第二章 随机变量及其分布

- □ § 2.1 随机变量及其分布函数
- □ <u>§ 2.2 概率函数</u>
- □ § 2.3 常用离散随机变量型
- □ § 2.4 二维随机变量及其分布
- □ § 2.5 随机变量的独立性与条件分布
- □ § 2.6 随机变量函数的分布

离散性随机变量及其分布

例: 抛一次硬币观察正反面出现情况。

§ 2.1 随机变量及其分布函数

定义2.1:

给定一个随机试验, Ω 是样本空间。如果对 Ω 中的每一个样本点,有一个实数 $X(\omega)$ 与它对应,那么就把这个定义域为 Ω 的单值实值函数 $X = X(\omega)$ 称为(一维)随机变量。常用大写字母X,Y,Z,…来表示随机变量。随机变量X的值域用 Ω_x 表示。 $\Omega_y \subset (-\infty, \infty)$

EX. 引入适当的随机变量描述下列事件:

- □ ①将3个球随机地放入三个格子中,事件A={有1个空格},C={全有球}。B={有2个空格},
- □ ②进行5次试验,

F={试验至少成功一次},事件D={试验成功一次},

③抛硬币观察正反面 G={至多成功3次}

EX. 引入适当的随机变量描述下列事件:

□ ①将3个球随机地放入三个格子中,事件A={有1个空格}, B={有2个空格}, C={全有球}。

§ 2.2 概率函数

- □ 如果一个随机变量只可能取有限个或可列无限个值(即它的值域 是一个有限集或可列无限集),那么便称这个随机变量为(一维) 离散型随机变量。
- 口 设随机变量**X**的值域为 $\Omega_x = \{a_1, a_2, \cdots\}$ 对于每一个i=1,2,...,X取值 为 a_i (即事件 $\{X = a_i\}$) 的概率为

$$P(X=a_i)=p_i$$

按照概率的定义与性质, P_1, P_2, \cdots 应该满足下列两个条件:

$$p_i \geq 0, i = 1, 2 \cdots$$

ii
$$\sum_{i} p_{i} = 1$$

其中 $\sum_{i} p_{i}$ 表示对一切求和。当满足这两个条件时,称:

$$P(X = a_i) = p_i, i = 1, 2, \dots$$

为随机变量X的<mark>概率(质量)函数</mark>(或分布律)。 常用表格表示概率函数

X	a_1	 a_i	
Р	P_1	 p_{i}	

从一批次品率为p的产品中,有放回抽样直到抽到次品为止。求抽到次品时,已抽取的次数X的分布律。

解 记A_i="第i次取到正品", i=1, 2, 3, ... 则 A_i, i=1, 2, 3, ... 是相互独立的! 且

(X=k)对应着事件
$$A_1A_2\cdots A_{k-1}A_k$$

X的所有可能取值为 1,2,3,...,k,...

$$P(X=k) = P(A_1A_2 \cdots A_{k-1}\overline{A_k}) = (1-p)^{k-1}p, k=1, 2, ...$$

例2.1某位足球运动员罚点球命中的概率为**0.8**, 今给他**4**次 罚球的机会,一旦命中即停止罚球。假定各次罚球是相互独立的记随机变量**X**为这位运动员罚球的次数。问**X**的概率函数。

对实数轴上的集合S:

$$P(X \in S) = \sum_{a_i \in S} P(X = a_i) = \sum_{a_i \in S} P_i$$

因此知道了随机变量的概率函数,就可以算出任意一个概率

$$P(X \in S)$$

称概率函数为离散型随机变量的(概率)分布。

§ 2.3 常用离散随机变量型

一、二项分布

 \square 在n重贝努利试验中,设随即变量X表示n次试验中事件A发生的次数。易见, $\Omega_x = \{0,1,\cdots,n\}$,由二项概率 $P_n(k)$ 推得,X的概率函数为

$$P(X=k) = {n \choose k} p^k (1-p)^{n-k}, k = 0,1,\dots,n$$

称这个随机变量X服从参数为n,p的二项分布,记作 $X \sim B(n, p)$ 其中0 。二项概率查表可得部分数值。

口 在二项分布中,当n=1时,称X服从参数为p的0-1分布,记作 $X \sim B(1,p)$ 这时, $\Omega_x = \{0,1\}$,X的概率分布函数为

