第一章 随机事件和概率

1.1 事件、样本空间、事件间的关系与运算

1.1.1 随机试验

> 对随机现象进行观察或实验成为随机试验，简称试验，记作E，它具有如下特点：
>
> 1. 可以在相同条件下重复进行。
> 2. 所得的结果可能不止一个，且所有可能结果都能事前已知。
> 3. 每次具体实验之前无法预知会出现哪个结果。

1.1.2 样本空间

> 随机试验的每一可能结果称为样本点，记作ω.由所有样本点全体组成的集合称为样本空间，记作Ω.

样本点是组成样本空间的元素，样本空间是样本点的全集，样本空间有以下三种类型：

1. 有限集合：样本空间中的样本点个数是有限的。
2. 无限可列集合：样本空间中的样本点个数是无限的，但是可以列出来。
3. 无限不可列集合：样本空间中的样本点个数是无限的，而且不能列出来。

1.1.3 随机事件

> 样本空间的子集称为随机事件，简称事件，常用字母A,B,C等表示

随机事件是由样本空间中的元素即样本点组成，由一个样本点组成的子集是最简单事件，称为基本事件。随机事件既然由样本点组成，那么也可以认为随机事件是由基本事件组成。

如果一次试验的结果为某一基本事件出现，就称该基本事件出现或发生。如果组成事件A的一个基本事件出现或发生，也称事件A出现或发生。

把Ω看成一个事件，则每次试验必有Ω中某一基本事件发生，也就是每次试验Ω必然发生，称Ω为必然事件。

把不包含任何样本点的空集ϕ看成一个事件，则每次试验ϕ必不发生，称ϕ为不可能事件。

1.1.4 事件的包含

> 如果事件A发生必然导致事件B发生，则称事件B包含事件A，或称事件A包含于事件B，记为$B \supset A$或$A \subset B$.

从集合关系来说，事件A的每一个样本点都属于事件B。

1.1.5 事件的相等

> 如果$A \supset B$和$B \supset A$同时成立，则称事件A与事件B相等。

1.1.6 事件的交

> 如果事件A与事件B同时发生，则称这样的一个事件为事件A与事件B的交或积，记为$A \cap B$或AB

集合$A \cap B$是由同时属于A与B的所有公共样本点构成。

事件的交可以推广到有限多个事件或可数无穷多个事件的情形：

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n \tag{1}$$
$$\bigcap_{i=1}^{+\infty} A_i = A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots$$

1.1.7 互斥事件

> 如果事件A与事件B的关系为$AB = \phi$，即A与B不能同时发生，则称事件A与事件B为互斥或互不相容。

互斥的两个事件没有公共样本点。

事件的互斥可以推广到有限多个事件或可数无穷多个事件的情形：

若n个事件$A_1, A_2, A_3 \cdots, A_n$中任意两个事件均互斥，则称这n个事件两两互斥或两两互不相容。

若可数无穷多个事件$A_1, A_2, A_3 \cdots, A_n, \cdots$中任意两个事件均互斥，则称这可数无穷多个事件两两互斥或两两互不相容。

1.1.8 事件的并

> 如果事件A与事件B至少有一个发生，则称这样一个事件为事件A与事件B的并或者和，记为$A \cup B$

集合$A \cup B$是由属于A与B的所有样本点构成。

事件的并可推广到有限多个事件或可数无穷多个事件的情形：

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots A_n \tag{2}$$
$$\bigcup_{i=1}^{+\infty} A_i = A_1 \cup A_2 \cup \cdots A_n \cup \cdots$$

1.1.9 对立事件

> 如果事件A与事件B有且仅有一个发生，即同时成立$A \cup B = \Omega$和$A \cap B = \phi$，则称事件A与事件B为对立事件或互逆事件，记为$\overline{A} = B$或$\overline{B} = A$.

1.1.10 事件的差

> 事件A发生而事件B不发生称为事件A与事件B的差，记为$A - B$.

