信号检测与估值

段江涛
机电工程学院

2019 年 8 月

ch2. 信号检测与估计理论的基础知识

概率论简单回顾

随机现象

随机现象:

落叶的运动轨迹,气泡的扩散,灯泡的寿命

随机现象特点

至少两种可能; 不确定性。

概率论中的三个组成部分

- 样本空间 Ω
- 事件域 $\mathcal{F}$
- 概率 P

- 样本空间 Ω:一个随机试验所有可能出现的结果的全体,称为随机事件的样本空间。
  每一个可能的结果称为基本事件,它们的全体就是样本空间。
- 样本点 ξ_k:随机试验的一个结果,就是某个基本事件,也就是 Ω 中的一个元素。
  $\Omega = \{\xi_k | k = 1, \ldots, n\}$
- 随机事件 A: 样本空间中的某个子集称为随机事件,简称事件 (事件是集合)。
- 事件域 $\mathcal{F}$: 样本空间中的某些子集构成的满足如下条件的集合,称为事件域 (又称 σ^- 域)。
  (1) $\Omega \in \mathcal{F}$
  (2) 若 $A \in \mathcal{F}$, 则 A 的补 $\overline{A} \in \mathcal{F}$
  (3) 若 $A_n \in \mathcal{F}, n = 1, 2, \ldots,$ 则 $\bigcap\limits_{n=1}^{\infty} A_n \in \mathcal{F}$

Example

一个盒子中有 10 个相同的球,5 各白色,5 个黑色,搅匀后从中任意摸取一球。

$\xi_1 = \{\text{取得白球}\}, \xi_2 = \{\text{取得黑球}\}$

$\Omega = \{\xi_1, \xi_2\}$

Example

一个盒子中有 10 个相同的球,编号 1,2,...,10, 从中取一球。

$\Omega = \{1, 2, \ldots, 10\}$

$A = \{6\text{ 号球}\} = \{6\}, B = \{\text{偶数编号球}\} = \{2, 4, 6, 8, 10\},$

$\overline{B} = \{\text{奇数编号球}\}, C = \{\text{编号小于等于 5 的球}\} = \{1, 2, 3, 4, 5\}$

事件 A 是基本事件,而 B 和 C 则由多个基本事件所组成,并且 $A, B, C \subset \Omega$。

空集 $\emptyset$ 可以看作 Ω 的子集,在任意一次试验中不可能有此事件发生,称为不可能事件。

事件 A 的概率 $P(A)$

事件域中的元素就是随机事件。如果这些事件的随机性能够由定义在 $\mathcal{F}$ 上的具有非负性,归一性和可列加性的实函数 $P(A)$ 来确定,则称 P 是定义在二元组 $(\Omega, \mathcal{F})$ 上的概率,而称 $P(A)$ 为事件 A 的概率。

(1) 非负性。$P(A) \geq 0$

(2) 归一性。$P(\Omega) = 1$

(3) 可列加性。$A_1, A_2, ..., A_n$ 互不相容 $(A_i \cap A_j = \emptyset, i \neq j)$,则
$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

古典概型

- 样本空间的元素 (即基本事件) 只有有限个,不妨设为 n 个, $\Omega = \{\xi_1, \xi_2, \dots, \xi_n\}$
- 每个基本事件出现的可能性是相等的,即有 $P(\xi_1) = P(\xi_2) = \cdots = P(\xi_n)$
- 事件域 $\mathcal{F}$ 为 Ω 的所有子集的全体,即是 $Pwr(\Omega)$, Ω 的幂集,共有 2^n 个事件, $\emptyset \in \mathcal{F}, \Omega \in \mathcal{F}$。
- 由概率的有限可加性知 $1 = P(\Omega) = \sum_{i=1}^{n} P(\xi_i) \implies P(\xi_i) = \frac{1}{n}, (i = 1, \dots, n)$
- 对任意一个随机事件 $A \subseteq \mathcal{F}$, 如果 A 是 k 个基本事件的和,即 $A = \{\xi_{i_1}\} \cup \{\xi_{i_2}\} \cup \cdots \cup \{\xi_{i_k}\}$, 则

$$P(A) = \frac{k}{n} = \frac{A\text{ 中所包含的基本事件数}}{\text{基本事件总数}}$$

样本空间的选取

为求一个事件的概率, 样本空间可以有不同的取法, 但一定要认真, 基本事件和求概事件数的计算都要在同一个样本空间中进行, 否则会导致谬误!

