随机变量的数学期望和方差的一些性质

邓启龙
（广东省中山纪念中学）

随机变量的数学期望和方差是高中数学概率统计的重要内容.本文通过探究得到了随机变量的数学期望和方差的一些性质.

1 数学期望

若离散型随机变量 X 的概率分布为$P(X=x_i)=p_i, i=1,2,\cdots,n$，其中 $p_i\geqslant 0, i=1,2,\cdots,n$ 且 $\sum\limits_{i=1}^{n}p_i=1$,则 X 的数学期望为$E(X)=\sum\limits_{i=1}^{n}x_ip_i$.

连续型随机变量 X 的概率密度函数为$f(x)$,若 $\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x<\infty$,则 X 的数学期望为

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

注：若离散型随机变量 X 有无穷多个取值,且其概率分布为 $P(X=x_i)=p_i, i=1,2,3,\cdots$,其中 $p_i\geqslant 0, i=1,2,3,\cdots$,且 $\sum\limits_{i=1}^{\infty}p_i=1$,若$\sum\limits_{i=1}^{\infty}|x_i|p_i<\infty$,则 X 的数学期望为$E(X)=\sum\limits_{i=1}^{\infty}x_ip_i$.

2 方差和协方差

随机变量 X 的数学期望为 $E(X)$,若 $E((X-E(X))^2)$存在,则 X 的方差为

$$D(X)=E((X-E(X))^2).$$

随机变量 X,Y 的数学期望分别为$E(X),E(Y)$,若$E((X-E(X))(Y-E(Y)))$存在,则 X,Y 的协方差为 $\mathrm{cov}(X,Y)=E((X-E(X))(Y-E(Y)))$.特别地,$\mathrm{cov}(X,X)=D(X)$.

3 基本性质

随机变量的数学期望、方差和协方差具有以下性质.

性质 1 $E(X+Y)=E(X)+E(Y)$, $E(aX+b)=aE(X)+b$.

推论 1 $\forall c_i\in\mathbf{R}, i=1,2,\cdots,n, n\in\mathbf{N}^*$ 有

$$E\left(\sum_{i=1}^{n}c_iX_i\right)=\sum_{i=1}^{n}c_iE(X_i).$$

性质 2 $D(X)=E(X^2)-(E(X))^2$, $D(aX+b)=a^2D(X)$.

性质 3 $\mathrm{cov}(X,Y)=\mathrm{cov}(Y,X)$, $\mathrm{cov}(aX+bY,Z)=a\,\mathrm{cov}(X,Z)+b\,\mathrm{cov}(Y,Z)$, $a,b\in\mathbf{R}$.

性质 4 $\mathrm{cov}(X,Y)=E(XY)-E(X)E(Y)$, $D(X+Y)=D(X)+D(Y)+2\mathrm{cov}(X,Y)$.

证明 由推论 1 得

$$\begin{gathered}\mathrm{cov}(X,Y)=E((X-E(X))(Y-E(Y)))=\\ E(XY-E(Y)X-E(X)Y+E(X)E(Y))=\\ E(XY)-E(Y)E(X)-E(X)E(Y)+E(X)E(Y)=\\ E(XY)-E(X)E(Y),\\ D(X+Y)=E((X+Y-E(X+Y))^2)=\\ E((X-E(X)+Y-E(Y))^2)=E((X-E(X))^2+\\ (Y-E(Y))^2+2(X-E(X))(Y-E(Y)))=\\ E((X-E(X))^2)+E((Y-E(Y))^2)+\\ 2E((X-E(X))(Y-E(Y)))=\\ D(X)+D(Y)+2\mathrm{cov}(X,Y).\end{gathered}$$

推论 2 $D(X+Y)+D(X-Y)=2(D(X)+D(Y))$, $D(X+Y)-D(X-Y)=4\mathrm{cov}(X,Y)$.