在样本空间中的集合$A - B$是由属于事件A而不属于事件B的所有样本点构成的集合。

显然$A - B = A\overline{B}$.

1.1.11 事件的运算规律

1. 交换律 $A \cup B = B \cup A, A \cap B = B \cap A$
2. 结合律 $A \cup (B \cup C) = (A \cup B) \cup C$
   $A \cap (B \cap C) = (A \cap B) \cap C$

3. 分配律 $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$
   $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$
4. 对偶律 $\overline{A\cup B}=\overline{A}\cap\overline{B},\overline{A\cap B}=\overline{A}\cup\overline{B}$
   $\overline{\bigcup_{i=1}^{n}}=\bigcap_{i=1}^{n}\overline{A_i},\overline{\bigcap_{i=1}^{n}}=\bigcup_{i=1}^{n}\overline{A_i}$

1.2 概率、条件概率、独立性和五大公式

1.2.1 概率公理

设试验E的样本空间为Ω，称实值函数P为概率，如果P满足如下三个条件：

1. 对于任意事件A，有$P(A)\geq 0$
2. 对于必然事件Ω，有$P(A)=A$
3. 对于两两互斥的可数无穷多个事件$A_1,A_2,\cdots,A_n,\cdots$有
   $P(A_1\cup A_2\cup\cdots\cup A_n\cup\cdots)=P(A_1)+P(A_2)+\cdots+P(A_n)+\cdots$，称$P(A)$为事件A的概率。

1.2.2 条件概率

> 设A,B为两个事件，且$P(A)>0$，称
>
> $$P(B|A)=\frac{P(AB)}{P(A)} \tag{3}$$
>
> 为在事件A发生的条件下，事件B发生的条件概率。

1.2.3 概率的性质

1. $P(\phi)=0$
2. $0\leq P(A)\leq 1$
3. $P(\overline{A})=1-P(A)$
4. $A\subset B$，则$P(A)\leq P(B)$
5. 对于两两互斥的有限个事件$A_1,A_2,\cdots,A_n$，有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n) \tag{4}$$

1.2.4 事件独立性

> 设A,B两事件满足等式
>
> $$P(AB)=P(A)P(B) \tag{5}$$
>
> 则称A与B相互独立。

设$A_1,A_2,\cdots,A_n$是n个事件，如果对于任意$k(1<k<n)$，任意$1\leq i_1<i_2<\cdots<i_k\leq n$满足等式

$$P(A_1A_2\cdots A_k)=P(A_1)P(A_2)\cdots P(A_k) \tag{6}$$

则称$A_1,A_2,\cdots,A_n$为相互独立的事件。

【注】：n个事件相互独立需要

$$C_n^2+C_n^3+\cdots+C_n^n=2^n-n-1\text{个等式成立} \tag{7}$$

1.2.5 相互独立的性质

1. A与B相互独立的充要条件是A与$\overline{B}$或$\overline{A}$与B或$\overline{A}$与$\overline{B}$相互独立。
2. 当$0<P(A)<1$时，A与B独立$\Leftrightarrow P(B|A)=P(B)$或$P(B|A)=P(B|\overline{A})$
3. 若$A_1,A_2,\cdots,A_n$相互独立，则$A_1,A_2,\cdots,A_n$必两两独立。反之，$A_1,A_2,\cdots,A_n$两两独立，则$A_1,A_2,\cdots,A_n$不一定相互独立。
4. 当$A_1,A_2,\cdots,A_n$相互独立时，它们的部分事件也是相互独立的。
5. 当$A_1,A_2,\cdots,A_n$相互独立时，任意几个事件换成它们的对立事件，这些事件也是相互独立的。

1.2.6 五大公式

1. 加法公式

$$\begin{aligned}&P(A\cup B)=P(A)+P(B)-P(AB)\\&P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)\end{aligned} \tag{8}$$