Example

一个盒子中有 10 个相同的球, 编号 1,2,...,10, 从中取一球, 求此球的号码为偶数的概率。

$\Omega = \{1, 2, \ldots, 10\}$

$A = \{\text{偶数编号球}\} = \{2\} \cup \{4\} \cup \{6\} \cup \{8\} \cup \{10\} = \{2, 4, 6, 8, 10\}$。

$P(A) = \frac{|A|}{n} = \frac{5}{10} = \frac{1}{2}$

另外一种解法:

$\Omega = \{A, \overline{A}\}, A = \{\text{编号为偶数的球}\}, \overline{A} = \{\text{编号为奇数的球}\}, \mathcal{F} = \{\emptyset, \Omega, A, \overline{A}\}$, 由 $A, \overline{A}$ 的对称性, 即得 $P(A) = \frac{1}{2}$

样本空间的选取

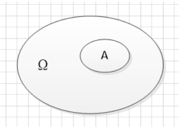
Notes

两种解法的样本空间 Ω 不同 (从而事件域 $\mathcal{F}$ 是不同的)。严格地说,两者所描述的随机试验是不同的。例如对于第二种解法来说,$B=\{$号码为 4 的球$\}$ 并不属于事件域 $\mathcal{F}$, 就是说 B 不是一个事件,从而也就没有概率可言。但对第一种解法, B 是事件,而且 $P(B)=\frac{1}{10}$。

几何概率

我们在一个面积为 S_Ω 的区域 Ω 中,等可能地任意投点,如果点落入小区域 S_A 中的可能性与 S_A 成正比,而与 A 的位置及形状无关。如果"点落入小区域 A"这个随机事件仍然记为 A, 则由 $P(\Omega)=1$ 可得

$$P(A)=\frac{S_A}{S_\Omega}$$

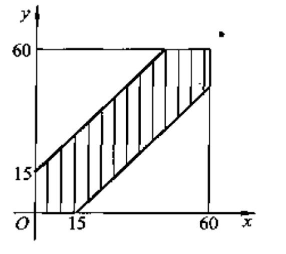


Example

(会面问题) 甲、乙两人约定在 6 时到 7 时之间在某处会面,并约定先到者应等候另一个人一刻钟,过时即可离去。求两人能会面的概率。

如图,以 x,y 表示甲乙两人,则两人能会面的充要条件是: $|x-y| \le 15$

$$P(A) = \frac{S_A}{S_\Omega} = \frac{60^2 - 45^2}{60^2} = \frac{7}{16}$$



条件概率

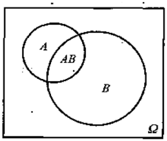
Definition

若 $\Omega, \mathcal{F}, P$ 是一个概率空间, $B \in \mathcal{F}$, 且 $P(B) > 0$, 则对任意的 $A \in \mathcal{F}$, 称

$$P(A|B) = \frac{P(AB)}{P(B)}$$

为在已知事件 B 发生的条件下, 事件 A 发生的条件概率。

$$P(A|B) = \frac{S_{AB}}{S_B} = \frac{S_{AB}/S_\omega}{S_B/S_\omega} = \frac{P(AB)}{P(B)}$$



条件概率的性质及其推论

条件概率 $P(\bullet|B)$ 的具备概率的三个基本性质

1. 非负性: 对任意的 $A \in F, P(A|B) \geq 0$;
2. 规范性: $P(\Omega|B) = 1$;
3. 可列加性: 对任意的一列两两互不相容的事件 $A_i (i = 1, 2, \ldots)$, 有

$$P\left[\bigcup_{i=1}^{+\infty} (A_i|B)\right] = \sum_{i=1}^{+\infty} P(A_i|B)$$

Corollary

概率的乘法公式: $P(AB) = P(B)P(A|B)$

Corollary

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})$$

Example

一个家庭中有两个小孩,已知其中有一个是女孩,问这时另一个小孩也是女孩的概率有多大?