性质 5 若 X,Y 相互独立,则

$$\begin{gathered}E(XY)=E(X)E(Y), \mathrm{cov}(X,Y)=0,\\ D(X+Y)=D(X)+D(Y).\end{gathered}$$

证明 假设 X,Y 都是离散型随机变量,X 的概率分布为 $P(X=x_i)=p_i, i=1,2,\cdots,n$, Y 的概率分布为 $P(Y=y_j)=q_j, j=1,2,\cdots,m$.

由 X,Y 相互独立得

$$\begin{gathered}P(X=x_i,Y=y_j)=P(X=x_i)P(Y=y_j)=p_iq_j,\\ i=1,2,\cdots,n, j=1,2,\cdots,m,\end{gathered}$$

则

$$E(XY)=\sum_{i=1}^{n}\sum_{j=1}^{m}x_iy_jP(X=x_i,Y=y_j)=$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m}x_iy_jp_iq_j=\left(\sum_{i=1}^{n}x_ip_i\right)\left(\sum_{j=1}^{n}y_jq_j\right)=E(X)E(Y).$$

对于其他情形同理可证 $E(XY)=E(X)E(Y)$.

由性质 4 得 $\mathrm{cov}(X,Y)=0$, $D(X+Y)=D(X)+D(Y)$.

推论 3 若 $X_1,X_2,\cdots,X_n(n\in\mathbf{N}^*,n\geqslant 2)$ 相互独立,则 $\forall c_i\in\mathbf{R},i=1,2,\cdots,n$,有

$$D(\sum_{i=1}^{n}c_iX_i)=\sum_{i=1}^{n}c_i^2D(X_i).$$

4 其他性质

性质 6 若 $X\leqslant Y$,则 $E(X)\leqslant E(Y)$.

推论 4 $|E(X)|\leqslant E|X|,D|X|\leqslant D(X)$.

证明 由 $-|X|\leqslant X\leqslant|X|$ 和性质 6 得

$$|E(X)|\leqslant E|X|.$$

由性质 2 可得 $D(X)=E(X^2)-(E(X))^2$,且 $D|X|=E(X^2)-(E|X|)^2$,由 $|E(X)|\leqslant E|X|$ 得 $(E(X))^2\leqslant(E|X|)^2$,所以 $D|X|\leqslant D(X)$.

性质 7 $(E|XY|)^2\leqslant E(X^2)E(Y^2)$.

证明 由均值不等式得

$$\left|\frac{X}{\sqrt{E(X^2)}}\right|\left|\frac{Y}{\sqrt{E(Y^2)}}\right|\leqslant\frac{1}{2}\left(\frac{X^2}{E(X^2)}+\frac{Y^2}{E(Y^2)}\right),$$

两边取数学期望由性质 6 即得.

推论 5 $(E(XY))^2\leqslant E(X^2)E(Y^2)$.

推论 6 $|\mathrm{cov}(X,Y)|\leqslant\sqrt{D(X)}\cdot\sqrt{D(Y)}$.

性质 8 $\forall c\in\mathbf{R},D(X)\leqslant E((X-c)^2)$,当且仅当 $c=E(X)$ 时取等号.

证明 由性质 2 得

$$D(X)=D(X-c)=E((X-c)^2)-(E(X-c))^2=$$
$$E((X-c)^2)-(E(X)-c)^2\leqslant E((X-c)^2),$$

当且仅当 $c=E(X)$ 时取等号.

性质 9 $(\sqrt{D(X)}-\sqrt{D(Y)})^2\leqslant D(X+Y)\leqslant(\sqrt{D(X)}+\sqrt{D(Y)})^2\leqslant 2(D(X)+D(Y))$.

