

MEM6810 工程系统建模与仿真

案例 软件

第三讲：一般随机变量与随机数

沈海辉

中美物流研究院
上海交通大学

🏠 shenhaihui.github.io/teaching/mem6810p
✉ shenhaihui@sjtu.edu.cn

2025年春 (MEM非全日制)



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

董浩云智能制造与服务管理研究院
CY TUNG Institute of Intelligent Manufacturing and Service Management
(中美物流研究院)
(Sino-US Global Logistics Institute)



1 离散与连续随机变量

- ▶ 基础定义
- ▶ 期望和方差
- ▶ 随机样本
- ▶ 点估计和区间估计

2 常用的分布

- ▶ 泊松分布
- ▶ 三角分布
- ▶ 指数分布
- ▶ 正态分布
- ▶ 经验分布

3 一般随机数的生成

- ▶ 逆变换法
- ▶ 接受-拒绝法
- ▶ Box-Muller 法



1 离散与连续随机变量

- ▶ 基础定义
- ▶ 期望和方差
- ▶ 随机样本
- ▶ 点估计和区间估计

2 常用的分布

- ▶ 泊松分布
- ▶ 三角分布
- ▶ 指数分布
- ▶ 正态分布
- ▶ 经验分布

3 一般随机数的生成

- ▶ 逆变换法
- ▶ 接受-拒绝法
- ▶ Box-Muller 法



- 对于任意的随机变量 X , 我们定义累积分布函数 (cumulative distribution function, CDF) $F(x)$ 为

$$F(x) := \mathbb{P}(X \leq x), \text{ 对于任意的 } x \in \mathbb{R}.$$

- $F(x)$ 具有如下性质:
 - $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1$;
 - $F(x)$ 为 x 的非减函数;
 - $F(x)$ 为右连续, 即, 对任意 $x_0 \in \mathbb{R}$,

$$\lim_{x \downarrow x_0} F(x) = F(x_0).$$



- 离散随机变量: 可能的取值是离散的.
- 若 X 为离散随机变量, 我们定义概率质量函数 (probability mass function, pmf) $p(x)$ 为

$$p(x) := \mathbb{P}(X = x), \text{ 对于任意的 } x \in \mathbb{R}.$$

- $p(x)$ 具有如下性质:
 - $p(x) \geq 0$, 对于任意的 $x \in \mathbb{R}$;
 - $\sum_{x \in \mathbb{R}} p(x) = 1$.
- 易知, $F(x) = \sum_{y \in (-\infty, x]} p(y)$, 对于任意的 $x \in \mathbb{R}$.



- 简单例子: 假设 X 是一个离散随机变量, 它的可能取值为 0, 1 和 2, 相应的概率为 0.5, 0.3 和 0.2.
- X 的 pmf 和 CDF 如下:

$$p(x) = \begin{cases} 0.5, & x = 0, \\ 0.3, & x = 1, \\ 0.2, & x = 2, \end{cases} \quad F(x) = \begin{cases} 0, & x < 0, \\ 0.5, & 0 \leq x < 1, \\ 0.8, & 1 \leq x < 2, \\ 1, & 2 \leq x. \end{cases}$$

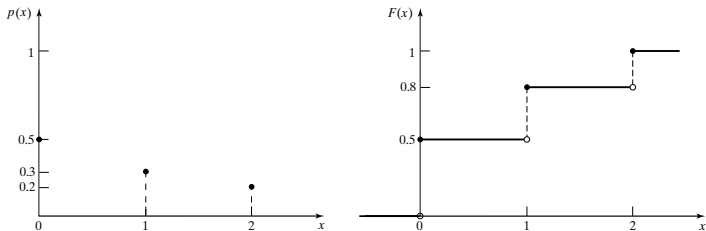


图: X 的 pmf 和 CDF 图像



- 连续随机变量: 可能的取值是连续的.
- 若 X 为连续随机变量, 我们无法定义 pmf, 因为对于任意的 $x \in \mathbb{R}$, $\mathbb{P}(X = x) = 0$.
- 我们定义概率密度函数 (probability density function, pdf) $f(x)$, 使其满足

