第二章：随机变量与概率分布

2.1 一维随机变量

赵俊舟

junzhou.zhao@xjtu.edu.cn

2025 年 3 月 5 日

目录

目录

随机变量

- 很多随机试验的样本本身就是数值，例如掷一颗筛子观察点数、每天进入某超市的顾客数、手机的寿命等；
- 还有很多随机试验的样本不是数值，例如抛硬币观察正反面、人的血型等，这时需要将样本映射为数值以便于数学处理。

例 (检测一个产品的质量是否合格)

$$\Omega = \{\text{合格}, \text{不合格}\}$$

样本	X
合格	1
不合格	0

- 可将 X 解释为“检测一个产品时合格品的数量”
- 如果用 p 表示产品合格率，则 X 取不同值的概率为

X	0	1
P	$1-p$	p

随机变量

例 (投掷一枚硬币)

$\Omega = \{$正面, 反面$\}$

样本	X
正面	1
反面	0

- 如果用 p 表示“出现正面”的概率，则 X 取不同值的概率为

X	0	1
P	$1-p$	p

例 (观察一次试验的结果)

$\Omega = \{$成功, 失败$\}$

样本	X
成功	1
失败	0

- 如果用 p 表示“试验成功”的概率，则 X 取不同值的概率为

X	0	1
P	$1-p$	p

随机变量

例 (检测三个产品中的合格情况)

用 X 表示“三个产品中的合格品数量”

样本	X
$\omega_1=(0,0,0)$	0
$\omega_2=(1,0,0)$	1
$\omega_3=(0,1,0)$	1
$\omega_4=(0,0,1)$	1
$\omega_5=(0,1,1)$	2
$\omega_6=(1,1,0)$	2
$\omega_7=(1,0,1)$	2
$\omega_8=(1,1,1)$	3

- X 取各种值就是如下互不相容事件

$$\{X=0\}=\{\omega_1\}$$
$$\{X=1\}=\{\omega_2,\omega_3,\omega_4\}$$
$$\{X=2\}=\{\omega_5,\omega_6,\omega_7\}$$
$$\{X=3\}=\{\omega_8\}$$

- 若产品的合格率为 p，则 X 取不同值的概率为

X	0	1	2	3
P	$(1-p)^3$	$3p(1-p)$	$3p^2(1-p)$	p^3

随机变量

定义 (随机变量)

随机变量是定义在样本空间 Ω 上的**实值函数** $X: \Omega \mapsto \mathbb{R}$，常用大写字母 X, Y, Z 等表示，其取值用小写字母 x, y, z 等表示。

- 如果一个随机变量仅取有限个或可列个值，则称为**离散型随机变量**；
- 如果一个随机变量的可能取值充满数轴上的区间 (a, b)（其中 a 可以是 $-\infty$， b 可以是 $+\infty$），则称为**连续型随机变量**。

- 随机变量是样本的函数，允许不同样本对应不同的实数, 也允许对应同一实数;
- 函数的自变量（样本）可以是数，也可以不是数，但因变量一定是实数。

举例：随机变量

例 (射击评估)

记录某人进行射击训练的情况，以对其射击能力进行评估。假设给其 5 发子弹，命中记为 H，否则记为 N。问：如果关心命中数量，怎么定义随机变量？

- 样本空间：$\Omega = \{\mathrm{NNNNN}, \mathrm{NNNNH}, \ldots, \mathrm{HHHHH}\}$.
- 随机变量 $X(\omega)$ 为样本点 ω 中 H 的数量。

例 (血糖值问题)

设空腹血糖高于 6.10mol/L 即为高血糖。随机抽一人检测血糖，研究事件“检测结果为高血糖”。

- 样本空间本身为实数集，可定义随机变量为 $X(\omega) = \omega$
- 事件“检测结果为高血糖”表示为：$A = \{\omega | X(\omega) > 6.10\}$

随机变量

随机现象（所关心的） $\xleftrightarrow{\text{随机变量}}$ **实值函数**（所熟悉的）

- 引入随机变量，可以在更抽象和一般的层次上研究随机现象。
- 有些问题虽然具体背景不同，但数学本质完全一样。
- 例如前面的射击评估问题，可用于类似的随机变量的研究：
  - n 次试验中成功的次数；
  - n 件产品合格的件数；
  - n 个病人治愈的数量，等等
- 随机变量 $X(\omega)$ 就是随着试验结果 ω 的不同而变化的量。
- 随机变量 $X(\omega)$ 可以简写为 X。

随机变量的分布函数

- 若 $B \subset \mathbb{R}$ 是某些实数的集合，则随机变量与随机事件的关系可以表示为 $\{\omega \in \Omega | X(\omega) \in B\}$，简写为 $\{X \in B\}$。
- 记 X 表示掷一颗筛子出现的点数，$X \in \{1, 2, \ldots, 6\}$。事件“点数不大于 3”可以表示为 $\{X \leq 3\}$
- 记 Y 表示一天内到达某超市的顾客数，$Y \in \{0, 1, 2, \ldots\}$。事件“至少来 1000 位顾客”可以表示为 $\{Y \geq 1000\}$
- 记 T 表示某手机的使用寿命，$T \in [0, +\infty)$，事件“使用寿命在 2 年到 3.5 年之间”可以表示为 $\{2 \leq T \leq 3.5\}$
- 注意到

$$\{X > x\} = \overline{\{X \leq x\}}$$

$$\{x_1 < X \leq x_2\} = \{X \leq x_2\} - \{X \leq x_1\}$$

因此，我们只需研究形如 $\{X \leq x\}$ 的事件。

随机变量的分布函数

定义 (分布函数)