证明 由性质 4 和推论 6 得

$$D(X+Y)=D(X)+D(Y)+2\mathrm{cov}(X,Y)\leqslant$$
$$D(X)+D(Y)+2\sqrt{D(X)}\cdot\sqrt{D(Y)}=$$
$$(\sqrt{D(X)}+\sqrt{D(Y)})^2\leqslant 2(D(X)+D(Y)),$$
$$D(X+Y)=D(X)+D(Y)+2\mathrm{cov}(X,Y)\geqslant$$
$$D(X)+D(Y)-2\sqrt{D(X)}\cdot\sqrt{D(Y)}=$$
$$(\sqrt{D(X)}-\sqrt{D(Y)})^2.$$

对于随机变量 X,Y,定义 $X\vee Y=\max\{X,Y\}$,$X\wedge Y=\min\{X,Y\}$.当 $Y=c$ 为常数时,$X\vee c,X\wedge c$ 分别是左截断和右截断数据.特别地,定义

$$X^+=X\vee 0=\max\{X,0\},$$
$$X^-=-X\wedge 0=-\min\{X,0\},$$

X^+,X^- 分别称为 X 的正部和负部.由定义易得 $X+Y=X\vee Y+X\wedge Y,|X-Y|=X\vee Y-X\wedge Y$ 和 $X=X^+-X^-,|X|=X^++X^-$.

性质 10 $D(X+Y)+D|X-Y|=2(D(X\vee Y)+D(X\wedge Y))$.

推论 7 已知 $c\in\mathbf{R}$ 是常数.

(1)$D(X)+D|X-c|=2(D(X\vee c)+D(X\wedge c))$.

(2)$D|X-c|\leqslant D(X\vee c)+D(X\wedge c)\leqslant D(X)$.

(3)$D(X)+D|X|=2(D(X^+)+D(X^-))$.

(4)$D|X|\leqslant D(X^+)+D(X^-)\leqslant D(X)$.

性质 11 $D(X)+D(Y)+\dfrac{D|X-Y|}{2}=D(X\vee Y)+D(X\wedge Y)+\dfrac{D(X-Y)}{2}$.

证明 由推论 2 得 $D(X+Y)=2(D(X)+D(Y))-D(X-Y)$,由性质 10 得 $D(X+Y)=2(D(X\vee Y)+D(X\wedge Y))-D|X-Y|$,于是 $2(D(X)+D(Y))-D(X-Y)=2(D(X\vee Y)+D(X\wedge Y))-D|X-Y|$,所以

$$D(X)+D(Y)+\frac{D|X-Y|}{2}=$$
$$D(X\vee Y)+D(X\wedge Y)+\frac{D(X-Y)}{2}.$$

推论 8 $\dfrac{D(X+Y)}{2}\leqslant D(X\vee Y)+D(X\wedge Y)\leqslant D(X)+D(Y)$.

证明 由推论 4 得 $D|X-Y|\leqslant D(X-Y)$,故结合性质 11 得

$$D(X\vee Y)+D(X\wedge Y)\leqslant D(X)+D(Y).$$

由性质 10 得

$$D(X\vee Y)+D(X\wedge Y)=$$
$$\frac{D(X+Y)+D|X-Y|}{2}\geqslant\frac{D(X+Y)}{2}.$$

5 典型例题

例 1 在独立重复试验中,设每次试验中事件 A 发生的概率为 $p\ (0<p<1)$,试验进行到事件 A 首次发生时停止,用 X 表示此时所进行的试验次数,求 X 的数学期望和方差.

解析 由已知得 X 的概率分布为 $P(X=k)=(1-p)^{k-1}p,k=1,2,3,\cdots$.

$$E(X)=\sum_{k=1}^{\infty}k(1-p)^{k-1}p.$$

恰当选择研究角度,巧妙破解排列组合问题

俞　纲[1]　周继波[2]

(1.云南省昆明市第三中学　2.云南省昆明市第十六中学)

排列组合与计数问题的计算是围绕着事件的完成来展开的,而完成事件的角度往往不止一种.从不同的角度来完成事件其复杂程度各不相同,这就需要我们多思考、多总结,学会选择恰当的角度来完成事件,以达到事半功倍、化繁为简的效果.

例1　学校安排全校六个年级的同学去春游,有四个公园可供选择,每个年级去一个公园,则共有多少种不同的安排方法?