Ω ={(男, 男),(男, 女),(女, 男),(女 , 女)}

A ={已知有一个是女孩}={(男, 女),(女, 男),(女, 女)}

B ={另一个也是女孩}={(女, 女)}

于是所求概率为

$P(B|A) = \frac{P(AB)}{P(A)} = \frac{1/4}{3/4} = \frac{1}{3}$

概率树/全概率公式

概率树思想:为了求解复杂事件的概率,往往可以先把它分解成两个 (或若干个) 互不相容的较简单的事件之并。求出这些较简单事件的概率,再利用加法公式即得所要求的复杂事件的概率。把这个方法一般化,便的到下述定理。

Theorem

设 $B_1, B_2, \cdots$ 是一列互不相容的事件,且有

$$\bigcup_{i=1}^{+\infty} B_i = \Omega, P(B_i) > 0$$

则对任一事件 A, 有

$$P(A) = \sum_{i=1}^{+\infty} P(B_i)P(A|B_i)$$

Example

某工厂有 4 条流水线生产同一种产品, 该 4 条流水线的产品分别占总产量的 15%, 20%, 30%, 35%, 又这 4 条流水线的不合格品率依次为 0.05, 0.04, 0.03 及 0.02. 现从出厂产品中任取一件, 问 (1) 恰好抽到不合格品的概率为多少? (2) 第 4 条流水线应承担的责任?

解: (1) 令

$$A = \{\text{任取一件,恰好抽到不合格品}\}$$

$$B = \{\text{任取一件,恰好抽到第 } i \text{ 条流水线的产品}, (i = 1, 2, 3, 4)\}$$

于是由全概率公式可得

$$P(A) = \sum_{i=1}^{4} P(B_i)P(A|B_i) = 0.15 \times 0.05 + 0.20 \times 0.04 + 0.30 \times 0.03 + 0.35 \times 0.02$$

$$= 0.0315 = 3.15\%$$

实际上,$P(A|B_i)$ 可以从过去生产的产品中统计出来,称为先验概率。

(2) 从概率论的角度考虑可以按 $P(B_i|A)$ 的大小来追究第 i 条 $i = 1, 2, 3, 4$ 流水线的责任。

$$P(AB_4) = P(B_4)P(A|B_4) = 0.35 \times 0.02 = 0.007$$

由条件概率的定义知

$$P(B_4|A) = \frac{P(AB_4)}{P(A)} = \frac{P(B_4)P(A|B_4)}{\sum\limits_{i=1}^{4} P(B_i)P(A|B_i)} = \frac{0.007}{0.0315} \approx= 0.222$$

贝叶斯 (Bayes) 公式

Theorem

若 $B_1, B_2, \ldots$ 为一系列互不相容的事件, 且

$$\bigcup_{i=1}^{+\infty} B_i = \Omega$$

$$P(B_i) > 0, i = 1, 2, \ldots$$

则对任一事件 A, 有

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{P(A)} = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{+\infty} P(B_j)P(A|B_j)}, \quad i = 1, 2, \ldots$$

$P(B_i)$ 是试验以前就已经知道的概率—**先验 (先于试验) 概率**。

条件概率 $P(B_i|A)$ 反映了试验以后, 对 A 发生的 “来源” 的各种可能性的大小—**后验概率**。

相互独立事件

条件概率: $P(B|A)=\frac{P(AB)}{P(A)}$

一般的概率乘法公式: $P(AB)=P(A)P(B|A)$

如果 “事件 B 发生与否不受事件 A 的影响”: $P(B)=P(B|A)$

乘法公式变为: $P(AB)=P(A)P(B)$

Definition

对任意的两个事件 A,B, 若

$$P(AB)=P(A)P(B)$$

成立, 则称事件 A,B 是相互独立的, 简称为独立的。

依这个定义, 不难验证:

若 A 与 B 相互独立, 则 $\{\emptyset, A, \overline{A}, \Omega\}$ 中的任意一个与 $\{\emptyset, B, \overline{B}, \Omega\}$ 中的任意一个仍相互独立。

分别掷两枚均匀的硬币,令

$$A = \{\text{硬币甲出现正面}\} \quad B = \{\text{硬币乙出现正面}\}$$

验证事件 A, B 是相互独立的。

Proof.

$$\text{样本空间} = \{(\text{正},\text{正}),(\text{正},\text{反}),(\text{反},\text{正}),(\text{反},\text{反})\}$$

共还有 4 个基本事件, 它们是等可能的,各有概率为 1/4, 而

$$A = \{(\text{正},\text{正}),(\text{正},\text{反})\}$$
$$B = \{(\text{正},\text{正}),(\text{反},\text{正})\}$$
$$AB = \{\text{正},\text{正}\}$$