2. 减法公式

$$P(A-B)=P(A)-P(AB) \tag{9}$$

3. 乘法公式

$$\begin{aligned}&\text{当}P(A)>0\text{时，}P(AB)=P(A)P(B|A)\\&\text{当}P(A_1A_2\cdots A_n)>0\text{时，}P\text{（}A_1)P(A_2|A_1)\cdots P(A_n|A_1A_2\cdots A_n)\end{aligned} \tag{10}$$

4. 全概率公式

$$\text{设}B_1,B_2,\cdots,B_n\text{满足}\bigcup_{i=1}^{n}B_i=\Omega,B_iB_j=\phi(i\neq j)\text{且}P(B_k)>0,k=1,2,\cdots,n, \tag{11}$$

$$\text{对于任意事件}A\text{有}$$

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$$

$$\text{称满足}\bigcup_{i=1}^{n}B_i=\Omega,B_iB_j=\phi(i\neq j)\text{的}B_1,B_2,\cdots,B_n\text{为}\Omega\text{的一个完备事件组}$$

5. 贝叶斯公式

设$B_1,B_2,\cdots,B_n$满足

$$\bigcup_{i=1}^{n}B_i=\Omega,B_iB_j=\phi(i\neq j)\text{且}P(A)>0,P(B_k)>0,k=0,1,2,\cdots,n \tag{12}$$

则

$$P(B_j|A)=\frac{P(B_j)P(A|B_j)}{\sum_{i=1}^{n}P(B_j)P(A|B_j)},j=1,2,\cdots,n \tag{13}$$

1.3 古典概型与伯努利概型

1.3.1 古典型概率

> 当试验结果为有限n个样本点，且每个样本点的发生具有相等的可能性，如果事件A由n_A个样本点组成，则事件A的概率为
>
> $$P(A)=\frac{n_A}{n}=\frac{A\text{所包含的样本点数}}{\text{样本点总数}} \tag{14}$$
>
> 称有限等可能试验中事件A的概率$P(A)$为古典型概率。

1.3.2 几何型概率

> 当试验的样本空间是某区域，以$L(\Omega)$表示其几何度量。事件A的样本点所表示的区域为Ω_A，则事件A的概率是
>
> $$P(A)=\frac{L(\Omega_A)}{L(\Omega)}=\frac{\Omega_A\text{的几何度量}}{\Omega\text{的几何度量}} \tag{15}$$
>
> 称这种样本点个数无限但是几何度量上的等可能试验中事件A的概率P(A)为几何型概率

1.3.3 n重伯努利试验

> 把一个随机试验独立重复做若干次，即各次试验所联系的事件之间相互独立，且同一事件在各个试验中出现的概率相同，称为独立重复试验。
>
> 如果每次试验只有两个结果$A\text{和}\overline{A}$，则称这种试验为伯努利试验。将伯努利试验独立重复进行n次，称为n重伯努利试验。

设在每次试验中，概率$P(A)=p(0<p<1)$，则在n重伯努利试验中事件A发生k次的概率，又称为二项概率公式：

$$C_n^k p^k(1-p)^{n-k}, k=0,1,2,\cdots,n \tag{16}$$

第二章 随机变量及其概率分布

2.1 随机变量及其分布函数

2.1.1 随机变量

> 在样本空间Ω上的实值函数$X=X(\omega),\omega\in\Omega$，称$X(\omega)$为随机变量，简记成$X$.
>
> $X(\omega)\text{的定义域是}\Omega$

2.1.2 分布函数

> 对于任意实数x，记函数$F(x)=P\{X\leq x\},-\infty<x<+\infty$，称$F(x)$为随机变量$X$的分布函数。
>
> 分布函数$F(x)$是定义在$(-\infty,+\infty)$上的一个实值函数，$F(x)$的值等于随机变量X在区间$(-\infty,x]$内取值的范围，即事件"$X\leq x$"的概率。

