第二章 随机变量及其分布

2.1 随机变量及分布函数

两种随机变量：离散型与连续型

分布函数：$F(x) = P\{X \leq x\}$,记为$X \sim F(x)$或$F_X(x)$.

分布函数的性质：
(1) 单调性:$\forall x_1 < x_2, F(x_1) \leq F(x_2)$
(2) 有界性:$\forall x, 0 \leq F(x) \leq 1$且$F(-\infty) = 0, F(+\infty) = 1$
(3) 右连续性:$F(x_0 + 0) = F(x_0)$.

> 由性质可判断一个函数是否为某一随机变量的分布函数。

> 柯西分布函数:$F(x) = \frac{1}{\pi} \arctan(x + \frac{\pi}{2})$

离散随机变量的分布列: $p_i = p(x_i) = P(X = x_i)$,记为$X \sim p_i$.

> 分布列的基本性质:
> (1) **非负性**：$p(x_i) \geq 0$
> (2) **正则性**：$\Sigma_{i \geq 1} p(x_i) = 1$

> 分布列得分布函数：$F(x) = \Sigma_{x_i \leq x} p(x_i)$
> 注：$F(x)$是递增的阶梯函数，间断点是右连续的，且间断点即是X的可能取值，间断点跳跃高度是对应的概率值。

连续随机变量的概率密度函数：$F(x) = \int_{-\infty}^{x} p(t) dt$

> $p(x) = F'(x)$(在导数存在的点上，不可导时令$p(x) = 0$)

> 概率密度函数的性质:
> (1) **非负性**：$p(x) \geq 0$
> (2) **正则性**：$\int_{-\infty}^{+\infty} p(x) dx = 1$ (含有$p(x)$的可积性)

> 随机变量的概率密度函数不唯一

2.2 随机变量的数学期望

数学期望：$E(X)=\begin{cases}\sum_{i=1} x_ip_i, 离散型\\ \int_{-\infty}^{+\infty} xp(x)dx, 连续型\end{cases}$

> 充分条件是:
> $\sum_{i\geq 1}|x_i|p(x_i)<\infty$
> $\int_{-\infty}^{+\infty}|x|p(x)dx<\infty$
>
> > 级数的绝对收敛保证了级数的和不随级数各项次序的改变而改变。

> 期望是一种加权平均

数学期望的性质

(1) $E[g(X)]=\begin{cases}\sum_i g(x_i)p_i, 离散型\\ \int_{-\infty}^{+\infty} g(x)p(x)dx, 连续型\end{cases}$

(2) 若 $c=const, E(c)=c$

(3) $\forall a=const, E(aX)=aE(X)$

(4) $E[g(X_1)\pm g(X_2)]=E[g(X_1)]\pm E[g(X_2)]$

2.3 随机变量的方差与标准差

方差：$Var(X)=E[(X-E(X))^2]$
标准差：$\sigma(X)=\sqrt{Var(X)}$

> 充分条件:$E(X^2)存在$

> 方差和标准差的区别在于量纲的不同
>
> > 随机变量的标准化：设$Var(X)>0$,令$Y=\frac{X-E(X)}{\sigma(X)}$

> 期望存在方差不一定存在，方差存在期望一定存在

方差的性质

(1) $Var(X)=E(X^2)-[E(X)]^2$
(2) $Var(c=const)=0$
(3) $Var(aX+b)=a^2Var(X)$

雪比切夫(Chebyshev)不等式
设随机变量X的期望和方差都存在，则对任意常数$\epsilon>0$,有

$$P(|X-E(X)|\geq\epsilon)\leq\frac{Var(X)}{\epsilon^2}$$

> 雪比切夫不等式给出了大偏差事件的上界,方差越大上界越大。

定理:若随机变量X的方差存在，则$Var(X)=0$的充要条件是X几乎处处为某个常数a,即$P(X=a)=1$.

2.4 常用离散分布

二项分布：$X\sim b(n,k)$
$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$
$E(X)=np$
$Var(X)=np(1-p)$

> 二点分布(0-1分布\伯努利分布):b(1,p)

泊松分布：$X\sim P(\lambda),\lambda>0$
$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$
$E(X)=\lambda$
$Var(X)=\lambda$

> **泊松定理**
> 在n重伯努利试验中，记事件A在一次试验中发生的概率为p_n,如果当$n\to\infty$,有$np_n\to\lambda$,则
>
> $$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^k=\frac{\lambda^k}{k!}e^{-\lambda}$$

超几何分布：$X\sim h(n,N,M)$

$$P(X=k)=\frac{\begin{pmatrix}M\\k\end{pmatrix}\begin{pmatrix}N-M\\n-k\end{pmatrix}}{\begin{pmatrix}N\\n\end{pmatrix}},\ \ k=0,1,...,r$$

其中$r=min\{M,n\},M\leq N,n\leq N$
$E(X)=n\frac{M}{N}$
$Var(X)=\frac{nM(N-M)(N-n)}{N^2(N-1)}$

> 超几何分布的二项近似
> 当$n\ll N,p=\frac{M}{N}$总体不合格律改变甚微，则
> $$\frac{\begin{pmatrix}M\\k\end{pmatrix}\begin{pmatrix}N-M\\n-k\end{pmatrix}}{\begin{pmatrix}N\\n\end{pmatrix}}\approx\binom{n}{k}p^k(1-p)^{n-k}$$

几何分布：$X\sim Ge(p)$
$P(X=k)=(1-p)^{k-1}p$
$E(X)=\frac{1}{p}$
$Var(X)=\frac{1-p}{p^2}$