由此知

$$P(A) = P(B) = \frac{1}{2}$$

这时有

$$P(AB) = \frac{1}{4} = P(A)P(B)$$

成立,所以 A, B 事件是相互独立的。 □

Definition

设 $(\Omega,\mathcal{F},P)$ 是一概率空间, $x(\xi)|\xi\in\Omega$ 是定义在 Ω 上的单值实函数, 如果对任一实数 x, 集合 $\{x(\xi)\le x\}\in\mathcal{F}$, 则称 $x(\xi)$ 为 $(\Omega,\mathcal{F},P)$ 上的一个**随机变量**。

随机变量 $x(\xi)$ 的定义域为样本空间 Ω, 它的值域是实数 R。所有随机变量 $x(\xi)$ 实际上是一个映射, 这个映射为每个来自概率空间的结果 (样本点)ξ 赋予一个实数 x。这种映射必须满足条件:

(1) 对任一 x, 集合 $\{x(\xi)\le x\}$ 是这个概率空间中的一个事件, 并有确定的概率 $P\{x(\xi)\le x\}$;

(2) $P\{x(\xi)=\infty\}=0, P\{x(\xi)=-\infty\}=0$

Notes

随机变量 $x(\xi)$ 就是试验结果 (即样本点) 和实数之间的一一对应关系。虽然在试验之前不能肯定随机变量 $x(\xi)$ 会取哪一个数值, 但是对于任一实数 a, 我们可以研究 $\{x(\xi)=a\}$ 发生的概率, 也就是 $x(\xi)$ 取值的统计规律。

随机变量的分布函数

Definition

设 $x(\xi)$ 是随机变量, 对 $\forall x \in \mathbb{R}$, 称函数

$$F(x) = P\{x(\xi) \leq x\}$$

为随机变量 $x(\xi)$ 的一维 (累积) 分布函数 [(cumulative) distribution function]。

分布函数性质

1. 单调不减性: 对 $\forall x_1 < x_2$, 恒有 $F(x_1) \leq F(x_2)$
2. 规范性: $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0, F(+\infty) = \lim\limits_{x \to +\infty} F(x) = 1$
3. 右连续性: 对 $\forall x_0$, 恒有 $F(x_0 + 0) = \lim\limits_{x \to x_0^+} F(x) = F(x_0)$

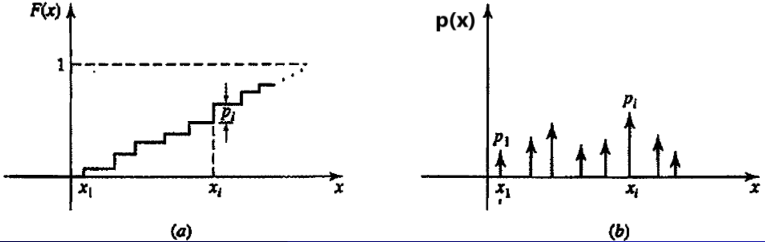
随机变量的概率密度函数

Definition

设连续随机变量 $x(\xi)$ 的一维累积分布函数为 $F(x)$, 如果 $F(x)$ 对 x 的一阶导数存在, 则有

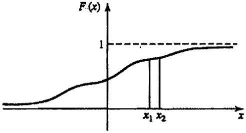
$$p(x) \overset{def}{=} \frac{dF(x)}{dx}$$

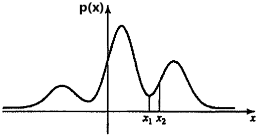
式中, $p(x)$ 称为随机变量 $x(\xi)$ 的一维概率密度函数, 简称概率密度函数 (probability density function,p.d.f)



随机变量概率密度函数性质

1. 根据随机变量 $x(\xi)$ 的 $p(x)$ 与 $F(x)$ 的关系, 有
$$F(x) = \int_{-\infty}^{x} p(u)du$$
2. 对所有 x, p(x) 是非负函数, 即
$$p(x) \geq 0, \quad -\infty < x < +\infty$$
3. $p(x)$ 对 x 的全域积分结果等于 1, 一般表示为
$$\int_{-\infty}^{\infty} p(x)dx = 1$$
4. 随机变量 $x(\xi)$ 落在区间 $[x_1, x_2]$ 内的概率为
$$P\{x_1 \leq x(\xi) \leq x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} p(x)dx$$

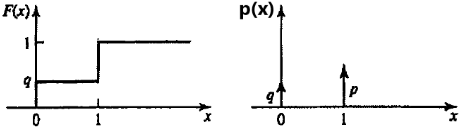




抛掷一枚硬币: 样本空间: $\Omega = \{h, t\}$, h 表示正面, t 表示反面。正面的概率 p, 反面的概率 q. 定义随机变量 $x(\xi), \xi \in \Omega$ 满足:

$$x(\xi = h) = x(h) = 1 \qquad x(\xi = t) = x(t) = 0,$$

求 $F(x)$, 其中: $-\infty < x < \infty$.



