2 一维随机变量及其分布

2020年7月31日 星期五 上午9:14

一维随机变量及其分布

- **一维随机变量**
  - 概念：在样本空间 Ω 上的单值实值函数 $X = X(\omega)$，$\omega \in \Omega$（ω 为样本点）称为随机变量，简记为 X．常用 X, Y, Z 等表示随机变量．随机变量的定义域是 Ω．
  - 分布函数
    - 概念：设 X 是随机变量，x 是任意实数，称函数 $F(x) = P\{X \le x\}$ $(x \in R)$ 为随机变量 X 的分布函数，或称 X 服从分布 $F(x)$，记为 $X \sim F(x)$．
    - 性质
      - ① $F(x)$ 是 x 的单调不减函数，即对任意的 $x_1 < x_2$，有 $F(x_1) \le F(x_2)$．
      - ② $F(x)$ 是右连续函数，即对任意的 $x_0 \in R$，有 $\lim\limits_{x \to x_0^+} F(x) = F(x_0+0) = F(x_0)$
      - ③ $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$，$F(+\infty) = \lim\limits_{x \to +\infty} F(x) = 1$
      - ⇒ 这三个性质也是判断 $F(x)$ 是否为某一随机变量的分布函数的充要条件
    - 分布函数的应用（求概率）
      - ① $P\{X \le a\} = F(a)$
      - ② $P\{X < a\} = F(a-0)$
      - ③ $P\{X = a\} = F(a) - F(a-0)$
- **一维离散型随机变量**
  - 离散型随机变量及其概率分布：如果一个随机变量的可能取值是有限多个或可数无穷多个，则称它为离散型随机变量，称

    $$P\{X = x_i\} = p_i,\quad i = 1, 2, \cdots$$

    为 X 的分布列、分布律或概率分布，记为 $X \sim p_i$，概率分布常常用表格形式或矩阵形式表示，即

X	x_1	x_2	……
P	p_1	p_2	……

    或 $X \sim \begin{bmatrix} x_1 & x_2 & \cdots \\ p_1 & p_2 & \cdots \end{bmatrix}$
  - 分布律性质（充要条件）
    - ① $p_k \ge 0$，$k = 1, 2, \cdots$；
    - ② $\sum\limits_{k=1}^{n} p_k = 1$
  - 分布函数：设离散型随机变量 X 的概率分布为 $P\{X = x_i\} = p_i$，则 X 的分布函数

    $$F(x) = P\{X \le x\} = \sum_{x_i \le x} P\{X = x_i\}$$
  - 五大分布
    - ① (0-1) 分布，$X \sim B(1, p)$：如果 X 的概率分布为 $X \sim \begin{bmatrix} 1 & 0 \\ p & 1-p \end{bmatrix}$，即 $P\{X=1\} = p$，$P\{X=0\} = 1-p$，则称 X 服从参数为 p 的 0-1 分布，记为 $X \sim B(1, p)$ $(0<p<1)$ ⇒ $E(X) = p$，$D(X) = p(1-p)$
    - ② 二项分布，$X \sim B(n, p)$：如果 X 的概率分布为 $P\{X=k\} = C_n^k p^k (1-p)^{n-k}$ $(k = 0, 1, \cdots, n;\ 0<p<1)$，则称 X 服从参数为 (n, p) 的二项分布，记为 $X \sim B(n, p)$ ⇒ $E(X) = np$，$D(X) = np(1-p)$
    - ③ 泊松分布，$X \sim P(\lambda)$：如果 X 的概率分布为 $P\{X=k\} = \frac{\lambda^k}{k!} e^{-\lambda}$ $(k = 0, 1, \cdots;\ \lambda > 0)$，则称 X 服从参数为 λ 的泊松分布，记为 $X \sim P(\lambda)$ ⇒ $E(X) = \lambda$，$D(X) = \lambda$
    - ④ 几何分布，$X \sim G(p)$：如果 X 的概率分布为 $P\{X=k\} = (1-p)^{k-1} p$ $(k = 1, 2, \cdots;\ 0<p<1)$，则称 X 服从参数为 p 的几何分布，记为 $X \sim G(p)$． ⇒ $E(X) = \frac{1}{p}$，$D(X) = \frac{1-p}{p^2}$
    - ⑤ 超几何分布，$X \sim H(n, N, M)$：如果 X 的概率分布为 $P\{X=k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$ ($\max\{0, n-N+M$ 则称 X 服从参数为 (n, N, M) 的超几何分布，记为 $X \sim H(n, N, M)$． ⇒