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx, \text{ 对于任意的 } a, b \in \mathbb{R} \text{ 且 } a < b.$$

- $f(x)$ 具有如下性质:
 - $f(x) \geq 0$, 对于任意的 $x \in \mathbb{R}$;
 - $\int_{-\infty}^{+\infty} f(t)dt = 1$.
- 易知, 对于任意的 $x \in \mathbb{R}$, $F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t)dt$,
且 $\frac{d}{dx}F(x) = f(x)$.



- 简单例子: 假设连续随机变量 $X \sim \text{uniform}(a, b)$, 那么它的 pdf 和 CDF 为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{其他}, \end{cases} \quad F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & b < x. \end{cases}$$

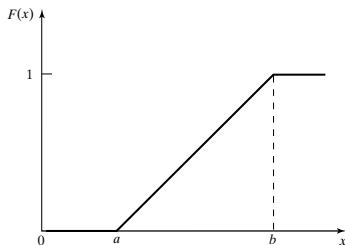
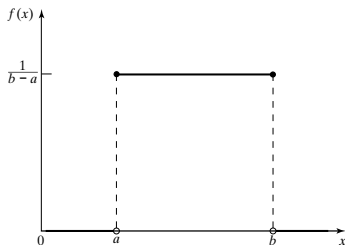


图: X 的 pdf 和 CDF 图像



- 随机变量 X 的期望 (又称为均值), 记为 $\mathbb{E}[X]$, 有时也简记为 μ .

- 离散: $\mathbb{E}[X] = \sum_{x \in \mathbb{R}} xp(x)$, $\mathbb{E}[h(X)] = \sum_{x \in \mathbb{R}} h(x)p(x)$;

- 连续: $\mathbb{E}[X] = \int_{-\infty}^{+\infty} xf(x)dx$, $\mathbb{E}[h(X)] = \int_{-\infty}^{+\infty} h(x)f(x)dx$.

- 随机变量 X 的方差, 记为 $\text{Var}(X)$, 有时也简记为 σ^2 :

$$\text{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

- 随机变量 X 与 Y 之间的线性关联:

- 协方差:

$$\text{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

- 相关系数: $\rho(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$

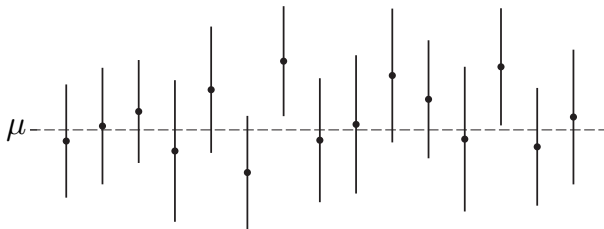
- $\rho(X, Y) = 0 \iff \text{Cov}(X, Y) = 0.$

- 一般地, X 与 Y 统计上独立 $\iff \rho(X, Y) = 0.$



- 若 x_1, x_2, \dots, x_n 为随机变量 X 的一组独立观测值, 即, 该分布下的一组随机数, 则也称它们为 X 的一组样本.
 - 该样本的样本均值为 $\bar{x} := \frac{1}{n} \sum_{i=1}^n x_i$.
 - 该样本的样本方差为 $s^2 := \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$.
- 注意到, 当样本点 x_1, x_2, \dots, x_n 被观测到之前, 它们也是随机的, 都服从 X 的分布, 并且相互独立.
- 记观测之前的样本点为 X_1, X_2, \dots, X_n , 则它们是 X 的一组 (尚未观测的) 随机样本.
 - 该随机样本的样本均值为 $\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i$.
 - 该随机样本的样本方差为 $S^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$.
- 可知, \bar{X} 和 S^2 也是随机变量, 服从一定的分布.
 - $\mathbb{E}[\bar{X}] = \mathbb{E}[X] = \mu$, $\mathbb{E}[S^2] = \text{Var}(X) = \sigma^2$.
 - $\text{Var}(\bar{X}) = \sigma^2/n$.
 - 当 n 很大时, \bar{X} 的随机性减弱, 最终趋向常数 μ . (大数定律)