设 X 是一个随机变量，记

$$F(x) \triangleq P(X \leq x) \qquad x \in \mathbb{R}$$

称 $F(x)$ 为随机变量 X 的分布函数。

- 与随机变量不同，分布函数是定义在 $\mathbb{R}$ 上的普通实函数。
- $X(\omega)$ 可以看成是样本 ω 在实数轴上的“坐标”，$F(x)$ 可以看成是坐标 $X(\omega)$ 落在区间 $(-\infty, x]$ 中的概率。
- 若已知随机变量 X 的分布函数, 就可以知道 X 落在任一区间 $(a, b]$ 的概率。
- 所以，分布函数完整地描述了随机变量的统计规律。

举例：随机变量的分布函数

例

用随机变量 X 的分布函数 $F(x)$ 表示下列事件的概率：
(a) $\{X > a\}$
(b) $\{a < X \le b\}$

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

$$\begin{aligned} P(a < X \le b) &= P(\{X \le b\} - \{X \le a\}) \\ &= P(X \le b) - P(X \le a) \\ &= F(b) - F(a) \end{aligned}$$

举例：随机变量的分布函数

例 (收入分布)

据有关研究资料，我国 2012 年家庭人均收入如下表：

x（千元）	1	2	4.5	9	15.9	25.8	34.3
收入不高于 x 的家庭比例	0.05	0.1	0.25	0.5	0.75	0.9	0.95

设 X 是随机抽取的一个家庭的人均收入，$F(x)$ 为其分布函数，试用分布函数表示下列事件的概率：
(a) $\{X \le 1\}$,　(b) $\{X > 34.3\}$,　(c) $\{2 < X \le 25.8\}$.

(a) $F(1) = 0.05$

(b) $1 - F(34.3) = 1 - 0.95 = 0.05$

(c) $F(25.8) - F(2) = 0.9 - 0.1 = 0.8$

分布函数的基本性质

性质一：单调性

$F(x)$ 是一个单调非降函数。

因为，当 $x_2 > x_1$ 时，$F(x_2) - F(x_1) = P(x_1 < X \leq x_2) \geq 0.$

性质二：有界性

对任意实数 x，$0 \leq F(x) \leq 1$，且

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \to +\infty} F(x) = 1$$

$$1 = P(-\infty < X < +\infty) = P(\bigcup_{i=-\infty}^{+\infty} \{i - 1 < X \leq i\}) = \sum_{i=-\infty}^{+\infty} P(i - 1 < X \leq i)$$

$$= \lim_{\substack{n \to +\infty \\ m \to -\infty}} \sum_{i=m}^{n} P(i - 1 < X \leq i) = \lim_{n \to +\infty} F(n) - \lim_{m \to -\infty} F(m) \Rightarrow \begin{cases} \lim_{n \to +\infty} F(n) = 1 \\ \lim_{m \to -\infty} F(m) = 0 \end{cases}$$

分布函数的基本性质

性质三：右连续性

$F(x)$ 是右连续函数，即 $\lim_{\Delta x\to 0^+} F(x+\Delta x)=F(x)$ 或 $F(x+0)=F(x)$。

因为 $F(x)$ 单调有界，则 $F(x)$ 在任意 x 的极限都存在。任取一个以 x 为极限的递减数列 $x_1>x_2>\cdots>x_n>\cdots>x$，且事件列 $\{X\le x_n\}$ 单调递减并收敛到可列交 $\bigcap_{n=1}^{\infty}\{X\le x_n\}=\{X\le x\}$，于是由概率的连续性得

$$\begin{aligned}F(x+0)&=\lim_{n\to\infty}F(x_n)=\lim_{n\to\infty}P(X\le x_n)\\&=P(\lim_{n\to\infty}\{X\le x_n\})=P(X\le x)=F(x)\end{aligned}$$

如果一个函数满足上述三条性质，那么这个函数一定是某个随机变量的分布函数。

举例：随机变量的分布函数

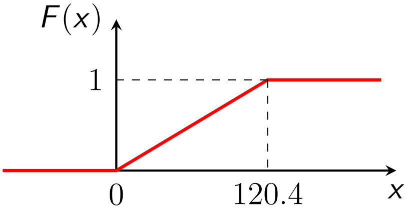
例 (污染问题)

已知空气中 PM2.5 一般在 $0.0 \sim 120.4\mu g/m^3$ 之间。根据有关指数标准，PM2.5 含量在 100.5 以上为对人体有害。设 PM2.5 的值在任一小区间 $[a, b] \subseteq [0, 120.4]$ 中的概率与区间长 $b - a$ 成正比，随机抽检空气质量，求

(a) PM2.5 值 X 的分布函数并作图；

(b) 空气质量正常的概率。

- 当 $0 \le x \le 120.4$ 时，$F(x) = \frac{x}{120.4}$
- 当 $x < 0$ 时，$F(x) = 0$
- 当 $x > 120.4$ 时，$F(x) = 1$
- 空气质量正常的概率：$F(100.5) \approx 0.83$



目录

离散型随机变量

例

1. 记 X 表示掷一颗骰子出现的点数，$X \in \{1, 2, \ldots, 6\}$。事件“点数不大于 3”可以表示为 $\{X \leq 3\}$
2. 记 Y 表示一天内到达某超市的顾客数，$Y \in \{0, 1, 2, \ldots\}$。事件“至少来 1000 位顾客”可以表示为 $\{Y \geq 1000\}$
3. 记 T 表示某手机的使用寿命，$T \in [0, +\infty)$，事件“使用寿命在 2 年到 3.5 年之间”可以表示为 $\{2 \leq T \leq 3.5\}$

- (1) 中的随机变量 X 取有限个值；(2) 中的随机变量 Y 取可数个值，这些均为离散型随机变量；
- (3) 中的随机变量 T 则连续取值，不属离散型随机变量。

离散型随机变量

定义 (离散型随机变量)

如果一个随机变量仅取有限个或可列个值，则称其为离散型随机变量。

定义 (分布律)

设 X 为离散型随机变量，其所有可能的取值为 $\{x_1, x_2, \ldots\}$，记

$$p_i \triangleq P(X = x_i), \quad i = 1, 2, \ldots$$

称为随机变量 X 的分布律，也称概率函数或概率分布。

有了分布律，可以求得分布函数：

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} p_i$$

举例：离散型随机变量

例 (掷硬币问题)