一维连续性随机变量

连续型随机变量及其概率密度：如果随机变量 X 的分布函数可以表示为 $F(x)=\int_{-\infty}^{x}f(t)dt\ (x\in R)$，其中 $f(x)$ 是非负可积函数，则称 X 为连续型随机变量，称 $f(x)$ 为 X 的概率密度函数，简称概率

概率密度 $f(x)$ 的性质（充要条件）
- ① $f(x)\geq 0$
- ② $\int_{-\infty}^{+\infty}f(x)dx=1$
- ③ 对任意实数 $x_1<x_2$，有 $P\{x_1<X\leq x_2\}=\int_{x_1}^{x_2}f(t)dt=F(x_2)-F(x_1)$
- ④ 在 $f(x)$ 的连续点处有 $F'(x)=f(x)$。

三大分布

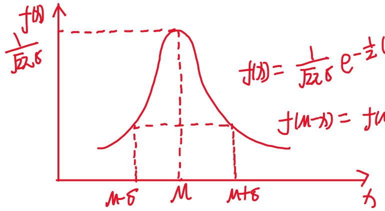
① 均匀分布，$X\sim U(a,b)$：如果随机变量 X 的概率密度或分布函数分别为 $f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b\\ 0, & 其他\end{cases}$，$F(x)=\begin{cases} \dots \end{cases}$

则称 X 在区间 (a,b) 上服从均匀分布，记为 $X\sim U(a,b)$ ⇒ $E(x)=\frac{a+b}{2}$，$D(x)=\frac{(b-a)^2}{12}$

② 指数分布，$X\sim E(\lambda)$：如果 X 的概率密度或分布函数分别为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & 其他\end{cases}$，$F(x)=\begin{cases}1-e\dots\\ 0\dots\end{cases}$

则称 X 服从参数为 λ 的指数分布，记为 $X\sim E(\lambda)$ ⇒ $E(x)=\frac{1}{\lambda}$，$D(x)=\frac{1}{\lambda^2}$

③ 正态分布，$X\sim N(\mu,\sigma^2)$：如果 X 的概率密度 $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ (-\infty<x<+\infty)$，则称 X 服从 … 或称 X 为正态变量，记为 $X\sim N(\mu,\sigma^2)$。



$x=\mu$ 为对称轴，且在 $x=\mu$ 处取唯一最大值 $f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$

$\mu=0,\ \sigma=1$ 时为标准正态分布 ⟶ 概率密度 $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$，偶函数

$X\sim N(\mu,\sigma^2)$ ⇒
- $F(x)=P\{X\leq x\}=\Phi\left(\frac{x-\mu}{\sigma}\right)$ ⟸⟹ $\frac{X-\mu}{\sigma}\sim N(0,1)$
- $F(\mu-x)+F(\mu+x)=1$
- $P\{a<X<b\}=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$
- $aX+b\sim N(a\mu+b,\ a^2\sigma^2)\ (a\neq 0)$

定义：设 X 为随机变量，函数 $y=g(x)$，则以随机变量 X 作为自变量的函数 $Y=g(X)$ 也是随机变量，称为随机变量 X 的函数。

离散型：设 X 为离散型随机变量，其概率分布为 $p_i=P\{X=x_i\}\ (i=1,2,\cdots)$，

则 X 的函数 $Y=g(X)$ 也是离散型随机变量，其概率分布为 $P\{Y=g(x_i)\}=p_i\ (i=\dots$

一维随机变量函数的分布

随机变量函数的分布

连续型：

(注：若有若干个 $y(x_k)$ 相同，则合并诸项为一项 $y(x_k)$，并将相应概率相加作为 Y

设 X 为连续型随机变量，其分布函数、概率密度分别为 $F_X(x)$ 与 $f_X(x)$，随机变

则 Y 的分布函数或概率密度可用分布函数法求得

$$F_Y(y) = P\{Y \le y\} = P\{y(X) \le y\} = \int_{y(x)\le y} f_X(x)\,dx.$$

如果 $F_Y(y)$ 连续，且除有限个点外，$F_Y'(y)$ 存在且连续，则 Y 的概率密度