2.1.3 分布函数性质

1. $0 \leq F(x) \leq 1$
2. $lim_{x \to -\infty} F(x) = 0$，记为$F(-\infty) = 0$. $lim_{x \to +\infty} F(x) = 1$，记为$F(+\infty) = 1$
3. $F(x)$是单调非减函数，即当$x_1 < x_2$时，$F(x_1) \leq F(x_2)$
4. $F(x)$是右连续的，即$F(x+0) = F(x)$
5. 对任意$x_1 < x_2$,有$P\{x_1 < X \leq x_2\} = F(x_2) - F(x_1)$
6. 对任意的x,$P\{X = x\} = F(x) - F(x-0)$

2.2 离散型随机变量和连续型随机变量

2.2.1 离散型随机变量

> 如果一个随机变量的可能取值是有限多个或可数无穷多个，则称它为离散型随机变量。

2.2.2 离散型随机变量X的概率分布

> 设离散型随机变量X的可能取值是$x_1, x_2, \cdots, x_n, \cdots$，$X$取各可能值的概率为
>
> $$P\{X = x_k\} = p_k, k = 1, 2, \cdots \tag{17}$$
>
> 称上式为离散型随机变量X的概率分布或分布律。

分布律也有用列表方式给出的(我只能说好难打😂)：

$$\begin{array}{c|cccccc} X & x_1 & x_2 & x_3 & \cdots & x_n & \cdots \\ \hline P & p_1 & p_2 & p_3 & \cdots & p_n & \cdots \end{array} \tag{18}$$

2.2.3 连续型随机变量及其概率密度

> 如果对随机变量X的分布函数$F(x)$，存在一个非负可积函数$f(x)$，使得对任意实数x，都有：
>
> $$F(x) = \int_{-\infty}^{x} f(t)dt, -\infty < x < +\infty \tag{19}$$
>
> 称X为连续型随机变量，函数$f(x)$称为X的概率密度。

2.2.4 概率分布的性质

1. $p_k \geq 0, k = 1, 2, \cdots$
2. $\sum_{k=1}^{+\infty} p_k = 1$

二者也是概率分布(分布律)的充要条件。

2.2.5 概率密度$f(x)$的性质

1. $f(x) \geq 0$
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$
3. 对任意实数$x_1 < x_2$，有$P\{x_1 < X \leq x_2\} = \int_{x_1}^{x_2} f(t)dt$
4. 在$f(x)$的连续点处有$F'(x) = f(x)$

1和2是函数$f(x)$成为某个连续型随机变量的概率密度的充要条件。

2.3 常用分布

2.3.1 0-1分布

> 如果随机变量X有分布律
>
> $$\begin{array}{c|cc} X & 0 & 1 \\ \hline P & 1-p & p \end{array}, 0<p<1 \tag{20}$$
>
> 则称X服从参数为p的0-1分布。

2.3.2 二项分布

> 如果随机变量X有分布律
>
> $$P\{X=k\}=C_n^k p^k q^{n-k}, k=0,1,2,\cdots,n \tag{21}$$
>
> 其中$0<p<1, q=1-p$，则称X服从参数为n, p的二项分布，记作$X \sim B(n, p)$

在n重伯努利试验中，若每次试验的成功率为$p(0<p<1)$，则在n次独立重复试验中成功的总次数X服从二项分布。

当$n=1$时，二项分布就退化成0-1分布。所以0-1分布记作$B(1, p)$

2.3.3 几何分布

> 如果随机变量X有分布律
>
> $$p\{X=k\}=pq^{k-1}, k=1,2,\cdots \tag{22}$$
>
> 其中$0<p<1, q=1-p$，则称X服从参数为p的几何分布。

在独立重复的伯努利试验中，若每次试验成功率为$p(0<p<1)$，则在第k次试验时才首次试验成功的概率服从几何分布。

2.3.4 超几何分布

> 如果随机变量X的分布律为
>
> $$p\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k=l_1,\cdots,l_2 \tag{23}$$
>
> 其中$l_1=max(0, n-N+M), l_2=min(M, n)$，则称随机变量$X$服从参数为n,N,M的超几何分布。

