第四周 常见随机变量

这一周我们介绍几种常见的随机变量。我们希望能够从各种随机变量产生的机理角度进行说明，从而使它们的性质展开更加自然，同时也能更深入地理解它们之所以常见的内在原因。本周学习的分布包括：二项分布，负二项分布，泊松分布，几何分布，指数分布，正态分布。

****

4.1 二项分布与负二项分布

伯努利（Bernoulli）试验

一个随机试验只有“成功”和“失败”两种可能的结果，其中出现“成功”的概率为 $p\left(0<p<1\right)$，则称此随机试验为一个参数为 p 的伯努利试验。

由参数为 p 的伯努利试验定义一个随机变量 X，

$$X=\begin{cases}1, & \text{伯努利试验成功}\\ 0, & \text{否则}\end{cases}$$

则称 X 是参数为 p 的伯努利随机变量，或称 X 服从参数为 p 的伯努利分布。

****

例 4.1.1 抛一颗均匀色子，如果出现偶数点称为试验“成功”，出现奇数点为试验“失败”，则随机变量

$$X=\begin{cases}1, & \text{抛出的点数为偶数},\\ 0, & \text{抛出的点数为奇数}.\end{cases}$$

是一个参数为 $p=\dfrac{1}{2}$ 的伯努利随机变量。

****

二项分布

将参数为 p 的伯努利试验独立地重复 n 次，定义随机变量 X 为试验成功的次数，则 X 的

分布律为：

$$\begin{pmatrix} 0 & 1 & 2 & \cdots & k & \cdots & n \\ p_0 & p_1 & p_2 & \cdots & p_k & \cdots & p_n \end{pmatrix},$$

其中 $p_k = P\left(X = k\right) = C_n^k\ p^k\left(1-p\right)^{n-k}$，$k = 0,1,\cdots,n$。

此分布即称为二项分布，记为 $X \sim B\left(n,p\right)$，也称 X 服从参数为 $\left(n,p\right)$ 的二项分布。

利用二项式定理可验证： $\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} C_n^k p^k\left(1-p\right)^{n-k} = \left[p+\left(1-p\right)\right]^n = 1$，

****

例 4.1.2 甲、乙两棋手约定进行 10 局比赛，每局棋甲获胜的概率是 0.6，乙获胜的概率为 0.4。如果各局比赛独立进行，试问甲获胜、战平和失败的概率？

X 表示甲获胜的局数，则 $X \sim b(10,0.6)$

$$P\left(\text{甲胜}\right) = P\left(X > 5\right) = \sum_{k=6}^{10} C_{10}^k 0.6^k 0.4^{10-k} = 0.6330,$$

$$P\left(\text{乙胜}\right) = P\left(X < 5\right) = \sum_{k=0}^{4} C_{10}^k 0.6^k 0.4^{10-k} = 0.1663,$$

$$P\left(\text{战平}\right) = P\left(X = 5\right) = C_{10}^5 0.6^5 0.4^5 = 0.2007。$$

****

例 4.1.3 一个通讯系统由 n 个部件组成，每个部件独立工作且能正常运行的概率均为 p，如果构成系统的部件中至少有一半以上能正常运行，则称系统是“有效”的。试问当 p 取何值时，由 5 个部件组成的系统要比由 3 个部件组成的系统更有效？

解　设 n 个部件能正常运行的数目为随机变量 X_n，则 $X_n \sim B\left(n,p\right)$

由 5 个部件组成的系统是“有效”的概率为： $P\left(X_5 > 2\right)$

$$P\left(X_5 > 2\right) = P\left(X_5 = 3\right) + P\left(X_5 = 4\right) + P\left(X_5 = 5\right) = C_5^3 p^3(1-p)^2 + C_5^4 p^4(1-p) + C_5^5 p^5$$

由 3 个部件组成的系统是“有效”的概率为： $P\left(X_3 > 1\right)$

$$P\left(X_3>1\right)=P\left(X_3=2\right)+P\left(X_3=3\right)=C_3^2 p^2(1-p)+C_3^3 p^3$$

