随机变量的分布、数字特征与极限定理

葛旭
哈尔滨工业大学（深圳）
自动化五班
广东，深圳

***Abstract*—**随机变量在概率论中是一个非常重要的关键点**.**利用随机变量的分布我们可以对一个或者多个随机变量预测它的取值情况来获得我们想要的结果。利用它的数字特征我们可以得出诸如方差、期望、协方差等特征量来描述随机变量。根据极限定理我们可以对较大的样本量进行取值的范围估计。因此这一部分内容十分重要，堪称核心内容。

I. 一维随机变量的分布

A. 离散型随机变量

1) 概率分布列*:* 只能取有限个值或者可列无穷多个值的随机变量X称为离散随机变量。设x的所有可能取值为x_1, x_2,···,x_k,···,我们知道对应的概率

$$P(X = x_k) = p_k(k = 1, 2, \cdots) \tag{1}$$

那么我们可以描述x的概率分布,式（1）被称为离散型随机变量x的概率分布列或简称分布列，又称分布律。也可以用如下表格形式描述:

TABLE I
概率分布列

X	x_1	x_2	$\ldots$	x_k	$\ldots$
P	p_1	p_2	$\ldots$	p_3	$\ldots$

对任意分布列都有如下两个性质:

$$p_k \geqslant 0(k = 1, 2, \cdots) \tag{2}$$

$$\sum_k p_k = 1 \tag{3}$$

反之，满足上述两条性质的数列p_k也可作为某一离散随机变量的分布列。

*2) 0-1*分布*:* 若随机变量X只可能取0和1两个值，它的分布列是

$$P(X = 1) = p, P(X = 0) = 1 - p = q, 0 < p < 1 \tag{4}$$

或
则称X服从0-1分布或伯努利分布，也称两点分布，记

TABLE II
0-1分布

X	0	1
P	1-p	p

为$X \sim B(1, p)$

3) 二项分布*:* 若随机变量X的分布列为

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

则称X服从二项分布(参数为n,p),常用记号$X \sim N(n, p)$表示。当n=1时即为0-1分布。

4) 泊松分布*:* 若随机变量X的分布列为

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \lambda > 0(k = 0, 1, 2, \cdots) \tag{5}$$

则称X服从参数为λ的泊松分布，并用记号$X \sim P(\lambda)$表示。由式(5)可知，$P(X = k) \geqslant 0(k = 0, 1, 2, \cdots)$且有

$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = 1$$

故泊松分布满足分布列的两个性质。以n,p为参数的二项分布，当n很大，p很小的时候，近似于$\lambda = np$为参数的泊松分布。

5) 几何分布*:* 若随机变量的分布列为

$$P(X = k) = q^{k-1}p(k = 1, 2,), 0 < p < 1, q = 1 - p \tag{6}$$

则称X服从参数为p的几何分布，并用记号$X \sim G(p)$表示。
由式(6)可知，$P(X = k) \geqslant 0(k = 0, 1, 2, \cdots)$，且有

$$\sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} q^{k-1}p = p \cdot \frac{1}{1-q} = 1$$

故几何分布满足分布列的两个性质。 设$X \sim G(p)$，n,m为任意的两个正整数，则

$$P(X > n + m | X > n) = P(X > m)$$

这个性质被称为几何分布的无记忆性。

6) 超几何分布*:* 若随机变量的分布列为

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}(k = 0, 1, \cdots, l) \tag{7}$$

其中$l = min(M, n)$,并规定当$i \geqslant m$时，$C_M^i = 0$。由式(7)确定的分布为超几何分布。 二项分布可用来描述有放回抽样，超几何分布可用来描述不放回抽样。

B. 连续型随机变量

1) 均匀分布: 若连续型随机变量X的概率密度

$$f(x)=\begin{cases}\frac{1}{b-a},a\leqslant x\leqslant b\\0,other\end{cases},a<b \tag{8}$$

则称X在区间[a,b]上服从均匀分布，记为$X\sim U[a,b]$。由式(8)，相应的分布函数为

$$F(x)=\begin{cases}0,x<a\\\frac{x-a}{b-a},a\leqslant x<b\\1,x\geqslant b\end{cases} \tag{9}$$