如果N件产品中含有M件次品，从中任意取出n件(或一件一件不放回的取n件)，令$X=$抽取的n件产品中次品的件数，则X服从参数为n,N,M的超几何分布。(有放回取n件，X服从$B(n, \frac{M}{N})$)

2.3.5 泊松分布

> 如果随机变量X的分布律为

> $$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda},k=0,1,2,\cdots \tag{24}$$
>
> 其中$\lambda>0$为常数，则称随机变量X服从参数为λ的泊松分布，记为$X\sim P(\lambda)$

2.3.6 均匀分布

> 如果连续型随机变量X的概率密度为
>
> $$f(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b\\ 0, & \text{其他}\end{cases} \tag{25}$$
>
> 则称X在区间$[a,b]$上服从均匀分布，记作$X\sim U[a,b]$
>
> 如果连续型随机变量X的概率密度为
>
> $$f(x)=\begin{cases}\frac{1}{b-a}, & a< x< b\\ 0, & \text{其他}\end{cases} \tag{26}$$
>
> 则称X在区间(a,b)上服从均匀分布，记作$X\sim U(a,b)$

无论$X\sim U[a,b]$或$X\sim U(a,b)$，它们的分布函数均为

$$F(x)=\begin{cases}0, & x<a\\ \frac{x-a}{b-a}, & a\le x<b\\ 1, & b\le x\end{cases} \tag{27}$$

2.3.7 指数分布

> 如果连续型随机变量X的概率密度为
>
> $$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases} \tag{28}$$
>
> 则称X服从参数为λ的指数分布，记作$X\sim E(\lambda)$

设$X\sim E(\lambda)$，则X的分布函数为

$$F(x)=\begin{cases}1-e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases} \tag{29}$$

2.3.8 正态分布

> 如果随机变量X的分布律为
>
> $$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty \tag{30}$$
>
> 其中μ,σ为常数且$\sigma>0$，则称X服从参数为μ,σ的正态分布，记作$X\sim N(\mu,\sigma^2)$

当$\mu=0,\sigma^2=1$时，即$X\sim N(0,1)$，称X服从标准正态分布，此时用$\varphi(x)$表示X的概率密度，即

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},-\infty<x<+\infty \tag{31}$$

$X\sim N(\mu,\sigma^2)$，其分布函数为

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt \tag{32}$$

当$X\sim N(0,1)$时，分布函数用$\Phi(x)$表示

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt \tag{33}$$

2.3.9 常用性质

1. 泊松定理：在伯努利试验中，p_n代表事件A在试验中出现的概率，它与试验总数n有关，如果$lim_{n\to\infty}=\lambda$

$$lim_{n\to\infty}C_n^kp_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda} \tag{34}$$

应用泊松定理的要求：n较大($n\geq 100$),p_n较小($p_n\leq 0.1$),np不太大，这时有近似公式

$$C_n^kp^k(1-p)^{n-k}\approx\frac{(np)^k}{k!}e^{-np} \tag{35}$$

2. 设$X\sim U[a,b]$，则对$a\leq c<d\leq b$，有

$$P\{c<X\leq d\}=\frac{d-c}{b-a} \tag{36}$$

即随机变量落入区间$[c,d]$的概率等于该区间长度与$[a,b]$长度之比。

3. 设$X\sim E(\lambda)$，则有
   1. $P\{X>t\}=\int_t^{+\infty}\lambda e^{-\lambda t}dt,e>0$
   2. $P\{X>t+s|X>s\}=\frac{P\{X>t+s\}}{P\{X>s\}}=\frac{e^{-\lambda(t+s)}}{e^{-\lambda s}}=e^{-\lambda t}=P\{X>t\},\ t,s>0$
4. 设$N\sim N(\mu,\sigma^2)$，其分布函数为$F(x)$，则
   1. $F(x)=\Phi(\frac{x-\mu}{\sigma})$
   2. $P\{a<X\leq b\}=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma}),\ a<b$
   3. 概率密度$f(x)$关于$x=\mu$对称，标准正态分布的概率密度$\phi(x)$是偶函数
   4. $\Phi(-x)=1-\Phi(x),\Phi(0)=\frac{1}{2}$
   5. $P\{|X|\leq a\}=2\Phi(a)-1$

