2.3 连续型随机变量

2.3.1 概率密度

定义 2.9 设X分布函数为F. 若存在非负可积函数 $f(x)$ s.t. $\forall x\in\mathbb{R}$,

$$F(x)=\int_{-\infty}^{x} f(t)\,dt,$$

则称X为连续型随机变量, $f(x)$称为X的概率密度.

注. (1) 改变 f 在"零测集"上的取值不影响 $\int_{-\infty}^{x} f(t)\,dt$,

∴ 概率密度不唯一.

(2) $F(x+h)-F(x)=\int_{x}^{x+h} f(t)\,dt \approx f(x)h+o(h)$

$\Rightarrow\ F'(x)=f(x)\quad x-a.e.$ (几乎处处)

(3) 若 f 连续, 则

$$F'(x)=f(x),\quad \forall x\in\mathbb{R}.$$

性质: (1) $f(x)\geqslant 0$.

(2) $\int_{-\infty}^{+\infty} f(x)\,dx=1.$

若X为连续型随机变量, 则

$$P(X=x)=F(x)-F(x^-)=0,\quad \forall x\in\mathbb{R}.$$

∴ $\forall\ a<b$,

$$P(a<X\leq b)=P(a\leq X<b)=P(a<X<b)$$

$$=P(a\leq X\leq b)=F(b)-F(a)=\int_{a}^{b} f(t)\,dt.$$

例. 设连续型随机变量X分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ A\sin x, & 0\le x<\frac{\pi}{2},\\ 1, & x\ge\frac{\pi}{2}.\end{cases}$$

等号在左边

求 (1) A. (2) $P(|X|<\frac{\pi}{6})$ (3) X的概率密度 $f(x)$.

解: (1) $F(0)=F(0-)$ $F(\frac{\pi}{2})=F(\frac{\pi}{2}-)$

$\Rightarrow A=1$

(2) $P(|X|<\frac{\pi}{6})=F(\frac{\pi}{6})-F(-\frac{\pi}{6})=\frac{1}{2}$

(3) $$f(x)=F'(x)=\begin{cases}0, & x\le 0\\ \cos x, & 0<x<\frac{\pi}{2}\\ 0, & x\ge\pi/2\end{cases}$$

等号位置可任取

□.

2.3.2. 常见的连续型分布

(1) 均匀分布 $a<b$.

概率密度 $$f(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b\\ 0, & \text{其他}\end{cases}$$

称X服从区间 $[a,b]$ 上的均匀分布. 记为 $X\sim U(a,b)$.

若 $X\in U[a,b]$, 则 $\forall\ a\le x<y\le b$,

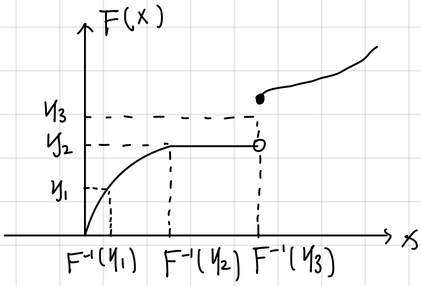
$$P(x\le X\le y)=\int_x^y f(t)\,dt=\frac{y-x}{b-a}.$$

广义逆* 设F满足单调非降性、右连续性、规范性. $\forall\ 0<y<1$,

定义

$$F^{-1}(y)=\inf\{x: F(x)\ge y\}.$$

F^{-1}性质:



(1) 单调非降

$\forall\ y_1 \le y_2,\quad F^{-1}(y_1) \le F^{-1}(y_2).$

(2) $F(F^{-1}(y)) \geq y$

$$\Rightarrow \{F(x) \geq y\} = \{x \geq F^{-1}(y)\}.$$

$\because$

若 $F(x) \geq y$，则由定义，$x \geq F^{-1}(y)$

若 $x \geq F^{-1}(y)$，则 $F(x) \geq F(F^{-1}(y)) \geq y$.

(3) 若 F 在 $F^{-1}(y)$ 处连续，则 $F(F^{-1}(y)) = y$.

命题. 1) 设 X 为连续型分布变量，分布函数为 F. 令 $Y = F(X)$，则

$$Y \sim U[0,1].$$

2) 若 $X \sim U[0,1]$，则 $F^{-1}(X)$ 分布函数为 F.

证: 1) $\forall\ y \in [0,1]$,

$$P(Y \le y) = P(F(X) \le y) = \lim_{n\to\infty} P(F(X) < y + \tfrac{1}{n})$$

$$= 1 - \lim_{n\to\infty} P(F(X) \geq y + \tfrac{1}{n}) = 1 - \lim_{n\to\infty} P(X \geq F^{-1}(y + \tfrac{1}{n}))$$

$$= 1 - \lim_{n\to\infty} \left\{ 1 - F\left(F^{-1}(y + \tfrac{1}{n})\right) \right\}$$

$$= \lim_{n\to\infty} (y + \tfrac{1}{n}) = y.$$

2) $\forall\ y \in \mathbb{R}$,

$$P(F^{-1}(X) \le y) = P(X \le F(y)) = F(y). \qquad \blacksquare$$

注: 1) 中 X 为连续型 r.v. 是必要的. 若 X 为离散型 r.v.，则 $F(X)$ 必为离散型 r.v.