将一枚硬币连抛三次，观察正、反面出现的情况，记 X 为正面出现的次数, 求 X 的分布律。

X 的取值范围为 $\{0,1,2,3\}$，样本空间为

$$\Omega = \{\mathrm{TTT}, \mathrm{TTH}, \mathrm{THT}, \mathrm{THH}, \mathrm{HTT}, \mathrm{HTH}, \mathrm{HHT}, \mathrm{HHH}\}$$

故 X 的分布律为

$$P(X=0) = \frac{1}{8}, \quad P(X=1) = \frac{3}{8}$$

$$P(X=2) = \frac{3}{8}, \quad P(X=3) = \frac{1}{8}$$

分布律的性质

性质 (分布律的性质)

离散型随机变量 X 的分布律有以下性质:

- $p_i \geq 0, i = 1, 2, \ldots$
- $\sum_{i=1}^{\infty} p_i = 1$

$$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} P(X = x_i) = P\left(\bigcup_{i=1}^{\infty}\{X = x_i\}\right) = P(\Omega) = 1$$

常用上式来验证所给出的随机变量的分布律是否正确。

分布律的表示方法

- 解析法:

$$p_i = P(X = x_i), \quad i = 1, 2, \ldots$$

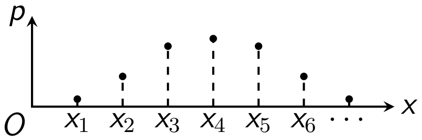
- 列表法:

X	x_1	x_2	$\cdots$	x_i	$\cdots$
P	p_1	p_2	$\cdots$	p_i	$\cdots$

- 矩阵法:

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_i & \cdots \\ p_1 & p_2 & \cdots & p_i & \cdots \end{bmatrix}$$

- 图示法:



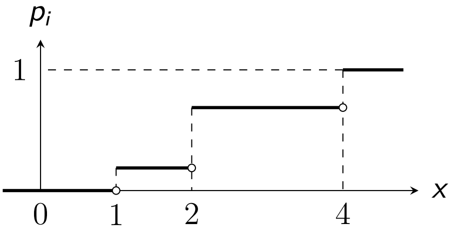
举例：离散型随机变量

例 (输血问题)

随机抽取一人观察血型，X 表示他可给别人输血的血型数。据有关资料，中国人 O，A，B，AB 血型的人分别占 41%，28%，24%，7%。(1) 写出 X 的分布律；(2) 写出其分布函数，并作图。

X	1	2	4
P	0.07	0.52	0.41

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.07, & 1 \le x < 2 \\ 0.59, & 2 \le x < 4 \\ 1, & x \ge 4. \end{cases}$$



离散型随机变量的分布函数

- 阶梯函数，阶梯数为有限或可列个。
- 一般地，设 X 的分布律为 $P(X=x_i)=p_i, i=1,2,\ldots$，则

$$F(x)=\begin{cases} 0, & x<x_1, \\ p_1, & x_1\le x<x_2 \\ p_1+p_2, & x_2\le x<x_3 \\ \ldots & \\ \sum_{j=1}^{i-1} p_i, & x_{i-1}\le x<x_i \\ \ldots & \end{cases}$$

- 分布函数又称为累积分布函数（Cumulative Distribution Function，CDF）。

举例：离散型随机变量

例 (路口红灯停车)

一汽车沿一街道行驶到终点，需要通过三个均设有红绿灯的路口，在每个路口遇到红灯的概率都为 $1/2$，且信号灯工作相互独立。以 X 表示该汽车首次停车时，在该街道上已经通过的路口数，求 X 的分布律及分布函数？

- 随机变量 $X \in \{0, 1, 2, 3\}$
- 设 A_i 表示事件“汽车在第 i 个路口遇到红灯”，$i = 1, 2, 3$
- 则 X 的分布律为

X	0	1	2	3
P	$1/2$	$1/4$	$1/8$	$1/8$

- X 的分布函数为

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \le x < 1 \\ \frac{3}{4}, & 1 \le x < 2 \\ \frac{7}{8}, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

几种常用的离散型随机变量及其分布

- 单点分布
- 两点分布
- 二项分布
- 泊松分布
- 几何分布

单点分布、两点分布、伯努利分布

单点分布

若随机变量 X 的分布律为 $P(X=a)=1$，则称 X 服从单点分布。

两点分布

若随机变量 X 的分布律为 $P(X=a_0)=1-p, P(X=a_1)=p$，或者统一表示为

$$P(X=a_i)=p^i(1-p)^{1-i} \qquad i=0,1$$

其中 $p\in[0,1]$，则称 X 服从两点分布。

$0-1$ 分布，伯努利分布，Bernoulli Distribution

特别地，当 $a_0=0, a_1=1$ 时，称 X 服从0-1 分布，也称为伯努利分布，记为 $X\sim B(1,p)$。

举例：伯努利分布

例 (收入分布)

据有关研究资料，我国 2012 年家庭人均收入如下表：

R（千元）	1	2	4.5	9	15.9	25.8	34.3
收入不高于 R 的家庭比例	0.05	0.1	0.25	0.5	0.75	0.9	0.95

若规定家庭人均收入 2 千元为贫困线。令

$$X = \begin{cases} 1, & R \leq 2 \\ 0, & R > 2 \end{cases}$$

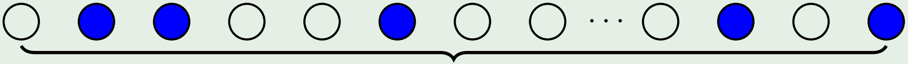
则 X 服从参数为 0.1 的伯努利分布，X 可以看成是某家庭是否是贫困家庭的示性函数（Indicator Function）。

