概率论与数理统计B

第一次习题课

3月25日　朱心远　PB17000015　zhuxinyuan@mail.ustc.edu.cn

第三周作业 参考答案

2-25.设某种昆虫单只每次产卵的数量服从参数为 λ的Poisson分布, 而每个虫卵能孵出幼虫的概率均为 $p\ (0<p<1)$且相互独立. 分别以 Y 和 Z 记一只昆虫一次产卵后幼虫和未能孵出幼虫的虫卵的个数. 试问Y 和Z 分别服从什么分布? 它们是否相互独立?

解:

> *Slides 第二章(上)Page 28, 29*
>
> *n重伯努利试验; 二项分布*
>
> *Slides 第一章(下)Page 52*
>
> *全概率公式: 如果事件$A_1, A_2, \ldots, A_n$互不相容, $B \subset \bigcup_{j=1}^{n} A_j$, 则*
>
> $$\mathrm{P}(B)=\sum_{j=1}^{n} \mathrm{P}\left(A_j\right) \mathrm{P}\left(B \mid A_j\right).$$

记X表示一只昆虫每次产卵的数量, 有$X \sim P(\lambda)$, 即$P(X=n)=\frac{\lambda^n}{n!}e^{-\lambda}$.

考虑该昆虫每次产卵后幼虫的个数Y,

$$\begin{aligned}
P(Y=k) &= \sum_{n=k}^{\infty} P(X=n) P(Y=k \mid X=n) \\
&= \sum_{n=k}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \\
&= \frac{e^{-\lambda} p^k}{k!} \sum_{n=k}^{\infty} \frac{\lambda^n (1-p)^{n-k}}{(n-k)!} \\
&= \frac{e^{-\lambda} p^k}{k!} \sum_{n=0}^{\infty} \frac{\lambda^{n+k} (1-p)^{n+k-k}}{(n+k-k)!} \\
&= \frac{e^{-\lambda} p^k \lambda^k}{k!} \sum_{n=0}^{\infty} \frac{\lambda^n (1-p)^n}{n!} \\
&= \frac{e^{-\lambda} (p\lambda)^k}{k!} e^{\lambda(1-p)} \\
&= \frac{(\lambda p)^k}{k!} e^{-\lambda p}, \quad k=0,1,2,\ldots.
\end{aligned}$$

也就是, $Y \sim P(\lambda p)$.

将上述过程中的p替换为$1-p$, 我们可以立刻得到关于Z的类似结论, 即$Z \sim P(\lambda(1-p))$.

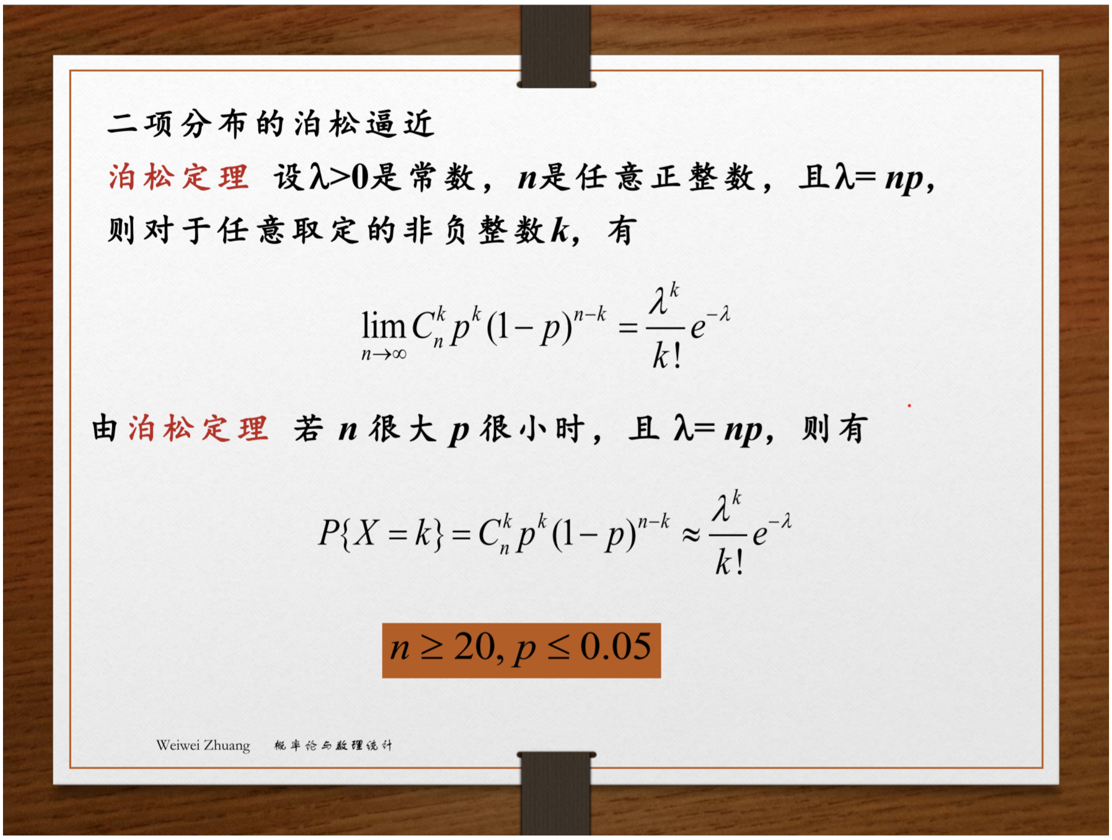
考虑Y和Z之间的独立性，应当从验证$P(Y=m,Z=n)=P(Y=m)P(Z=n)$是否成立出发.

