设$F(x,y)$及$F_X(x),F_Y(y)$分别是二元随机变量(X,Y)的分布函数及边际分布函数，若对所有x,y有：

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$$

即 $F(x,y) = F_X(x)F_Y(y)$

称随机变量X,Y相互独立。

➢ 离散型、连续型

若(X,Y)是离散型随机变量，则X,Y相互独立的条件等价于：$P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)$

即 $p_{ij}=p_{i\bullet}p_{\bullet j}$ 对一切i,j都成立。

若(X,Y)是连续型随机变量，$f(x,y),f_X(x),f_Y(y)$分别是(X,Y)的概率密度和边际概率密度，则X,Y相互独立的条件等价于：$\boxed{f(x,y)=f_X(x)f_Y(y)}$几乎处处成立；

即在平面上除去“面积”为零的集合以外，处处成立。

➢ 多元随机变量之间的独立性

设$(X_1,X_2,\cdots,X_m)$的分布函数为$F_1(x_1,x_2,\cdots x_m)$,

$(Y_1,Y_2,\cdots,Y_n)$的分布函数为$F_2(y_1,y_2,\cdots y_n)$,

$(X_1,X_2,\cdots,X_m,Y_1,Y_2,\cdots,Y_n)$的分布函数为

$$F(x_1,x_2,\cdots x_m,y_1,y_2,\cdots y_n)$$

若$F(x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_n)=F_1(x_1,x_2,\cdots x_m)F_2(y_1,y_2,\cdots y_n)$

称$(X_1,X_2,\cdots,X_m)$与$(Y_1,Y_2,\cdots,Y_n)$相互独立。

定理

设$(X_1,X_2,\cdots,X_m)$与$(Y_1,Y_2,\cdots,Y_n)$相互独立，

若$h(x_1,x_2,\cdots,x_m)$和$g(y_1,y_2,\cdots,y_n)$是连续函数，

则$h(X_1,X_2,\cdots,X_m)$和$g(Y_1,Y_2,\cdots,Y_n)$相互独立。

§1.4 数字特征

➢ 一元情形

定理：设Y是随机变量X的函数：

$$Y=g(X)\left(g\text{是连续函数}\right),$$

X是离散型随机变量，它的分布律为：

$$P(X=x_k)=p_k,\quad k=1,2,\cdots$$

若$\sum_{k=1}^{\infty}g(x_k)p_k$绝对收敛，则有

$$E(Y)=E[g(X)]=\sum_{k=1}^{\infty}g(x_k)p_k$$

X是连续型随机变量，它的概率密度为$f(x)$

若$\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛

则有$E(Y)=E(g(X))=\int_{-\infty}^{+\infty} g(x)f(x)dx$

➤ 二元情形

定理：设Z是随机变量X,Y的函数：$Z=h(X,Y)\big(h$是连续函数$\big)$，

若二维离散型随机变量(X,Y)的分布律为：

$P(X=x_i,Y=y_j)=p_{ij}$， $i,j=1,2,\cdots$

则有$E(Z)=E[h(X,Y)]=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}h(x_i,y_j)p_{ij}$