> 几何分布的无记忆性
> $P(X>m+n|X>n)=P(X>n)$

负二项分布(帕斯卡分布) $X\sim Nb(r,p)$
$P(X=k)=\binom{k-1}{r-1}p^r(1-p)^{k-r}$
$E(X)=\frac{r}{p}$
$Var(X)=\frac{r(1-p)}{p^2}$

2.5 常用连续分布

正态分布(高斯分布)
$X\sim N(\mu,\sigma^2),-\infty<\mu<+\infty,\sigma>0$
$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ \ -\infty<x<+\infty$
$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$

$E(X) = \mu$
$Var(X) = \sigma^2$

> μ为位置参数，σ为尺度参数

> **标准正态分布** $U \sim N(0,1)$
> $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$
> $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$

> **正态变量的标准化**
> 定理:若$X \sim N(\mu, \sigma)$,则$Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

> 正态分布的3σ原则

均匀分布 $X \sim U(a,b)$
$p(x) = \begin{cases} \frac{1}{b-a}, a < x < b \\ 0, 其他 \end{cases}$
$F(x) = \begin{cases} 0, x < a \\ \frac{x-a}{b-a}, a \le x \le b \\ 1, x > b \end{cases}$
$E(X) = \frac{a+b}{2}$
$Var(X) = \frac{(b-a)^2}{12}$

指数分布 $X \sim Exp(\lambda)$
$p(x) = \begin{cases} \lambda e^{-\lambda x}, x \ge 0 \\ 0, x < 0 \end{cases}$
$F(x) = \begin{cases} 1 - e^{-\lambda x}, x \ge 0 \\ 0, x < 0 \end{cases}$
$E(X) = \frac{1}{\lambda}$
$Var(X) = \frac{1}{\lambda^2}$

> 指数分布的无记忆性:
> $P(X > s+t | X > s) = P(X > t)$

> 泊松分布与指数分布的关系

伽马分布 $X \sim Ga(\alpha, \lambda)$

$p(x)=\begin{cases}\frac{\lambda^a}{\Gamma(a)}x^{a-1}e^{-\lambda x},x\geq 0\\0,x<0\end{cases}$
$E(X)=\frac{\alpha}{\lambda}$
$Var(X)=\frac{\alpha}{\lambda^2}$

> 伽马函数: $\Gamma(a)=\int_0^{+\infty}x^{a-1}e^{-x}dx$
> (1) $\Gamma(1)=1,\Gamma(\frac{1}{2})=\sqrt{\pi}$
> (2) $\Gamma(a+1)=a\Gamma(a)$

> 常用特例:
> (1) 指数分布:$Ga(1,\lambda)=Exp(\lambda)$
> (2) 卡方分布:$Ga(\frac{n}{2},\frac{1}{2})=\chi^2(n)$

> 密度函数图像
>
> > α越大，越接近正态密度函数

> 伽马分布与泊松分布

贝塔分布 $X\sim Be(a,b)$
$p(x)=\begin{cases}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1},0<x<1\\0,其他\end{cases}$
$E(X)=\frac{a}{a+b}$
$Var(X)=\frac{ab}{(a+b)^2(a+b+1)}$

> 贝塔函数:
> $B(a,b)=\int_0^1x^{a-1}(1-x)^{b-1}dx$
>
> > $B(a,b)=B(b,a)$
> > $B(a,b)=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

2.6 随机变量函数的分布

离散随机变量函数的分布
当X是离散随机变量时，$Y=g(X)$也是离散随机变量，将$g(x_i)$一一列出，相等的值相加即

可。

连续随机变量函数的分布

1. 当 $Y = g(X)$ 为离散随机变量(处理方法与离散随机变量函数相似)

2. 当 $Y = g(X)$ 为严格单调函数

$p_Y(y) = \begin{cases} p_X[h(y)]h'(y), a < y < b \\ 0, 其他 \end{cases}$

其中 $a = min\{g(-\infty), g(+\infty)\}$， $b = max\{g(-\infty), g(+\infty)\}$, h 为 g 的反函数

> 定理：正态变量的线性变换也是正态变量

> **对数正态分布**

> 定理：$X \sim Ga(\alpha, \lambda)$,则$Y = kX \sim Ga(\alpha, \frac{\lambda}{k})$
>
> > 任意伽马分布可以变化为卡方分布

> 若X的分布函数严格单增连续，则$Y = F_X(X) \sim U(0,1)$
>
> > 均匀分布在连续分布类中占有特殊地位，蒙特卡洛法的基础。

3. 当 $Y = g(X)$ 为其他形式

直接从分布函数的定义出发。

> $X \sim N(0,1), X^2 \sim \chi^2(1)$

2.7 分布的其它特征数

k阶原点矩：$\mu_k = E(X^k)$

k阶中心矩：$\upsilon_k = E[(X - E(X))^k]$

> 前提：期望存在

> $\upsilon_k = \sum_{i=0}^{k} \binom{k}{i} \mu_i (-\mu_1)^{k-i}$

变异系数：$C_v = \frac{\sigma(X)}{E(X)}$

> 前提：二阶矩存在

下侧分位数：满足$F(x_p) = \int_{-\infty}^{x_p} p(x)dx = p$的$x_p$称为下侧$p$分位数。

中位数：下侧0.5分位数

偏度系数：$\beta_S = \frac{v_3}{v_2^{3/2}}$

> 前提：前三阶矩都存在

> 偏度系数是描述分布偏离对称性程度的一个特征数

峰度系数：$\beta_S = \frac{v_4}{v_2^2} - 3$

> 前提：前四阶矩都存在

> 峰度系数是描述分布尖峭程度或尾部粗细的一个特征数

> 常见分布的偏度与峰度