概率论与数理统计

何志坚

华南理工大学数学学院

hezhijian@scut.edu.cn

第二章：一维随机变量及其分布

目录

为什么引进随机变量?

- 一些随机现象的样本点用数量表示，如骰子的点数、产品的寿命、测量误差
- 但有些随机现象的样本点本身不是数，如$\Omega = \{合格品, 不合格品\}$
- 随机变量将所有的样本点（基本事件）映射到实数域$\mathbb{R}$，这样便于定量化地分析和研究随机现象。

随机变量的概念

定义：考虑概率空间$(\Omega, \mathcal{F}, P)$, 称映射$X : \Omega \to \mathbb{R}$为**随机变量**，如果对任意$x \in \mathbb{R}$，有

$$\{\omega | X(\omega) \leq x\} \in \mathcal{F}.$$

也就是$\{\omega | X(\omega) \leq x\}$为一个随机事件。

注:

- 一般$X(\omega)$简记为X，$\{\omega | X(\omega) \leq x\}$简记为$\{X \leq x\}$
- X取值为有限或可列则称为**离散随机变量**
- X取值充满某个区间(a, b)则称为**连续随机变量**
- 对于同个概率空间可以构造多种随机变量
- 证明：$\{X > x\}, \{X \geq x\}, \{X < x\}, \{X = x\}$都为事件。

思考：满足什么条件的集合$A \subset \mathbb{R}$使得$\{X \in A\} \in \mathcal{F}$成立?

分布函数的定义

定义：设X是一个随机变量，对任意实数x，称

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

为随机变量X的分布函数(cumulative distribution function, CDF)。为了强调与X的对应关系，有时写成$F_X(x)$。

注：分布函数将概率和函数建立起联系，由此可计算与随机变量X有关事件的概率。

分布函数的性质

- 单调性: 分布函数$F(x)$是$\mathbb{R}$上单调非减函数，即对任意$x_1 < x_2$, $F(x_1) \leq F(x_2)$.
- 有界性: 对任意$x \in \mathbb{R}$, $F(x) \in [0, 1]$, 且

$$F(\infty) = \lim_{x \to \infty} F(x) = 1$$

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0.$$

- 右连续性: 对任意$x_0 \in \mathbb{R}$,

$$F(x_0) = \lim_{x \to x_0^+} F(x) = F(x_0 + 0).$$

注：满足这三个基本性质成为判别某个函数是否成为分布函数的充要条件。

分布函数的性质

对任意$a < b, x \in \mathbb{R}$,

$$P(X = x) = F(x) - F(x-0)$$
$$P(X > x) = 1 - F(x)$$
$$P(X \geq x) = 1 - F(x-0)$$
$$P(X < x) = F(x-0)$$
$$P(a < X < b) = F(b-0) - F(a)$$
$$P(a \leq X \leq b) = F(b) - F(a-0)$$
$$P(a \leq X < b) = F(b-0) - F(b-0)$$
$$P(a < X \leq b) = F(b) - F(a)$$

若$F(x)$在a与b处连续，则$F(a-0) = F(a)$, $F(b-0) = F(b)$

目录

一维离散型随机变量

定义：设离散型随机变量X的所有可能取值为$x_1, x_2, \cdots, x_n, \cdots$. 则称

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

为X的概率分布列（简称分布列），也可写成

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$\cdots$

或者写成

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n & \cdots \\ p(x_1) & p(x_2) & \cdots & p(x_n) & \cdots \end{pmatrix}$$

注意：分布列必须考虑所有的情况，即满足

- 非负性：对于所有的i, $p(x_i) \geq 0$
- 正则性：$\sum p(x_i) = 1$

离散随机变量的分布函数

由分布列不难求出离散随机变量X的分布函数

$$F(x) = \sum_{x_i \le x} p(x_i).$$

该函数为右连续的阶梯函数。

例：设离散随机变量X的分布列为

X	-1	2	3
P	0.25	0.5	0.25

试求$P(X \le 0.5), P(1.5 < X \le 2.5)$，并写出$X$的分布函数。

解：

$$P(X \le 0.5) = P(X = -1) = 0.25$$

$$P(1.5 < X \le 2.5) = P(X = 2) = 0.5$$

$$F(x) = \begin{cases} 0, & x < -1, \\ 0.25, & -1 \le x < 2, \\ 0.75, & 2 \le x < 3, \\ 1, & x \ge 3 \end{cases}$$