$$\begin{aligned}
P(Y=m,Z=n) &= P(Y=m|X=m+n)P(X=m+n) \\
&= \binom{m+n}{m}p^m(1-p)^n\cdot\frac{\lambda^{m+n}}{(m+n)!}e^{-\lambda} \\
&= \frac{(m+n)!}{m!n!}p^m(1-p)^n\cdot\frac{\lambda^{m+n}}{(m+n)!}e^{-\lambda} \\
&= \frac{(\lambda p)^m}{m!}e^{-\lambda p}\cdot\frac{(\lambda(1-p))^n}{n!}e^{-\lambda(1-p)} \\
&= P(Y=m)\cdot P(Z=n).
\end{aligned}$$

所以我们有结论：Y与Z相互独立.

2-27. 保险公司的资料表明持某种人寿保险单的人在保险期内死亡的概率为2%. 利用Poisson分布，试求在400份保单中最终至少赔付两份保单的概率(确到小数点后三位).

解：

> *《概率论与数理统计》(陈希孺) Page46.*
>
> *"...泊松分布可作为二项分布的极限而得到. 一般地说，若$X\sim B(n,p)$，其中n很大，p很小，而$np=\lambda$不太大时，则X的分布接近于泊松分布$P(\lambda)$. 这个事实在所述条件下可将较难计算的二项分布转化为泊松分布去计算...."*
>
> *Slides 第二章(上) Page35.*
>
> 
>

我们用X表示赔付保单的数目.

$np = 400 \cdot 2\% = 8$. 我们近似地认为$X \sim P(8)$, 也就是

$P(X = k) = \frac{8^k}{k!}e^{-8}, k = 0, 1, \cdots$.

$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - e^{-8} - 8e^{-8} = 1 - 9e^{-8} \approx 0.997$.

2-32. 设随机变量X的分布函数为

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x}{4}, & 0 \leq x < 1 \\ \frac{1}{2} + \frac{x-1}{4}, & 1 \leq x < 2 \\ \frac{5}{6}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

试求(1) $P(X = k), k = 1, 2, 3$; (2) $P(\frac{1}{2} < X < \frac{3}{2})$.

请注意, 原题目中分布函数在$x = 1$处有误, 这里已经修正.

解:

> *分布函数的右连续性, 由分布函数的定义 $F(x) = P(X \leq x)$要求.*
>
> $P(X = x) = P(X \leq x) - P(X < x) = F(x) - P(X < x) = F(x) - \lim_{x\prime \to x} F(x\prime)$
>
> .
>
> *所以, 如果$F(x)$在$x = x_0$处连续, 我们有$P(X = x_0) = 0$.*

(1). 考虑到$P(X = k) = F(k) - \lim_{x \to k^-} F(x)$, 我们有

$P(X = 1) = F(1) - \lim_{x \to 1^-} F(x) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$;

$P(X = 2) = F(2) - \lim_{x \to 2^-} F(x) = \frac{5}{6} - \frac{3}{4} = \frac{1}{12}$;

$P(X = 3) = F(3) - \lim_{x \to 3^-} F(x) = 1 - \frac{5}{6} = \frac{1}{6}$.

(2). 注意到分布函数在$X = \frac{1}{2}$和$X = \frac{3}{2}$附近是连续的, 所以无需考虑边界取值开闭的问题.

