第五章 随机变量的数字特征

分布函数能够完整地描述随机变量的统计特性，但在一些实际问题中，只需知道随机变量的某些特征，因而不需要求出它的分布函数.

例如：

评定某企业的经营能力时，只要知道该企业**人均赢利水平**；

研究水稻品种优劣时，我们关心的是稻穗的**平均粒数**及每粒的**平均重量**；

检验棉花的质量时，既要注意纤维的**平均长度**，又要注意**纤维长度与平均长度的偏离程度**，平均长度越长、偏离程度越小，质量就越好;

考察一射手的水平，既要看他的平均环数是否高，还要看他弹着点的范围是否小，即数据的波动是否小.

由上面例子看到，与随机变量有关的某些数值，虽不能完整地描述随机变量，但能清晰地描述随机变量在某些方面的重要特征，这些数字特征在理论和实践上都具有重要意义.

随机变量某一方面的概率特性

可用数字来描写

本章内容

- 随机变量的平均取值 —— 数学期望
- 随机变量取值平均偏离平均值的情况 —— 方差
- 描述两个随机变量之间的某种关系的数 —— 协方差与相关系数

§5.1 随机变量的数学期望

引例1 测量 50 个圆柱形零件直径（见下表）

尺寸（cm）	8	9	10	11	12	
数量（个）	8	7	15	10	10	50

则这 50 个零件的平均直径为

$$\frac{8\times 8+9\times 7+10\times 15+11\times 10+12\times 10}{50}$$

$$=8\times\frac{8}{50}+9\times\frac{7}{50}+10\times\frac{15}{50}+11\times\frac{10}{50}+12\times\frac{10}{50}\quad=10.14cm$$

换一个角度看，从一堆零件中任取一个零件，它的尺寸为随机变量X，且X的概率分布为

X	8	9	10	11	12
P	$\frac{8}{50}$	$\frac{7}{50}$	$\frac{15}{50}$	$\frac{10}{50}$	$\frac{10}{50}$

则加 权 平 均

$$8\times\frac{8}{50}+9\times\frac{7}{50}+10\times\frac{15}{50}+11\times\frac{10}{50}+12\times\frac{10}{50}$$

$$=\sum_{k=8}^{12}k\times P(X=k)=\sum_{k=8}^{12}kp_k=10.14$$

可以表示这堆零件的平均直径， 称之为数学期望

数学期望的定义

定义1 设 X 为离散型随机变量，其概率分布为

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

若无穷级数

$$\sum_{k=1}^{+\infty} x_k p_k$$

绝对收敛，则称其和为随机变量 X 的数学期望，记作 $E(X)$

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

定义2 设 X 为连续型随机变量, 其密度函数为 $f(x)$

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

则称此积分为随机变量 X 的数学期望, 记作 $E(X)$

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

随机变量的数学期望的本质 —— 加权平均, 它是一个数不再是随机变量

例1 $X \sim B(n, p)$，求 $E(X)$.

解

$$E(X) = \sum_{k=0}^{n} k C_n^k p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

$$= np \sum_{k=0}^{n-1} C_{n-1}^k p^k (1-p)^{(n-1)-k}$$

$$= np$$

例2 $X \sim N(\mu, \sigma^2)$, 求 $E(X)$.

解 $$E(X) = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$\overset{\text{令}\frac{x-\mu}{\sigma}=y}{=} \int_{-\infty}^{+\infty} (y\sigma + \mu) \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2}} \sigma dy$$

$$= \int_{-\infty}^{+\infty} y\sigma \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

$$+ \int_{-\infty}^{+\infty} \mu \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

$$= \mu$$

常见随机变量的数学期望

分布	概率分布	期望
参数为p 的0-1分布	$P(X=1)=p$ $P(X=0)=1-p$	p
$B(n,p)$	$P(X=k)=C_n^k p^k (1-p)^{n-k}$ $k=0,1,2,\cdots,n$	np
$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!}$ $k=0,1,2,\cdots$	λ

分布	概率密度	期望
区间(a,b)上的均匀分布	$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & 其它\end{cases}$	$\dfrac{a+b}{2}$
$E(\lambda)$	$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & 其它\end{cases}$	$\dfrac{1}{\lambda}$
$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ

注意：不是所有的随机变量都有数学期望

例如：Cauchy分布的密度函数为

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

但 $\int_{-\infty}^{+\infty}|x|f(x)dx=\int_{-\infty}^{+\infty}\frac{|x|}{\pi(1+x^2)}dx$ 发散