如果 $x \geq 1$, 则 $x(h) = 1 \leq x$, 且 $x(t) = 0 \leq x$, 有

$$F(x) = P\{x(\xi) \leq x\} = P\{h, t\} = 1 \qquad x \geq 1$$

如果 $0 \leq x < 1$, 则 $x(h) = 1 > x$, 且 $x(t) = 0 \leq x$, 有

$$F(x) = P\{x(\xi) \leq x\} = P\{t\} = q \qquad 0 \leq x < 1$$

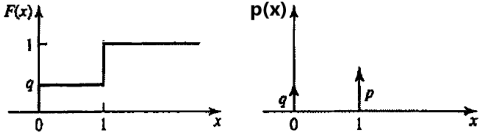
如果 $x < 0$, 则 $x(h) = 1 > x$, 且 $x(t) = 0 > x$, 有

$$F(x) = P\{x(\xi) \leq x\} = P\{\emptyset\} = 0 \qquad x < 0$$

事件 A, 试验的样本空间: $\Omega = \{A, \overline{A}, \emptyset\}$. 定义随机变量 $x(\xi)$, 满足:

$$x(\xi) = 1, \quad \xi \in A$$

$$x(\xi) = 0, \quad \xi \in \overline{A}$$



$P(A) = p, P(\overline{A}) = q = 1 - p$

如果 $x \geq 1$, 则 $\{x(\xi) \leq x\} = \{\Omega\}$, 有

$$F(x) = P\{x(\xi) \leq x\} = P\{\Omega\} = 1 \qquad x \geq 1$$

如果 $0 \leq x < 1$, 则 $\{x(\xi) \leq x\} = \{\overline{A}\}$, 有

$$F(x) = P\{x(\xi) \leq x\} = P\{\overline{A}\} = q \qquad 0 \leq x < 1$$

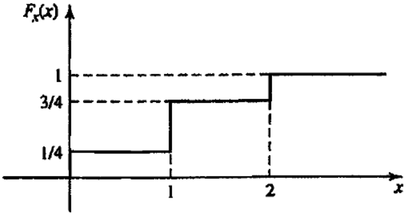
如果 $x < 0$, 则 $\{x(\xi) \leq x\} = \{\emptyset\}$, 有

$$F(x) = P\{x(\xi) \leq x\} = P\{\emptyset\} = 0 \qquad x < 0$$

抛掷两枚硬币: 随机变量 $x(\xi)$ 表示正面数目。求 $F(x)$.

样本空间: $\Omega = \{HH, HT, TH, TT\}$, H 表示正面, T 表示反面。

随机变量 $x(\xi)$: $x(HH) = 2, \quad x(HT) = 1, \quad x(TH) = 1, \quad x(TT) = 0$



如果 $x \geq 2$, $\{x(\xi) \leq x\} = \Omega \Rightarrow F(x) = 1$

如果 $1 \leq x < 2$,

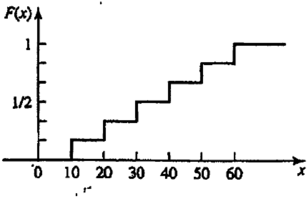
$\{x(\xi) \leq x\} = \{TT, HT, TH\} \Rightarrow F(x) = P\{TT\} + P\{HT\} + P\{TH\} = \frac{3}{4}$

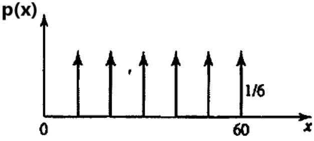
如果 $0 \leq x < 1$, $\{x(\xi) \leq x\} = \{TT\} \Rightarrow F(x) = P\{TT\} = P(T)P(T) = \frac{3}{4}$

如果 $x < 0$, $\{x(\xi) \leq x\} = \emptyset \Rightarrow F(x) = 0$

当 $x = 1$, $P\{x(\xi) = 1\} = F(1) - F(1^-) = 3/4 - 1/4 = 1/2$

掷一枚骰子: 样本空间: $\Omega = \{f_1, f_2, f_3, f_4, f_5, f_6\}$。定义随机变量 $x(\xi), \xi \in \Omega$ 满足:$x(f_i) = 10i$ 求 $F(x)$, 其中: $-\infty < x < \infty$.