$P(\frac{1}{2} < X < \frac{3}{2}) = F(\frac{3}{2}) - F(\frac{1}{2}) = \frac{5}{8} - \frac{1}{8} = \frac{1}{2}$.

2-45. 若随机变量X服从区间$(-5, 5)$上的均匀分布, 求方程$x^2 + Xx + 1 = 0$有实根的概率.

解:

> *Slides 第二章(中) Page 42.*

第二章 随机变量及概率分布 §2.2 连续型随机变量 §2.3 随机变量函数的分布

§2.2 连续型随机变量

下面介绍几个常见的连续型随机变量.

(1) 均匀分布(Uniform 分布) $\mathcal{U}(a,b)$: 对$a<b$, 如果X的概率密度是

$$f(x)=\begin{cases}\frac{1}{b-a}, & x\in(a,b),\\ 0, & x\notin(a,b),\end{cases} \qquad (2.2.3)$$

则称X服从区间(a,b)上的均匀分布, 记做$X\sim\mathcal{U}(a,b)$. 这里$\mathcal{U}$ 是Uniform的缩写.

明显, 表达式(2.2.3)中的区间(a,b)也可以写成$(a,b]$, $[a,b)$或$[a,b]$. 为了方便, 还可以把X的概率密度简写成

$$f(x)=\frac{1}{b-a},\quad x\in(a,b).$$

Weiwei Zhuang 概率论与数理统计

$X\sim U(-5,5)$, 即$f(x)=\begin{cases}\frac{1}{10}, & x\in(-5,5)\\ 0, & else\end{cases}$.

该方程有实根$\Leftrightarrow X^2-4\geq 0$.

记A表示该方程有实根, 也就是说, $P(A)=P(x^2\geq 4)=P(x\in(-5,-2]\cup[2,5))$

$$=\int_{-5}^{-2}\frac{1}{10}\mathrm{d}\,x+\int_{2}^{5}\frac{1}{10}\mathrm{d}\,x=\frac{3}{5}.$$

2-50. 设顾客在某银行的窗口等待服务的时间X服从参数为$\lambda=1/5$的指数分布(单位: 分钟). 假设某顾客一旦等待时间超过**10**分钟他就立即离开且一个月内要到该银行**5**次, 试求他在一个月内至少有一次未接受服务而离开的概率.

解:

> *指数分布的概率密度函数* $f(x)=\lambda e^{-\lambda x}, x>0$;
>
> *指数分布的分布函数* $F(x)=1-e^{-\lambda x}, x>0$.

记A表示他一个月内至少有一次未接受服务而离开的事件. 记$\bar{A}$表示他一个月内到该银行5次均接受了服务.

对于顾客一次到银行的等待时间, 我们有

$$P(X\leq 10)=F(10)=1-e^{-2}$$

所以

$$P(A)=1-P(\bar{A})=1-(P(X\leq 10))^5=1-(1-e^{-2})^5\approx 0.517.$$

2-55. 由学校到飞机场有两条路线可供选择：第一条要穿过市区，路程短但塞车现象严重，所需时间(单位：分钟)服从正态分布$N(30,100)$，另一条是环城公路，路程长但很少塞车，所需时间服从正态分布$N(40,16)$如果要求**(1)**在**50**分钟内达到机场；**(2)**在**45**分钟内达到机场. 试问各应该选择哪条路线？

解:

> *在实际应用上，常考虑一组数据具有近似于正态分布的概率分布。若其假设正确，则约**68.3%**数值分布在距离平均值有1个标准差之内的范围，约**95.4%**数值分布在距离平均值有2个标准差之内的范围，以及约**99.7%**数值分布在距离平均值有3个标准差之内的范围。*
>
> *查表可在教材page376 附表1. 或者下载*
>
> *https://www.math.arizona.edu/~rsims/ma464/standardnormaltable.pdf*

记X_1, X_2分别从学校分别经过市区路和环城公路到达机场所需要的时间, 即
$X_1 \sim N(30,100), X_2 \sim N(40,16)$.

(1) $P(0 < X_1 \le 50) \approx P(X_1 \le 50) = \Phi(\frac{50-30}{10}) = \Phi(2) \approx 0.9772$.

$P(0 < X_2 \le 50) \approx P(X_2 \le 50) = \Phi(\frac{50-40}{4}) = \Phi(2.5) \approx 0.9938$.

所以应该选择环城公路.

(2) $P(0 < X_1 \le 45) \approx P(X_1 \le 45) = \Phi(\frac{45-30}{10}) = \Phi(1.5) \approx 0.9332$.