2.4 随机变量的函数的分布

2.4.1 离散型随机变量的函数分布

设X的分布律为$P\{X=x_i\}=p_i,i=1,2,\cdots$，则$X$的函数$Y=g(X)$的分布律为

$$P(Y=y_i)=\sum_{g(x_j)=y_i} p_j, i=1,2,\cdots \tag{37}$$

2.4.2 连续型随机变量的函数分布

1. 公式法

   设X是一个具有概率密度$f_X(x)$的随机变量，又设$y=g(x)$是单调、导数不为0的可导函数，$x=h(y)$是它的反函数，则$Y=g(X)$的概率密度为

$$f_Y(y)=\begin{cases}|h'(y)|f_X(h(y)), & \alpha<y<\beta \\ 0, & 其他\end{cases} \tag{38}$$

  其中(α,β)是函数$g(X)$在X的定义域上的值域。

2. 定义法

   先求Y的分布函数

$$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=\int_{g(x)\le y} f_X(x)dx \tag{39}$$

  再求

$$f_Y(y)=F_Y'(y) \tag{40}$$

第三章 多维随机变量及其分布

3.1 二维随机变量及其分布

3.1.1 二维随机变量

> 设$X=X(\omega),Y=Y(\omega)$是定义在样本空间Ω上的两个随机变量，则称向量(X,Y)为二维随机变量。

3.1.2 二维随机变量(X,Y)的分布

> $F(x,y)=P\{X\le x,Y\le y\},x,y\in(-\infty,+\infty)$

3.1.3 二维随机变量的边缘分布

二维随机变量(X,Y)的分布函数为$F(x,y)$，分别称$F_X(x)=P\{X\le x\}$和$F_Y(y)=P\{Y\le y\}$为关于X和关于Y的边缘分布。

显然$F_X(x)=P\{X\le x\}=P\{X\le x,y<+\infty\}=F(x,+\infty)=lim_{y\to+\infty}F(x,y)$

3.1.4 二维随机变量的条件分布

准确定义看书吧$\cdots$懒得打了$\cdots$

$F_{X|Y}(x|y)=P\{X\le x|Y=y\}$相反同理

3.1.5 二维离散型随机变量

> 如果随机变量(X,Y)的可能取值为有限个或可数无穷多个$(x_i,y_i),i,j=1,2,\cdots$，则称(X,Y)为二维离散型随机变量。

3.1.6 二维离散型随机变量的概率分布

> 二维离散型随机变量(X,Y)的可能取值为$(x_i,y_i),i,j=1,2,\cdots$，则
>
> $$P\{X=x_i,Y=y_i\}=p_{ij},\quad i,j=1,2,\cdots \tag{41}$$
>
> 称为二维离散型随机变量(X,Y)的概率分布或分布律。

3.1.7 二维离散型随机变量的边缘分布

> $$p_{i\cdot}=P\{X=x_i\}=\sum_{j=1}^{+\infty}p_{ij},i=1,2,\cdots \tag{42}$$
> $$p_{\cdot j}=P\{Y=y_j\}=\sum_{i=1}^{+\infty}p_{ij},j=1,2,\cdots$$

3.1.8 二维离散型随机变量的条件分布

> 对与给定的j，如果$P\{Y=y_i\}>0,j=1,2,\cdots$
>
> 则称$P\{X=x_i|Y=y_i\}=\frac{P\{X=x_i,Y=y_i\}}{P\{Y=y_i\}}=\frac{p_{ij}}{p_{\cdot j}}$，$i=1,2,\cdots$
>
> 为在$Y=y_j$条件下随机变量X的条件分布。