当$C_5^3 p^3(1-p)^2+C_5^4 p^4(1-p)+C_5^5 p^5 > C_3^2 p^2(1-p)+C_3^3 p^3$时，

由 5 个部件组成的系统要比由 3 个部件组成的系统更有效。

整理后得，p应满足$3(1-p)^2(2p-1)>0$，

即当$p>\dfrac{1}{2}$时，5 个部件组成的系统要更有效些。

****

例 4.1.4 某人将一枚均匀的硬币随机抛了 10 次，已知有 6 次抛出正面，问他是在前 6 次抛出正面的概率？

解 设随机变量X为 10 次抛掷中正面出现的次数，则$X\sim B\left(10,\dfrac{1}{2}\right)$。

记事件A为“前 6 次抛出正面且后 4 次抛出反面”，由题意，要求的概率为

$$P(A\mid X=6)=\frac{P(A\cap\{X=6\})}{P(X=6)}=\frac{P(A)}{P(X=6)}=\frac{(\frac{1}{2})^{10}}{C_{10}^6(\frac{1}{2})^6(\frac{1}{2})^4}=\frac{1}{C_{10}^6}\approx 0.0048$$

****

负二项分布

连续不断且独立地重复进行一个参数为p的伯努利试验，记X为第r次“成功”出现时所需的试验次数，则事件$\{X=k\}$等价于{第k次试验“成功”且前$k-1$次试验中恰好“成功”$r-1$次}，故由试验的独立性以及二项分布的性质可得

$$P(X=k)=p\cdot P\left(B\left(k-1,p\right)=r-1\right)=pC_{k-1}^{r-1}p^{r-1}q^{k-1-(r-1)}=C_{k-1}^{r-1}p^r q^{k-r},$$

其中$q=1-p$，$k=r,r+1,\cdots$.

若随机变量X的分布律为

$$P(X=k)=C_{k-1}^{r-1}p^r q^{k-r},\quad q=1-p,\ k=r,r+1,\cdots.$$

则称X服从参数为r,p的负二项分布，记为$X\sim NB(r,p)$.

****

例 4.1.5　甲、乙两人进行比赛，直到某一人先赢到 5 局为止，假设每局比赛独立，且每局甲胜的概率为 0.58，乙胜的概率为 0.42。求

（1） 比赛在第 7 局结束的概率？ （2） 在第 7 局结束的条件下，获胜方为甲的概率？

解　设 X 为甲赢 5 局时所需的比赛局数，Y 为乙赢 5 局时所需的比赛局数，则

$$X \sim NB(5, 0.58);\quad Y \sim NB(5, 0.42)$$

设事件 A 为{甲最终获的比赛胜利} ，事件 B 为{比赛在第 7 局结束}，则

（1） $P(B) = P(X=7) + P(Y=7)$

$$= C_6^4 (0.58)^5 (0.42)^2 + C_6^4 (0.42)^5 (0.58)^2 = 0.17 + 0.066 = 0.24,$$

（2） $P(A \mid B) = \dfrac{P(AB)}{P(B)} = \dfrac{P(X=7)}{P(X=7)+P(Y=7)} = \dfrac{0.17}{0.24} = 0.71$。

****

负二项的名称来自广义的牛顿二项式公式

Newton 二项式公式 （$\forall a \in R$，$|t| < 1$）

$$(1+t)^a = \sum_{k=0}^{\infty} C_a^k t^k = 1 + C_a^1 t + C_a^2 t^2 + \cdots,$$

C_a^k 是整数组合数的推广，$C_a^k = \dfrac{a \cdot (a-1) \cdots (a-k+1)}{k!}$。

整理负二项分布随机变量 $X \sim Nb(r, p)$ 的分布律

$$P(X = r+k) = C_{r+k-1}^{r-1} (1-p)^k p^r = C_{r+k-1}^{r-1} q^k p^r = C_{-r}^k p^r (-q)^k \quad q = 1-p,\ k = 0,1,2,\cdots$$

$$\sum_{k=0}^{\infty} P(X = r+k) = p^r \sum_{k=0}^{\infty} C_{-r}^k (-q)^k = p^r (1-q)^{-r} = 1$$