- 基于随机变量 X 的一组随机样本 X_1, X_2, \dots, X_n , 如何估计 μ (即, $\mathbb{E}[X]$)?
- 点估计: 用 \bar{X} 来估计 μ .
- 区间估计: 用 $[\bar{X} - H, \bar{X} + H]$ 来估计 μ , 确保 $100(1 - \alpha)\%$ 置信水平 (confidence level).



- 试验一下! <http://www.rossmanchance.com/applets/ConfSim.html>



1 离散与连续随机变量

- ▶ 基础定义
- ▶ 期望和方差
- ▶ 随机样本
- ▶ 点估计和区间估计

2 常用的分布

- ▶ 泊松分布
- ▶ 三角分布
- ▶ 指数分布
- ▶ 正态分布
- ▶ 经验分布

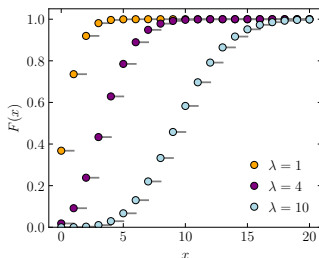
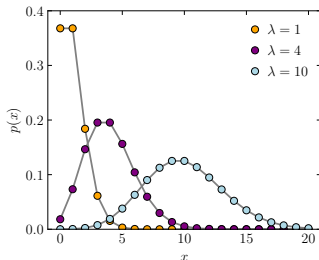
3 一般随机数的生成

- ▶ 逆变换法
- ▶ 接受-拒绝法
- ▶ Box-Muller 法



- 泊松 (Poisson) 分布, 常被用来建模给定时间段内某事件发生的次数, 如
 - 电话客服系统每分钟收到的呼叫次数;
 - 每小时到达公共汽车站的乘客数.
- 记 $X \sim \text{Poisson}(\lambda)$, 其中 $\lambda > 0$, 如果

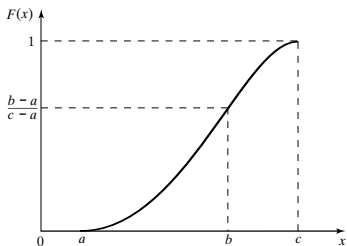
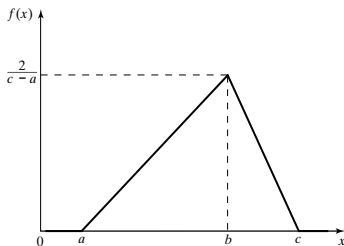
$$p(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$



- $\mathbb{E}[X] = \lambda, \text{Var}(X) = \lambda.$



- 三角 (triangular) 分布, 是工程中常用的分布, 它用于当只知道某个随机变量的“最小值”、“最大值”和“最可能值”时.
- 若 X 服从三角分布, 且最小值为 a , 最大值为 c , 最可能值为 b , 记为 $\text{triangular}(a, b, c)$, 则它的 pdf 和 CDF 如下图所示.

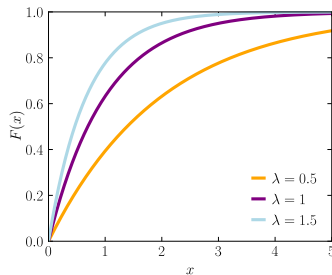
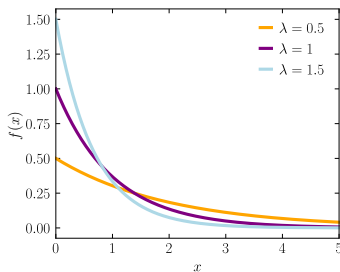


- $\mathbb{E}[X] = \frac{a+b+c}{3}$, $\text{Var}(X) = \frac{a^2+b^2+c^2-ab-ac-bc}{18}$.