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

也可以用表2.3来表示:

表2.3

X	0	1
p_r	1 - p	p

二超几何分布(产品检验问题):

在N件产品中有M件不合格品,即不合格率为p=M/N,现从这批产品中随机地抽取n件作检查(不放回抽取),发现X件是不合格品,求X的概率函数。此时X服从参数为N,M,n的超几何分布。

若采取有放回抽取, X服从参数为n, p的二项分布

$$\lim_{N\to\infty} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \binom{n}{k} p^k (1-p)^{n-k}$$

在实际应用中只要N≥10n可用二项分布来 近似描述产品抽样检查中的不合格品的个数。

结合实际例子应用服从二项分布的随机变量

- 例2.2从积累的资料看,某条流水线生产的产品中,一级品率为0.9,今从某天生产的1000件产品中,随机地抽取20件作检查,试求:
 - (1) 恰有18件一级品的概率;
 - (2) 一级品不超过18件的概率。
- 例2.3 分析病史资料表明,因患感冒而最终导致死亡 (相互独立)比例占0.2%。试求,目前正在患感冒的 1000个病人中:
 - (1) 最终恰有4个人死亡的概率;
 - (2) 最终死亡人数不超过2个人的概率。

泊松分布和泊松定理

三、泊松分布:

随机变量X的概率函数为 $P(X=k)=e^{-\lambda}\frac{\lambda^{\kappa}}{k!}, k=0,1,2,\cdots$

称这个随机变量X服从参数为 λ 的泊松分布,记作 $X \sim P(\lambda)$,其中, $\lambda > 0$ 。

定理2.1(泊松定理)

设
$$\lambda = np_n > 0, 0 < p_n < 1$$
 对于任意一个非负函数k:
$$\lim_{n \to \infty} \binom{n}{k} p_n^k \left(1 - p\right)^{n-k} = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

当 $n \ge 10$, $p \le 0.1$, 近似效果很好。

□ 证明②(板书)

- 例2.3 (续) 现在,n=1000,p=0.002,因此, $\lambda=np=2$ 。
- 例2.4 某物业管理公司负责10000户居民的房屋维修工作。假定每户居民是否报修是相互独立的,且每一时段内报修的概率都是0.04%。另外,一户居民住房的维修只需1名修理工来处理。在某个时段报修的居民数 X~B(10000,0.0004)。按波松定理,可以近似的认为X~P(4),试问:
 - (1)该物业管理公司至少需要配备多少名维修工,才能使居民报修后得到及时维修的概率不低于99%?这里不考虑维修时间的长短。

- (2) 如果物业管理公司现有4名修理工,那么居民保修后不能得到及时维修的概率有多大?
- (3) 如果该物业管理公司现有的4名修理工采用承包的方式:每两个人负责5000户居民房屋的维修,那么居民报修后不能得到及时维修的概率有多大?

四、几何分布

例: 从一批次品率为p的产品中,有放回抽样直到抽到次品为止。求抽到次品时,已抽取的次数 X的分布律。

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

离散型均匀分布

五、古典概型可以用具有下列概率函数的随机变量X(表2.5)来描写

X	a_1	•	•	•	a_n
P_r	1	•	•	•	1_
,	$\mid n \mid$				n

称这个随机变量服从集合 $\{a_1, \dots a_n\}$ 上的(<mark>离散型)均匀分布</mark>。易见,均匀分布的基本特征是:样本空间(或值域)是一个有限集,且每种试验结果以相等的概率出现。

§ 2.4 二维随机变量及其分布

例: 抛两次硬币,观察正反面出现情况。

样本空间:

$$\Omega = \{ (E, E), (E, E), (E, E), (E, E) \}$$
定义
$$X = \begin{cases} 1, & \text{第一次正面} \\ 0, & \text{第一次反面} \end{cases} Y = \begin{cases} 1, & \text{第二次正面} \\ 0, & \text{第二次反面} \end{cases}$$

