§5 随机变量的数字特征

一、问题的提出

例如：考察射击手的水平，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。

又如：有一大批产品，其中15% 为一等品，75% 为二等品，10% 为三等品。一、二、三等产品的单价分别为10 元、8 元和6 元。有人要采购一批这种产品，但来不及检验，该如何定价？

随机变量的分布函数能够完整地描述随机变量的统计特征。但是在实际中，一方面得到随机变量分布函数（分布律或分布密度）相当困难；另一方面，许多实际问题只需要知道随机变量的某些特征即可。此时，我们关心的是描写随机变量的某些特征的数字，即数字特征。

就上两例来说，我们需要引入能反映随机变量重要特征的统计指标，随机变量的数学期望（均值）、方差（均值的偏离程度）。

二、数学期望

设随机变量 ξ 所有可能取的值为 x_1, x_2, x_3 ，但预期 ξ 取值的平均数一般不等于 $\dfrac{x_1+x_2+x_3}{3}$ ，因为 ξ 取每个值的概率一般不同，概率大的值取到的机会多，概率小的值取到的机会少，所以不能对随机变量所能取的每个值同等对待，必须考虑其相应的概率。

例：假定对 ξ 作 n 次独立观察，ξ 取值 x_1, x_2, x_3 的次数分别为 m_1, m_2, m_3 ，$m_1+m_2+m_3=n$ ，那么 ξ 取值的平均数为多少呢？

定义1：设离散型随机变量 ξ 的分布律为

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

若级数 $\sum_{i=1}^{\infty} x_i p_i$ 绝对收敛，则称此级数的收敛值为

随机变量 ξ 的数学期望（简称**期望**）或**均值**。

记为 $$E\xi = \sum_{i=1}^{\infty} x_i p_i$$

就上例，ξ 取值的平均数为 $x_1 \dfrac{m_1}{n} + x_2 \dfrac{m_2}{n} + x_3 \dfrac{m_3}{n}$

其中 $\dfrac{m_i}{n}$ 是事件 $\{\xi = x_i\}$ 在 n 次观察中出现的频率，由概率

的统计定义，当 n 充分大时，$\dfrac{m_i}{n} \approx p_i = p(\xi = x_i)$

所以，随机变量 ξ 真正的平均数为 $x_1 p_1 + x_2 p_2 + x_3 p_3$

定义2：设连续型随机变量 ξ **的分布密度为** $\varphi(x)$，

若 $\int_{-\infty}^{+\infty} x\varphi(x)dx$ **绝对收敛，则称** $\int_{-\infty}^{+\infty} x\varphi(x)dx$ **为** ξ

的数学期望。记为 $E\xi = \int_{-\infty}^{+\infty} x\varphi(x)dx$

注意：1）一个随机变量的期望存在与否取决于级数 $\sum_{i=1}^{\infty} x_i p_i$

或广义积分 $\int_{-\infty}^{+\infty} x\varphi(x)dx$ **是否绝对收敛，所以并**

不是任何随机变量都存在数学期望。

例：设离散型随机变量 ξ 的概率函数为

$$P(\xi = k) = \frac{1}{k(k+1)} \qquad k = 1, 2, \cdots$$

由于级数 $\sum_{k=1}^{\infty} |x_k p_k| = \sum_{k=1}^{\infty} \left| \frac{k}{k(k+1)} \right| = \sum_{k=1}^{\infty} \frac{1}{k+1}$ 发散，

所以，$E\xi$ 不存在。

2）一个随机变量的期望是一个常数，它表示的是随机变量取值的平均，是以概率为权的加权平均值；它反映了随机变量得一大特征，即随机变量取值集中在期望值附近。

3）计算时需要已知随机变量的分布律或分布密度。

随机变量函数的数学期望：

1）离散型

设 r.v. X 的分布律为 $P(X=x_k)=p_k$， $k=1,2,\cdots$

r.v. $Y=f(X)$ ，如级数 $\sum_{k=1}^{\infty} f(x_k)p_k$ 绝对收敛，

则 r.v. $Y=f(X)$ 的数学期望 $EY=\sum_{k=1}^{\infty} f(x_k)p_k$.