常见的离散型分布

- 单点分布$P(X=c)=1$，唯一取值。又称退化的分布。
- 二项分布$X\sim B(n,p)$，取值范围为$1,\ldots,n$
- 泊松分布$X\sim \mathrm{Pois}(\lambda)$，取值范围为非负整数
- 几何分布$X\sim \mathrm{Geo}(p)$，取值范围为正整数
- 超几何分布$h(n,N,M)$，取值范围$0,1,\ldots,n$
- 负二项分布$Nb(r,p)$，取值范围$r,r+1,\cdots$

二项分布

定义： 记X为n重伯努利试验中事件A出现的次数，$P(A)=p$，则有

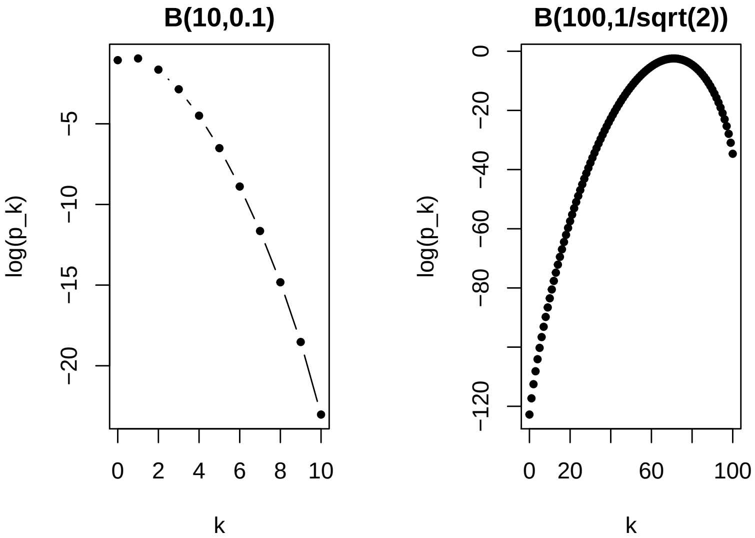
$$P(X=k)=b(k;n,p)=C_n^k p^k(1-p)^{n-k}, k=0,\ldots,n.$$

这样的分布称为二项分布，记为$X\sim B(n,p)$.

特别地，当$n=1$时，二项分布也称为两点分布，其分布列为

$$\begin{pmatrix} 0 & 1 \\ p & 1-p \end{pmatrix}.$$

二项分布


B(10,0.1)
log(p_k)
−5
−10
−15
−20
0
2
4
6
8
10
k
B(100,1/sqrt(2))
log(p_k)
0
−20
−40
−60
−80
−120
0
20
60
100
k

二项分布

定理：令$m = (n+1)p$.

$$\arg\max_{k=0,\dots,n} b(k;n,p) = \begin{cases} \lfloor m \rfloor, \text{如果}m\text{不为整数} \\ \{m, m-1\}, \text{如果}m\text{为整数} \end{cases}$$

证明：

$$r = \frac{b(k;n,p)}{b(k-1;n,p)} = \frac{C_n^k p^k q^{n-k}}{C_n^{k-1} p^{k-1} q^{n-k+1}} = \frac{(n-k+1)p}{kq} = 1 + \frac{m-k}{kq}$$

如果$k < m$，则$r > 1$，即$b(k;n,p) > b(k-1;n,p)$；如果$k > m$，$b(k;n,p) < b(k-1;n,p)$. 所以当m不为整数时，$k = \lfloor m \rfloor$时，$b(k;n,p)$取得最大值。若m为整数，则$b(m;n,p) = b(m-1;n,p)$.

思考：当k, n固定，p取何值时$b(k;n,p)$最大?

泊松分布

定义： 泊松分布是1837年由法国数学家泊松(Poisson，1781–1840)首次提出的，泊松分布的概率分布列为

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

其中参数$\lambda>0$，记为$X\sim \mathrm{Pois}(\lambda)$.