第一次为正面的概率 第一次为反面的概率 第二次为正面的概率 第二次为反面的概率

Y	0	1
0	0.25	0.25
1	0.25	0.25

§ 2.4 二维随机变量及其分布

一、联合概率函数

定义2.2

给定一个随机试验, Ω 是它的样本空间,如果对 Ω 中的每一个样本点 ω ,有一对有序的实数 $(X(\omega),Y(\omega))$ 与它对应,那么就把这样一个定义域为 Ω ,取值为有序实数 $(X,Y)=(X(\omega),Y(\omega))$ 的变量为二维随机变(向)量。

如果一个二维随机变量只可能取有限个或可列无限个向量值(即它的值域是一个二维有限集或可列无限集),那么我们便称这个随机变量为二维离散型随机变量。

假定(X,Y)的值域 $\Omega_{(X,Y)} = \{(a_i,b_j): i=1,2,\cdots,j=1,2,\cdots\}$, 称 $P(X=a_i,Y=b_j) \triangleq P(\{X=a_i\} \cap \{Y=b_j\}) = p_{ij}, i,j=1,2,\cdots$

例 一个班级中,数学为优的占**0.2**, 语文为优的占**0.1**, 都为优的占**0.08**。设

$$X =$$
 $\begin{cases} 1, & \text{数学为优}; \\ 0, & \text{数学不为优}; \end{cases}$ $Y =$ $\begin{cases} 1, & \text{语文为优}; \\ 0, & \text{语文不为优}; \end{cases}$

试求:

- (1) X与Y的联合概率函数
- (2) $P(X \ge Y)$.
- (3) X与Y的边缘概率函数(后)

为二维随机变量(**X,Y**)的概率(质量)函数(或分布律,或(概率)分 布),或者称它为随机变量**X**与**Y**的联合概率(质量)函数(或联合分布律, 或联合(概率)分布)。

易见, $p_{ij}(i, j=1, 2, \cdots)$ 应该满足下列两个条件:

(i)
$$p_{ij} \ge 0, i, j = 1, 2, \cdots$$

(ii)
$$\sum_{i} \sum_{j} p_{ij} = 1$$

X与Y的联合概率函数常用表2.6的格式表示

X	b_1	b_2	
a_1	p_{11}	p_{12}	
a_2	p_{21}	p_{22}	
	i	į	<u> </u>

用联合概率函数计算概率:对平面上任意一个集合D,

$$P((X,Y) \in D) = \sum_{(i,j): (a_i,b_j) \in D} P(X = a_i, Y = b_j) = \sum_{(i,j): (a_i,b_j) \in D} p_{ij}, i, j = 1, 2, \cdots$$

事件 $\{(X,Y) \in D\}$ 可用一个或若干个不等式表达。

例2.5 一个口袋中装有5只球,其中4只是红球,1只是白球,采用无放回抽样,接连摸两次。设

$$X = \begin{cases} 1, 第 - 次摸到红球; \\ 0, 第 - 次摸到白球; \end{cases}$$
 $Y = \begin{cases} 1, 第 - 次摸到红球; \\ 0, 第 - 次摸到白球; \end{cases}$

试求:

- (1) X与Y的联合概率函数
- (2) $P(X \ge Y)$.
- (3) X与Y的边缘概率函数(后)

注: 若采取有放回抽样情形, 讨论上述三个问题。

二、边缘概率函数

□ 设(X,Y)的概率函数为

$$P(X = a_i, Y = b_j) = p_{ij}, i, j = 1, 2, \dots$$

X的值域为 $\Omega_X = \{a_1, a_2, \cdots\}$ 按概率的可加性,

$$P(X = a_i) = P(\bigcup_{j} \{X = a_i, Y = b_j\})$$

$$= \sum_{i} P(X = a_i, Y = b_j) = \sum_{j} p_{ij} \triangleq p_{i}, i = 1, 2, \dots$$

