Distribution Functions, Probability Density Functions, and Statistical Quantity

October 20, 2023

随机变量 – 分布函数

定义 0.26 给定任意随机变量 X 和实数 x , 函数

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

称为随机变量 X 的分布函数, 分布函数的本质是概率.

- 分布函数 $F(x)$ 定义在 $(-\infty, +\infty)$ 的普通函数, 将概率与普通函数联系起来, 有利于利用数学分析的知识来研究随机变量;
- 分布函数不限制随机变量的类型, 无论时离散型随机变量还是非离散型随机变量, 都有各自的分布函数;
- 对任意实数 $x_1 < x_2$, 有

$$P(x_1 < X \leq x_2) = P(X \leq x_2) - P(X \leq x_1) = F(x_2) - F(x_1)$$

分布函数的性质

分布函数 $F(x)$ 具有如下性质:

- 单调性: 若 $x_1 < x_2$, 则 $F(x_1) < F(x_2)$
- 规范性: $F(x) \in [0, 1]$ 且 $F(-\infty) = 0$, $F(+\infty) = 1$
- 右连续性: $F(x + 0) = x$

任何分布函数都需要满足上述三性质, 满足上述三性质的函数必是某随机变量的分布函数. 分布函数可由上述三性质完全刻画.

概率的计算

有了分布函数 $F(x)$, 就很容易计算随机变量 X 的概率, 如:

$$P(X > a) = 1 - F(a)$$

$$P(X < a) = F(a-0) = \lim_{x \to a} F(x)$$

$$P(X = a) = F(a) - F(a-0)$$

$$P(X \geq a) = 1 - F(a-0)$$

$$P(a \leq X \leq b) = F(b) - F(a-0)$$

分布函数：例 0.61

对离散型随机变量 X , 设其分布列为 $p_k = P(X = x_k)(k = 1, 2, \dots)$, 可得 X 的分布函数为:

$$F(x) = P(X \leq x) = \sum_{k: x_k \leq x} p_k$$

例 0.61 随机变量 X 的分布列如下, 求 X 的分布函数.

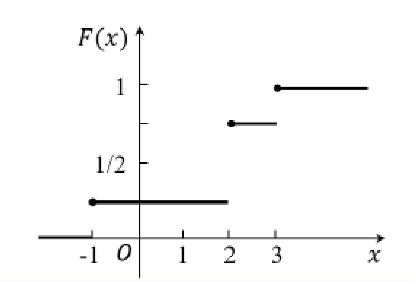
X	-1	2	3
P	1/4	1/2	1/4

解答：例 0.61

题目: 随机变量 X 的分布列为 $P(X=-1)=P(X=3)=1/4$ 和 $P(X=2)=1/2$, 求 X 的分布函数

解答:

- 对离散型随机变量 X , 设其分布列为 $p_k=P(X=x_k)(k=1,2,\dots)$, X 的分布函数为 $F(x)=P(X\le x)=\sum_{k:x_k\le x}p_k$.
- 当 $x<-1$, 有 $F(x)=P(X\le x)=P(\emptyset)=0$; 当 $-1\le x<2$, 有 $F(x)=P(X\le x)=P(X=-1)=\frac{1}{4}$; 当 $2\le x<3$, 有 $F(x)=P(X\le x)=P(X=-1)+P(X=2)=\frac{3}{4}$; 当 $x\ge 3$, 有 $F(x)=P(X\le x)=P(X=-1)+P(X=2)+P(X=3)=1$.
- 如下图所示, 分布函数 $F(x)$ 是一条阶梯形的曲线, 在 $x=-1,2,3$ 处有跳跃点.



分布函数：例 0.62

例 0.62 随机变量 X 的分布函数为 $F(x) = A + B\arctan x, x \in (-\infty, +\infty)$, 求 $P(X \leq 1)$.

解答：例 0.62

题目: 随机变量 X 的分布函数为 $F(x) = A + B\arctan x$,$x \in (-\infty, +\infty)$, 求 $P(X \le 1)$.