在实际问题中，有很多随机变量都近似服从泊松分布。例如:

- 在任给一段固定的时间间隔内，来到公共设施（公共汽车站、商店、电话交换台等）要求给予服务的顾客个数
- 炸弹爆炸后落在平面上某区域的碎弹片个数
- 落在显微镜片上的某种细菌个数

泊松定理——二项分布的泊松近似

定理： 设随机变量X服从二项分布$B(n,p_n)(n=1,2,\cdots)$，其中概率p_n与n有关，并且满足

$$\lim_{n\to\infty} np_n=\lambda>0,$$

则

$$\lim_{n\to\infty} P(X=k)=\lim_{n\to\infty} C_n^k p_n^k(1-p_n)^{n-k}=\frac{e^{-\lambda}\lambda^k}{k!},k=0,1,2,\cdots.$$

证明： 令$\lambda_n=np_n$,则

$$\begin{aligned} b(k;n,p_n) &= \frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k} \\ &= \frac{\lambda_n^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda_n}{n}\right)^{n-k} \\ &\to \frac{e^{-\lambda}\lambda^k}{k!}. \end{aligned}$$

其中用到，$\lim_{n\to\infty}(1-\lambda_n/n)^{n-k}=\lim_{n\to\infty}[(1-\lambda_n/n)^{-n/\lambda_n}]^{-\lambda_n}=e^{-\lambda}$.

泊松定理的应用

在应用中，当n很大($n \geq 10$)，p很小($p \leq 0.1$)时，可以用以下近似

$$b(k; n, p) \approx \frac{e^{-\lambda}\lambda^k}{k!}, \lambda = np.$$

例：设每次击中目标的概率为0.001，且各次射击是否中目标可看作相互没有影响，如果射击5000次，试求：（1）击中12弹的概率；（2）至少击中12弹的概率。

解：设击中次数为X,则$X \sim B(5000, 0.001)$. $\lambda = 5000 \times 0.001 = 5$.

(1)

$$P(X = 12) \approx \frac{e^{-\lambda}\lambda^{12}}{12!} = \frac{e^{-5}5^{12}}{12!} = 0.00343424(0.00342154).$$

(2)

$$P(X < 12) \approx \sum_{k=0}^{11} \frac{e^{-5}5^k}{k!} = 0.9945469.$$

则有$P(X \geq 12) \approx 0.0054531(0.0054284)$.

例

设有同类设备80台，各台工作相互独立的，发生故障的概率都是0.01，并且一台设备的故障可由一个人来处理，试求

- 一个人负责维修20台设备时，设备发生故障而不能及时维修的概率；
- 由三个人共同负责维修80台设备时，设备发生故障而不能及时维修的概率。

解： (1)设同一时刻发生故障的数目为X. $X \sim B(20, 0.01)$. 求得

$$P(X \geq 2) \approx 0.0176$$

(2)设同一时刻发生故障的数目为Y. $Y \sim B(80, 0.01)$. 求得

$$P(Y \geq 4) \approx 0.0091$$

计算结果表明，由三人共同负责维修80台，每人平均约维修27台，比一个单独维修20台更好，既节约了人力又提高了工作效率。

几何分布

定义： 在"成功"概率是p的伯努利试验中，若以X记首次出现"成功"的试验次数。则X所服从的分布便是几何分布，分布列为

$$P(X=k)=g(k;p)=(1-p)^{k-1}p, k=1,2,\cdots.$$

记为$X\sim \mathrm{Geo}(p)$.

例： 一个人要开门，他共有n把钥匙，其中仅有一把是能开此门的，现随机地从中取出一把钥匙来试开门，在试开时每一把钥匙均以$1/n$的概率被取用，问此人直到第k次试开时方才成功的概率是多少?

$$g(k;1/n)=\left(1-\frac{1}{n}\right)^{k-1}\frac{1}{n}.$$

几何分布的特征——无记忆性

定理： 设$X \sim \mathrm{Geo}(p)$, 则有对于任意正整数k, n,

$$P(X = k + n | X > k) = P(X = n).$$

注： 已知第k次还未成功，那么从第$k+1$次开始，首次成功出现在哪一次与k无关。

证明：

$$\begin{aligned} P(X = k + n | X > k) &= \frac{P(X = k + n)}{P(X > k)} \\ &= \frac{(1-p)^{k+n-1} p}{\sum_{i=k+1}^{\infty} (1-p)^{i-1} p} \\ &= \frac{(1-p)^{k+n-1} p}{(1-p)^k} \\ &= (1-p)^{n-1} p = P(X = n). \end{aligned}$$

推论： $P(X > k + n | X > k) = P(X > n)$.