□ 因此,X的概率函数如表2.9所示,称这个概率函数为X的边缘概率(质量) 函数(或边缘分布律,或边缘(概率)分布)。

X	a_1	a_1	
P_r	$P_{1.} = \sum_{j} p_{1j}$	$P_{2\cdot} = \sum_{j} p_{2j}$	

二 类似的,**Y**的值域为 $\Omega_Y = \{b_1, b_2, \cdots\}$,按概率的可加性,有 $Pig(Y = b_jig) = Pig(ig(X = a_i, Y = b_jig)ig)$ $= \sum_i P(X = a_i, Y = b_j) = \sum_i p_{ij} \triangleq p_{\cdot j}, j = 1, 2, \cdots$

这个概率函数为Y的边缘概率函数。

Υ	b_1	b_2	
P_r	$P_{\cdot 1} = \sum_{i} p_{i1}$	$P_{\cdot 2} = \sum_{i} p_{i2}$	

例2.5 一个口袋中装有5只球,其中4只是红球,1只是白球,采用无放回抽样,接连模两次。设

$$X = \begin{cases} 1, 第 - 次摸到红球; \\ 0, 第 - 次摸到白球; \end{cases}$$
 $Y = \begin{cases} 1, 第 - 次摸到红球; \\ 0, 第 - 次摸到白球; \end{cases}$

试求:

- (1) X与Y的联合概率函数
- (2) $P(X \ge Y)$.
- (3) X与Y的边缘概率函数

注: 若采取有放回抽样情形,讨论上述三个问题。

例2.6 设二维随机变量(X,Y)的概率函数为

X	b_1	b_2	
a_1	0.1	X	
a_2	У	0.4	

已知 $P(X = a_2 | Y = b_2) = 2/3, i = 1, 2, \cdots$, 求常数x.y的值 (0.2,0.3)

例2.5 一个班级中,数学为优的占0.2,语文为优的占0.1,都为优的占0.08。设

$$X = \begin{cases} 1, & \text{数学为优}; \\ 0, & \text{数学不为优}; \end{cases}$$
 $Y = \begin{cases} 1, & \text{语文为优}; \\ 0, & \text{语文不为优}; \end{cases}$

试求:

- (1) X与Y的联合概率函数
- (2) $P(X \ge Y)$.
- (3) X与Y的边缘概率函数(后)

注: 若采取有放回抽样情形,讨论上述三个问题。

§ 2.5 随机变量的独立性与条件分布

一、随机变量的独立性

定义2.3 设随机变量X与Y的联合分布律函数为

$$P(X = a_i, Y = b_j) = p_{ij}, i, j = 1, 2, \dots$$

如果联合概率函数恰为两个边缘概率函数的乘积,即

$$p_{ij} = p_{i\cdot}p_{\cdot j}$$
,对于一切 $i, j = 1, 2, \cdots$

那么,就称随机变量X与Y相互独立。见例2.5

定理2.2

随机变量X与Y相互独立的充分必要条件是:对实数轴上任意两个集合 S_1, S_2 ,总有

$$P(X \in S_1, Y \in S_2) = P(X \in S_1)P(Y \in S_2)$$

定义2.4

如果随机变量 X_1, \cdots, X_n 的联合概率函数恰为n个边缘概率函数的乘积,即对 X_i 的值域 Ω_{X_i} 中任意一个值 $x_i, i=1,\cdots,n$,总有

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n P(X_i = x_i)$$

那么,就称这n个随机变量相互独立。

注: 当这n个随机变量相互独立时,其中任意2≤k(<n)个也相互独立。 当n个随机变量相互独立时,各个分量的边缘概率函数唯一确定 随机变量的概率函数

二、条件概率函数

定义2.5 设随机变量(X,Y)的概率函数为

$$P(X = a_i, Y = b_j) = p_{ij}, i, j = 1, 2, \dots$$

对任意一个固定的 $i, j = 1, 2, \cdots$, 称

$$P(X = a_i | Y = b_j) = \frac{p_{ij}}{p_{ij}}, i = 1, 2, \dots$$

为已知 $\{Y=b_j\}$ 发生的条件下X(记作 " $X|Y=b_j$ ")的条件概率(质量)函数(或条件分布律,或条件(概率)分布)。 **J**固定

 $\frac{p_{1j}}{p_{.j}}, \frac{p_{2j}}{p_{.j}}, \ldots$ 满足作为概率函数的两个条件:

$$(1)\frac{p_{ij}}{p_{\cdot j}} \ge 0, i = 1, 2, \dots$$

$$(2)\sum_{i} \frac{p_{ij}}{p_{\bullet j}} = \frac{1}{p_{\bullet j}} \sum_{i} p_{ij} = 1$$