****

4.2. 泊松（Poisson）分布

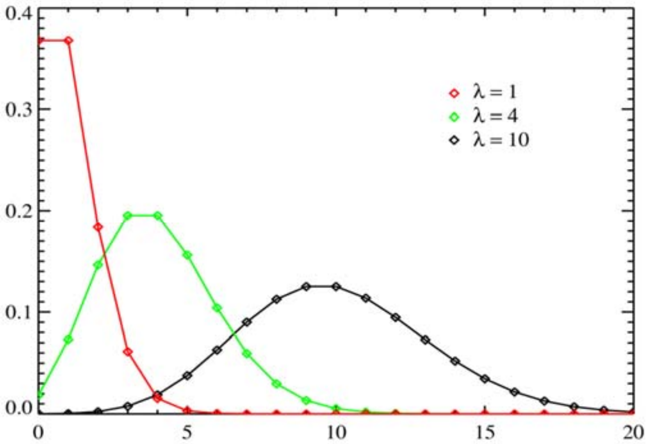
泊松分布（刻画稀有事件的发生规律）

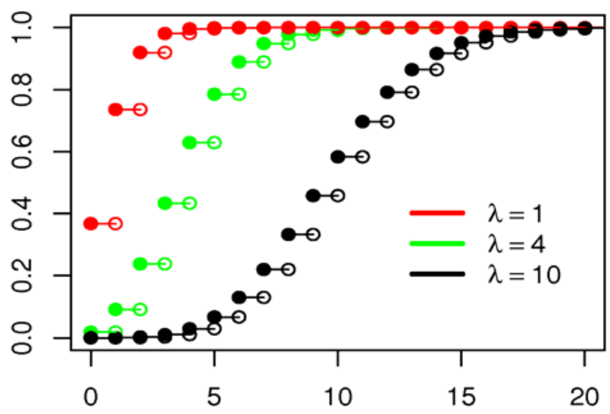
对给定的常数 $\lambda > 0$，如果随机变量 X 的分布律为

$\begin{pmatrix} 0 & 1 & 2 & \cdots \\ p_0 & p_1 & p_2 & \cdots \end{pmatrix}$， 其中 $p_k = P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}$， $k=0,1,2,\cdots$

则称随机变量 X 服从参数为 λ 的泊松分布，记为 $X \sim P(\lambda)$

验证：$\forall k \in Z^+, p_k > 0$， $\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$





****

泊松分布与二项分布的关系

考虑二项分布 $B(n,p)$，当 p 很小 n 很大时，$B(n,p)$ 与 $P(np)$ 非常接近，可相互近似

若 $X \sim B(n,p)$，$Y \sim P(np)$， 则 $P(X=k) \approx P(Y=k)$

令 $\lambda = np$， 则 $\frac{B_k(n,p)}{B_{k-1}(n,p)} = \frac{\lambda-(k-1)p}{k(1-p)} \approx \frac{\lambda}{k}$，

同时 $P(X=0) = B_0(n,p) = (1-p)^n = \left(1-\frac{\lambda}{n}\right)^n = \left(1+\left(-\frac{\lambda}{n}\right)\right)^{\left(-\frac{n}{\lambda}\right)\cdot(-\lambda)} \approx e^{-\lambda}$，所以有

$B_1(n,p) \approx \lambda B_0(n,p) \approx \lambda e^{-\lambda}$，$B_2(n,p) \approx \frac{\lambda}{2}B_1(n,p) \approx \frac{\lambda^2}{2}e^{-\lambda}, \cdots$， $B_k(n,p) \approx \frac{\lambda^k}{k!}e^{-\lambda}, \cdots$

一个业余选手射箭射中靶心的概率是 0.0001，重复 30000 次，射中靶心 3 次的概率大约为 $B_3(30000, 0.0001) \approx P_3(3) = e^{-3}\frac{3^3}{3!} = 0.224$。

泊松分布刻画了小概率事件，或是说稀有事件，多次重复时，其发生概率的规律。

****

泊松分布与二项分布严格意义下的相互关系。

泊松定理 设 $X_n \sim B(n, p_n)$，其中 p_n 与 n 有关，且满足极限关系 $\lim\limits_{n\to+\infty} np_n = \lambda$，$\lambda$ 是一个与 n 无关的常数。则对任意固定的非负整数 k，$\lim\limits_{n\to+\infty} P(X_n = k) = \dfrac{\lambda^k}{k!}e^{-\lambda}$。