- 指数 (exponential) 分布, 常用于建模独立事件之间的时间间隔, 或者一个“**无记忆的**”过程的时间长度.
- 记 $X \sim \text{exponential}(\lambda)$, 其中 $\lambda > 0$, 如果

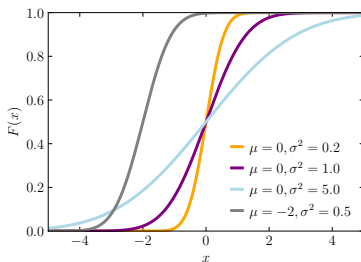
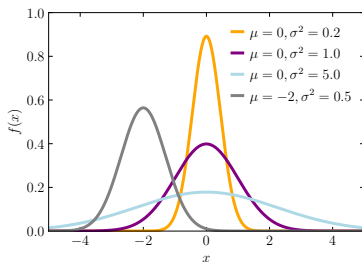
$$f(x) = \lambda e^{-\lambda x}, \quad F(x) = 1 - e^{-\lambda x}, \quad x \in [0, \infty).$$



- $\mathbb{E}[X] = 1/\lambda$, $\text{Var}(X) = 1/\lambda^2$.
- 无记忆性: $\mathbb{P}(X > s | X > t) = \mathbb{P}(X > s - t)$.

- 正态 (normal) 分布, 也叫高斯分布, 是统计中**最重要**的一个分布.
- 记 $X \sim \mathcal{N}(\mu, \sigma^2)$, 其中 $\sigma > 0$, 如果

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



- $\mathbb{E}[X] = \mu, \text{Var}(X) = \sigma^2.$



- $\mathcal{N}(0, 1)$ 称为标准正态分布.
- 如果 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则 $(X - \mu)/\sigma \sim \mathcal{N}(0, 1)$.
- 如果 $Z \sim \mathcal{N}(0, 1)$, 则 $\mu + \sigma Z \sim \mathcal{N}(\mu, \sigma^2)$.
- 正态分布的重要性来自于**中心极限定理**!
- 假设 $X_1, \dots, X_n \sim$ 任意一个分布, 且它们相互独立. 记

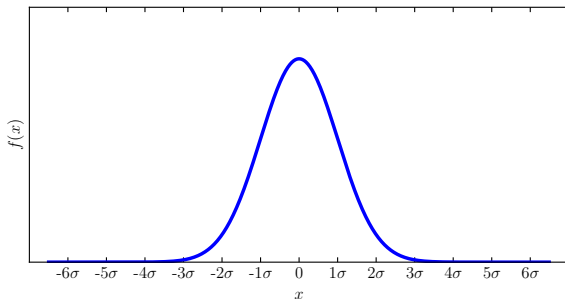
$$\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i.$$

那么中心极限定理告诉我们, 当 n 很大时,

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \stackrel{\text{近似}}{\sim} \mathcal{N}(0, 1), \quad \bar{X} \stackrel{\text{近似}}{\sim} \mathcal{N}(\mu, \sigma^2/n).$$



- 正态分布 $\mathcal{N}(\mu, \sigma^2)$ 下的 6σ



| | | |
|--------------------|------------|-------------------------------------|
| 在 $\pm 1\sigma$ 之间 | 68.27% | 317300 PPM (parts per million) 落在外面 |
| 在 $\pm 2\sigma$ 之间 | 95.45% | 45500 PPM |
| 在 $\pm 3\sigma$ 之间 | 99.73% | 2700 PPM |
| 在 $\pm 4\sigma$ 之间 | 99.9937% | 63 PPM |
| 在 $\pm 5\sigma$ 之间 | 99.999943% | 0.57 PPM |
| 在 $\pm 6\sigma$ 之间 | 99.999998% | 0.002 PPM |

