第五周　随机变量函数的分布及随机变量的数字特征

5.1 随机变量函数的分布

同学们好！随机变量的函数仍然是随机变量，本周我们先学习随机变量函数的分布，然后我们学习期望、方差等随机变量的数字特征。

这一讲我们介绍一下随机变量函数的分布的基本计算方法，从而使我们能够掌控的随机变量范围有进一步的扩充。一般而言，随机变量经过初等函数的作用，仍然是一个随机变量。如果我们掌握了随机变量函数的概率分布有效的计算方法，那么我们就可以从初始分布出发，比较方便地得到新的随机变量的分布规律，而不用每个新的随机变量的分布都从重头计算，为概率计算带来便利。

$X \sim N\left(\mu,\sigma^2\right)$，则 $Y=\dfrac{X-\mu}{\sigma} \sim N(0,1)$

$X \sim N(0,1)$，则 $X^2 \sim ?$

$X \sim Exp(2)$，则 $e^{-2X} \sim ?$

更一般地，已知随机变量 X 的分布，求 $Y=g(X)$ 的分布

****

离散随机变量函数的分布

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix} \overset{Y=g(X)}{\Rightarrow} Y \sim \begin{pmatrix} g(x_1) & g(x_2) & \cdots & g(x_n) & \cdots \\ p_1 & p_2 & \cdots & p_n & \cdots \end{pmatrix}$$

例 5.1.1 已知随机变量 $X \sim \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 0.2 & 0.1 & 0.1 & 0.3 & 0.3 \end{pmatrix}$，求 $Y=X^2+X$ 的分布。

解：$X \sim \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 0.2 & 0.1 & 0.1 & 0.3 & 0.3 \end{pmatrix} \overset{Y=g(X)}{\Rightarrow} Y \sim \begin{pmatrix} 2 & 0 & 0 & 2 & 6 \\ 0.2 & 0.1 & 0.1 & 0.3 & 0.3 \end{pmatrix}$

$$Y \sim \begin{pmatrix} 2 & 0 & 0 & 2 & 6 \\ 0.2 & 0.1 & 0.1 & 0.3 & 0.3 \end{pmatrix} \Rightarrow Y \sim \begin{pmatrix} 0 & 2 & 6 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

****

连续型随机变量函数的分布

$$Y=g(X) \quad \Rightarrow \quad F_Y(y)=P(Y\le y)=P(g(X)\le y)$$

例 5.1.2. 设 $X\sim N(\mu,\sigma^2)$，求 $Y=\dfrac{X-\mu}{\sigma}$ 的分布。

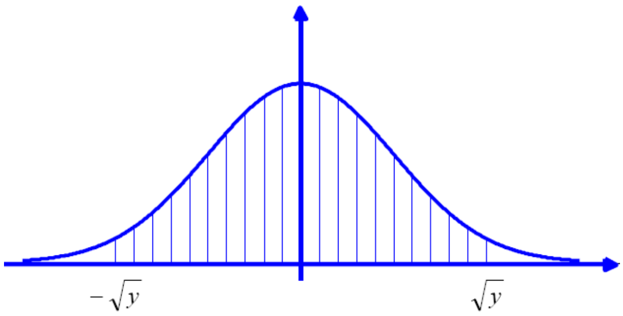
$$F_Y(y)=P(Y\le y)=P\left(\frac{X-\mu}{\sigma}\le y\right)=P(X\le \sigma y+\mu)$$

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

做变量代换 $t=\dfrac{x-\mu}{\sigma}$，　$F_Y(y)=\int_{-\infty}^{y}\dfrac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$，　即 $Y\sim N(0,1)$。

****

例 5.1.3 设 $X\sim N(0,1)$，求 $Y=X^2$ 的分布



解：　$F_Y(y)=P(Y\le y)=P(X^2\le y)$，

当 $y<0$ 时，$F_Y(y)=0$

当 $y\ge 0$ 时，$F_Y(y)=P(-\sqrt{y}\le X\le \sqrt{y})=2\left[\Phi(\sqrt{y})-\Phi(0)\right]=2\left[\Phi(\sqrt{y})-\frac{1}{2}\right]=2\Phi(\sqrt{y})-1$

当 $y\ge 0$ 时，$f_Y(y)=\dfrac{dF_Y(y)}{dy}=\dfrac{d\left(2\Phi(\sqrt{y})-1\right)}{dy}=2\varphi(\sqrt{y})\dfrac{d\sqrt{y}}{dy}=\varphi(\sqrt{y})y^{\frac{-1}{2}}$

所以 $f_Y(y)=\dfrac{dF_Y(y)}{dy}=\begin{cases}\varphi(\sqrt{y})y^{\frac{-1}{2}}, & y\ge 0\\ 0, & y<0\end{cases}$