这里设上式右边的级数绝对收敛，

若二维连续型随机变量(X,Y)的概率密度为$f(x,y)$,

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

这里设上式右边的积分绝对收敛

一、数学期望

定义：设离散型随机变量X的分布律为：

$$P(X = x_k) = p_k \quad k = 1, 2, \cdots$$

若级数$\sum_{k=1}^{+\infty} x_k p_k$绝对收敛，则称

级数$\sum_{k=1}^{+\infty} x_k p_k$的值为随机变量$X$

的数学期望，记为$E(X)$,即

$$E(X) = \sum_{k=1}^{+\infty} x_k p_k$$

设连续型随机变量X的概率概率为$f(x)$，

若积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛

则称积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 的值为随机变量X的

数学期望，记为$E(X)$

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

➢数学期望的性质：

1. 设C是常数，则有$E(C)=C$

2. 设X是一个随机变量，C是常数，则有$E(CX)=CE(X)$

3. 设X,Y是两个随机变量，则有$E(X+Y)=E(X)+E(Y)$

将上面三项合起来就是：$E(aX+bY+c)=aE(X)+bE(Y)+c$

这一性质可以推广到任意有限个随机变量线性组合的情况

$$E(c_0+\sum_{i=1}^{n}c_iX_i)=c_0+\sum_{i=1}^{n}c_iE(X_i)$$

4. 设X,Y是相互独立的随机变量，则有$E(XY)=E(X)E(Y)$

这一性质可以推广到任意有限个相互独立的随机变量之积的情况

$$E(\prod_{i=1}^{n}X_i)=\prod_{i=1}^{n}E(X_i),\text{其中}X_i\text{相互独立.}$$

二、方差

定义：设X是一个随机变量，若$E\left\{[X-E(X)]^2\right\}$存在，

则称其为X的方差，记为$D(X)$或$Var(X)$，即

$$D(X)=Var(X)=E\left\{[X-E(X)]^2\right\}$$

$$D(X)=E(X^2)-[E(X)]^2$$

➢方差的性质：

1. 设C是常数，则$D(C)=0$

2. 设X是随机变量，C是常数，则有$D(CX)=C^2D(X)$

3. 设X,Y是两个随机变量，

   则有$D(X+Y)=D(X)+D(Y)+2E\{[X-E(X)][Y-E(Y)]\}$

   特别，若X,Y相互独立，则有$D(X+Y)=D(X)+D(Y)$

综合上述三项，设X,Y相互独立，a,b,c是常数，

则$D(aX+bY+c)=a^2D(X)+b^2D(Y)$

这一性质可以推广到任意有限个独立随机变量线性组合的情况

$$D(c_0+\sum_{i=1}^{n}c_iX_i)=\sum_{i=1}^{n}c_i^2D(X_i)$$

4. $D(X)=0\Leftrightarrow P(X=C)=1$ 且$C=E(X)$

➢常用分布的均值与方差

分布	分布律	期望	方差
0-1分布	$P(X=1)=p,\ P(X=0)=q,$ $0<p<1,\ p+q=1$	p	pq
二项分布	$P(X=k)=C_n^k p^k q^{n-k},$ $0<p<1,\ p+q=1,\ k=0,1,\cdots,n$	np	npq
泊松分布	$P(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda},\ \lambda>0,\ k=0,1,\cdots$	λ	λ
几何分布	$P(X=k)=pq^{k-1},$ $0<p<1,\ p+q=1,\ k=0,1,\cdots$	$1/p$	q/p^2
负二项分布	$P(X=j)=C_{j-1}^{k-1}p^k q^{n-k},$ $0<p<1,\ p+q=1,\ j\ge k$	k/p	kq/p^2
离散均匀分布	$P\left(X=a+i\dfrac{b-a}{n}\right)=\dfrac{1}{n+1},\quad i=0,1,\cdots,n$	$(a+b)/2$	$\dfrac{(n+2)(b-a)^2}{12n}$

分布	概率密度	期望	方差
均匀分布	$f(x)=\begin{cases}1/(b-a), & a<x<b\\ 0, & \text{其它}\end{cases}$	$(a+b)/2$	$(b-a)^2/12$
正态分布	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-(x-a)^2/2\sigma^2}$	a	σ^2
指数分布	$f(x)=\lambda e^{-\lambda x}, \quad \lambda>0$	$1/\lambda$	$1/\lambda^2$
瑞利分布	$f(x)=\frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right), \quad \sigma>0$	$\sqrt{\pi/2}\sigma$	$(2-\pi/2)\sigma^2$
Γ 分布	$f(x)=\frac{x^{\alpha-1}}{\beta^\alpha\Gamma(\alpha)}\exp\left(-\frac{x}{\beta}\right), \quad \alpha,\beta>0$	$\beta\alpha$	$\beta^2\alpha$
χ^2 分布	$f(x)=\frac{x^{(N/2)-1}}{2^{N/2}\Gamma(N/2)}\exp\left(-\frac{x}{2}\right), \ N>0$	N	$2N$
β 分布	$f(x)=\begin{cases}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, & 0<x<1\\ 0, & \text{其它}\end{cases}, \quad \alpha,\beta>0$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

三、协方差和相关系数

量$E\{[X-E(X)][Y-E(Y)]\}$称为随机变量X与Y的协方差，

记为：$Cov(X,Y)$，即

$Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$. 称

$$\rho_{XY}=\frac{Cov(X,Y)}{\sqrt{D(X)D(Y)}}$$

为随机变量X与Y的相关系数. ρ_{XY}是一个无量纲的量

➢ 相关公式

$$Cov(X,Y)=E(XY)-E(X)E(Y)$$

$$D(X+Y)=D(X)+D(Y)+2Cov(X,Y)$$

$$D(\sum_{i=1}^{n}X_i)=\sum_{i=1}^{n}D(X_i)+2\sum_{1\le i<j\le n}Cov(X_i,X_j)$$

➢ 协方差的性质：

1. $Cov(X,Y)=Cov(Y,X)$;
2. $Cov(X,X)=D(X)$;
3. $Cov(aX,bY)=ab\cdot Cov(X,Y)$,其中$a,b$为两个实数；
4. $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$;

➢ 相关系数的性质

1. $|\rho_{XY}| \le 1$

2. $|\rho_{XY}| = 1 \Leftrightarrow$ 存在常数a,b，使$P(Y = a + bX) = 1$

特别的，$\rho_{XY} = 1$时，$b > 0$; $\rho_{XY} = -1$时，$b < 0$

定义：$\rho_{XY}=0$，称X与Y不相关或零相关.