伯努利试验

- 当随机试验的结果只有两个时，称为**伯努利试验**。
- 将伯努利试验独立地做 n 次，称为 **n 重伯努利试验**。
  - 两个要素：(1) 成功率 p 为常数；(2) 试验相互独立。

例 (n 重伯努利试验)

- 事件“n 次试验中有某特定的 k 次试验成功”的概率为：$p^k(1-p)^{n-k}$。



n 次试验

- 事件“n 次试验中有 k 次试验成功”的概率为 $\binom{n}{k}p^k(1-p)^{n-k}$，因为一共有 $\binom{n}{k}$ 种情况，彼此互斥。

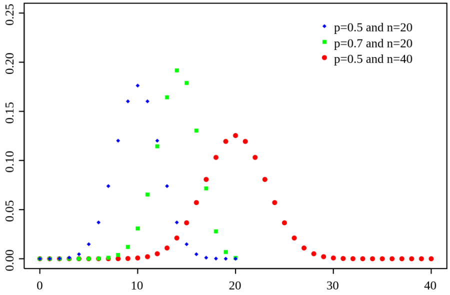
二项分布

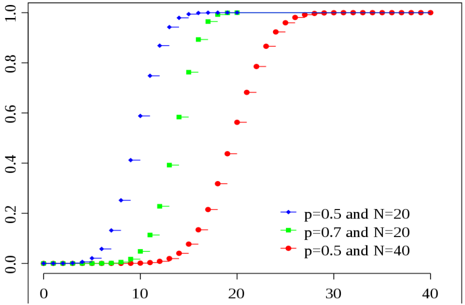
二项分布（Binomial Distribution）

若随机变量 X 的分布律为

$$P(X=i)=\binom{n}{i}p^i(1-p)^{n-i}\qquad i=0,1,\ldots,n$$

其中 $p\in[0,1]$，则称 X 服从参数为 n,p 的**二项分布**，记为 $X\sim B(n,p)$。





举例：二项分布

例 (电子元件)

一大批电子元件有 10% 已损坏，若从这批元件中随机选取 20 只来组成一个线路，问这个线路能正常工作的概率是多少？

因为元件的数量很大，所以取 20 只元件可看作是有放回抽样。令 X 表示“取出的 20 只元件中未损坏元件的数量”，则 $X \sim B(20, 0.9)$，于是

$$P(\text{线路正常}) = P(X = 20) = \binom{20}{20} \times 0.9^{20} \times 0.1^{0} \approx 0.1216$$

注：二项分布理论上只适用于有放回抽样，但当总体规模很大时，也可近似用于无放回抽样。

举例：二项分布

例 (汽车故障)

有一繁忙的汽车站，每天有大量汽车通过，设每辆汽车在一天的某时段内出故障的概率为 0.001，在每天的该时段内有 $1,000$ 辆汽车通过，问出事故的次数不小于 2 的概率是多少？

设 1000 辆车通过，出事故的次数为 X。由于每辆汽车是否发生事故相互独立，故 X 服从二项分布 $B(1000, 0.001)$。所求概率为

$$
\begin{aligned}
P(X \geq 2) = 1 - P(X < 2) &= 1 - P(X = 0) - P(X = 1) \\
&= 1 - 0.999^{1000} - \binom{1000}{1} \times 0.001 \times 0.999^{999} \\
&\approx 0.264
\end{aligned}
$$

应用二项分布时，往往计算比较繁琐，使用起来不太方便。

二项分布的泊松近似

定理（泊松定理）

设随机变量 X_n 服从二项分布 $B(n, p_n)$，其中 p_n 与 n 有关且满足 $\lim_{n\to\infty} np_n = \lambda > 0$，则

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda} \qquad k = 0, 1, \ldots$$

- 实际应用中，当 n 很大（指 $n \geq 10$）且 p 较小（指 $p \leq 0.1$）时，有以下泊松近似公式

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!}e^{-\lambda} \qquad k = 0, 1, \ldots, n.$$

- 也称为二项分布的泊松近似。

二项分布的泊松近似

证明.

令 $np_n = \lambda_n$，则 $p_n = \lambda_n/n$，而

$$\binom{n}{k} p_n^k (1-p_n)^{n-k}$$

$$= \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda_n}{n}\right)^k \left(1-\frac{\lambda_n}{n}\right)^n \left(1-\frac{\lambda_n}{n}\right)^{-k}$$

$$= \left[1(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})\right] \frac{\lambda_n^k}{k!} \left(1-\frac{\lambda_n}{n}\right)^n \left(1-\frac{\lambda_n}{n}\right)^{-k}$$

对任意固定的 k，当 $n \to \infty$ 时，因为

$$1(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n}) \to 1, \lambda_n^k \to \lambda^k, \left(1-\frac{\lambda_n}{n}\right)^{-k} \to 1$$

二项分布的泊松近似

及

$$\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^n=\lim_{n\to\infty}\left(1-\frac{\lambda_n}{n}\right)^{\frac{n}{\lambda_n}\lambda_n}=e^{-\lambda}$$

所以

$$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,\ldots$$

□

泊松分布

泊松分布（Poisson Distribution）

若随机变量 X 的分布律为

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}\qquad k=0,1,\ldots$$

其中 $\lambda>0$ 为常数，则称 X 服从参数为 λ 的泊松分布，记为 $X\sim P(\lambda)$。

- 泊松分布由法国数学家泊松提出，泊松分布是二项分布的极限分布；
- 故障、不幸事件、自然灾害等稀有事件在独立重复试验中出现的次数、固定时间段内到达的顾客数、显微镜下某区域中的细胞或微生物数量等，往往服从泊松分布。

泊松分布

(1) $p_k = \frac{\lambda^k}{k!}e^{-\lambda} > 0, k = 0, 1, \ldots$
(2) 根据指数函数 e^x 的泰勒展开，有

$$\begin{aligned}\sum_{k=0}^{\infty} p_k &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}e^{-\lambda} \\ &= e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots\right) \\ &= e^{-\lambda}e^{\lambda} \\ &= 1\end{aligned}$$