证明： $$P(X_n = k) = C_n^k p_n{}^k (1-p_n)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} p_n{}^k (1-p_n)^{n-k}$$

$$= \frac{(np_n)^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)(1-p_n)^{n-k}$$

令 $np_n = \lambda_n$，则 $\lim\limits_{n\to+\infty}(1-p_n)^{n-k} = \lim\limits_{n\to+\infty}\left(1-\dfrac{\lambda_n}{n}\right)^{n-k} = \lim\limits_{n\to+\infty}\left(1-\dfrac{\lambda_n}{n}\right)^{n} = e^{-\lambda}$，

又 $\lim\limits_{n\to+\infty}\left(1-\dfrac{1}{n}\right)\left(1-\dfrac{2}{n}\right)\cdots\left(1-\dfrac{k-1}{n}\right) = 1$，所以 $\lim\limits_{n\to+\infty} C_n^k p_n{}^k (1-p_n)^{n-k} = \dfrac{\lambda^k}{k!}e^{-\lambda}$。

****

例 4.2.1 卢瑟福和盖革在 1910 年观察了放射性物质放出 α 粒子的个数的情况，共观察 $N = 2608$ 次，每次观察间隔 7.5 秒，记录到达指定区域的 α 粒子数，共记录下 10094 个粒子，N_k 表示恰好记录到 k 个 α 粒子的观察次数。

k	0	1	2	3	4	5	6	7	8	9	≥ 10
N_k	57	203	383	525	532	408	273	139	45	27	16
$N \cdot p_k(3.87)$	54	211	407	525	508	394	254	140	68	29	17

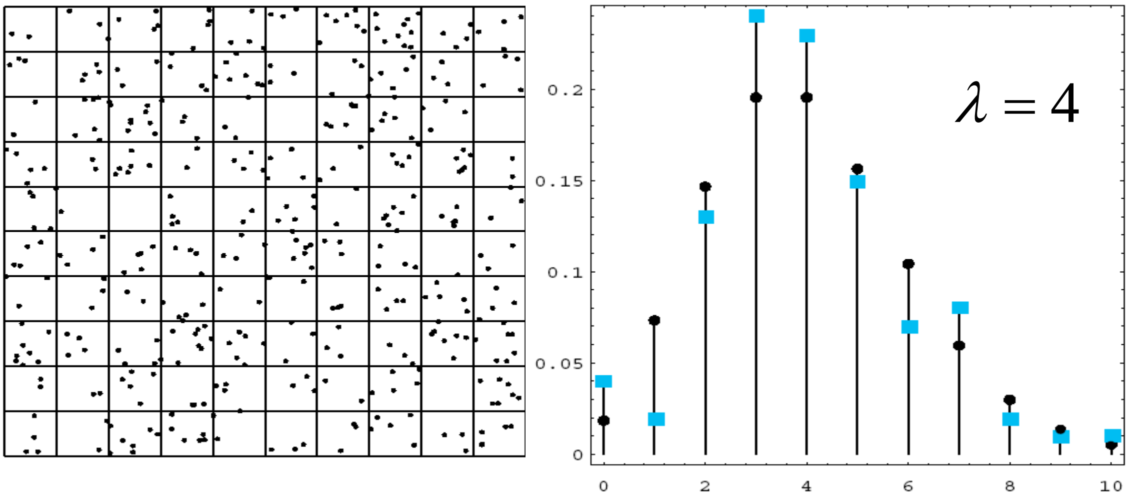
设 X 为 1 次观察中到达的粒子数，则 $X \sim B\left(10094, \dfrac{1}{2608}\right)$，　$10094 \times \dfrac{1}{2608} \approx 3.87$

X 近似服从 $P(3.87)$。

例 4.2.2 伦敦飞弹。二战时伦敦遭到很多次炸弹袭击，将整个面积分为 $N = 567$ 小块，其中发现 k 枚炸弹的小块数为 N_k，总共发现炸弹 537 枚。

k	0	1	2	3	4	≥ 5
N_k	229	211	93	35	7	1
$N \cdot p(k, 0.9323)$	226.7	211.4	98.6	31.6	7.1	1.6