2) 指数分布: 若连续型随机变量X的概率密度

$$f(x)=\begin{cases}0,x\leqslant 0\\\lambda e^{-\lambda x},x>0\end{cases}$$

其中λ是正常数，则称X服从参数为λ的指数分布。记为$X\sim E(\lambda)$。 其分布函数为

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

3) 正态分布: 若连续型随机变量X的概率密度

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

μ,σ为常数，且$\sigma>0$，则称X服从参数为μ,σ的正态分布，也称X为正态变量，记作$X\sim N(\mu,\sigma^2)$

II. 二维随机变量的分布

A. 二维离散随机变量

若二维随机变量(X,Y)所有可能取的值是有限多对或可列无限多对，则称(X,Y)为二维离散随机变量。有

$$p_{ij}=P(X=x_i,Y=y_i)(i,j=1,2,\cdots) \tag{10}$$

则称$p_{ij}(i,j=1,2,\cdots)$为(X,Y)的分布列，或称为X和Y的联合分布列。 二维离散随机变量的分布函数可表示如下:

$$F(x,y)=P(X\leqslant x,Y\leqslant y)=\sum_{x_i\leqslant x}\sum_{y_i\leqslant y}P(X_i=x,Y_i=y)$$

$$=\sum_{x_i\leqslant x}\sum_{y_i\leqslant y}P_{ij}$$

二维离散随机变量的分布列及边缘分布列可用表格表示，如下表所示:

TABLE III
二维分布列

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

B. 二维连续随机变量

1) 二维均匀分布: 设G是平面上的有界区域，其面积为$S(G)$。若二维随机变量(X,Y)具有概率密度

$$f(x,y)=\begin{cases}\frac{1}{S(G)},(x,y)\in G\\0,other\end{cases} \tag{11}$$

则称(X,Y)在G上服从均匀分布。

2) 二维正态分布: 设二维随机变量(X,Y)具有概率密度

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{-\frac{1}{2(1-\rho^2)}\cdot$$
$$[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]\}$$
$$(-\infty<x<+\infty,-\infty<y<+\infty)$$

称(X,Y)服从参数为$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$的二维正态分布，记为$(X,Y)\sim N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2;\rho)$

C. 二维随机变量函数的分布

1) 和的分布: 设X与Y是相互独立的随机变量，分布列分别为

$$P(X=i)(i=0,1,2,\cdots)$$
$$P(Y=j)(j=0,1,2,\cdots)$$

则

$$P(Z=k)=\sum_{i=0}^{k}P(X=i,Y=k-i)$$
$$=\sum_{i=0}^{k}P(X=i)P(Y=k-i)$$

对于连续二维随机变量,令$z=x+y$,有

$$f_z(z)=\int_{-\infty}^{+\infty}f(x,z-x)dx \tag{12}$$

或

$$f_z(z)=\int_{-\infty}^{+\infty}f(z-y,y)dx \tag{13}$$

2) *max*和*min*的分布: 先求$M=\max(X,Y)$的分布:

$$F_M(z)=P(M\leqslant z)=P(\max(X,Y)\leqslant z)$$
$$=P(X\leqslant z,Y\leqslant z)=P(X\leqslant z)P(Y\leqslant z)$$
$$=F_X(z)F_Y(z)$$

再求$N=\min(X,Y)$的分布:

$$F_N(z)=P(N\leqslant z)=P(\min(X,Y)\leqslant z)$$
$$=1-P(\min(X,Y)>z)=1-P(X>z,Y>z)$$
$$=1-P(X>z)P(Y>z)$$
$$=1-[1-P(X\leqslant z)][1-P(Y\leqslant z)]$$
$$=1-[1-F_X(z)][1-F_Y(z)]$$

特别地，当$X_1, X_2, \cdots, X_n$是相互独立的且具有相同分布函数$F(z)$的n个随机变量时，有

$$F_{\max}(z) = [F(z)]^n \tag{14}$$

$$F_{\min}(z) = 1 - [1 - F(z)]^n \tag{15}$$

III. 随机变量的数字特征

A. 数学期望

1) 离散型随机变量的数学期望： 设离散型随机变量X的分布列

$$P(X = x_i) = p_i (i = 1, 2, \cdots)$$

若级数$\sum_{k=1}^{\infty} x_k p_k$绝对收敛，即$\sum_{k=1}^{\infty} |x_k| p_k < +\infty$，则称$\sum_{k=1}^{\infty} x_k p_k$为$X$ 的数学期望或者均值，记为$E(X)$,即