超几何分布

定义：在一箱N件装的产品中混进了M件次品，今从中抽取n件($n \leq M$)，求从中查出次品的件数X的概率分布.

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k = 0, 1, \ldots, n.$$

记为$X \sim h(n, N, M)$.

当$n \ll N$, 即抽取个数n远小于总数N,每次抽取后总体的不合格率$p = M/N$改变甚微，所以可以视为不放回抽样，这是超几何分布可以用二项分布近似：

$$P(X = k) \approx b(k; n, M/N)$$

负二项分布(帕斯卡分布)

定义： 在"成功"概率是p的伯努利试验中，出现第r次成功时所作的试验次数X所服从的分布称为负二项分布，分布列为

$$P(X=k)=f(k;r,p)=C_{k-1}^{r-1}p^{r}(1-p)^{k-r},k=r,r+1,\cdots$$

记为$X\sim Nb(r,p)$.

由于$f(k;r,p)$是负指数二项式$(1/p-q/p)^{-r}$展开式中的项$(q=1-p)$，故X所服从的分布称为负二项分布。由此也可以证明

$$\begin{aligned}\sum_{k=r}^{\infty}f(k;r,p)&=p^{r}\sum_{k=r}^{\infty}C_{k-1}^{r-1}q^{k-r}=p^{r}\sum_{k=0}^{\infty}C_{r+k-1}^{r-1}q^{k}\\&=p^{r}\sum_{k=0}^{\infty}C_{-r}^{k}(-q)^{k}=p^{r}(1-q)^{-r}=1.\end{aligned}$$

注： $C_{-r}^{k}=(-r)(-r-1)\ldots(-r-k+1)/k!=(-1)^{r}C_{r+k-1}^{k}=(-1)^{r}C_{r+k-1}^{r-1}$.

几何分布与负二项分布的关系

若$X_i, i = 1, \ldots, r$独立同分布，且$X_i \sim \mathrm{Geo}(p)$. 则

$$X_1 + X_2 + \cdots + X_r \sim Nb(r, p).$$

即负二项分布随机变量可以分解成r个独立同分布的几何分布随机变量之和。

目录

一维连续型随机变量

定义：设随机变量X的分布函数为$F(x)$，若存在非负函数$f(x)$，使得对一切实数x，恒有

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

则称X为连续型随机变量，$f(x)$称为X的密度函数(Probability Distribution Function, PDF)。

密度函数的最基本性质：

- 非负性

$$f(x) \geq 0$$

- 正则性

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$$

连续型随机变量的性质

1. 连续型随机变量的分布函数是连续函数
2. $$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x)\,\mathrm{d}x$$
3. 若$f(x)$在x的邻域连续，则$f(x) = F'(x)$
4. 对于任意实数a, $P(X = a) = 0$
5. $$P(a < X \leq b) = P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b)$$

例1

设随机变量X的分布函数为：

$$F(x) = \begin{cases} A + \frac{1}{3}e^{x}, & x < 0 \\ B - \frac{1}{3}e^{-2x}, & x \geq 0 \end{cases}$$

1. 求常数A, B
2. 判断X是否为连续型随机变量
3. 求$P(-1 < X \leq 1/2)$

证明： $F(-\infty) = A = 0$，$F(\infty) = B = 1$. 因为$F(0) = 2/3 \neq F(0-) = 1/3$，所以$F(x)$在$x = 0$处不连续，也就意味着$X$不为连续型随机变量。

$$P(-1 < X \leq 1/2) = F(1/2) - F(-1) = 1 - \frac{2}{3e}$$

注：X既不是离散型也不是连续型随机变量。

例2

设随机变量X的密度函数为：

$$f(x)=\frac{a}{1+x^2},x\in(-\infty,\infty).$$

1. 试确定a的值
2. 试求X的分布函数
3. 试求$P(X^2\le 1)$

证明： 由正则性得

$$\int_{-\infty}^{\infty}\frac{a}{1+x^2}\,\mathrm{d}x=a\arctan(x)|_{-\infty}^{\infty}=a\pi=1$$