解答:

- 直接代入定义: $F(1) = A + B\arctan 1 = A + B\pi/4$.
- 由分布函数的性质可知, $F(-\infty) = 0$ 和 $F(+\infty) = 1$.
- 求解下列式子:

$$0 = F(-\infty) = \lim_{x \to -\infty} A + B\arctan x = A - \pi B/2$$

$$1 = F(+\infty) = \lim_{x \to +\infty} A + B\arctan x = A + \pi B/2$$

可得 $A = 1/2$ 和 $B = 1/\pi$, 从而得到 $P(X \le 1) = 3/4$.

回顾：随机变量

根据取值类型, 可以将随机变量进行分类:

- 离散型随机变量: X 的取值是有限的、无限可列的
  - 抛一枚骰子的点数: $1, 2, \ldots, 6$ ——有限的

X	1	2	3	4	5	6

  - 国家一年出生的婴儿数: $1, 2, \ldots, n$ ——有限的或者无限可列的

X	1	2	$\ldots$	n	$\ldots$

- 非离散型随机变量: X 的取值是无限不可列的

连续型随机变量 – 概率密度函数

连续型随机变量: 随机变量的取值充满整个区间 $[a,b]$ 或 (a,∞) , 例如火车的到站时间、或一盏灯泡的寿命等.

定义 0.27 设随机变量 X 的分布函数为 $F(x)$, 如果存在可积函数 $f(x)$, 使得对任意实数 x 有

$$F(x)=\int_{-\infty}^{x} f(x)dx$$

成立, 则称 X 为连续型随机变量, 函数 $f(x)$ 为随机变量 X 的概率密度函数, 简称密度函数.

Remarks:

- 随机变量的分布函数由一元积分表示, why ?
- 随机变量的取值 x 的取值为连续型

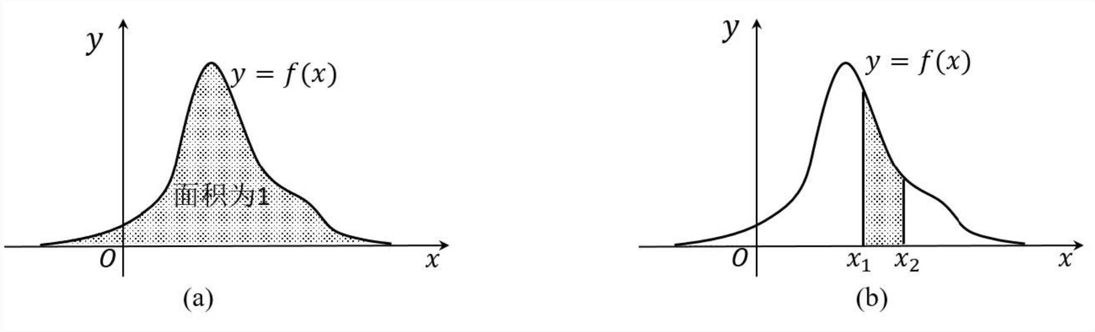
密度函数的几何解释

密度函数 $f(x)$ 满足非负性 $f(x) \geq 0$ 和规范性 $\int_{-\infty}^{+\infty} f(x)dt = 1$.

- 根据规范性可知曲线 $y = f(x)$ 与 x 轴所围成的面积为 1. 对任意 $x_1 \leq x_2$, 有

$$P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dt$$

- 几何解释: X 落入区间 $(x_1, x_2]$ 的概率等于 x 轴, $x = x_1$, $x = x_2$ 和 $y = f(x)$ 所围成的曲边梯形的面积. $f(x)$ 指向 x 点处的可能性.



概率密度函数相关定理

定理 0.9 对连续随机变量 X, 其分布函数 $F(x)$ 在整个实数域上连续; 若 $f(x)$ 在 x 点连续, 则 $F(x)$ 在 x 点可导, 且 $F^{'}(x) = f(x)$.

分布函数的导数是概率密度函数，概率密度函数的积分是分布函数?

定理 0.10 对连续型随机变量 X 和常数 x , 有 $P(X = x) = 0$.

- 事件是孤点的：一个事件的概率为 0, 不能推出该事件是不可能事件; 一个事件的概率为 1, 也不能推出该事件是必然事件.
- $P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = P(a < X \leq b)$.
- 概率密度函数不是概率: $P(X = x) = 0 \neq f(x)$.