解析　此题常见的研究角度有两种:如果从公园"接纳"学生的角度来完成事件,由于每个公园"接纳"了哪几个年级不确定,而第一个公园"接纳"年级的每一种情况对下一个公园"接纳"哪些年级会产生不同的影响,这样研究就会比较复杂,难以用一个算式直接描述;如果从学生选择公园的角度来完成此事件,由于每一个年级都是从四个不同公园中选出一个,且相互没有影响,则运用乘法原理,每个年级逐一选择,所有情况数共有 $4\times4\times4\times4\times4\times4=4^6$ 种.

点评　如果完成一个事件有多个步骤,则研究事件的角度应该尽量让每个步骤之间相互独立或具体明确前一步骤对后一步骤的影响.

对 $\dfrac{1}{1-x}=1+x+x^2+x^3\cdots$,$|x|<1$ 两边求导得

$\dfrac{1}{(1-x)^2}=1+2x+3x^2+4x^3+\cdots$,$|x|<1$.

令 $x=1-p$,得 $\dfrac{1}{p^2}=1+2(1-p)+3(1-p)^2+4(1-p)^3+\cdots$,所以

$$E(X)=\sum_{k=1}^{\infty}k(1-p)^{k-1}p=p\sum_{k=1}^{\infty}k(1-p)^{k-1}=\frac{1}{p}.$$

由性质2得 $D(X)=E(X^2)-(E(X))^2$,其中 $E(X^2)=\sum\limits_{k=1}^{\infty}k^2(1-p)^{k-1}p$,且 $\dfrac{1}{(1-x)^2}=1+2x+3x^2+4x^3+\cdots$,$|x|<1$,于是 $\dfrac{x}{(1-x)^2}=x+2x^2+3x^3+4x^4+\cdots$,$|x|<1$,两边求导得

$$\frac{x+1}{(1-x)^3}=1+2^2x+3^2x^2+4^2x^3+\cdots,|x|<1.$$

令 $x=1-p$,得

$$\frac{2-p}{p^3}=1+2^2(1-p)+3^2(1-p)^2+4^2(1-p)^3+\cdots,$$

则 $E(X^2)=\sum\limits_{k=1}^{\infty}k^2(1-p)^{k-1}p=p\sum\limits_{k=1}^{\infty}k^2(1-p)^{k-1}=\dfrac{2-p}{p^2}$,所以

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

故 X 的数学期望和方差分别为 $\dfrac{1}{p}$ 和 $\dfrac{1-p}{p^2}$.

注:实际上,例1中的 X 服从几何分布,即

$$X\sim G(p),E(X)=\frac{1}{p},D(X)=\frac{1-p}{p^2}.$$

例2　流水线上生产的每个产品为不合格品的概率为 p,当生产出 $n(n\in\mathbf{N}^*)$ 个不合格品时,即停工检修一次.求在两次检修之间的产品总数 X 的数学期望和方差.

解析　记第 $i-1$ 个不合格品之后到第 i 个不合格品之间的产品数(包括第 i 个不合格品)为 X_i,$i=1,2,\cdots,n$,则 $X=X_1+X_2+\cdots+X_n$,且 X_1,X_2,$\cdots$,X_n 相互独立,都服从几何分布 $G(p)$,即对 $i=1,2,\cdots,n$,有 $P(X_i=k)=(1-p)^{k-1}p$,$k=1,2,3,\cdots$,由例1得 $E(X_i)=\dfrac{1}{p}$,$D(X_i)=\dfrac{1-p}{p^2}$,$i=1,2,\cdots,n$.由推论1得 $E(X)=\sum\limits_{i=1}^{n}E(X_i)=\dfrac{n}{p}$.

由 $X_1,X_2,\cdots,X_n$ 相互独立和推论3得

$$D(X)=\sum_{i=1}^{n}D(X_i)=\frac{n(1-p)}{p^2}.$$

(完)