注：当 $X\sim N(0,1)$时，X^2 称为一个自由度的 χ^2 分布，χ^2 分布是统计学中一类有重要应用的分布。

****

例 5.1.4 设 $X \sim Exp(2)$，求 $Y = e^{-2X}$ 的分布

解： $F_Y(y) = P(Y \le y) = P\left(e^{-2X} \le y\right)$，

当 $y < 0$ 时，$F_Y(y) = 0$； 当 $y \ge 1$ 时，$F_Y(y) = 1$，

当 $0 \le y < 1$ 时，

$$F_Y(y) = P\left(e^{-2X} \le y\right) = P(-2X \le \ln y) = P\left(X \ge -\frac{1}{2}\ln y\right) = e^{-2\left(-\frac{1}{2}\ln y\right)} = y$$

所以，随机变量 e^{-2X} 服从 $(0,1)$ 区间的均匀分布。

****

5.2 随机变量的数学期望

例 5.2.1

项目 1：投资 10 万元

60%可能回收 10 万元保本；40%可能回收 15 万元，盈利 5 万元

平均收益为 $0 \times \frac{3}{5} + 5 \times \frac{2}{5} = 2$ 万元

项目 2：投资 10 万元

60%可能回收 0 万元，亏损 10 万元；40%可能回收 30 万元，盈利 20 万元

平均收益为 $-10 \times \frac{3}{5} + 20 \times \frac{2}{5} = 2$ 万元

两项投资预期的平均收益都是 2 万元，若要从中做出决策，如何决定？

可能不少人会选择第一种最差也能保本的项目，但也有人更愿意尝试第二个看似风险更大的项目。所谓的风险大就是不同可能性对应的收益差别大，收益的波动大。一般而言，人们会依据平均收益和风险程度两个方面进行判断。

上述例子反映了随机或不确定情况下，平均值是人们常用的参考量，分散程度或波

动程度的不同又会带来差异。在概率论中，描述随机变量这两方面特征的标准概念是期望和方差，下面分别给出它们在数学上的定义。

****

随机变量 X 的数学期望：（加权）平均值

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

备注：离散型随机变量需要满足 $\sum_{i=1}^{\infty}|x_i|P(X=x_i)<\infty$，

连续型随机变量需要满足 $\int_{-\infty}^{+\infty}|x|f(x)dx<\infty$ 才称其数学期望存在。

****

例 5.2.2 投掷一颗均匀的色子，求掷出点数的数学期望。

解：设投出的点数为随机变量 X，则 X 服从下面分布

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6\\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6}\end{pmatrix},$$

$$E(X)=\sum_{i=1}^{6}x_i\cdot P(X=x_i)=\sum_{i=1}^{6}i\cdot\frac{1}{6}=\frac{1}{6}\frac{6\times 7}{2}=3.5$$

****

随机变量 X 函数的数学期望

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

例 5.2.3 已知随机变量 $X\sim\begin{pmatrix}-2 & -1 & 0 & 1 & 2\\ 0.2 & 0.1 & 0.1 & 0.3 & 0.3\end{pmatrix}$，求 $Y=X^2+X$ 的期望。

$$E(Y)=\sum_{k=1}^{5}\left(x_k^2+x_k\right)\cdot P\left(X=x_k\right)$$

$$=\left[(-2)^2+(-2)\right]\times 0.2+\left[(-1)^2+(-1)\right]\times 0.1+\left(0^2+0\right)\times 0.1+\left(1^2+1\right)\times 0.1+\left(2^2+2\right)\times 0.1$$

$$=1.2$$

****

数学期望的几个基本性质

$E(c)=c$ （c 为常数，常值分布）　　$E(cX)=cE(X)$

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

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

$E(g_1(X)+g_2(X))=E(g_1(X))+E(g_2(X))$

前两个结论很容易验证，后面几个关于随机变量求和的期望的结论，证明有些麻烦，我们暂且先不加证明地引入这几个结论。

****

例 5.2.4（匹配问题） n 封写给不同人的信随机放入 n 个写好收信人姓名的信封，求平均有几封信会装对信封？

解：将 n 封不同的信分别编号 $1,2,\cdots,n$， n 个对应的信封同样编号 $1,2,\cdots,n$，

定义随机变量 $X_k=\begin{cases}1, & \text{编号为}k\text{的信件装入了编号为}k\text{的信封}\\ 0, & \text{其他}\end{cases}$， $k=1,2,\cdots,n$，

$P\left(X_k=1\right)=\dfrac{1}{n}$， $P\left(X_k=0\right)=1-\dfrac{1}{n}$， 所以 $E\left(X_k\right)=1\cdot\dfrac{1}{n}+0\cdot\left(1-\dfrac{1}{n}\right)=\dfrac{1}{n}$。

装对了信封的信件总数为：$X_1+X_2+\cdots+X_n$，

所以，装对信封的信件的平均数为

$$E\left(X_1+X_2+\cdots+X_n\right)=E\left(X_1\right)+E\left(X_2\right)+\cdots+E\left(X_n\right)=n\cdot\frac{1}{n}=1。$$

****

5.3 随机变量的方差

方差：随机变量偏离期望的程度（随机变量分布的分散程度）

$$Var(X)=E\left(\left(X-E(X)\right)^2\right),$$

$$\begin{aligned}Var(X)=E\left(\left(X-E(X)\right)^2\right)&=E\left(X^2-2XE(X)+E(X)^2\right)\\&=E\left(X^2\right)-2E\left(XE(X)\right)+E(X)^2=E\left(X^2\right)-2E(X)E(X)+E(X)^2\\&=E\left(X^2\right)-E(X)^2\end{aligned}$$