2）连续型

设 r.v. X 的概率密度函数 $\varphi_X(x)$， r.v. $Y=f(X)$ ，

如广义积分 $\int_{-\infty}^{+\infty} f(x)\varphi_X(x)dx$ 绝对收敛，

则 r.v. $Y=f(X)$ 的数学期望 $EY=\int_{-\infty}^{+\infty} f(x)\varphi_X(x)dx$.

随机变量数学期望的性质：

1）设 $\xi = c$ ，则 $E\xi = Ec = c$

2）若 k 为常数 ，则 $E(k\xi) = kE\xi$

3）若 b 为常数 ，则 $E(\xi + b) = E\xi + b$

由 2）、3）得 $E(a\xi + b) = aE\xi + b$

4）若两个随机变量 ξ、η ，则 $E(\xi + \eta) = E\xi + E\eta$

可推广至有限个

$$E(X_1 + X_2 + \cdots + X_n) = EX_1 + EX_2 + \cdots + EX_n$$

5）若随机变量 ξ 、η 相互独立，则 $E(\xi \cdot \eta) = E\xi \cdot E\eta$

可推广至有限个：若 X_1、X_2、… X_n 相互独立，则

$$E(X_1 \cdot X_2 \cdots X_n) = EX_1 \cdot EX_2 \cdots EX_n$$

常见分布的数学期望

1）二点分布（0-1 分布）

分布律：$P\{\xi=1\}=p$

$P\{\xi=0\}=1-p=q$ $\quad 0<p,q<1$

$$E\xi=1\cdot p+0\cdot(1-p)=p$$

2）二项分布

分布律：$P\{\xi=k\}=C_n^k p^k (1-p)^{n-k} \quad k=0,1,2,\cdots,n$

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

$$=\sum_{k=0}^{n}\frac{k\,n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

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

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

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

$$=np$$

3）Poisson 分布 $P(\lambda)$

分布律：$P\{\xi = k\} = \dfrac{\lambda^k}{k!}e^{-\lambda} \qquad k = 0, 1, 2, \cdots \quad \lambda > 0$

$$E\xi = \sum_{k=0}^{n} k \cdot P(\xi = k) = \sum_{k=0}^{n} k \cdot \frac{\lambda^k}{k!}e^{-\lambda}$$

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

4）指数分布 $E(\lambda)$

密度函数: $\varphi(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$

$$E\xi = \int_{-\infty}^{+\infty} x\varphi(x)dx = \int_{-\infty}^{0} x \cdot 0 dx + \int_{0}^{+\infty} x\lambda e^{-\lambda x} dx$$

$$= -\int_{0}^{+\infty} x d(e^{-\lambda x}) = -[x e^{-\lambda x} |_{0}^{+\infty} - \int_{0}^{+\infty} e^{-\lambda x} dx]$$

$$= -\frac{1}{\lambda} e^{-\lambda x} |_{0}^{+\infty} = \frac{1}{\lambda}$$

例1、设十只同种电器元件中有二只废品，装配仪器时，从这批元件中任取一只，若是废品，则扔掉重新任取一只，若仍是废品，则再扔掉还取一只。求：在取到正品之前，已取出的废品数 X 的概率分布及其数学期望 EX .

解：设 $\bar{A}_i$： 从10只中任取一只，第 i 次取到废品。

由题意，在取到正品前， X 的可能取值为 0, 1, 2 .

$$p(X=0)=p(A_1)=\frac{C_8^1}{C_{10}^1}=\frac{4}{5}$$

$$p(X=1)=p(\bar{A}_1A_2)=p(\bar{A}_1)p(A_2/\bar{A}_1)$$

$$=\frac{C_2^1}{C_{10}^1}\cdot\frac{C_8^1}{C_9^1}=\frac{8}{45}$$

返回

$$p(X=2)=p(\bar{A}_1\bar{A}_2A_3)$$

$$=p(\bar{A}_1)p(\bar{A}_2/\bar{A}_1)p(A_3/\bar{A}_1\bar{A}_2)$$

$$=\frac{C_2^1}{C_{10}^1}\cdot\frac{C_1^1}{C_9^1}\cdot\frac{C_8^1}{C_8^1}$$

$$=\frac{1}{45}$$

$$EX=\sum_{k=0}^{2}x_k p_k$$

$$=0\cdot\frac{4}{5}+1\cdot\frac{8}{45}+2\cdot\frac{1}{45}$$

$$=\frac{2}{9}$$

返回

例2、某产品的次品率为0.1，检验员每天检验 4 次。每次随机地取10 件产品进行检验，如发现其中的次品数多于1，就去调整设备，以 X 表示一天中调整设备的次数，试求 EX .（设产品是否为次品是相互独立的）