- 六西格玛 (σ) DPMO (defects per million opportunity)

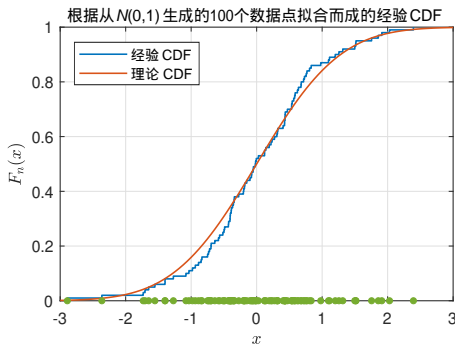
| 西格玛水平 | DPMO | 废品率 | 合格率 |
|-------|---------|----------|-----------|
| 1 | 691,462 | 69.77% | 30.23% |
| 2 | 308,538 | 30.88% | 69.12% |
| 3 | 66,807 | 6.68% | 93.32% |
| 4 | 6,210 | 0.62% | 99.38% |
| 5 | 233 | 0.023% | 99.977% |
| 6 | 3.4 | 0.00034% | 99.99966% |

- 3.4 DPMO vs 0.002 PPM? (Reason: 1.5 σ shift.)



- 经验分布 (empirical distribution), 常用于当理论分布均不适用时; 它的 CDF 为

$$F(x) = \frac{n \text{ 个点中小于或等于 } x \text{ 的点的数量}}{n}.$$



- 经验分布是离散的, 它的 CDF 是一个阶梯状的函数.

1 离散与连续随机变量

- ▶ 基础定义
- ▶ 期望和方差
- ▶ 随机样本
- ▶ 点估计和区间估计

2 常用的分布

- ▶ 泊松分布
- ▶ 三角分布
- ▶ 指数分布
- ▶ 正态分布
- ▶ 经验分布

3 一般随机数的生成

- ▶ 逆变换法
- ▶ 接受-拒绝法
- ▶ Box-Muller 法

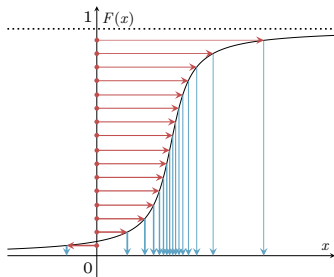


一般随机数的生成

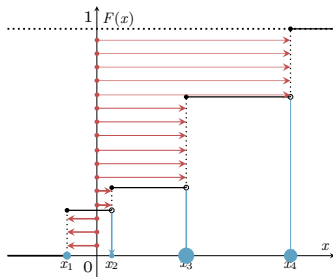
- 假设我们已经有了优良的 $\text{uniform}(0, 1)$ 随机数发生器, 即, 我们可以有一个序列的 $\text{uniform}(0, 1)$ 随机数.
- 如何生成一个给定的一般的分布 (如, 泊松、三角、指数、正态等) 下的随机数?
- 常用的技术
 - 逆变换法 (一种通用的方法)
 - 接受-拒绝法 (一种通用的方法)
 - 其他为某些分布专门设计的方法 (如, 用 Box-Muller 法生成 $\mathcal{N}(0, 1)$ 随机数)



- 逆变换法 (Inverse-Transform Technique)



图：连续随机变量



图：离散随机变量

- 步骤

- 1 生成所需数量的 $\text{uniform}(0, 1)$ 随机数 (于纵坐标).
- 2 反向映射至横坐标, 所得的点即为采样自 $F(x)$ 的随机数.