它的数学期望不存在

随机变量函数的数学期望

- 设X为离散型随机变量，概率分布为

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

$Y = g(X)$,

若级数 $\sum_{i=1}^{\infty} g(x_i) p_i$ 绝对收敛，则

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

$$E(g(X)) = \sum_{i=1}^{\infty} g(x_i) p_i \qquad E(Y) = \sum_{i=1}^{\infty} y_i p(Y = y_i)$$

$$E(Y)=\sum_{i=1}^{\infty} y_i p(Y=y_i)$$

$$=\sum_{i=1}^{\infty}\left\{g(x_j)\sum_{g(x_j)=y_i} p(x_j)\right\}$$

$$=\sum_{i=1}^{\infty}\left\{\sum_{g(x_j)=y_i} g(x_j)p(x_j)\right\}$$

$$=\sum_{k=1}^{\infty} g(x_k)p(x_k)$$

$$=\sum_{i=1}^{\infty} g(x_i)p_i$$

❑ 设X为连续型随机变量，密度函数为$f(x)$

$Y = g(X)$,

若积分 $\int_{-\infty}^{+\infty} g(x) f(x) dx$ 绝对收敛，则

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

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

- 设(X,Y)为二维离散型随机变量，概率分布为

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

$Z=g(X,Y)$,

若级数 $\sum_{i,j=1}^{\infty} g(x_i,y_j)p_{ij}$ 绝对收敛，则

$$E(Z)=\sum_{i,j=1}^{\infty} g(x_i,y_j)p_{ij}$$

$$E(g(X,Y))=\sum_{i,j=1}^{\infty} g(x_i,y_j)p_{ij}$$

- 设(X,Y)为二维连续型随机变量，密度函数为$f(x,y)$

$$Z = g(X,Y),$$

若积分 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$ 绝对收敛，则

$$E(Z) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$$

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

几个重要的随机变量函数的数学期望

$E(X^k)$ —— X 的 k 阶原点矩

$E(X)$ —— X 的 数学期望

$E(|X|^k)$ —— X 的 k 阶绝对原点矩

$E((X-E(X))^k)$ —— X 的 k 阶中心矩

$E((X-E(X))^2)=D(X)$ —— X 的 方差

$E(X^k Y^l)$ —— X,Y 的 $k+l$ 阶混合原点矩

$E\left((X-E(X))^k (Y-E(Y))^l\right)$

—— X,Y 的 $k+l$ 阶混合中心矩

$E(XY)$ —— X,Y 的 二阶原点矩

$E\left((X-E(X))(Y-E(Y))\right)$

—— X,Y 的二阶混合中心矩

X,Y 的协方差

$$E\left(\frac{(X-E(X))(Y-E(Y))}{\sqrt{D(X)}\sqrt{D(Y)}}\right)=\rho_{XY}$$

—— X,Y 的相关系数

例 设$X \sim N(0,1)$, $Y \sim N(0,1)$, X, Y相互独立，求 $E(\max(X,Y))$.

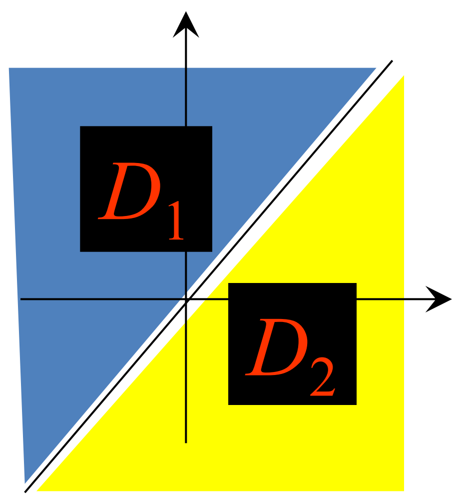
解 $f(x,y) = f_X(x) f_Y(y) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \dfrac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$

$$= \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

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

$$= \iint_{D_1} \max\{x,y\} f(x,y)dxdy$$

$$+ \iint_{D_2} \max\{x,y\} f(x,y)dxdy$$



$$=\iint_{D_1} y\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy+\iint_{D_2} x\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}}dx\int_{x}^{+\infty}ye^{-\frac{y^2}{2}}dy+\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-\frac{y^2}{2}}dy\int_{y}^{+\infty}xe^{-\frac{x^2}{2}}dx$$

$$=\frac{1}{\pi}\int_{-\infty}^{+\infty}e^{-\frac{x^2}{2}}dx\boxed{\int_{x}^{+\infty}ye^{-\frac{y^2}{2}}dy}=\int_{x}^{+\infty}e^{-\frac{y^2}{2}}d\frac{y^2}{2}=-e^{-\frac{y^2}{2}}\Big|_{x}^{+\infty}$$