$P(0 < X_2 \le 45) \approx P(X_2 \le 45) = \Phi(\frac{45-40}{4}) = \Phi(1.25) \approx 0.8944$.

所以应该选择穿过市区的路.

第四周作业 参考答案

2-63. 设随机变量X的分布律为

X	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$
P	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

试求随机变量$Y = cos(2X-\pi)$和$Z = |X - \frac{\pi}{2}|$的分布律.

解:

$P(Y=1) = P(X=\frac{\pi}{2}) + P(X=\frac{3\pi}{2}) = \frac{1}{2}$.

$P(Y=-1) = P(X=0) + P(X=\pi) = \frac{1}{2}$.

所以

Y	-1	1
P	$\frac{1}{2}$	$\frac{1}{2}$

$P(Z=0) = P(X=\frac{\pi}{2}) = \frac{1}{3}$.

$P(Z=\frac{\pi}{2})=P(X=0)+P(X=\pi)=\frac{1}{2}$.

$P(Z=\pi)=P(X=\frac{3\pi}{2})=\frac{1}{6}$.

所以

Z	0	$\frac{\pi}{2}$	π
P	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

2-65. **设随机变量 $X\sim U(0,1)$, 试求下列随机变量的密度函数. (1) $Y_1=e^X$; (2) $Y_2=X^{-1}$; (3)$Y_3=-\frac{1}{\lambda}\ln X$, 其中$\lambda>0$为常数.**

解:

> *概率论与数理统计(陈希孺) Page 79.*
>
> *设 X 有概率密度$f(x)$. 设$Y=g(X)$, g是一个严格单调的函数, 其反函数 $X=h(Y)$ 存在.*
>
> $$f_Y(y)=f(h(y))\,|h'(y)|.$$
>
> *或者用*
>
> $$P(X=h(y))=|f(h(y))\mathrm{d}h(y)|=f(h(y))\,|h'(y)|\,\mathrm{d}y.$$

$X\sim U(0,1)$, 也就是$f(x)=\begin{cases}1, & x\in(0,1)\\ 0 & else\end{cases}$.

(1) $P(Y_1=y)=P(e^X=y)=P(X=\ln y)=f_X(\ln y)|\mathrm{d}\ln y|=\frac{1}{y}dy, \ln y\in(0,1)$.

$$f_{Y_1}(y)=\begin{cases}\frac{1}{y}, & y\in(1,e)\\ 0, & else\end{cases}.$$

(2)
$P(Y_2=y)=P(X^{-1}=y)=P(X=y^{-1})=f_X(y^{-1})|\mathrm{d}\,y^{-1}|=\frac{1}{y^2}\mathrm{d}\,y, y^{-1}\in(0,1)$.

$$f_{Y_2}(y)=\begin{cases}\frac{1}{y^2}, & y>1\\ 0, & else\end{cases}.$$

(3)
$P(Y_3=y)=P(-\frac{1}{\lambda}\ln X=y)=P(X=e^{-\lambda y})=f_X(e^{-\lambda y})|\mathrm{d}\,e^{-\lambda y}|=\lambda e^{-\lambda y}\,\mathrm{d}\,y, e^{-\lambda y}\in(0,1)$
.

$$f_{Y_3}(y)=\begin{cases}\lambda e^{-\lambda y}, & y>0\\ 0, & else\end{cases}.$$

2-71. **设随机变量X服从参数为λ的指数分布, 且随机变量Y 定义为**

$$Y=\begin{cases}X, & \text{若} X\geq 1,\\ -X^2, & \text{若} X<1.\end{cases}$$

试求Y的密度函数$p(y)$.

解:

从Y的定义中我们可以看出,Y可能的取值范围包括$[1,+\infty)\cup(-1,0)$.

$X\sim E(\lambda)$, 也就是$f_X(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & else\end{cases}$

如果$Y\geq 1, P(Y=y)=P(X=y)=f_X(y)|\mathrm{d}\,y|=\lambda e^{-\lambda y}\,\mathrm{d}\,y, y>1.$

如果$Y\in(-1,0), P(Y=y)=P(-X^2=y)=P(X=\sqrt{-y})=f_X(\sqrt{-y})|\mathrm{d}\sqrt{-y}|$
$$=\lambda e^{-\lambda\sqrt{-y}}\cdot\frac{1}{2}\frac{1}{\sqrt{-y}}\mathrm{d}\,y=\frac{\lambda}{2\sqrt{-y}}e^{-\lambda\sqrt{-y}}\,\mathrm{d}\,y,-1<y<0.$$

所以

$$f_Y(y)=\begin{cases}\lambda e^{-\lambda y}, & y\geq 1\\ \frac{\lambda}{2\sqrt{-y}}e^{-\lambda\sqrt{-y}}, & -1<y<0\\ 0, & else\end{cases}.$$

2-72. 设随机变量X服从标准正态分布, 试求下列随机变量的密度函数. **(1)** $Y_1=e^X$; **(2)** $Y_2=|X|$; **(3)**$Y_3=2X^2+1$.