$Var(X)=E\left(X^2\right)-E(X)^2$，　　　$Var(aX+b)=a^2Var(X)$

$\sigma(X)=\sqrt{Var(X)}$，　标准差，也记作σ_X

$Var(X+Y)\neq Var(X)+Var(Y)$

方差通常缩写为 $Var(X)$（varience）或 $D(X)$（deviation）。

****

例 5.3.1 项目 1：投资 10 万元

可能回收 10 万元保本；40%可能回收 15 万元，盈利 5 万元

$X_1\sim\begin{pmatrix}0&5\\\frac{3}{5}&\frac{2}{5}\end{pmatrix}$，平均收益为 $E\left(X_1\right)=0\times\frac{3}{5}+5\times\frac{2}{5}=2$万元，

项目 2：投资 10 万元

60%可能回收 0 万元，亏损 10 万元；40%可能回收 30 万元，盈利 20 万元

$X_2 \sim \begin{pmatrix} -10 & 20 \\ \frac{3}{5} & \frac{2}{5} \end{pmatrix}$，平均收益为 $E\left(X_2\right)=-10\times\frac{3}{5}+20\times\frac{2}{5}=2$ 万元

$E\left(X_1^{\ 2}\right)=0\times\frac{3}{5}+5^2\times\frac{2}{5}=10$， $Var\left(X_1\right)=E\left(X_1^{\ 2}\right)-E\left(X_1\right)^2=6$；

$E\left(X_2^{\ 2}\right)=(-10)^2\times\frac{3}{5}+20^2\times\frac{2}{5}=220$， $Var\left(X_2\right)=E\left(X_2^{\ 2}\right)-E\left(X_2\right)^2=216$。

两项投资的期望相等，均为 2 万元，但它们的方差一个是 6，一个是 216，差异非常大。期望刻画平均收益，而方差则刻画收益的波动，反映了投资的风险程度。

****

5.4 原点矩与中心矩

随机变量的原点矩与中心矩

定义 $E(X^n)$ 称为随机变量 X 的 n 阶（原点）矩； $E\left[\left(X-E\left(X\right)\right)^n\right]$ 称为随机变量 X 的 n 阶中心矩。

期望 $E(X)$ 即为随机变量 X 的 1 阶原点矩；

方差 $Var\left(X\right)=E\left[\left(X-E\left(X\right)\right)^2\right]$ 即为随机变量 X 的 2 阶中心矩。

期望和方差都是特殊的矩。

期望为随机变量 X 的 1 阶原点矩，方差为随机变量 X 的 2 阶中心矩。

****

例 5.4.1 若连续型随机变量 X 的概率密度函数为 $f(x)=\begin{cases} 3x^2, & 0<x<1, \\ 0, & 其他. \end{cases}$

试求随机变量 X 的 3 阶矩 $E\left(X^3\right)$ 和 3 阶中心矩 $E\left[\left(X-E\left(X\right)\right)^3\right]$。

解 X 的 n 阶原点矩 $E(X^n)=\int_{-\infty}^{+\infty}x^n f(x)dx=\int_0^1 x^n\cdot 3x^2dx=3\int_0^1 x^{n+2}dx=\dfrac{3}{n+3}$

故 $E(X^3)=\dfrac{1}{2}$，$E(X)=\dfrac{3}{4}$。 X 的 3 阶中心矩为

$$E\left[\left(X-E(X)\right)^3\right]=\int_{-\infty}^{+\infty}\left[x-E(X)\right]^3\cdot f(x)dx=3\int_0^1\left(x-\frac{3}{4}\right)^3x^2dx$$

$$=3\int_0^1\left(x^3-\frac{9}{4}x^2+\frac{27}{16}x-\frac{27}{64}\right)x^2dx=-\frac{1}{160}。$$

****

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

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

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

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

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

****

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

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

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 $\varepsilon>0$，$P\left(|X-E(X)|\ge\varepsilon\right)\le\frac{Var(X)}{\varepsilon^2}$。

证明：我们只对连续型随机变量给出证明，

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

（考虑到 $|x-E(X)|\ge\varepsilon \Leftrightarrow [x-E(X)]^2\ge\varepsilon^2 \;\Rightarrow\; \frac{[x-E(X)]^2}{\varepsilon^2}\ge 1$ ）

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

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

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

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。

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

****

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 $X\sim B\left(100,\frac{1}{2}\right)$，

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

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

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

****

$Var(X)=0\ \Leftrightarrow\ P(X=c)=1$，$c$ 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 $n\in Z^+$，

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

****

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 $X \sim C(\lambda, \mu)$， $f(x) = \dfrac{\lambda}{\pi\left(\lambda^2 + (x-\mu)^2\right)}, x \in R$。

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

所以 $E(X)$ 不存在，同理 $Var(X)$ 也不存在。

****

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

$k = 1, 2, 3, \cdots$。

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

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

****