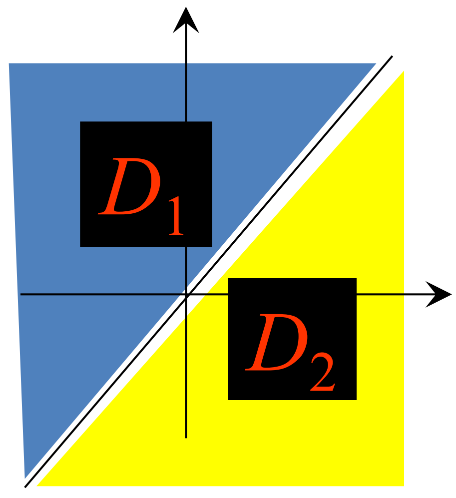
$$=e^{-\frac{x^2}{2}}$$

$$=\frac{1}{\pi}\int_{-\infty}^{+\infty}e^{-x^2}dx=\frac{1}{\pi}\cdot\sqrt{\pi}$$

$$=\frac{1}{\sqrt{\pi}}$$

其中$\int_{-\infty}^{+\infty}e^{-x^2}dx=\sqrt{\pi}$

称为泊松-欧拉积分



$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

$$\left(\int_{-\infty}^{+\infty} e^{-x^2} dx\right)^2 = \int_{-\infty}^{+\infty} e^{-x^2} dx \int_{-\infty}^{+\infty} e^{-y^2} dy$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dxdy$$

$$= \int_0^{2\pi} d\theta \boxed{\int_0^{+\infty} e^{-r^2} rdr} \quad = \int_0^{+\infty} e^{-r^2} \frac{1}{2} dr^2$$

$$= 2\pi \times \frac{1}{2} = \pi \qquad = -\frac{1}{2} e^{-r^2} \Big|_0^{+\infty}$$

$$= \frac{1}{2}$$

所以 $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$

一般地，若 $X \sim N(\mu,\sigma^2), Y \sim N(\mu,\sigma^2)$, X, Y 相互独立，则

$$E(\max\{X,Y\}) = \mu + \frac{\sigma}{\sqrt{\pi}}$$

$$E(\min\{X,Y\}) = \mu - \frac{\sigma}{\sqrt{\pi}}$$

数学期望的性质

1、 $E(C)=C$

2、 $E(aX)=aE(X)$

3、 $E(X+Y)=E(X)+E(Y)$

$$E\left(\sum_{i=1}^{n}a_iX_i+C\right)=\sum_{i=1}^{n}a_iE(X_i)+C$$

4、当X,Y相互独立时，$E(XY)=E(X)E(Y)$.

证明：$E(X+Y)=E(X)+E(Y)$

证：以离散情况为例

$$E(X+Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(x_i+y_j)p_{ij}$$

$$=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}x_i p_{ij}+\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}y_j p_{ij}$$

$$=\sum_{i=1}^{\infty}x_i\sum_{j=1}^{\infty}p_{ij}+\sum_{j=1}^{\infty}y_j\sum_{i=1}^{\infty}p_{ij}$$

$$=\sum_{i=1}^{\infty}x_i p_{i\cdot}+\sum_{i=1}^{\infty}y_j p_{\cdot j}$$

$$=E(X)+E(Y)$$

证明：当X,Y相互独立时，$E(XY)=E(X)E(Y)$.