$$F(100) = P\{x(\xi) \le 100\} = P(\Omega) = 1$$

$$F(35) = P\{x(\xi) \le 35\} = P\{f_1, f_2, f_3\} = \frac{3}{6}$$

$$F(30.01) = P\{x(\xi) \le 30.01\} = P\{f_1, f_2, f_3\} = \frac{3}{6}$$

$$F(30) = P\{x(\xi) \le 30\} = P\{f_1, f_2, f_3\} = \frac{3}{6}$$

$$F(29.9) = P\{x(\xi) \le 35\} = P\{f_1, f_2\} = \frac{2}{6}$$

均匀分布随机变量 $x(t) = t, 0 < t < T$

$P\{t_1 \le t \le t_2\} = \frac{t_2 - t_1}{T}, 0 < t_1 < t_2 < T$

如果 $x > T$, 有

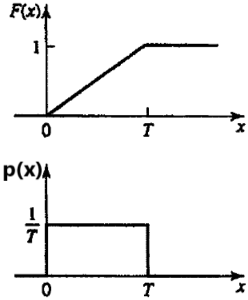
$$F(x) = P\{x(t) \le x\} = P\{0 \le t \le T\} = P(\Omega) =$$

如果 $0 \le x \le T$, 有

$$F(x) = P\{x(t) \le x\} = P\{0 \ge t \le x\} = \frac{x}{T}$$

如果 $x < 0$, 有

$$F(x) = P\{x(t) \le x\} = P\{\emptyset\} = 0 \qquad x < 0$$



定义随机变量 $x(\xi)$, 满足

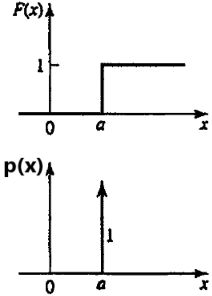
$\forall xi \in \Omega, x(\xi) = a$

如果 $x \geq a$, 则,$\forall \xi \in \Omega, x\xi = a \leq x$, 有

$$F(x) = P\{x(\xi) \leq x\} = P(\Omega) = 1 \qquad x \geq a$$

如果 $x < a$, 有

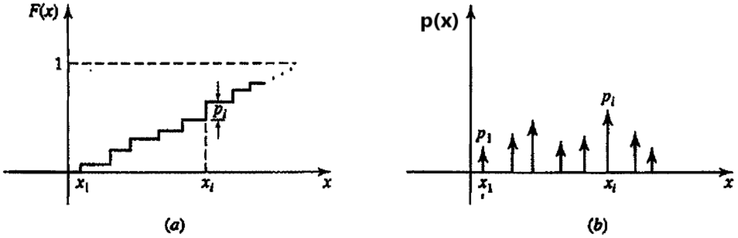
$$F(x) = P\{x(t) \leq x\} = P\{\emptyset\} = 0 \qquad x < a$$



非负实数集合 $\{p_i\}$,$\forall i, i = 1, 2, \ldots, \infty$ 满足,

1. $P\{x(\xi) = x_i\} = p_i$
2. $\sum_{i=1}^{\infty} p_i = 1$

求 $F(x)$, 其中: $-\infty < x < \infty$.



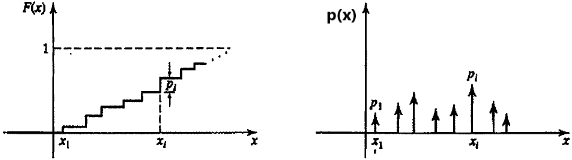
对于 $x_i \leq x < x_{i+1}$, 我们有 $\{x(\xi) \leq x\} = \bigcup_{x_k \leq x} \{x(\xi) = x_k\} = \bigcup_{k=1}^{i} \{x(\xi) = x_k\}$, 因此

$$F(x) = P\{x(\xi) \leq x\} = \sum_{k=1}^{i} p_k \qquad x_i \leq x < x_{i+1}$$

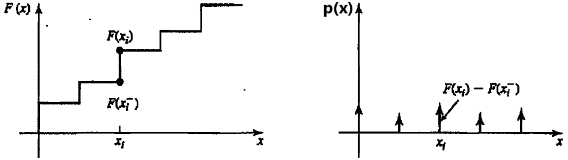
离散型随机变量 $x(\xi)$ 的 $p(x)$

$$P\{x(\xi)=x_i\}=F(x_j)-F(x_i^-)=p_i$$

$$p(x)=\sum_i p_i\delta(x-x_i)$$



(a) (b)



欢迎批评指正！