概率与密度函数的关系

若 $f(x)$ 在点 x 连续, 根据连续性有

$$\begin{aligned}\lim_{\Delta x\to 0}\frac{P(x-\Delta x\le X\le x+\Delta x)}{\Delta x} &= \lim_{\Delta x\to 0}\frac{\int_{x-\Delta x}^{x+\Delta x} f(t)dt}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{2\Delta x f(\xi)}{\Delta x}\\ &= 2f(x)\Delta x\end{aligned}$$

其中 $\xi\in(x-\Delta x, x+\Delta x)$, 由此可得

$$P(x-\Delta x\le X\le x+\Delta x)\approx 2f(x)\Delta x$$

概率密度 $f(x)$ 越大, 则 X 在 x 附近取值的概率越大.

密度函数：例 0.63

例 0.63 设连续随机变量 X 的密度函数

$$f(x)=\begin{cases} x, & 0<x<1 \\ a-x, & 1<x\leq 2 \\ 0, & \text{others} \end{cases}$$

求分布函数 $F(x)$.

解答：例 0.63

题目: 如上所述.

解答:

- 考察密度函数的规范性及分布函数与密度函数的函数关系.
- 根据概率密度的规范性有

$$1 = \int_{-\infty}^{+\infty} f(t)dt = \int_0^1 tdt + \int_1^2 (a-t)dt = a-1$$

从而求解出 $a = 2$, 于是得到具体的密度函数 $f(x)$.

- 当 $x \leq 0$ 时, 有 $F(x) = 0$; 当 $0 < x \leq 1$ 时, 有

$$F(x) = \int_0^x f(t)dt = x^2/2;$$

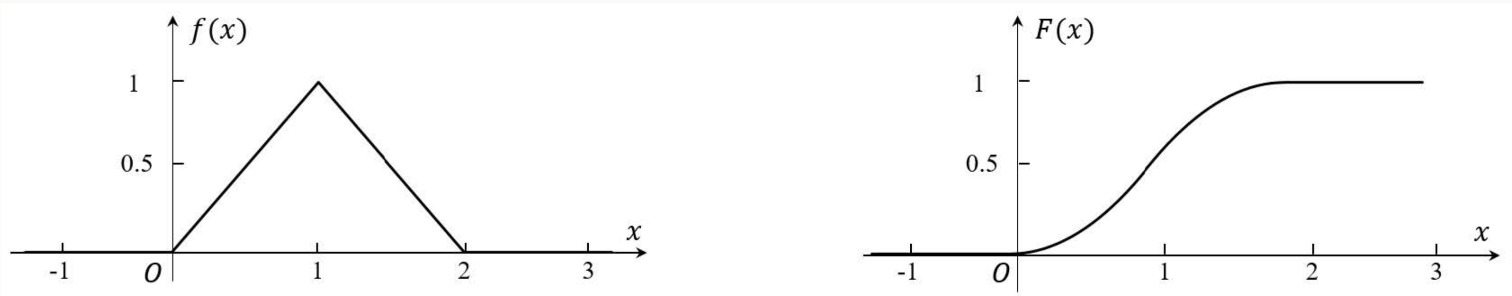
当 $1 < x \leq 2$ 时, 有

$$F(x) = \int_0^1 f(t)dt + \int_1^x f(t)dt = 1/2 + \int_1^x (2-t)dt = -x^2/2 + 2x - 1;$$

当 $x \geq 2$ 时, 有 $F(x) = 1$. 综合可得

$$f(x) = \begin{cases} 0 & x \leq 0, \\ x^2/2 & 0 < x \leq 1, \\ -x^2/2 + 2x - 1 & 1 < x \leq 2, \\ 1 & x \geq 2. \end{cases}$$

随机变量 X 的密度函数和分布函数如下图所示.



(a) 概率密度函数　　(b) 分布函数

密度函数：例 0.64

例 0.64 对连续随机变量 X , 当 $x \in (0,3)$ 时密度函数 $f(x) = cx^2$, 在其它点的密度函数 $f(x) = 0$. 设随机变量

$$Y = \begin{cases} 2, & X \leq 1 \\ X, & X \in (1,2) \\ 1, & X \geq 2 \end{cases}$$

求概率 $P(Y \geq X)$.