证：以离散情况为例

$$E(XY)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(x_i\cdot y_j)p_{ij}$$

$$=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}x_i p_{i\cdot}\cdot y_j p_{\cdot j}$$

$$=\sum_{i=1}^{\infty}x_i p_{i\cdot}\cdot\sum_{j=1}^{\infty}y_j p_{\cdot j}$$

$$=E(X)\cdot E(Y)$$

注　性质 4 的逆命题不成立，即

若$E(XY)=E(X)E(Y)$，X,Y不一定相互独立

反例 1

$Y \backslash X$ (p_{ij})	-1	0	1	$p_{\bullet j}$
-1	$1/8$	$1/8$	$1/8$	$3/8$
0	$1/8$	0	$1/8$	$2/8$
1	$1/8$	$1/8$	$1/8$	$3/8$
$p_{i\bullet}$	$3/8$	$2/8$	$3/8$	

XY	-1	0	1
P	$2/8$	$4/8$	$2/8$

$$E(X)=E(Y)=0; \qquad E(XY)=0;$$

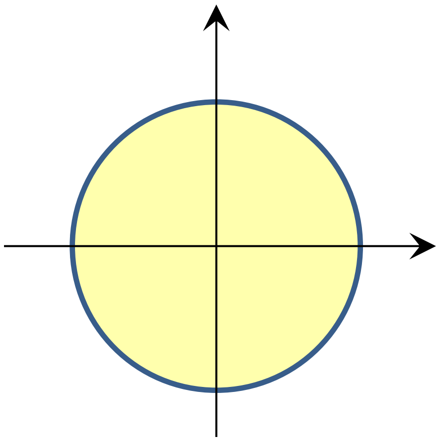
$$E(XY)=E(X)E(Y)$$

但 $$P(X=-1,Y=-1)=\frac{1}{8}$$

$$\neq P(X=-1)P(Y=-1)=\left(\frac{3}{8}\right)^2$$

反例 2 $(X,Y)\sim U(D),\quad D=\{(x,y)\,|\,x^2+y^2\le 1\}$

$$f(x,y)=\begin{cases}\dfrac{1}{\pi}, & x^2+y^2\le 1,\\ 0, & 其它\end{cases}$$



$$f_X(x)=\begin{cases}\dfrac{2\sqrt{1-x^2}}{\pi}, & -1<x<1,\\ 0, & 其它\end{cases}$$

$$E(X)=\int_{-1}^{1}x\frac{2\sqrt{1-x^2}}{\pi}dx=0;\quad f_Y(y)=\begin{cases}\dfrac{2\sqrt{1-y^2}}{\pi}, & -1<y<1,\\ 0, & 其它\end{cases}$$

$$E(XY)=\iint\limits_{x^2+y^2\le 1}xy\frac{1}{\pi}dxdy=0;$$

$$E(XY)=0\ =E(X)E(Y)$$

$$f(x,y)\neq f_X(x)f_Y(y)$$

例 将 4 个可区分的球随机地放入 4 个盒子中，每盒容纳的球数无限，求空着的盒子数的数学期望.

解一 设 X 为空着的盒子数, 则 X 的概率分布为

X	0	1	2	3
P	$\dfrac{4!}{4^4}$	$\dfrac{C_4^1C_4^2P_3^3}{4^4}$	$\dfrac{C_4^2(C_4^2+C_2^1C_4^3)}{4^4}$	$\dfrac{C_4^1}{4^4}$

$$E(X)=\frac{81}{64}$$

解二 引入 $X_i, i = 1,2,3,4$

$$X_i = \begin{cases} 1, & \text{第}i\text{盒空}, \\ 0, & \text{其它}, \end{cases}$$

$$X = X_1 + X_2 + X_3 + X_4$$

X_i	1	0
P	$\left(\frac{3}{4}\right)^4$	$1-\left(\frac{3}{4}\right)^4$

$$E(X_i) = \left(\frac{3}{4}\right)^4$$

$$E(X) = 4 \cdot \left(\frac{3}{4}\right)^4 = \frac{81}{64}$$

例：将100只铅笔随机的分给80个孩子，如果每支铅笔分给哪个孩子是等可能的，问：平均有多少孩子得到铅笔？

解 引入 $X_i, i = 1,2,\ldots\ldots, 80$

$$X_i = \begin{cases} 1, & \text{第}i\text{个孩子得到了铅笔}, \\ 0, & \text{其它}, \end{cases}$$

$$X = X_1 + X_2 + \cdots + X_{80}$$

X_i	1	0
P	$1-\left(\frac{79}{80}\right)^{100}$	$\left(\frac{79}{80}\right)^{100}$

$$E(X_i) = 1-\left(\frac{79}{80}\right)^{100}$$

$$E(X) = E\left(\sum_{i=1}^{80} X_i\right) = 80\left[1-\left(\frac{79}{80}\right)^{100}\right]$$