(2) 指数分布　$\lambda>0$.

概率密度　$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geq 0.\\ 0, & x<0.\end{cases}$$

称X服从参数为λ的指数分布. 记为 $X\sim E(\lambda)$.

若 $X\sim E(\lambda)$, 则 $\forall x>0$,

$$P(X>x)=\int_x^{\infty}\lambda e^{-\lambda t}\,dt=e^{-\lambda x}.$$

定理2.5 (无记忆性)　X取非负实值. 则X服从指数分布

$\Leftrightarrow$　$\forall x, y>0$

$$P(X>x+y\mid X>x)=P(X>y).$$

证: "$\Rightarrow$" 显然.

"$\Leftarrow$" 令 $g(x)=P(X>x)$. 则

$$g(x+y)=g(x)g(y).$$

$\Rightarrow \forall\ m, n$正整数,

$$g\left(\frac{m}{n}\right)=g(1)^{\frac{m}{n}}$$

g右连续 $\Rightarrow$　$g(x)=g(1)^x$, $\forall x\in\mathbb{R}$.

令 $\lambda:=-\ln g(1)$. 则 $P(X>x)=e^{-\lambda x}$.　□.

实际应用: 某仪器的寿命, 银行人员的服务时间…

(3) 正态分布　$\mu\in\mathbb{R}$, $\sigma^2>0$.

$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x\in\mathbb{R}.$$

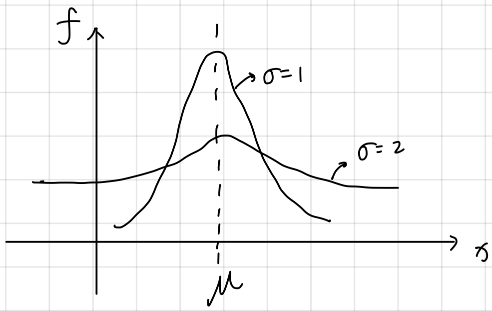
称X服从参数为 μ, σ^2 的正态分布或高斯分布. 记为 $X\sim N(\mu,\sigma^2)$.

下证 $\int_{-\infty}^{+\infty} f(x)\,dx = 1.$

只需证 $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx = 1.$

$$\left(\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx\right)^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2}}\,dx\,dy$$

$$= \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_0^{+\infty} e^{-\frac{r^2}{2}}\, r\,dr = 1.$$



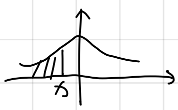
性质. (1) $f(\mu+x) = f(\mu-x).$

(2) f 在 $x=\mu$ 处取最大值 $\frac{1}{\sqrt{2\pi\sigma^2}}.$

(3) f 拐点：$x = \mu \pm \sigma.$

若 $X \sim N(0,1)$（标准正态分布），记

$$\Phi(x) = P(X \le x).$$



则 $\Phi(x) = 1 - \Phi(-x).$

若 $X \sim N(\mu, \sigma^2)$，则 $\frac{X-\mu}{\sigma} \in N(0,1)$，且

$$P(X \le x) = \phi\left(\frac{x-\mu}{\sigma}\right).$$

上侧 α 分位点 u_α： $P(X > u_\alpha) = \alpha.$

即 $\Phi(u_\alpha) = 1-\alpha.$ $u_\alpha = \Phi^{-1}(1-\alpha).$

可查标准正态分布表.

例.（3σ 原则）若 $X \sim N(\mu, \sigma^2)$，则

$$P(|X-\mu| < 3\sigma) = \Phi(3) - \Phi(-3) \approx 0.9974.$$

实际应用：测量误差，农作物产量，…

2.3.3 混合型随机变量.

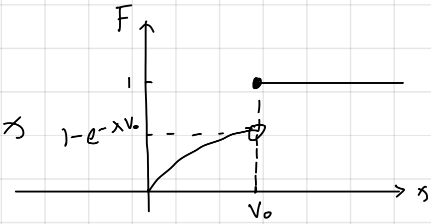
例. 设电压 $V \sim E(\lambda)$. 现用电压表测量, 最大读数为 V_0. 记 X 为电压表读数. 求 X 分布函数.

解: $X = \min\{V, V_0\}$.

$F(x) = P(X \leq x)$

$x < 0$ 时, $F(x) = 0$

$0 \leq x < V_0$ 时, $F(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$



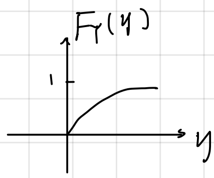
$x \geq V_0$ 时, $F(x) = 1$.

$\square$

练习.

1. 设 $X \sim U[0,1]$. 求函数 g s.t. $Y = g(X) \sim E(1)$.

解: $F_Y(y) = P(Y \leq y) = \begin{cases} 0, & y < 0. \\ 1 - e^{-y}, & y \geq 0. \end{cases}$



令 $x = 1 - e^{-y}$, $0 < x < 1$ 则

$y = -\log(1-x)$

$\therefore F^{-1}(x) = -\log(1-x)$, $0 < x < 1$

取 $g(x) = -\log(1-x)$, $0 < x < 1$ 即可.

$\square$

注: 取 $g(x) = -\log x$ 也可.