解答：例 0.64

题目: 如上所述.

解答:

- 考察密度函数的规范性及分布函数与密度函数的函数关系.
- 根据概率密度的规范性有 $1 = \int_{-\infty}^{+\infty} f(t)dt = 9c$, 由此可得 $c = 1/9$.
- 用 $F_Y(y)$ 表示随机变量 Y 的分布函数. 当 $y < 1$ 时, 有 $F_Y(y) = P(Y \leq y) = 0$; 当 $y \geq 2$ 时, 有 $F_Y(y) = P(Y \leq y) = 1$; 当 $1 \leq y < 2$ 时, 有

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) = P(Y = 1) + P(1 < Y \leq y) \\
&= P(X \geq 2) + P(1 < X \leq y) = \int_2^3 t^2/9dt + \int_1^y t^2/9dt = (18 + y^3)/27
\end{aligned}
$$

由此可得随机变量 Y 的分布函数为

$$F_Y(y) = \begin{cases} 0 & y < 1, \\ (18 + y^3)/27 & y \in [1, 2), \\ 1 & y \geq 2. \end{cases}$$

可以观察到随机变量 Y 不是连续型随机变量, 也不是离散型随机变量. 最后计算概率

$$P(X \leq Y) = P(X < 2) = \int_0^2 t^2/9dt = 8/27$$

连续型随机变量的数字特征 – 期望

定义 0.28 设连续随机变量 X 的密度函数为 $f(x)$, 若积分 $\int_{-\infty}^{+\infty}|x|f(x)dx$ 收敛, 称 $\int_{-\infty}^{+\infty}xf(x)dx$ 为随机变量 X 的期望, 记为 $\mathbb{E}(X)$, 即

$$\mathbb{E}(X)=\int_{-\infty}^{+\infty}xf(x)dx$$

期望：例 0.65

例 0.65 物理学中用到的柯西分布为: 随机变量 X 的密度函数是

$$f(x) = 1/\pi(1+x^2) \quad (x \in \mathbb{R}) \, ,$$

求期望 $\mathbb{E}(X)$.

解答：例 0.65

题目: 物理学中用到的柯西分布为: 随机变量 X 的密度函数是 $f(x) = 1/\pi(1+x^2)(x \in \mathbb{R})$, 求期望 $\mathbb{E}(X)$.

解答:

- 根据随机变量 X 的期望 $\mathbb{E}(X) = \int_{-\infty}^{+\infty} |x| f(x) dx$, 有
$$\int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)} dx = 2\int_{0}^{+\infty} \frac{x}{\pi(1+x^2)} dx = \frac{1}{\pi}\left[\ln(1+x^2)\right]_0^{+\infty} = +\infty$$
- 由上可知柯西分布的期望不存在.

连续期望的性质

- 对任意常数 a, b 和连续随机变量 X, 有 $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$.
- 对常数 $c_1, \ldots, c_n$ 和连续函数 $g_1(x), \ldots, g_n(x)$, 有

$$\mathbb{E}\left(\sum_{i=1}^{n} c_i g_i(X)\right) = \sum_{i=1}^{n} c_i \mathbb{E}\left(g_i(X)\right)$$

- 对连续随机变量 X 和凸函数 $f(x)$, 有 $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$.
- 对连续随机变量 X 和凹函数 $f(x)$, 有 $f(\mathbb{E}(X)) \geq \mathbb{E}(f(X))$.

非负随机变量期望的等价定义

定义 0.29 对非负随机变量 X, 有

$$\mathbb{E}[X] = \int_0^{+\infty} P(X > t)dt$$

证明

推论: 对非负随机变量 $g(X)$, 有 $\mathbb{E}[g(X)] = \int_0^{+\infty} P(g(X) > t)dt$.