因此 $p_k = \frac{\lambda^k}{k!}e^{-\lambda}, k = 0, 1, \ldots$ 确实为随机变量的分布律。

举例：泊松分布

例 (维修工人配备问题)

为保证设备正常工作，需配备适量的维修工人。现有同类设备 300 台，各台工作相互独立，故障率都是 0.01，一台设备的故障由一个人处理。问至少需配备多少工人，才能保证设备发生故障但不能及时维修的概率小于 0.005？

设需要配备 N 人，记发生故障的设备数为 X，易知 $X \sim B(300, 0.01)$。所需解决的问题是确定最小的 N，使得 $P(X \leq N) \geq 0.995$。由泊松定理得

$$P(X \leq N) \approx \sum_{k=0}^{N} \frac{3^k e^{-3}}{k!}$$

故希望上式 ≥ 0.995

举例：泊松分布

即希望

$$1-\sum_{k=0}^{N}\frac{3^k e^{-3}}{k!}=\sum_{k=N+1}^{\infty}\frac{3^k e^{-3}}{k!}$$
$$\leq 0.005$$

查泊松分布表得，满足上述不等式得最小值 N 为 8，故最少应配备 8 名工人。

$$1-F(x-1)=\sum_{r=x}^{+\infty}\frac{e^{-\lambda}\lambda^r}{r!}$$

附表 2 （续）

x	$\lambda=2.5$	$\lambda=3.0$	$\lambda=3.5$	$\lambda=4.0$	$\lambda=4.5$	$\lambda=5.0$
0	1.000 000	1.000 000	1.000 000	1.000 000	1.000 000	1.000 000
1	0.917 915	0.950 213	0.969 803	0.981 684	0.988 891	0.993 262
2	0.712 703	0.800 852	0.864 112	0.908 422	0.938 901	0.959 572
3	0.456 187	0.576 810	0.679 153	0.761 897	0.826 422	0.875 348
4	0.242 424	0.352 768	0.463 367	0.566 530	0.657 704	0.734 974
5	0.108 822	0.184 737	0.274 555	0.371 163	0.467 896	0.559 507
6	0.042 021	0.083 918	0.142 386	0.214 870	0.297 070	0.384 039
7	0.014 187	0.033 509	0.065 288	0.110 674	0.168 949	0.237 817
8	0.004 247	0.011 905	0.026 739	0.051 134	0.086 586	0.133 372
9	0.001 140	0.003 803	0.009 874	0.021 363	0.040 257	0.068 094
10	0.000 277	0.001 102	0.003 315	0.008 132	0.017 093	0.031 828
11	0.000 062	0.000 292	0.001 019	0.002 840	0.006 669	0.013 695
12	0.000 013	0.000 071	0.000 289	0.000 915	0.002 404	0.005 453
13	0.000 002	0.000 016	0.000 076	0.000 274	0.000 805	0.002 019
14		0.000 003	0.000 019	0.000 076	0.000 252	0.000 698
15		0.000 001	0.000 004	0.000 020	0.000 074	0.000 226

举例：泊松分布

例 (疾病分布律)

设某地区患某种疾病的人数 $X \sim P(\lambda)$，λ 未知，若已知该地区患此病的概率为 0.001，求 X 的分布律。

由 $1 - e^{-\lambda} = 0.001$ 得 $\lambda = \ln \frac{1000}{999} \approx 0.001$。故分布律为

$$P(X = k) = 0.999 \frac{0.001^k}{k!}, \quad k = 0, 1, 2, \ldots$$

几何分布

几何分布（Geometric Distribution）

若随机变量 X 的分布律为

$$P(X=k)=(1-p)^{k-1}p \qquad k=1,2,\ldots$$

其中 $p\in(0,1]$ 为常数，则称 X 服从参数为 p 的几何分布，记为 $X\sim \mathrm{Geom}(p)$。

- 令 $A_i, i\geq 1$，为彼此独立的伯努利试验，$P(A_i)=p$，则试验首次成功需要的总的试验次数 $X\sim \mathrm{Geom}(p)$：

$$P(X=k)=P(\overline{A_1}\,\overline{A_2}\cdots\overline{A_{k-1}}A_k)=(1-p)^{k-1}p$$

- 有时会关心试验失败的总次数 Y，随机变量 Y 同样被称为服从几何分布：

$$P(Y=k)=(1-p)^{k}p \qquad k=0,1,\ldots$$

几何分布

- 若 $X \sim \mathrm{Geom}(p)$，则
$$P(X=k+n|X>k)=P(X=n), \quad k,n=1,2,\ldots$$
称为**几何分布的无记忆性**。

$$P(X=k+n|X>k)=\frac{P(X=k+n,X>k)}{P(X>k)}=\frac{P(X=k+n)}{P(X>k)}$$

$$P(X=k+n)=(1-p)^{k+n-1}p$$

$$P(X>k)=\sum_{i=k+1}^{\infty}(1-p)^{i-1}p=\frac{p(1-p)^k}{1-(1-p)}=(1-p)^k$$

$$P(X=k+n|X>k)=\frac{(1-p)^{k+n-1}p}{(1-p)^k}=(1-p)^{n-1}p=P(X=n).$$

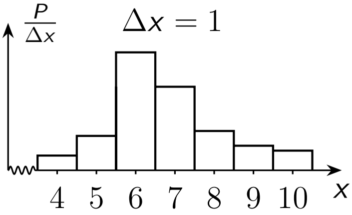
目录

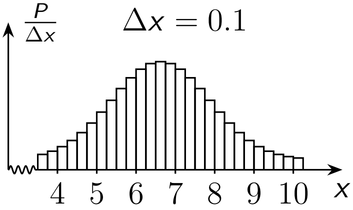
连续型随机变量

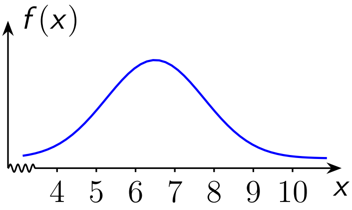
- 如果随机变量的可能取值充满某个实数区间，包含无穷不可列个实数，此时不能用分布律刻画这类**连续型随机变量**。

