第五周　随机变量函数的分布及随机变量的数字特征

5.5 期望和方差的一些补充性质

期望的最小二乘性质

X 为一随机变量，则 $c=E(X)$ 时，$E\left((X-c)^2\right)$ 达到最小。

证明：$E\left((X-c)^2\right)=E\left(\left[(X-E(X))+(E(X)-c)\right]^2\right)$

$$=E\left(\left[(X-E(X))\right]^2+\left[(E(X)-c)\right]^2+2(X-E(X))(E(X)-c)\right)$$

$$=E\left((X-E(X))^2\right)+(E(X)-c)^2+E\left(2(X-E(X))(E(X)-c)\right)$$

$$=E\left((X-E(X))^2\right)+(E(X)-c)^2+2(E(X)-c)(E(X)-E(X))$$

$$=E\left((X-E(X))^2\right)+(E(X)-c)^2$$

****

例 5.5.1 求区间 $[a,b]$ 上取值的所有随机变量可能达到的最大方差，并给出取到最大方差的随机变量。

解：$E\left((X-E(X))^2\right)\le E\left(\left(X-\frac{a+b}{2}\right)^2\right)\le E\left(\left(b-\frac{a+b}{2}\right)^2\right)=E\left(\left(a-\frac{a+b}{2}\right)^2\right)$

$$=\left(\frac{b-a}{2}\right)^2;$$

第一个小于等于号，等式成立的条件是 $E(X)=\frac{a+b}{2}$

第二个小于等于号，等式成立的条件是 $P(X=a)+P(X=b)=1$，

所以达到最大方差 $\left(\frac{b-a}{2}\right)^2$ 的条件是 $P(X=a)=P(X=b)=\frac{1}{2}$。

****

切比雪夫(Chebyshev)不等式： 对任意 $\varepsilon>0$，$P(|X-E(X)|\geq\varepsilon)\leq\dfrac{Var(X)}{\varepsilon^2}$。

证明：我们只对连续型随机变量给出证明，

$$P\left(\left|X-E(X)\right|\geq\varepsilon\right)=\int_{x\in D,D=\{x:|x-E(X)|\geq\varepsilon\}}f(x)dx$$

（考虑到 $|x-E(X)|\geq\varepsilon\Leftrightarrow[x-E(X)]^2\geq\varepsilon^2 \Rightarrow \dfrac{[x-E(X)]^2}{\varepsilon^2}\geq1$ ）

$$\leq\int_{x\in D,D=\{x:|x-E(X)|\geq\varepsilon\}}\left(\frac{x-E(X)}{\varepsilon}\right)^2f(x)dx$$

$$\leq\int_{-\infty}^{+\infty}\left(\frac{x-E(X)}{\varepsilon}\right)^2f(x)dx$$

$$=\frac{E\left(\left(X-E(X)\right)^2\right)}{\varepsilon^2}=\frac{Var(X)}{\varepsilon^2}$$

切比雪夫不等式可以给出随机变量的取值与其均值不同程度偏离的概率的估计。例如，由切比雪夫不等式可知，对任意随机变量，其取值在期望的正负 2 倍标准差之内的概率一定不小于 3/4。

$$P\left(\left|X-E(X)\right|<2\sigma_X\right)=1-P\left(\left|X-E(X)\right|\geq2\sigma_X\right)\geq1-\frac{Var(X)}{(2\sigma_X)^2}=\frac{3}{4}。$$

****

例 5.5.2 将一枚均匀的硬币独立地抛掷 100 次，用切比雪夫不等式估计，得到正面次数在 40-60 次之间的概率至少为多少？

解： 设投掷 100 次硬币得到正面的次数为随机变量 X ，则 $X\sim B\left(100,\dfrac{1}{2}\right)$，

$$E(X)=100\times\frac{1}{2}=50,\qquad Var(X)=100\times\frac{1}{2}\times\left(1-\frac{1}{2}\right)=25$$

$$P(40<X<60)=P\left(\left|X-E(X)\right|<10\right)=1-P\left(\left|X-E(X)\right|\geq10\right)$$

$$\geq 1-\frac{Var(X)}{10^2}=1-\frac{25}{10^2}=0.75$$

****

$Var(X)=0 \iff P(X=c)=1$，c 为某常数，即 X 以概率 1 等于一个常数；

说明： 若 $Var(X)=0$，则对任意 $n\in Z^+$，

$$P\left(\left|X-E(X)\right|\geq \frac{1}{n}\right)\leq n^2 Var(X)=0。$$

****

随机变量的期望或方差可能不存在

$$E(X)=\begin{cases}\sum_{i=1}^{\infty} x_i\cdot P(X=x_i), & X\text{为离散随机变量}\\ \int_{-\infty}^{+\infty} xf(x)dx, & X\text{为连续随机变量}\end{cases}$$

$\sum_{i=1}^{\infty}|x_i|P(X=x_i)<\infty$， $\int_{-\infty}^{+\infty}|x|f(x)dx<\infty$ 才称其数学期望存在。

例 5.5.2 柯西(Cauchy)分布 $X\sim C(\lambda,\mu)$， $f(x)=\frac{\lambda}{\pi\left(\lambda^2+(x-\mu)^2\right)}, x\in R$ 。

计算 $X\sim C(1,0)$ 的期望。 由于 $\int_{-\infty}^{+\infty}|x|\cdot f(x)dx=\int_{-\infty}^{+\infty}\frac{1}{\pi}\cdot\frac{|x|}{1+x^2}dx \ \rightarrow \ \infty$

所以 $E(X)$ 不存在，同理 $Var(X)$ 也不存在。

****

例 5.5.3 随机变量 $X\sim\begin{pmatrix}1 & 2 & 3 & \cdots\\ \frac{6}{\pi^2} & \frac{6}{4\pi^2} & \frac{6}{9\pi^2} & \cdots\end{pmatrix}$，其中 $P(X=k)=\frac{6}{k^2\pi^2}$，

$k=1,2,3,\cdots$ 。

因为$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots=\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}$，所以$\sum_{k=1}^{\infty}P\left(X=k\right)=1$

$\sum_{k=1}^{\infty}k\cdot P\left(X=k\right)=\sum_{k=1}^{\infty}k\cdot\frac{6}{k^2\pi^2}=\frac{6}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k}$发散，因此随机变量$X$的期望不存在。

****