所以$a=1/\pi$.

$$F(x)=\frac{1}{\pi}\int_{-\infty}^{x}\frac{1}{1+t^2}\,\mathrm{d}t=\frac{1}{2}+\frac{1}{\pi}\arctan(x)$$

$$P(X^2\le 1)=F(1)-F(-1)=[\arctan(1)-\arctan(-1)]/\pi=0.5$$

注： 此分布为柯西(Cauchy)分布

常见的连续型分布

1. 均匀分布
2. 指数分布
3. 正态分布
4. 卡方分布，t分布，F分布(pp.104-106)
5. 伽玛分布
6. 贝塔分布
7. 对数正态分布
8. 柯西分布

1. 均匀分布

定义：设a、b为有限数，且$a < b$。如果随机变量X分布密度为

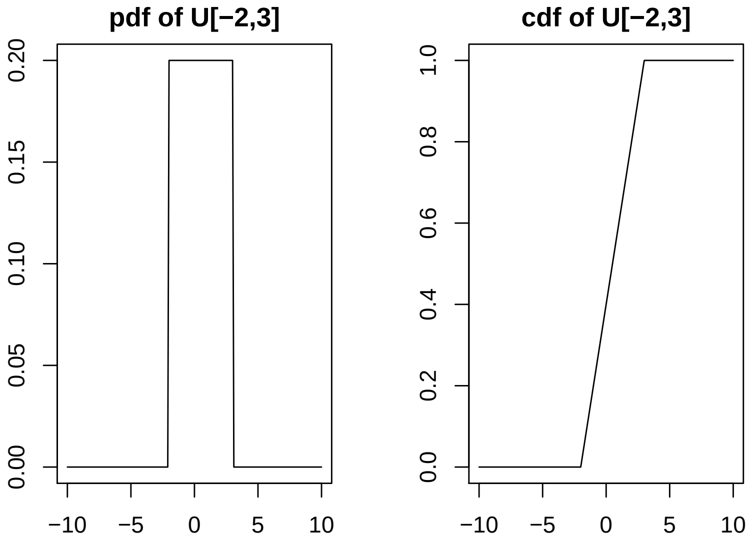
$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$$

则称X服从$[a, b]$上的均匀分布，记$X \sim U[a, b]$. 其分布函数为

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & x \in [a, b] \\ 1, & x > b. \end{cases}$$

注：$U[0, 1]$是最基本的随机变量，由此可以生成任意的分布。

均匀分布密度与分布函数示意图



例

向区间$(-1,1)$均匀地投掷一随机点，以X表示随机点的落点坐标，试求关于t的二次方程$t^2+3Xt+1=0$有实根的概率。

解： 该二次方程有实根的充要条件为$9X^2-4\geq 0$. 依题意$X\sim U(-1,1)$. 则

$$\begin{aligned}P(9X^2-4\geq 0)=P(|X|\geq 2/3)&=\int_{|x|\geq 2/3}\frac{1}{2}\mathbf{1}\{-1<x<1\}\,\mathrm{d}x\\&=\int_{-1}^{-\frac{2}{3}}\frac{1}{2}\,\mathrm{d}x+\int_{\frac{2}{3}}^{1}\frac{1}{2}\,\mathrm{d}x=\frac{1}{3}.\end{aligned}$$

2. 指数分布

定义：若随机变量X具有分布密度为

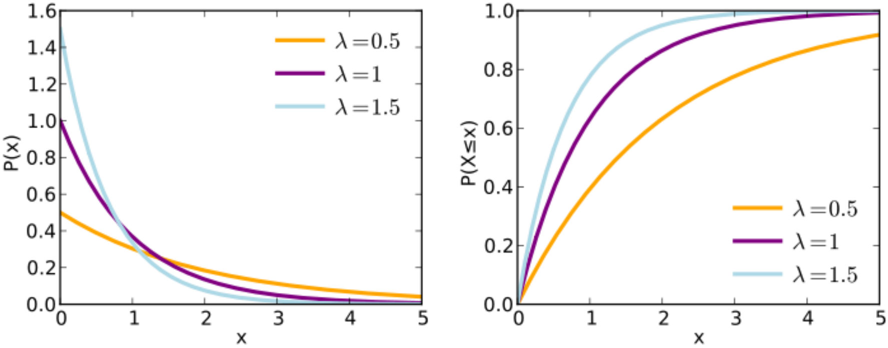
$$f(x) = \begin{cases} \lambda \exp(-\lambda x), & x \geq 0 \\ 0, & x < 0. \end{cases}$$