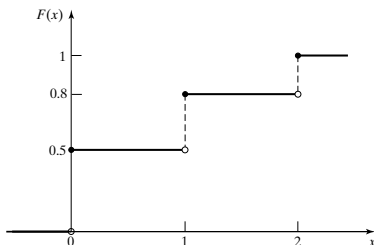
- 当一个随机变量或分布的 CDF 函数的反函数可以被**解析地求解**或者**容易地计算**时, 逆变换法是一种很有用的方法.
- 它可被用于从许多连续分布中生成随机数, 如
 - 均匀 (uniform) 分布
 - 指数 (exponential) 分布
 - 三角 (triangular) 分布
 - 威布尔 (Weibull) 分布
 - 柯西 (Cauchy) 分布
 - 帕累托 (Pareto) 分布
- 从原则上说, 它可被用于从任意的离散分布中生成随机数, 如
 - 离散均匀 (discrete uniform) 分布
 - 泊松 (Poisson) 分布
 - 经验 (empirical) 分布

- 简单例子: 假设离散随机变量 X 具有如下的 pmf 和 CDF:

$$p(x) = \begin{cases} 0.5, & x = 0, \\ 0.3, & x = 1, \\ 0.2, & x = 2, \end{cases} \quad F(x) = \begin{cases} 0, & x < 0, \\ 0.5, & 0 \leq x < 1, \\ 0.8, & 1 \leq x < 2, \\ 1, & 2 \leq x. \end{cases}$$

如何生成满足该分布的随机数?

- 求解 $F(x)$ 的反函数, $F^{-1}(y)$: (注: $F^{-1}(y) := \min\{x : F(x) \geq y\}$.)



$$F^{-1}(y) = \begin{cases} 0, & 0 < y \leq 0.5, \\ 1, & 0.5 < y \leq 0.8, \\ 2, & 0.8 < y < 1. \end{cases}$$

与最开始的直觉一致!

- 简单例子: 生成 $\text{uniform}(a, b)$ 随机数.
- 直觉: 先生成 $\text{uniform}(0, 1)$ 随机数 u , 然后输出 $x = a + (b - a)u$, 即为所需随机数.
- 已知 $\text{uniform}(a, b)$ 随机变量的 pdf 和 CDF 如下:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{其他}, \end{cases} \quad F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & b < x. \end{cases}$$

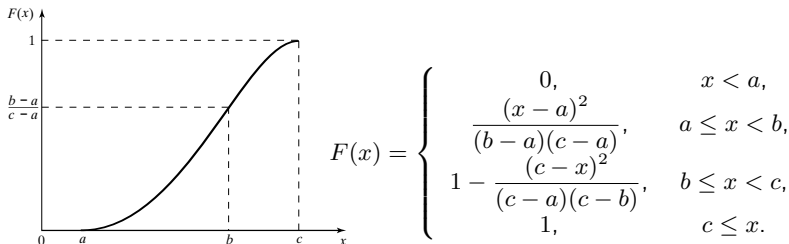
- 求解 $F(x)$ 的反函数:

$$F^{-1}(y) = a + (b - a)y, \quad 0 < y < 1.$$

- 与最开始的直觉一致!



- 例子: 生成 $\text{triangular}(a, b, c)$ 随机数.
- 已知 $\text{triangular}(a, b, c)$ 随机变量的 CDF 图像如下图所示



- 求解 $F(x)$ 的反函数:

$$F^{-1}(y) = \begin{cases} a + \sqrt{(b-a)(c-a)}\sqrt{y}, & 0 < y < \frac{b-a}{c-a}, \\ c - \sqrt{(c-b)(c-a)}\sqrt{1-y}, & \frac{b-a}{c-a} \leq y < 1. \end{cases}$$

- 在 Excel 中实施.



- 例子: 生成 $\text{exponential}(\lambda)$ 随机数.
- 已知 $\text{exponential}(\lambda)$ 随机变量的 pdf 和 CDF 如下:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0, \end{cases} \quad F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

- 求解 $F(x)$ 的反函数:

$$F^{-1}(y) = -\frac{1}{\lambda} \ln(1 - y), \quad 0 < y < 1.$$

- 注: 若 $U \sim \text{uniform}(0, 1) \implies 1 - U \sim \text{uniform}(0, 1)$, 因此就生成随机数而言, 只需计算 $-\frac{1}{\lambda} \ln(y)$ 即可.
- 在 Excel 中实施.