3.1.9 二维连续型随机变量及其概率密度

> 如果随机变量(X,Y)的分布函数$F(x,y)$存在非负函数$f(x,y)$使得对于任意实数x和y都有
>
> $$F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dudv,\quad x,y\in(-\infty,+\infty) \tag{43}$$
>
> 则称(X,Y)为二维连续型随机变量，函数$f(x,y)$称为(X,Y)的概率密度。

对于二维连续型随机变量(X,Y)，设它的概率密度$f(x,y)$，由

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}[\int_{-\infty}^{+\infty}f(x,y)dy]dx \tag{44}$$

可知X也是一个连续型变量，而且它的概率密度为 $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy$

3.1.10 二维连续型随机变量的边缘密度

> $f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy$和 $f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx$
>
> 被分别称为(X,Y)关于X和关于Y的边缘密度

3.1.11 二维连续型随机变量的条件密度

> 设$f(x,y)$在点(x,y)连续，$f_Y(y)$连续且$f_Y(y)>0$，则条件分布为
> $$F_{X|Y}=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}du \tag{45}$$
> 其中$\frac{f(u,y)}{f_Y(y)}$被称为在条件$Y=y$下的条件密度，记作
> $$f_{X|Y}(x|y)=\frac{f(u,y)}{f_Y(y)},\ f_Y(y)>0 \tag{46}$$
> 相反同理。

3.1.12 分布函数$F(x,y)$的性质

1. 对任意的x,y,均有$0\le F(x,y)\le 1$
2. $F(-\infty,y)=F(x,-\infty)=F(-\infty,-\infty)=0,F(+\infty,+\infty)=1$
3. F(x,y)关于x和关于y均单调不减，且是右连续的
4. $P\{a<X\le b,c<X\le d\}=F(b,d)+F(a,c)-F(b,c)-F(a,b)$

3.1.13 二维离散型随机变量的分布律$P\{X=x_i,Y=y_j\}=p_{ij}$的性质

1. $p_{ij}\ge 0,\ \ i,j=1,2,\cdots$
2. $\sum_i\sum_j p_{ij}=1$

3.1.14 二维连续型随机变量的概率密度$f(x,y)$的性质

1. $f(x,y)\ge 0$
2. $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)dxdy=1$
3. 随机变量(X,Y)落在区域D内的概率

$$P\{(X,Y)\in D\}=\iint_D f(x,y)dxdy \tag{47}$$

3.2 随机变量的独立性

3.2.1 随机变量的独立性

> 如果对任意的x,y都有$F(x,y)=F_X(x)F_Y(y)$则称随机变量X与Y相互独立。

3.2.2 随机变量相互独立的充要条件

1. 离散型随机变量X和Y相互独立的充要条件：对任意的$i,j=1,2,\cdots$有$p_{ij}=p_{i\cdot}p_{\cdot j}$
2. 连续型随机变量X和Y相互独立的充要条件：对任意的$x,y,$有$f(x,y)=f_X(x)f_Y(y)$

3.2.3 二维随机变量函数的分布

1. 离散型

$$P(Z=z_k)=\sum_{g(x_i,y_j)=z_k}P(X=x_i,Y=y_j) \tag{48}$$

2. 连续型

$F_Z(z)$的求法，可以用公式 (49)

$$F_Z(z) = P\{Z \le z\} = P\{g(x,y) \le z\} = \iint_{g(x,y)\le z} f(x,y)dxdy$$

特别的，当$Z = X + Y$时

$$F_Z(z) = P\{X + Y \le z\} = \iint_{x+y\le z} f(x,y)dxdy = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} f(x,y)dy$$

由此可得$Z = X + Y$的概率密度为

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

特别的，当X和Y相互独立时

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

这两个公式称为卷积公式，记为$f_X * f_Y$