其中$\lambda > 0$，则称X服从指数分布，记为$X \sim Exp(\lambda)$. 其分布函数为

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

注：指数分布常被用作各种“寿命”分布，譬如电子元件的寿命，动物的寿命，电话的通话时间，随机服务系统的服务时间。在可靠性与排队论中广泛应用。

指数分布的密度与分布函数示意图



指数分布的无记忆性

定理： 如果$X \sim Exp(\lambda)$, 则对任意$s, t > 0$,有

$$P(X > s + t | X > s) = P(X > t)$$

证明：

$$\begin{aligned} P(X > s + t | X > s) &= \frac{P(X > s + t, X > s)}{P(X > s)} = \frac{P(X > s + t)}{P(X > s)} \\ &= \frac{\exp(-\lambda(s + t))}{\exp(-\lambda(s))} = \exp(-\lambda t) = P(X > t) \end{aligned}$$

注： 还可以证明指数分布是连续型随机变量中唯一的具有无记忆性的分布。

例

设到某服务窗口办事，需排队等候。若等待的时间X服从参数为1/10的指数分布随机变量（单位:分钟），若等待超过15分钟后仍未能得到接待时，他就愤然离去。试求:

- 10位顾客有2位愤然离去的概率;
- 10位顾客最多有2位愤然离去的概率;
- 10位顾客至少有2位愤然离去的概率。

解: 顾客愤然离去的概率

$$p = P(X > 15) = 1 - F(15) = \exp(-15/10) = 0.2231$$

设10位顾客离开的个数为Y, 则$Y \sim B(10, 0.2231)$. (1)$P(Y = 2) = 0.2937$;(2)$P(Y \leq 2) = 0.6735$;(3)$P(Y \geq 2) = 1 - P(Y < 2) = 0.6238$.

3. 正态分布

定义： 若随机变量X的分布密度

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

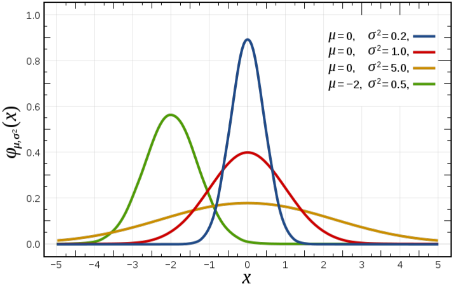
其中$\mu \in \mathbb{R}, \sigma > 0$, 则称$X$服从参数为$\mu, \sigma$的正态分布，简记为$X \sim N(\mu, \sigma^2)$。其分布函数为

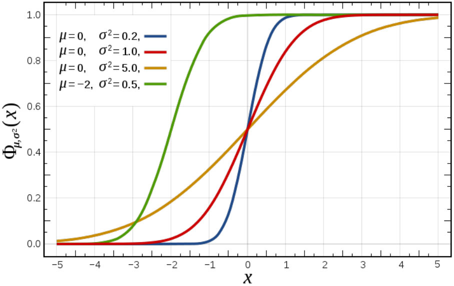
$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \mathrm{d}t.$$

特别地，$\mu = 0$, $\sigma = 1$, $N(0, 1)$称标准正态分布，其密度和分布函数常用ϕ, Φ表示。

注： 正态分布的分布函数没有解析解，可通过查表得到，常用的数学软件使用的是它的近似值。很多随机变量可以用正态分布来描述或者近似描述，譬如测量误差、产品重量、人的身高、年降雨量。