 \square 类似的,对任意一个固定的 $i, j = 1, 2, \dots$,称

$$P(Y = b_j | X = a_i) = \frac{p_{ij}}{p_{i}}, j = 1, 2, \dots$$

为已知 $\{X = a_i\}$ 发生的条件下Y(记作" $Y \mid X = a_i$ ")的条件概率(质量)函数(或条件分布律,或条件(概率)分布)。

注意下标的不同

由其中一个随机变量的边缘概率函数,及这个随机变量取任意一个固定值时另一个随机变量的条件概率函数,则可唯一确定联合概率函数

例2.7 设随机向量(X,Y)的联合概率函数为

XX	1	2	3	
1	1/4	1/8	1/8	1/2
2	1/12	1/12	1/6	1/3
3	1/18	1/18	1/18	1/6
	7/18	19/72	25/72	1

问(1)已知事件 {Y=1} 发生时X的条件概率函数(9,3,2)/14

(2) 已知事件 {X=2} 发生时Y的条件概率函数(1,1,2)/4

例2.8 某地公安部门经过调查后发现,交通事故由自 行车造成的(记为"X=1")占1/2,由汽车造成的 (记为"X=2") 占1/3, 由其它原因造成的(记为 "X=3") 占1/6。由自行车造成的交通事故引起轻 伤的(记为"Y=1")占1/2,引起重伤的(记为 "Y=2")与死亡的(记为"Y=3")各占1/4。由 汽车造成的交通事故引起轻伤的与重伤的个占1/4, 引起死亡的占1/2。由其它原因造成的交通事故引 起轻伤、重伤、死亡的比例相同。试求X与Y的联 合概率函数。

- 例:把一颗骰子独立地上抛两次,设X表示第一次出现的点数,Y表示两次出现点数的最大值。试求:
- (1) X与Y的联合概率函数;
- (2) P(X=Y);
- (3) X,Y的边缘概率函数;
- (4) {Y=3}发生的条件下X的条件概率函数
- (5) X与Y是否独立。

§ 2.6 随机变量函数的分布

- 一、一维随机变量的函数的概率函数
- \square 例:测量某个零件的横截面的半径 \mathbf{R} ,从而知道横截面 πR^2
- □ 当X的概率函数为 $P(X = a_i) = p_i, i = 1, 2, \cdots$ 时,随机变量 Y = g(X)的函数如表2.26所示,如果有若干个 $g(a_i)$ 的值相等,那么,必须把相应的概率 p_i 相加后合并成一项。

Y	$g(a_1)$	$g(a_2)$	
P_r	P_1	P_2	

例: ⁻	X	-1	0	1				
_	Р	0.5	0.25	0.25				
求了	X^2 的	概率函数	女。					
	X^2	1	0	<u>1</u>	_			
	Р	0.5	0.25	0.25	_			