解:

$X\sim N(0,1), f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$

(1)
$P(Y_1=y)=P(e^X=y)=P(X=\ln y)=f_X(\ln y)|\mathrm{d}\ln y|=\frac{1}{\sqrt{2\pi}y}e^{-\frac{(\ln y)^2}{2}}\,\mathrm{d}\,y,\ln y\in\mathbb{R}$
.

所以

$$f_{Y_1}(y)=\begin{cases}\frac{1}{\sqrt{2\pi}y}e^{-\frac{(\ln y)^2}{2}}, & y>0\\ 0, & else\end{cases}.$$

(2) $P(Y_2=y)=P(|X|=y)=P(X=-y)+P(X=y)$.
$$=\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}|\mathrm{d}(-y)|+\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}|\mathrm{d}\,y|=\sqrt{\frac{2}{\pi}}e^{-\frac{y^2}{2}}\,\mathrm{d}\,y,y>0$$

所以

$$f_{Y_2}(y)=\begin{cases}\sqrt{\frac{2}{\pi}}e^{-\frac{y^2}{2}}, & y>0\\ \frac{1}{\sqrt{2\pi}}, & y=0\\ 0, & else\end{cases}\cdot$$

(3) $P(Y_3=y)=P(2X^2+1=y)=P(X=\sqrt{\frac{y-1}{2}})+P(X=-\sqrt{\frac{y-1}{2}})$

$=\frac{1}{\sqrt{2\pi}}e^{-\frac{y-1}{4}}\,|\,\mathrm{d}\sqrt{\frac{y-1}{2}}|+\frac{1}{\sqrt{2\pi}}e^{-\frac{y-1}{4}}\,|\,\mathrm{d}(-\sqrt{\frac{y-1}{2}})|$

$=\frac{1}{\sqrt{2\pi}}e^{-\frac{y-1}{4}}\,\frac{1}{4}\sqrt{\frac{2}{y-1}}\,\mathrm{d}\,y+\frac{1}{\sqrt{2\pi}}e^{-\frac{y-1}{4}}\,\frac{1}{4}\sqrt{\frac{2}{y-1}}\,\mathrm{d}\,y$

$=\frac{1}{2\sqrt{\pi(y-1)}}e^{-\frac{y-1}{4}}\,\mathrm{d}\,y,y>1.$

所以

$$f_{Y_3}(y)=\begin{cases}\frac{1}{2\sqrt{\pi(y-1)}}e^{-\frac{y-1}{4}}, & y>1\\ \frac{1}{\sqrt{2\pi}}, & y=1\\ 0, & else\end{cases}\cdot$$

3-4. **箱中装有6个球,其中红、白、黑球的个数分别为1, 2, 3个.现从箱中随机地取出2个球,记X为取出的红球个数,Y为取出的白球个数.求随机变量(X, Y)的概率分布.**

解:

$P(X=0,Y=0)=\frac{\binom{3}{2}}{\binom{6}{2}}=\frac{1}{5}.$ // 取2个黑球

$P(X=0,Y=1)=\frac{\binom{2}{1}\binom{3}{1}}{\binom{6}{2}}=\frac{2}{5}.$ //取1个白球, 1个黑球

$P(X=0,Y=2)=\frac{\binom{2}{2}}{\binom{6}{2}}=\frac{1}{15}.$ //取2个白球

$P(X=1,Y=0)=\frac{\binom{1}{1}\binom{3}{1}}{\binom{6}{2}}=\frac{1}{5}.$ //取1个红球, 1个黑球

$P(X=1,Y=1)=\frac{\binom{1}{1}\binom{2}{1}}{\binom{6}{2}}=\frac{2}{15}.$ //取1个红球, 1个白球

所以

	$X=0$	$X=1$
$Y=0$	$\frac{1}{5}$	$\frac{1}{5}$
$Y=1$	$\frac{2}{5}$	$\frac{2}{15}$
$Y=2$	$\frac{1}{15}$	0