正态分布的密度与分布函数示意图





- 密度函数关于$X=\mu$对称，μ称为位置参数
- σ称为尺度参数，当μ不变时，σ越小，曲线呈现高瘦，σ越大，曲线呈现矮胖。

标准正态分布

标准正态分布$X \sim N(0,1)$, 密度函数

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

分布函数

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

- $\phi(x) = \phi(-x)$
- $\Phi(-x) = 1 - \Phi(x)$, 查表只给出分布函数在$x \geq 0$的近似值，由此可得$x < 0$时$\Phi(x)$的值
- $P(|X| \leq c) = 2\Phi(c) - 1$，其中$c \geq 0$.
- $P(|X| \geq c) = 2(1 - \Phi(c))$，其中$c \geq 0$.
- $P(|X| \leq 1) \approx 0.6827$, $P(|X| \leq 2) \approx 0.9545$, $P(|X| \leq 3) \approx 0.9973$

一般正态分布的性质

定理： 若$X \sim N(\mu, \sigma^2)$, 则

$$\frac{X-\mu}{\sigma} \sim N(0,1).$$

由此可得：

- $f(x+\mu) = f(\mu - x)$
- $F(x) = \Phi((x-\mu)/\sigma)$
- $F(\mu - x) = 1 - F(\mu + x)$
- $P(|X-\mu| \leq \sigma) \approx 0.6827$
- $P(|X-\mu| \leq 2\sigma) \approx 0.9545$
- $P(|X-\mu| \leq 3\sigma) \approx 0.9973$, 说明$X$的99.73%的值落在$(\mu - 3\sigma, \mu + 3\sigma)$内。这个性质被实际工作者称为正态分布的"$3\sigma$原则"。

例1

设$X \sim N(-1,4)$, 试求$P(-5 \le X < 1)$, $P(|X| < 1)$, $P(|X| \ge 3/2)$.

解:

$$
\begin{aligned}
P(-5 \le X < 1) &= \Phi((1-(-1))/2) - \Phi((-5-(-1))/2) = \Phi(1) - \Phi(-2) \\
&= \Phi(1) - (1 - \Phi(2)) = \Phi(1) + \Phi(2) - 1 = 0.8186.
\end{aligned}
$$

$$
P(|X| < 1) = \Phi((1-(-1))/2) - \Phi((-1-(-1))/2) = \Phi(1) - \Phi(0) = 0.3413
$$

$$
\begin{aligned}
P(|X| \ge 3/2) &= 1 - P(|X| < 3/2) = 1 - \Phi(5/4) + \Phi(-1/4) \\
&= 2 - \Phi(5/4) - \Phi(1/4) = 0.5070
\end{aligned}
$$

例2

公共汽车车门的高度是按男子与车门顶碰头的机会在0.01以下来设计的，设男子身高X服从$\mu = 170cm, \sigma = 6cm$的正态分布，即$X \sim N(170, 6^2)$，试确定车门的高度。

解： 设车门的高度为h, 依题意有

$$P(X > h) < 0.01$$

则

$$P(X \leq h) = \Phi((h - 170)/6) > 0.99$$

即

$$\frac{h - 170}{6} \geq \Phi^{-1}(0.99) = 2.33$$

所以

$$h \geq 170 + 6 \times 2.33 = 184$$

例3：股价变化幅度的估计

设某只股票的初始价格为$S_0 = 40$元，预期年收益率$\mu = 16\%$，年波动率为$\sigma = 20\%$. 在Black-Scholes模型下（1997年诺贝尔经济学奖得主），股票在每个时刻t的价格S_t为随机变量，且

$$S_t = S_0 \exp\{(\mu - \sigma^2/2)t + \sigma B_t\},$$

其中$B_t \sim N(0, t)$. 试估计六个月后这只股票的价格范围（允许出错的概率为5%）

解：令$Y_t = [\log(S_t) - (\mu - \sigma^2/2)t]/(\sigma\sqrt{t}) \sim N(0,1)$。求$y$使得

$$P(|Y_t| \le y) = 0.95$$

因为$P(|Y_t| \le y) = 2\Phi(y) - 1 = 0.95$, 所以$\Phi(y) = 0.975$, $y = \Phi^{-1}(0.975) = 1.96$
令$t = 1/2$, 由$P(\left|Y_{1/2}\right| \le 1.96) = 0.95$得到$P(32.51 \le S_{1/2} \le 56.6) = 0.95$.