例 (新生儿体重)

随机变量 X 表示新生儿体重。统计大量新生儿体重，得到体重频率直方图，图中所有矩形面积之和为 1。当 $\Delta x \to 0$，体重频率图趋于光滑曲线 $f(x)$。此时 X 取值充满整个区间，成为**连续型随机变量**，$f(x)$ 成为**概率密度函数**。







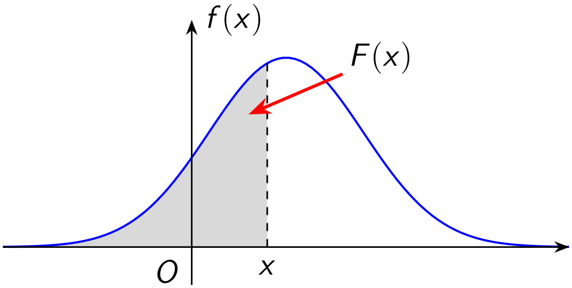
连续型随机变量

定义 (连续型随机变量及其概率密度)

设随机变量 X 的分布函数为 $F(x)$，如果实数轴上的一个非负可积函数 $f(x)$ 使得对任意实数 x 有

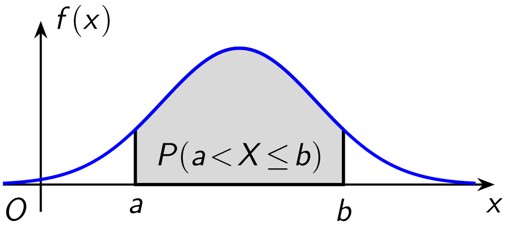
$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

则称 X 是**连续型随机变量**，$f(x)$ 是 X 的**概率密度函数**。



概率密度函数 $f(x)$ 的性质

- $f(x) \geq 0$;
- $\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$;
- 对任意实数 a, b 且 $a \leq b$，有
$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(t)\,\mathrm{d}t.$$
- 若 $f(x)$ 在 x 点处连续，则有 $F'(x) = f(x)$.



连续型随机变量的性质

- 连续型随机变量的分布函数总是连续的，因为对任意点 x，分布函数的增量总为

$$F(x+\Delta x) - F(x) = \int_{x}^{x+\Delta x} f(t)\,\mathrm{d}t \to 0 \quad \text{as} \quad \Delta x \to 0$$

- 对于连续型随机变量 X 有 $P(X=x)=0, \forall x \in \mathbb{R}$，因为

$$P(X=x) = \int_{x}^{x} f(t)\,\mathrm{d}t = 0$$

- 由于连续型随机变量 X 仅取一点的概率恒为 0，因此事件 $\{a \le X \le b\}$ 中去掉 a 或去掉 b 不影响其概率，即

$$P(a \le x \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$$

- 概率密度函数不唯一，改变 $f(x)$ 的“个别值”不影响积分值。

举例：连续型随机变量

例 (概率密度函数的性质)

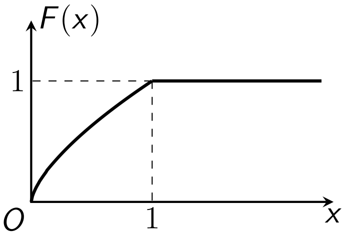
设 $f(x) = \begin{cases} kx^{-1/3}, & 0 < x < 1 \\ 0, & \text{其他} \end{cases}$ 为一概率密度函数，求：

(a) 常数 k；(b) 分布函数；(c) $P(1/4 < X \leq 2)$.

- $k = 2/3$
- 分布函数

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x^{2/3}, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

- $P(1/4 < X \leq 2) = 1 - \frac{1}{2\sqrt[3]{2}}$



反之，对分布函数逐段微分，也可得到原概率密度函数。

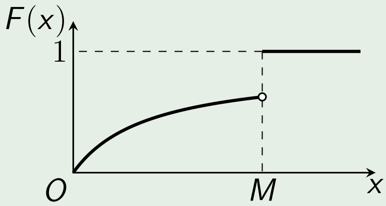
举例：连续型随机变量

- 并非所有随机变量非离散型即连续型，也有两者皆不是的。
- 例如下例的随机变量，分布函数既不是阶梯，又不连续，不属于本课程研究范围。

例 (电压问题)

前述电压问题，若电压表的量程到某 M 为止，设 X 为表测量值，则其分布函数应为

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{x}{1+x}, & 0 < x < M \\ 1, & x \ge M \end{cases}$$



X 既不能用概率密度函数，也不能用分布律刻画，但能用分布函数刻画。

几种常用的连续型随机变量概率分布

- 正态分布
- 均匀分布
- 指数分布

正态分布

- 正态分布是一种重要的分布：它所揭示的统计规律“两头小，中间大”在自然界与人类社会普遍存在：
  - 测量值与实际值的误差
  - 分子热运动时每个分子的运动速率
  - 大气中污染物的浓度

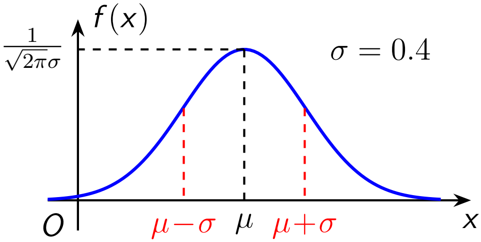
定义 (正态分布, Normal Distribution, Gaussian Distribution)