3-9. **从一副扑克牌(共52张)中任取13张,以X和Y分别记其中的黑桃和红桃张数,试求:**
(1) **(X,Y)的联合概率分布;**
(2) **已知取出的只有一张黑桃,求此时Y的条件概率分布.**

解:

(1)

$$P(X=i,Y=j)=\frac{\binom{13}{i}\binom{13}{j}\binom{26}{13-i-j}}{\binom{52}{13}},\ \ i,j\in\mathbb{N},\ i,j\le 13, i+j\le 13.$$

(2)

$$\begin{aligned}P(Y=j|X=1)&=\frac{P(X=1,Y=j)}{P(X=1)}\\&\frac{\binom{13}{1}\binom{13}{j}\binom{26}{12-j}}{\binom{52}{13}}\Big/\frac{\binom{13}{1}\binom{39}{12}}{\binom{52}{13}}\\&=\frac{\binom{13}{j}\binom{26}{12-j}}{\binom{39}{12}},\ \ j\in\mathbb{N},\ j\le 12.\end{aligned}$$

3-10. 设二维随机向量的概率分布为

$Y \backslash X$	-1	1
-1	0.2	b
1	a	0.3

已知事件$\{X=-1\}$和$\{X+Y=0\}$相互独立, 求**a, b.**

解:

$P(X=-1)=a+0.2;$

$P(X+Y=0)=P(X=-1,Y=1)+P(X=1,Y=-1)=a+b.$

$P(X=-1,X+Y=0)=P(X=-1,Y=1)=a.$

由于事件$\{X=-1\}$和$\{X+Y=0\}$相互独立, 我们有

$a=(a+0.2)(a+b),$

又因为$a+b+0.2+0.3=1,$

我们可以得到$\begin{cases}a=0.2\\b=0.3\end{cases}.$

Slides 第二章（下）**Page 26**：使用**Page 16**提供的定义计算例**2.4.3:** 计算X_1和(X_1,X_2)的边缘分布.

§2.4.1 离散型随机向量

如果X, Y都是离散型随机变量, 则称(X, Y)是离散型随机向量. 设离散型随机向量(X, Y) 有联合概率分布

$$p_{ij} = P(X = x_i, Y = y_j), \quad i, j \geqslant 1, \tag{2.4.1}$$

则X和Y分别有概率分布

$$\begin{aligned} p_i \equiv P(X = x_i) &= \sum_{j=1}^{\infty} P(X = x_i, Y = y_j) = \sum_{j=1}^{\infty} p_{ij}, \ i \geqslant 1, \\ q_j \equiv P(Y = y_j) &= \sum_{i=1}^{\infty} P(X = x_i, Y = y_j) = \sum_{i=1}^{\infty} p_{ij}, \ j \geqslant 1. \end{aligned}$$

这时称X的分布$\{p_i\}$, Y的分布$\{q_j\}$为(X, Y)的边缘分布.

例2.4.2(三项分布)　设A, B, C是试验S的完备事件组, $P(A) = p_1$, $P(B) = p_2$, $P(C) = p_3$. 对试验S进行n次独立重复试验时, 用X_1, X_2, X_3分别表示A, B, C发生的次数, 则(X_1, X_2, X_3) 的联合分布是

$$P(X_1 = i, X_2 = j, X_3 = k) = \frac{n!}{i!\, j!\, k!} p_1^i p_2^j p_3^k,$$

其中$i, j, k \geqslant 0$, $i + j + k = n$.

解:

$$P\left(X_1=i, X_2=j\right)=P\left(X_1=i, X_2=j, X_3=n-i-j\right)=\frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j}.$$

$$\begin{aligned} & P\left(X_1=i\right)=\sum_{j+k=n-i} P\left(X_1=i, X_2=j, X_3=k\right)=\sum_{j+k=n-i} \frac{n!}{i!j!k!} p_1^i p_2^j p_3^k \\ & =\sum_{j=0}^{n-i} \frac{n!}{i!j!(n-i-j)!} p_1^i p_2^j p_3^{n-i-j}=\frac{n!}{i!(n-i)!} p_1^i \sum_{j=0}^{n-i} \frac{(n-i)!}{j!(n-i-j)!} p_2^j p_3^{n-i-j} \\ & =\binom{n}{i} p_1^i\left(p_2+p_3\right)^{n-i} \\ & =\binom{n}{i} p_1^i\left(1-p_1\right)^{n-i} \end{aligned}.$$