目录

一维随机变量函数的分布

已知随机变量X的分布，如何求另一个随机变量$g(X)$的分布？先考虑离散的情形。

设离散型随机变量X的分布列为：

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$\cdots$

则$Y = f(X)$也为离散型随机变量，其分布列为

Y	$g(x_1)$	$g(x_2)$	$\cdots$	$g(x_n)$	$\cdots$
P	$p(x_1)$	$p(x_2)$	$\cdots$	$p(x_n)$	$\cdots$

注：当$g(x_1), g(x_2), \cdots, g(x_n), \cdots$中有某些值相等时，则把相等的值分别合并。

例

设X的分布列为：

X	-2	-1	0	1	2
P	0.15	0.2	0.2	0.2	0.25

求$Y=2X-1, Z=X^2$的分布列。

解:

$Y=2X-1$	-5	-3	-1	1	3
P	0.15	0.2	0.2	0.2	0.25

$Z=X^2$	4	1	0	1	4
P	0.15	0.2	0.2	0.2	0.25

$Z=X^2$	0	1	4
P	0.2	0.4	0.4

连续随机变量函数的分布

定理： 设X为连续随机变量，其密度函数为$p_X(x)$. $Y = g(X)$是另一个随机变量。若$y = g(x)$严格单调，其反函数$h(y)$有连续的导函数，则$Y = g(X)$的密度函数为

$$f_Y(y) = \begin{cases} f_X(h(y))\,|h'(y)|, & y \in (a, b) \\ 0, & \text{其他}, \end{cases} \tag{1}$$

其中$a = \min\{g(\infty), g(-\infty)\}, b = \max\{g(\infty), g(-\infty)\}$.

证明： 不妨设g严格单调递增。

$$F_Y(y) = P(g(X) \le y) = P(X \le h(y)) = F_X(h(y)).$$

$$f_Y(y) = F'_Y(y) = \frac{\mathrm{d}F_X(h(y))}{\mathrm{d}y} = f_X(h(y))h'(y)$$

$g(x)$为线性函数

特别地，当$g(x) = ax + b$, $a \neq 0$, 则$Y = g(X)$的密度函数为

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$

定理： 设$X \sim N(\mu, \sigma^2)$。则

- $Y = (X - \mu)/\sigma \sim N(0, 1)$
- $Z = aX + b \sim N(a\mu + b, a^2\sigma^2)$, 其中$a \neq 0$.

证明： (1)$a = 1/\sigma$, $b = -\mu/\sigma$,

$$f_Y(y) = \sigma f_X((\sigma y + \mu)) = \sigma \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\sigma y + \mu - \mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

(2)由(1)知，$X = \mu + \sigma X_0$, 其中$X_0 \sim N(0, 1)$. 所以，

$$Z = aX + b = a(\mu + \sigma X_0) + b = a\mu + b + a\sigma X_0 \sim N(a\mu + b, a^2\sigma^2)$$

$g(x)$为分布函数

定理： 若随机变量X的分布函数$F_X(x)$为严格单调递增的连续函数，其反函数F_X^{-1}存在，则$Y = F_X(X) \sim \mathbb{U}(0,1)$.

证明： 当$y \le 0$时，$P(Y \le y) = 0$; 当$y \ge 1$时，$P(Y \le y) = 1$.
当$y \in (0,1)$时，

$$P(Y \le y) = P(F_X(X) \le y) = P(X \le F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y.$$

注： 已知F_X^{-1}，则可以通过变换$X = F_X^{-1}(U)$来生成X的样本，其中，$U \sim \mathbb{U}(0,1)$。这种方式称为逆变换方法。由此可以看出，任何分布可以由均匀分布（在计算机中为伪随机数，如，Mersenne–Twister算法）变换产生。

一般的$g(x)$

一般情况下$Y = g(X)$的密度没有通用的表达式，通常的做法是按定义求分布函数，转化成计算以下概率：

$$F_Y(y) = P(g(X) \le y).$$

然后再求导得到密度。

例：如果$X \sim N(0,1)$，试求$Y = X^2$的密度函数。

解：

$$F_Y(y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) = 2\Phi(\sqrt{y}) - 1$$

$$f_Y(y) = F_Y'(y) = \frac{\phi(\sqrt{y})}{\sqrt{y}} = \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, y > 0.$$

第三次作业

第二章习题：4, 7, 9, 13, 16, 19, 21