例2.9 设随机变量X的概率函数如表所示,

✔ 试求下列随机变量的概率函数

(1)
$$Y = \sin X$$

(2)
$$Z = \cos X$$

X	$-\frac{\pi}{2}$	0	$\frac{\pi}{2}$
P_r	0.2	0.3	0.5

例2.9 设随机变量X的概率函数如表所示,

✓ 试求下列随机变量的概率函数

(1)
$$Y = \sin X$$

(2)
$$Z = \cos X$$

X	π	0	$\underline{\pi}$
	2		2
P	0.2	0.3	0.5
<i>r</i>			

二、二维随机变量函数的概率函数

□ 设(X,Y)的概率函数为

$$P(X = a_i, Y = b_j) = p_{ij}, i, j = 1, 2, \dots$$

随机变量 Z = g(X,Y) 的概率函数如表2.27所示,如果若干个 $g(a_i,b_i)$ 的值相等,那么必须把相应的概率 P_{ij} 相加后合成一项。

Z	 $g(a_i,b_j)$	
P_r	 P_{ij}	

作业2.26

例2.10 设二维随机变量(X,Y)的两个边缘概率函数分布 如表所示

X	0	1	-	Y	-1	0	1
P_r	1/2	1/2	-	P_r	1/6	1/3	1/2

已知X与Y相互独立,试求下列随机变量的概率函数

(1)
$$Z = X + Y^2$$

(2)
$$U = \max(X, Y)$$

分布的可加性

定理2.3

设 X_1,\cdots,X_n 是独立同分布的随机变量(即 X_1,\cdots,X_n 相互独立且它们同分布),且 $X_i \sim B(1,p), i=1,\cdots,n$ 记 $Y=X_1+\cdots+X_n$,那么, $Y\sim B(n,p)$ 。

证:

由于每一个 X_i 的值域都是 $\{0,1\}$,i=0,1,...,n。因此,Y的值域 $\Omega_Y = \{0,1,...,n\}$ 。事件 $\{Y=k\}$ 表示 $X_1,...,X_n$ 中恰有k个是

1,n-k个是0,且 $P(X_i=1)=p, i=1, \dots, n$ 因此可以把它看作是n重贝努力试验。按二项概率计算公式:

$$P(X=k) = {n \choose k} p^k (1-p)^{n-k}, k = 0, 1, \dots, n$$

这表明 $Y \sim B(n, p)$ 。

定理2.4(分布的可加性)

设X与Y相互独立

(i) 当
$$X \sim B(m,p), Y \sim B(n,p)$$
时, $X+Y \sim B(m+n,p)$,

(ii) 当
$$X \sim P(\lambda_1), Y \sim P(\lambda_2)$$
 时, $X + Y \sim P(\lambda_1 + \lambda_2)$ 。

证明:略

例2.11 设X与Y相互独立,且 $X \sim P(\lambda_1), Y \sim P(\lambda_2)$ 。 试求 $P(X = k \mid X + Y = n)$ 其中,k,n是整数, $0 \le k \le n$

定理**2.5** 设 X_1, \dots, X_n 是相互独立的随机变量。对于任意一个整数 m, $1 \le m \le n-1$,随机变量 $g(X_1, \dots, X_m)$ 与 $h(X_{m+1}, \dots, X_n)$ 相互 独立,其中函数g,h都是单值函数。

证: 为了不是记号过于复杂,下面仅就n=2,m=1进行证明,按习惯,记 $(X_1,X_2)=(X,Y)$,且设

$$P(X = a_i, Y = b_j) = p_{ij}, i, j = 1, 2, \dots$$

不失一般性,假定 $g(a_1), g(a_2), \cdots$ 两两不相等, $h(b_1), h(b_2), \cdots$ 两两不相等,于是,对任意的i,j=1,2,…,总有

$$P(g(X) = g(a_i), h(Y) = h(b_j))$$

$$= P(X = a_i, Y = b_j)$$

$$= P(X = a_i)P(Y = b_j)$$

$$= P(g(X) = g(a_i))P(h(Y) = h(b_j))$$

□ 由独立性的定义推得g(X)与h(Y)相互独立。

注意: 定理2.5的逆定理不成立,例如 X^2 与 Y^2 相互独立不能得到X 与Y相互独立。

XY	-1	0	1	$P_{i.}$	Y ² X ² 0 1	Ο	1	$\mid P$
-1	1/4	0	0	1/4	$\frac{X^2}{0}$	1/,	1/4	1/2
0	0	1/4	1/4	1/2	1	/4 1/ ₄	/4 1/ ₄	1/2
1	0	1/4	0	1/4	- D	1/	1/	1
$P_{.i}$	1/4	1/2	1/4	1	$P_{.j}$	/2	/2	1

 X^2 与 Y^2 独立,但X与Y不独立.当 X^2 =1时,X=1或-1