随机变量函数的函数期望

定理 0.11 设随机变量 X 的密度函数为 $f(x)$、且 $\int_{-\infty}^{+\infty} g(t)f(t)dt$ 绝对可积, 则随机变量 $Y = g(X)$ 的期望

$$\mathbb{E}\left((g(X)\right) = \int_{-\infty}^{+\infty} g(t)f(t)dt$$

函数期望：例 0.66

例 0.66 古人运送粮草，如果早到每天需要的存储费用 c 元，如果晚到每天需要的延期费用为 C 元. 粮草在运送过程中存在天气、路况等不确定因素，因此运送需要的天数是随机的，概率密度函数为 $f(x)$，问什么时候出发才能使费用的期望值最小?

解答：例 0.66

解答:

- 先列出所需费用的分布函数, 再求期望的表达式及其最小值.
- 用随机变量 X 表示实际的运送天数, 分布函数为 $F(x)$. 不妨假设提前了 t 天出发 (t 也表示运粮约定时间) , 那么所需费用为

$$\ell_t(X) = \begin{cases} c(t-X) & X \le t, \\ C(X-t) & X > t. \end{cases}$$

- 因此可得

$$\begin{aligned} \mathbb{E}\left[\ell_t(X)\right] &= \int_0^{+\infty} \ell_t(x) f(x) dx = \int_0^t c(t-x) f(x) dx + \int_t^{+\infty} C(x-t) f(x) dx \\ &= ctF(t) - c\int_0^t x f(x) dx + C\int_t^{+\infty} x f(x) dx - Ct(1-F(t)) \end{aligned}$$

- 该函数是一个关于运粮约定时间或者提前出发事件 t 的函数

- 对上式中的 t 求导、并令导数为零可得

$$\frac{d}{dt}\mathbb{E}\left[\ell_t(X)\right] = cF(t) - C(1 - F(t)) = (c + C)F(t) - C$$

求解可得期望最小的天数 t^* 满足

$$F(t^*) = C/(c + C)$$

连续型随机变量的数字特征 – 方差

定义 0.30 设连续随机变量 X 的密度函数为 $f(x)$, 称

$$\mathbb{VAR}(X) = \mathbb{E}(X - \mathbb{E}(X))^2 = \int_{-\infty}^{+\infty} (t - \mathbb{E}(X))^2 f(t)dt$$

为随机变量 X 的方差.

<u>等价定义</u>:

$$\mathbb{VAR}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \int_{-\infty}^{+\infty} t^2 f(t)dt - \left(\int_{-\infty}^{+\infty} t f(t)dt\right)^2$$

连续函数方差的性质

- 对任意常数 a, b 和连续随机变量 X, 有 $\mathbb{VAR}(aX+b) = a^2 \mathrm{VAR}(X)$.
- 对连续型随机变量 X 和常数 a,

$$\mathbb{VAR}(X) = \mathbb{E}(X - \mathbb{E}(X))^2 \leq (X-a)^2$$

- 对连续型随机变量 $X \in [a, b]$, 有

$$\mathbb{VAR}(X) \leq (b - \mathbb{E}(X))(\mathbb{E}(X) - a) \leq \frac{(b-a)^2}{4}$$

方差：例 0.67

例 0.67 设随机变量 X 的密度函数为

$$f(x) = \begin{cases} \frac{2}{3}x, & 1 < x < 2 \\ 0, & \text{其它} \end{cases}$$

求 X 的方差 $\mathbb{VAR}(X)$ 和 $-2X+3$ 的方差 $\mathbb{VAR}(-2X+3)$.

解答：例 0.67

题目: 设随机变量 X 的密度函数为

$$f(x) = \begin{cases} \frac{2}{3}x, & 1 < x < 2 \\ 0, & \text{其它} \end{cases}$$

求 X 的方差 $\mathbb{VAR}(X)$ 和 $-2X+3$ 的方差 $\mathbb{VAR}(-2X+3)$.

解答:

- 根据方差的定义可知, 因为 $\mathbb{VAR}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$, 且

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_1^2 x \cdot \frac{2}{3}xdx = \frac{14}{9}$$

$$\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_1^2 x^2 \cdot \frac{2}{3}xdx = \frac{5}{2}$$

  由此可得 $\mathbb{VAR}(X) = \frac{13}{162}$.
- 根据方差的性质得, $\mathbb{VAR}(-2X+3) = 4 \cdot \mathbb{VAR}(X) = \frac{26}{81}$