- 例子: 生成 $\text{Poisson}(\lambda)$ 随机数.
- 已知 $\text{Poisson}(\lambda)$ 随机变量的 pmf 和 CDF 如下:

$$p(x) = \mathbb{P}(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots,$$

$$F(x) = \mathbb{P}(X \leq x) = \sum_{i=0}^x p(i) = \sum_{i=0}^x \frac{e^{-\lambda} \lambda^i}{i!}, \quad x = 0, 1, 2, \dots$$

- 尽管 $F(x)$ 的反函数无法写成解析形式, 但由于其离散性, 可以较简单地通过反函数的定义计算:

$$F^{-1}(y) = \min\{x : F(x) \geq y\}, \quad 0 < y < 1.$$

- 在 Excel 中, 可以用

`POISSON.DIST(x,mean,cumulative)`

分别计算 $p(x)$ (`cumulative` 为 `FALSE`) 和 $F(x)$ (`cumulative` 为 `TRUE`).

- 经验分布随机数也可这样生成, 此时 $F(x)$ 是根据数据得出.



- 当一个随机变量或分布的 CDF 函数的反函数无解析形式, 且数值计算比较复杂时, 逆变换法的效率便会下降.
- 例如, 生成 $\mathcal{N}(\mu, \sigma^2)$ 随机数. 已知 $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ 以及 $F(x) = \int_{-\infty}^x f(t)dt$, 且 $F(x)$ 的反函数无解析形式, 只能通过复杂的数值计算得到 $F^{-1}(y)$.

- 在 Excel 中, 可以用

`NORM.DIST(x,mean,standard_dev,cumulative)`

分别计算 $f(x)$ (cumulative 为 FALSE) 和 $F(x)$ (cumulative 为 TRUE).

- 在 Excel 中, 还可以用

`NORM.INV(probability,mean,standard_dev)`

计算 $F^{-1}(y)$.

- 除了逆变换法之外, 还有其他的随机数生成的方法可以考虑, 如, 接受-拒绝法 (Acceptance-Rejection Technique).



- 若我们想产生某个分布下的随机数, 已知它的概率密度函数 $f(x)$ 只在 $x \in [a, b]$ 时为正值, 且 $f(x) \leq M$.

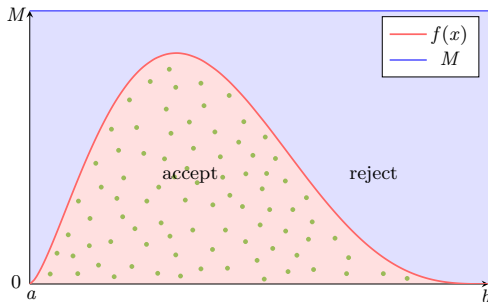
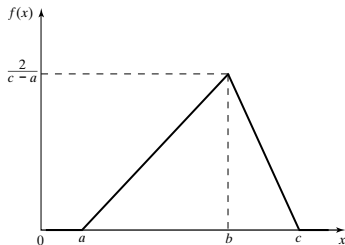


图: 有界的定义域 (original image from ZHANG Xiaowei)

- 生成 $\text{uniform}\{(y, z) : a \leq y \leq b, 0 \leq z \leq M\}$ 随机数对 $(y_1, z_1), (y_2, z_2), \dots$
 - y_i 来自 $\text{uniform}(a, b)$, z_i 来自 $\text{uniform}(0, M)$.
- 如果 $z_i < f(y_i)$, 接受这个随机数对, 并且输出 y_i .