解：设

例3、据统计65岁的人在10年内正常死亡的概率为0.98，因事故死亡概率为0.02. 保险公司开办老人事故死亡保险，参加者需缴纳保险费100元。若在10年内因事故死亡，公司赔偿 a 元，应如何定 a ，才能使公司可期望获益；若有1000人投保，公司期望总收益多少？

例4、已知某型号电子管的使用寿命 X 为连续型随机变量，其密度函数为

$$\varphi(x)=\begin{cases}\dfrac{C}{x^2} & x>1000\\ 0 & \text{其它}\end{cases}$$

1）求常数 C ； 2）计算 $P(X>2000)$

3）已知一设备装有3个这样的电子管，每个电子管能否正常工作相互独立，求在使用的最初 1500 天只有一个损坏的概率。

三、随机变量的方差与标准差

在许多实际问题中，只知道随机变量的均数是不够的，还需知道随机变量的取值在均数左右波动的大小（即离散程度）。

随机变量 ξ 与其数学期望 $E\xi$ 之差 $(\xi - E\xi)$，称为 ξ 的**离差**。

因为
$$\begin{aligned} E(\xi - E\xi) &= E\xi - E(E\xi) \\ &= E\xi - E\xi \\ &= 0 \end{aligned}$$

所以不宜用离差的数学期望来表示随机变量取值的离散程度。

1、定义：

随机变量ξ的离差的平方的数学期望称为ξ 的 **方差，**

记为 $D\xi = E(\xi - E\xi)^2$,

且称 $\sqrt{D\xi}$ (σ) 为 ξ 的 **均方差**或**标准差**。

1）离散型随机变量ξ ：

$$p(\xi = x_i) = p_i \quad i = 1, 2, \cdots$$

$$D\xi = \sum_{i=1}^{\infty}(x_i - E\xi)^2 p_i$$

2）连续型随机变量ξ ：

$$D\xi = \int_{-\infty}^{+\infty}(x - E\xi)^2 \varphi(x)dx$$

注意： **由定义公式计算方差是繁复的，常常用下重要公式：**

$$D\xi = E(\xi^2) - (E\xi)^2$$

证：
$$\begin{aligned} D\xi &= E(\xi - E\xi)^2 \\ &= E(\xi^2 - 2\xi E\xi + (E\xi)^2) \\ &= E\xi^2 - 2E\xi E\xi + (E\xi)^2 \\ &= E(\xi^2) - (E\xi)^2 \end{aligned}$$

计算 $E\xi^2$ ：

离散型随机变量： $E\xi^2 = \sum_{i=1}^{\infty} x_i^{\,2} p(\xi = x_i)$

连续型随机变量： $E\xi^2 = \int_{-\infty}^{+\infty} x^2 \varphi(x) dx$

2、方差的性质：

1) $D(c)=0$

2) $D(k\xi)=k^2D\xi$

3) $D(\xi+c)=D\xi$

由 2 ）、3 ）得 $D(k\xi+c)=k^2D\xi$.

4) 若随机变量ξ 、η 相互独立，则 $D(\xi+\eta)=D\xi+D\eta$.

可推广至有限个相互独立的随机变量 X_1 、X_2 、 … X_n ，

则 $D(X_1+X_2+\cdots+X_n)=DX_1+DX_2+\cdots+DX_n$

常见分布的方差：

1）二点分布（0-1 分布）

分布律：$P\{\xi=1\}=p$，

$$P\{\xi=0\}=1-p=q, \quad 0<p,q<1.$$

$$E\xi=1\cdot p+0\cdot(1-p)=p$$

$$E\xi^2=1^2\cdot p+0^2\cdot(1-p)=p$$

$$\begin{aligned} D\xi &= E(\xi^2)-(E\xi)^2 \\ &= p-p^2 \\ &= pq \end{aligned}$$

2）二项分布

分布律：$P\{\xi = k\} = C_n^k p^k (1-p)^{n-k} \quad k = 0, 1, 2, \cdots, n$

二项分布的 r.v. ξ 可看作是 n 个独立的 0-1 分布的 r.v. 的和，即 r.v. ξ 是重复独立地做 n 次 Bernoulli 试验中事件 A 出现的次数。 所以设事件 A 每次出现的概率为 p，