可知，当X与Y相互独立$\Rightarrow X$与Y一定不相关

反之，若X与Y不相关，X与Y却不一定相互独立

四、多元正态变量的重要性质

1. n元正态变量$(X_1, X_2, \cdots X_n)^T$中的任意子向量$(X_{i_1}, X_{i_2}, \cdots, X_{i_k})^T (1 \le k \le n)$也服从$k$元正态分布.

特别地，每一个分量$X_i, i = 1, 2, \cdots n$都是正态变量;

反之，若$X_1, X_2, \cdots X_n$都是正态变量，且相互独立，

则$(X_1, X_2, \cdots X_n)$是n元正态变量;

2. n元随机变量$(X_1, X_2, \cdots X_n)$服从n元正态分布

$\Leftrightarrow X_1, X_2, \cdots X_n$的任意线性组合

$l_1X_1 + l_2X_2 + \cdots + l_nX_n$服从一元正态分布

其中$l_1, l_2, \cdots l_n$不全为零

3. 若$(X_1,X_2,\cdots X_n)$服从n元正态分布，

设$Y_1,Y_2,\cdots Y_k$是$X_j\,(j=1,2,\cdots n)$的线性函数，

则$(Y_1,Y_2,\cdots Y_k)$也服从多元正态分布；

这一性质称为正态变量的线性变换不变性

4. 设$(X_1,X_2,\cdots X_n)$服从n元正态分布，

则$X_1,X_2,\cdots X_n$相互独立$\Leftrightarrow X_1,X_2,\cdots X_n$两两不相关

$\Leftrightarrow$协方差矩阵为对角矩阵.

§1.5 极限定理

一、依概率收敛

随机变量序列$Y_1, Y_2, Y_3, \cdots$，若存在某常数a，使得$\forall \varepsilon > 0$，均有：$\lim\limits_{n \to +\infty} P\left\{\left|Y_n - a\right| \geq \varepsilon\right\} = 0$，

则称$\{Y_n\}$依概率收敛于常数a，

记为：$Y_n \xrightarrow{P} a$，当$n \to +\infty$时.

➢依概率收敛的性质

$X_n \xrightarrow{P} a, Y_n \xrightarrow{P} b, g$在$(a,b)$连续，则

$$g(X_n, Y_n) \xrightarrow{P} g(a,b)$$

➢ 大数定律

定理（切比雪夫大数定律）：

设$X_1, X_2, \cdots, X_n, \cdots$相互独立，具有相同的数学期望$\mu$和相同的方差$\sigma^2$，则当$n \to +\infty$时，

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu.$$

定理（辛钦大数定律）：

设$X_1, X_2, \cdots, X_n, \cdots$独立同分布，$EX_i = \mu$,

则当$n \to +\infty$,

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{P} \mu.$$

二、中心极限定律

定理(独立同分布的中心极限定理)

设随机变量$X_1, X_2, \cdots, X_n, \cdots$相互独立同分布，

$$E(X_i)=\mu, D(X_i)=\sigma^2, i=1,2,\cdots$$

则前n个变量的和的标准化变量为： $Y_n=\dfrac{\sum_{i=1}^{n}X_i-n\mu}{\sqrt{n}\sigma}$

$\forall x\in R$,有：

$$\lim_{n\to+\infty}P(Y_n\le x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x)$$

此定理表明，当n充分大时，Y_n近似服从$N(0,1)$，

$$\sum_{i=1}^{n} X_i \text{（近似）} \sim N(n\mu, n\sigma^2),$$

从而，$P(a < \sum_{i=1}^{n} X_i \le b) \approx \Phi(\frac{b-n\mu}{\sqrt{n}\sigma}) - \Phi(\frac{a-n\mu}{\sqrt{n}\sigma}).$

德莫佛—拉普拉斯定理

推论：设n_A为n重贝努里试验中A发生的次数，

$P(A)=p(0<p<1)$,则对任何实数x，有：

$$\lim_{n\to+\infty} P\left(\frac{n_A-np}{\sqrt{np(1-p)}}\le x\right)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\Phi(x).$$

即，$B(n,p)\overset{\text{近似}}{\sim}N(np,np(1-p))$，当$n$充分大时.

$$n_A(\text{近似})\sim N(np,np(1-p)).$$