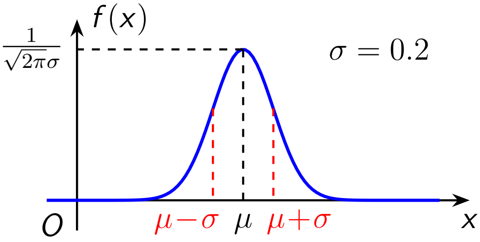
连续型随机变量 X 如果有如下形式的概率密度

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \mu \in \mathbb{R}, \sigma > 0$$

则称 X 服从参数为 μ, σ 的正态分布，也称作高斯分布，记为 $X \sim N(\mu, \sigma^2)$。

正态分布的性质





- 关于 μ 对称，在 $x = \mu$ 处达到极大值 $\frac{1}{\sqrt{2\pi}\sigma}$，越远离 μ，概率密度值越小。
- $x = \mu \pm \sigma$ 是 $f(x)$ 的拐点。

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \bigg|_{x=\mu\pm\sigma} = \frac{1}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left(\frac{(x-\mu)^2}{\sigma^2} - 1 \right)^2 \bigg|_{x=\mu\pm\sigma} = 0$$

正态分布

- 它是否是一个概率密度？需要验证 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$。
- 因为 μ 只对 $f(x)$ 左右平移，不影响积分大小，故令 $\mu = 0$。
- 令 $I = \displaystyle\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x$，则

$$I^2 = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2\sigma^2}}\,\mathrm{d}y = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2\sigma^2}}\,\mathrm{d}x\,\mathrm{d}y$$

令 $x = \rho\cos\theta, y = \rho\sin\theta$，转化为极坐标，得

$$I^2 = \int_0^{2\pi}\int_0^{+\infty} e^{-\frac{\rho^2}{2\sigma^2}}\rho\,\mathrm{d}\rho\,\mathrm{d}\theta = 2\pi\int_0^{+\infty} e^{-\frac{\rho^2}{2\sigma^2}}\,\mathrm{d}\frac{\rho^2}{2} = 2\pi\sigma^2$$

- 所以 $I = \sqrt{2\pi}\sigma$，从而 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$

举例：正态分布

例（男性身高）

设成年男子身高 $X \sim N(1.7, 0.1^2)$，随机观察一男子，求其身高超过 1.8 的概率。

$$P(X > 1.8) = \int_{1.8}^{+\infty} \frac{1}{\sqrt{2\pi} \times 0.1} e^{-\frac{(x-1.7)^2}{2\times 0.1^2}} \, \mathrm{d}x = ?$$

一般地，形如 $\int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \, \mathrm{d}t$ 的积分得不到解析表达式，只能作数值积分，对每个 x 制表，列出结果。

若 μ, σ 也变化怎么办？

举例：正态分布

考虑变换 $Z=\frac{X-\mu}{\sigma}$，称为 X 的标准化。

定理（正态分布标准化）

设 $X\sim N(\mu,\sigma^2)$，令 $Z=\frac{X-\mu}{\sigma}$，则 $Z\sim N(0,1)$。

证明.

令 Z 的分布函数和概率密度函数分别为 $\Phi(x)$ 和 $\phi(x)$，则

$$\Phi(x)=P(Z<x)=P(X<\sigma x+\mu)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\sigma x+\mu}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,\mathrm{d}t$$

令 $\frac{t-\mu}{\sigma}=u$，则

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-u^2/2}\,\mathrm{d}u,\quad \phi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

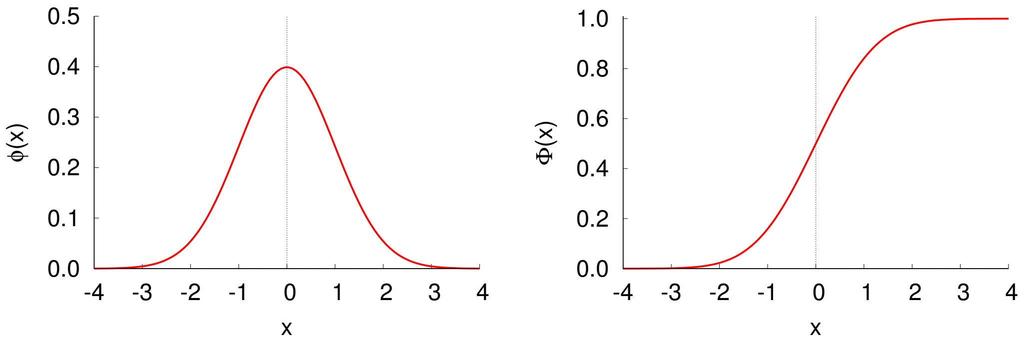
所以 $Z\sim N(0,1)$. $\square$

举例：正态分布

故只需对 $\Phi(x)$ 制表，列出结果 $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} \, \mathrm{d}u$.

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.0251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767

标准正态分布



- 称 $N(0,1)$ 为**标准正态分布**，其概率密度函数 $\phi(x)$ 和分布函数 $\Phi(x)$ 形状如上图。
- 实际问题中，常有先给出 $\Phi(x_\alpha)=\alpha$，求反函数 $\Phi^{-1}(\alpha)=x_\alpha$，称 x_α 为 α **分位数**（quantile）。

举例：正态分布

例（男性身高，查表求解法）

令 $Z = \frac{X-1.7}{0.1}$，得 $X = 0.1Z + 1.7$，故

$$P(X > 1.8) = P(0.1Z + 1.7 > 1.8) = P(Z > 1) = 1 - \Phi(1)$$

查表得 $1 - \Phi(1) = 1 - 0.8413 = 0.1587$。

例（男性身高，续）

城市公交车门高度设计中，受结构限制不能太高。但为了保证 95% 的男性都能顺畅登车，应保证车门的高度不低于多少？

设门的最低高度为 x 时可使 $P(X < x) \geq 0.95$，代入 $X = 0.1Z + 1.7$ 得

$$P(0.1Z + 1.7 < x) = P(Z < 10x - 17) = \Phi(10x - 17) \geq 0.95$$

查表得 $\Phi(1.65) = 0.9505$，故 $x = 1.865$.