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

当$\sum_{k=1}^{\infty} |x_k| p_k$发散时，称$X$的数学期望不存在。

a) 0-1分布的数学期望： 设随机变量X的分布列

TABLE IV
0-1分布

X	0	1
P	1-p	p

则$E(X) = p$

b) 二项分布的数学期望：

$$E(x) = \sum_{k=0}^{n} k C_n^k p^k q^{n-k}$$
$$= \sum_{k=0}^{n} k \frac{n(n-1)(n-2)\cdots[n-(k-1)]}{k!} p^k q^{n-k}$$
$$= np \sum_{k=1}^{n} \frac{(n-1)(n-2)\cdots[n-1-(k-2)]}{(k-1)!} p^{k-1} q^{n-1-(k-1)}$$
$$= np \sum_{k-1=0}^{n-1} C_{n-1}^{k-1} p^{k-1} q^{n-1-(k-1)}$$
$$= np(p+q)^{n-1}$$
$$= np$$

c) 泊松分布的数学期望：

$$E(x) = \sum_{k=0}^{n} k \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

2) 连续型随机变量的数学期望：

a) 均匀分布： 设随机变量X的概率密度

$$f(x) = \begin{cases} \frac{1}{b-a}, a \leqslant x \leqslant b \\ 0, other \end{cases}$$

则

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

b) 指数分布： 设随机变量X的概率密度

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, x > 0 \\ 0, other \end{cases}$$

则

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx = \int_0^{+\infty} x \lambda e^{-\lambda x} dx$$
$$= -\int_0^{+\infty} x d e^{-\lambda x} = \int_0^{+\infty} e^{-\lambda x} dx = \frac{1}{\lambda}$$

c) 正态分布： 设$X \sim N(\mu, \sigma^2)$ 则

$$E(x) = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$
$$\xrightarrow{t=\frac{x-\mu}{\sigma}} \int_{-\infty}^{+\infty} \frac{\mu + \sigma t}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$
$$= \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} dt$$
$$= \mu$$

d) 柯西分布： 设随机变量X的概率密度

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2} (-\infty < x < +\infty)$$

由于

$$\int_{-\infty}^{+\infty} |x| \frac{dx}{\pi(1+x^2)} = 2 \int_0^{+\infty} \frac{x dx}{\pi(1+x^2)} = \frac{1}{\pi} \ln(1+x^2)|_0^{+\infty} = +\infty$$

故$E(x)$不存在

e) 随机变量函数的数学期望：

f) 离散型随机变量：

$$E(Y) = E[g(X)] = \sum_{i=1}^{\infty} g(x_i) p_i$$

g) 连续型随机变量：

$$E(Y) = E[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

3) 数学期望的性质：

(1)$E(C) = C$, C为常数

(2)$E(CX) = CE(X)$, C为常数

(3)$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$

(4)若$X_1, X_2, \cdots, X_n$独立，则

$$E(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} E(X_i)$$

B. 方差

设X是一个随机变量，若$E[X - E(X)]^2$存在，则称$E[X - E(X)]^2$是X的方差，记作$D(X)$,即

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

同时，称$\sqrt{D(X)}$是X的标准差或者均方差，记作σ_X,即

$$\sigma_X = \sqrt{D(X)}$$

关于方差的计算，常利用如下的公式：

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

1) 一些常用的方差:

a) 0-1分布: 若X的分布列

TABLE V
0-1分布

X	0	1
P	1-p	p

则$D(X)=p(1-p)$

b) 泊松分布: 设$X\sim P(\lambda)$, 则

$$D(X)=\lambda$$

c) 几何分布: 设$X\sim G(\lambda)$, 则

$$D(X)=\frac{1-p}{p^2}$$

d) 均匀分布: 设$X\sim U(\lambda)$, 则

$$D(X)=\frac{(b-a)^2}{12}$$

e) 指数分布: 设$X\sim E(\lambda)$, 则

$$D(X)=\frac{1}{\lambda}$$

2) 方差的性质:

(1)$D(C)=0, C$为常数
(2)$D(CX)=C^2D(X), C$为常数
(3)若$X_1,X_2,\cdots,X_n$相互独立
$D(X_1+\cdots+X_n)=D(X_1)+D(X_2)+\cdots+D(X_n)$
(4)若$X_1,X_2,\cdots,X_n$相互独立
$D(XY)=D(X)D(Y)+D(X)[E(Y)]^2+D(Y)[E(X)]^2$
(5)$P(X=a)=1,E(X)=a,D(X)=0$