- 例子: 生成 $\text{triangular}(a, b, c)$ 随机数.
- 已知 $\text{triangular}(a, b, c)$ 随机变量的 pdf 图像如下图所示



- ① 生成随机数对 (y, z) , 其中 y 来自 $\text{uniform}(a, c)$, z 来自 $\text{uniform}(0, 2/(c-a))$.
 - ② 如果 $z < f(y)$, 接受这个随机数对, 并且输出 y ; 否则回到第 1 步.
- **注意:** 为了生成 1 个 $\text{triangular}(a, b, c)$ 随机数, 需要生成多个其他分布的随机数!

- 若我们想产生某个分布下的随机数, 已知它的概率密度函数 $f(x)$ 具有上界 $Mg(x)$, 其中 $g(x)$ 是另一个概率密度函数.

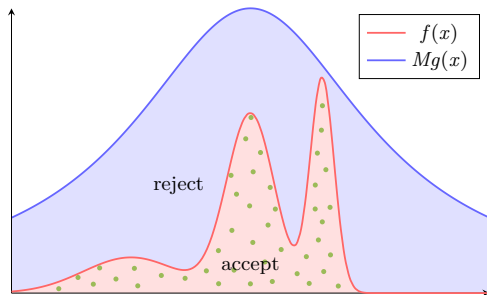


图: 无界的定义域 (original image from [ZHANG Xiaowei](#))

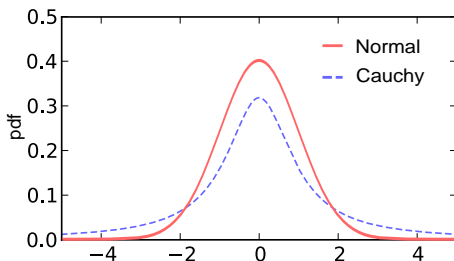
- 生成 $\text{uniform}\{(y, z) : y \in g(\cdot) \text{ 定义域}, 0 \leq z \leq Mg(y)\}$ 随机数对 $(y_1, z_1), (y_2, z_2), \dots$

• y_i 来自 $Y \sim g(\cdot)$, z_i 来自 $Z \sim \text{uniform}(0, Mg(y_i))$. (why?)

- 如果 $z_i < f(y_i)$, 接受这个随机数对, 并且输出 y_i .



- 例子: 生成 $\mathcal{N}(0, 1)$ 随机数.
 - 不难发现, 若 x 为 $\mathcal{N}(0, 1)$ 随机数, 则 $\mu + \sigma x$ 为 $\mathcal{N}(\mu, \sigma^2)$ 随机数.
- 采用 $\text{Cauchy}(0)$ 的概率密度函数作为辅助, 该函数形式为 $g(x) = \frac{1}{\pi(1+x^2)}$, $x \in (-\infty, \infty)$.



- 需要使 $M \geq \sqrt{\frac{2\pi}{e}}$ 才可以.



- 使用 Box-Muller 法生成 $\mathcal{N}(0, 1)$ 随机数。
 - 生成独立 $\text{uniform}(0, 1)$ 随机数 u_1 和 u_2 。
 - 令 $z_1 = \sqrt{-2 \ln u_1} \cos(2\pi u_2)$ 及 $z_2 = \sqrt{-2 \ln u_1} \sin(2\pi u_2)$ 。
- 可以证明, z_1 和 z_2 都是 $\mathcal{N}(0, 1)$ 随机数, 且它们相互独立。

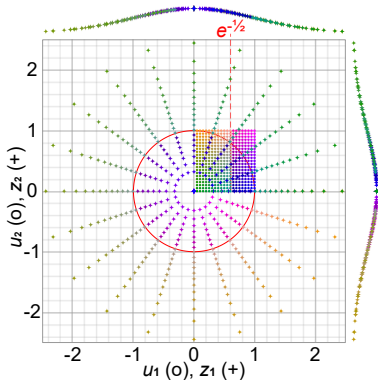


图: Box-Muller 法的可视化 (image by Cmglee / CC BY 3.0)