设 ξ_i 取值为：

$$\xi_i = \begin{cases} 1 & \text{第 } i \text{ 次试验中 } A \text{ 出现} \\ 0 & \text{第 } i \text{ 次试验中 } A \text{ 不出现} \end{cases} \qquad i = \overline{1, n}$$

$\therefore \xi_i \sim 0-1$ 分布，且 ξ_i 相互独立，$\therefore \xi = \xi_1 + \xi_2 + \cdots + \xi_n$

表示在 n 次独立试验中事件 A 出现的次数，

$$\therefore E\xi = E(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} E\xi_i = np$$

$$D\xi = D(\sum_{i=1}^{n} \xi_i) = \sum_{i=1}^{n} D\xi_i = npq$$

2）均匀分布

设 r.v. $\xi \sim U[a,b]$，则 ξ 的概率密度为

$$\varphi(x)=\begin{cases}\dfrac{1}{b-a} & a\le x\le b\\ 0 & \text{其他}\end{cases}$$

则 $$E\xi=\int_{-\infty}^{+\infty}x\varphi(x)dx=\int_a^b x\frac{1}{b-a}dx=\frac{a+b}{2}$$

$$E\xi^2=\int_{-\infty}^{+\infty}x^2\varphi(x)dx=\int_a^b x^2\frac{1}{b-a}dx=\frac{a^2+ab+b^2}{3}$$

$$D\xi=E(\xi^2)-(E\xi)^2$$

$$=\frac{a^2+ab+b^2}{3}-\left(\frac{a+b}{2}\right)^2$$

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

分布	概率密度	期望	方差
参数为 p 的 0-1分布	$P(X=1)=p$ $P(X=0)=1-p$	p	pq
$B(n,p)$	$P(X=k)=C_n^k p^k (1-p)^{n-k}$ $k=0,1,2,\cdots,n$	np	npq
$P(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!} \quad k=0,1,2,\cdots$	λ	λ
区间[a,b]上的均匀分布	$f(x)=\begin{cases} \dfrac{1}{b-a}, & a\le x\le b \\ 0, & \text{其它} \end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$E(\lambda)$	$f(x)=\begin{cases} \lambda e^{-\lambda x}, & x>0 \\ 0, & \text{其它} \end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

例5、求例 1、2 的 DX .

解： 例1 已求得

$$p(X=0)=\frac{4}{5} \quad p(X=1)=\frac{8}{45} \quad p(X=2)=\frac{1}{45}$$

$$EX=\frac{2}{9}$$

$$EX^2=\sum_{k=0}^{2}x_k^2p_k=0^2\cdot\frac{4}{5}+1^2\cdot\frac{8}{45}+2^2\cdot\frac{1}{45}=\frac{4}{15}$$

$$DX=E(X^2)-(EX)^2$$

$$=\frac{4}{15}-(\frac{1}{45})^2=\frac{88}{405}$$

返回

例6、设连续型 r.v. X 的概率密度 $\varphi_X(x)$ 为：

$$\varphi_X(x)=\begin{cases}\dfrac{1}{2} & 0<x<2\\ 0 & \text{其它}\end{cases}$$

求 $Y=1-X$ 的概率密度函数，EY，DY.

解：$F_Y(y)=P(Y\le y)=P(1-X\le y)$

$$=P(X\ge 1-y)=1-P(X<1-y)$$

$$=1-\int_0^{1-y}\frac{1}{2}dx=\frac{1}{2}(1+y)\qquad -1<y<1$$

当 $y\le -1$ 时，$F_Y(y)=1$；当 $y\ge 1$ 时，$F_Y(y)=1$.

$$\therefore \varphi_Y(y)=\begin{cases}\dfrac{1}{2} & -1<y<1\\ 0 & \text{其它}\end{cases}.$$

$$EY=\int_{-\infty}^{+\infty} y\varphi_Y(y)\,dy=\int_{-1}^{1}\frac{1}{2}y\,dy=0$$

$$E(Y^2)=\int_{-\infty}^{+\infty} y^2\varphi_Y(y)\,dy=\int_{-1}^{1}\frac{1}{2}y^2\,dy=\frac{1}{3}$$

$$DY=E(Y^2)-(EY)^2=\frac{1}{3}$$