C. 协方差和相关系数

1) 协方差: 设(X,Y)是一个二维随机变量。如果$E\{[X-E(X)][Y-E(Y)]\}$存在，则称它为X与Y的协方差，记作$Cov(X,Y)$,即

$$Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}$$

对于任意两个随机变量X与Y，下式成立。

$$D(X+Y)=D(X)+D(Y)+2Cov(X,Y) \tag{16}$$

由协方差的定义，可得下面的性质：

(1)$Cov(X,Y)=Cov(Y,X)$
(2)$Cov(aX,bY)=abCov(X,Y),a,b$是常数
(3)$Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

2) 相关系数: 设(X,Y)为一个二维随机变量,若X与Y的协方差$Cov(X,Y)$存在，且$D(X)>0,D(Y)>0$,则称$\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(X)}}$为$X$与$Y$的相关系数，记为$\rho$,即

$$\rho=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(X)}}=\frac{E\{[X-E(X)][Y-E(Y)]\}}{\sqrt{D(X)}\sqrt{D(X)}} \tag{17}$$

若X与Y的相关系数$\rho=0$,则称X与Y不相关。
随机变量X与Y不相关与下面的每一个结论都是等价的：

(1)$Cov(X,Y)=0$
(2)$D(X+Y)=D(X)+D(Y)$
(3)$E(XY)=E(X)E(Y)$

X与Y不相关和X与Y相互独立是两个不相同的概念。X与Y不相关是指X与Y不存在线性关系，不是说没有其他关系。若X与Y不相关，X与Y也可能独立也可能不独立。

IV. 随机变量的极限定理

A. 大数定律

1) 切比雪夫不等式: 对任意随机变量X,若它的方差$D(X)$存在，则对任意$\epsilon>0$,有

$$P(|X-E(X)|\geqslant\varepsilon)\leqslant\frac{D(X)}{\varepsilon^2} \tag{18}$$

成立。 同理有

$$P(|X-E(X)|\leqslant\varepsilon)\geqslant 1-\frac{D(X)}{\varepsilon^2} \tag{19}$$

成立。

2) 大数定律:

a) 伯努利大数定律: 设在n重伯努利实验中，成功的次数为Y_n,而在每次实验中成功的概率为$p(0<p<1)$,则对任意$\epsilon>0$,有

$$\lim_{n\to\infty}P(|\frac{Y_n}{n}-p|\geqslant\varepsilon)=0 \tag{20}$$

b) 切比雪夫大数定律: 设$X_1,X_2,\cdots,X_n,\cdots$是相互独立的随机变量序列，若存在常数$C$,使得$D(X_i\leqslant C)$,则对任意$\epsilon>0$有

$$\lim_{n\to\infty}P[|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)|\geqslant\varepsilon]=0 \tag{21}$$

或

$$\lim_{n\to\infty}P[|\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}E(X_i)|<\varepsilon]=1 \tag{22}$$

c) 辛钦大数定律: 设$X_1,X_2,\cdots,X_n,\cdots$是独立同分布的随机变量序列，且具有有限的数学期望$[E(X_i)=\mu(i=1,2,\cdots)$，使得$D(X_i\leqslant C)$,则对任意$\epsilon>0$有

$$\lim_{n\to\infty}P[|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu|\geqslant\varepsilon]=0 \tag{23}$$

或

$$\lim_{n\to\infty}P[|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu|<\varepsilon]=1 \tag{24}$$

成立

B. 中心极限定理

1) Lindeberg-Levi定理: 设$X_1, X_2, \cdots, X_n, \cdots$是独立同分布的随机变量序列，且具有有限的数学期望$[E(X_i) = \mu (i = 1, 2, \cdots)$和方差$D(X_i) = \sigma^2 > 0 (i = 1, 2, \cdots)$,则对一切$x \in \mathbb{R}$有

$$\lim_{n \to \infty} P(\frac{1}{\sqrt{n}\sigma}(\sum_{i=1}^{n} X_i - n\mu) \leqslant x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \quad (25)$$

2) 棣莫弗-拉普拉斯定理: 设在n重伯努利试验中，成功的次数为Y_n,而在每次实验中成功的概率为$p (0 < p < 1), q = 1 - p$,则对一切x有

$$\lim_{n \to \infty} P(\frac{Y_n - np}{\sqrt{npq}} \leqslant x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \Phi(X) \quad (26)$$

当n充分大时，

$$P(a < Y_n \leqslant b) = \Phi(\frac{b - np}{\sqrt{npq}}) - \Phi(\frac{a - np}{\sqrt{npq}})$$

References

[1] 王勇主编；方茹，周永春，李朝艳，田波平，王勇编. 大学数学 概率论与数理统计 第2版[M]. 北京：高等教育出版社, 2014.12.