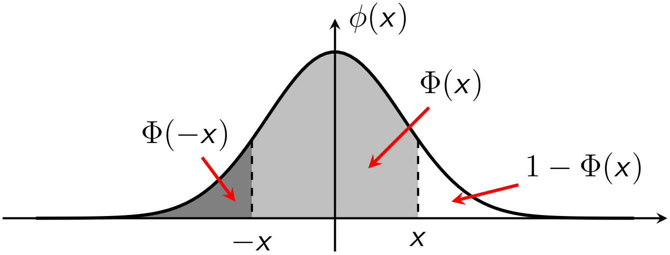
正态分布与“3σ 法则”

- $\Phi(x)$ 表中，通常只列到 $x=3.9$，为什么？

$$P(|X-\mu|\le\sigma)=0.6827 \quad P(|X-\mu|\le 2\sigma)=0.9545$$

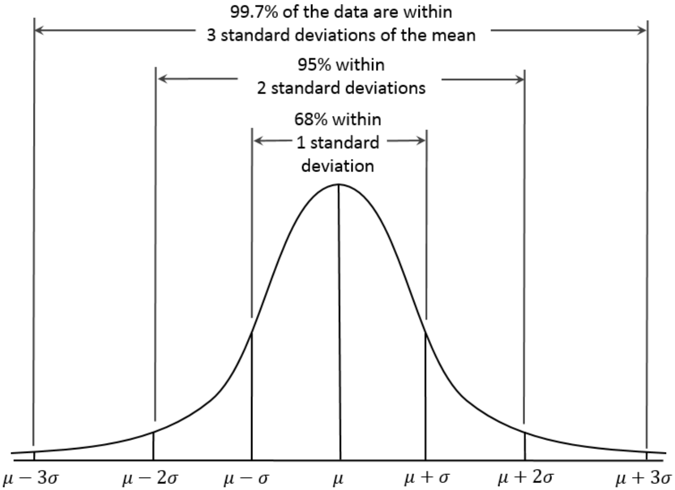
$$P(|X-\mu|\le 3\sigma)=0.9973$$

- 一个标准正态分布随机变量的取值几乎全部集中在区间 $(\mu-3\sigma,\mu+3\sigma)$ 内，超出这个范围的可能性不到 0.3%。
- 正态分布表中没有给出 $x<0$ 的情况，为什么？



正态分布与“3σ 法则”

- 工程上采用的 3σ 法则：要求正态变量（例如产品精度指标等）的取值位于 $\mu \pm 3\sigma$ 之间。



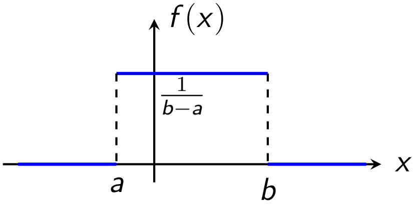
均匀分布

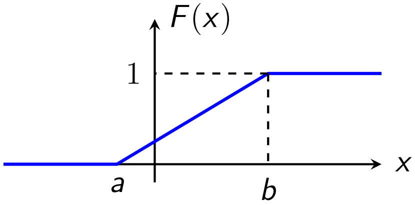
定义 (均匀分布，Uniform Distribution)

设随机变量 X 具有如下形式的概率密度函数

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

则称 X 服从区间 $[a, b]$ 上的**均匀分布**，记为 $X \sim U(a, b)$。





在 $[a, b]$ 内任意小区间取值的概率只与小区间的长度有关，与其位置无关。

指数分布

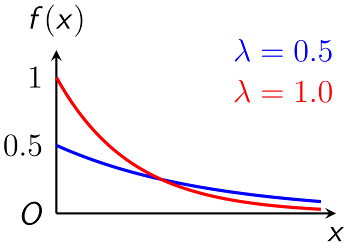
定义 (指数分布, Exponential Distribution)

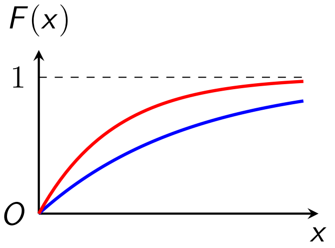
若随机变量 X 具有如下形式的概率密度函数

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

称 X 服从参数为 λ 的**指数分布**，记为 $X \sim \exp(\lambda)$。其分布函数为

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





指数分布

- 指数分布通常用来描述生命周期（生物、产品等）。
- 无记忆性：设 $X \sim \exp(\lambda)$，则对于任意 $t, s > 0$，有
$$P(X > s + t | X > s) = P(X > t).$$
这是因为
$$P(X > s + t | X > s) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t).$$

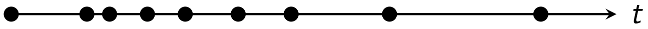
指数分布与泊松流

定义（泊松流）

对于源源不断依次随机而来的质点（粒子、顾客等），考虑其数量，如果满足

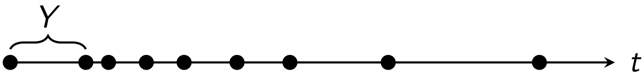
- 在任意时间 t 内质点数 $X_t \sim P(\lambda t)$；
- 在不同时段中的质点数相互独立；

就称 $\{X_t : t > 0\}$ 是个**泊松流**。



问题：泊松流中质点到达间隔时间服从什么分布？

指数分布与泊松流



第一个质点出来的时刻为 0，下一质点出来时刻为 Y，对于任意的 $t > 0$，考虑

$$P(Y > t) = P\{[0, t]\text{时间段内有 } 0 \text{ 个质点出现}\} = e^{-\lambda t}$$

故 Y 的分布函数为

$$F_Y(t) = P(Y \le t) = 1 - e^{-\lambda t}$$

即 $Y \sim \exp(\lambda)$。

小结

1. 随机变量与分布函数
2. 离散型随机变量
3. 连续型随机变量