设 X 为 1 个小块区域的炸弹个数，则 $X \sim B\left(537, \frac{1}{567}\right) \sim P\left(\frac{537}{567}\right) = P(0.9323)$



图中给出我们做的一项模拟试验，将 400 个点随机地投入到一个 10 乘 10 等分的包含 100 个小正方形的正方形区域中，右侧图中横坐标表示点数，彩色点表示落入相应点数的小正方形的比例，黑色点则表示参数为 4 的泊松分布的相应概率值。

****

例 4.2.3 上海市在 1875—1955 年中间有 63 年的夏季（5 月—9 月）暴雨记录，共计 180 次。每年夏季共有 $n = 31 + 30 + 31 + 31 + 30 = 153$ 天，每次暴雨如果以一天计算，则每天发生暴雨的概率为 $p = \frac{180}{63 \times 153} = 0.0187$，其值很小，同时 $n = 153$ 较大。如果暴雨可看成服从二项分布 $X \sim B(153, 0.0187)$， $\lambda = np = \frac{180}{63} \approx 2.9$， $X \sim P(2.9)$

暴雨次数	0	1	2	3	4	5	6	7	≥ 8
实际年份数	4	8	14	19	10	4	2	1	1
理论年数$(63 \times p_k(2.9))$	3.5	10.1	14.6	14.1	10.2	6.0	2.9	1.2	0.6

1898 年一位英国学者给出了一个应用泊松分布的经典例子，今天说起来多少有些荒诞。他列出一组数据表明一年中被马踢死的骑兵人数近似服从泊松分布。这也是一个典型的小概率事件被大量重复的例子。一个骑兵在一年中或者被马踢死或者不被马踢死，假定这一小概率事件发生的机会对所有骑兵来说都是一样的，并且士兵们被踢死的机会是独立的。这一假定如果合理的，则一年中被踢死的骑兵数是一个二项分布随机变量。但是被踢死的概率 p 很小，而试验次数，即骑兵人数很大，因此泊松分布对这些数据给出一个很好的描述。泊松分布刻画了稀有事件多次重复时的发生规律。书籍的印刷错误，在给定时间内某耐用设备大批量使用时的故障次数等等都近似服从泊松分布。

****

4.3 几何分布与指数分布

几何分布

连续不断独立地重复进行一个参数为 p 的伯努利试验，若记 X 为首次出现“成功”时所需的试验次数，X 是一个离散型随机变量，取值为全体正整数。则

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

若随机变量 X 的分布律为

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

这里 $0<p<1$，则称 X 服从参数为 p 的几何分布，记为 $X\sim Ge(p)$。

****

例 4.3.1 父亲要孩子们去后院整理杂物，于是他的 3 个孩子就用每人同时抛一个硬币来决定谁去整理，他们规定，谁抛出的面与另外两人的不同就谁去整理，若三人抛出的面相同则需重抛，直到选出为止，假设硬币出现正面的概率为 p，出反面为 q，求

(1) 他们抛了不到 n 轮就能选出人的概率；

(2) 若 $p=\dfrac{1}{2}$，最少要抛多少轮，才能以 0.95 以上的概率可以选出人来。

解 （1）设投出正面的次数为随机变量Y，则$Y \sim B(3, p)$，

在一轮中，选出人选的概率为 $P(Y=2)+P(Y=1)=C_3^2 p^2 q + C_3^1 p q^2 = 3pq$.

记X为选出人时所需要抛的轮数，故$X \sim Ge(3pq)$。

因此所求概率为 $P(X<n)=1-P(X \ge n)=1-(1-3pq)^{n-1}$。

（2）求最小的n，使得$P(X \le n) > 0.95$， $P(X \le n)=1-P(X>n)$，

$P(X>n)=(1-3pq)^n=(\frac{1}{4})^n$，故

$P(X \le n)=1-P(X>n)=1-(\frac{1}{4})^n > 0.95$，解出，$n>2.16$。

故最少要抛3轮，才能以0.95以上的概率可以选出人。

****

几何分布的无记忆性：

若随机变量X服从几何分布，对任意$s>0$和$t>0$，有$P(X>s+t \mid X>s)=P(X>t)$。

$$P(X>j)=\sum_{k=j+1}^{\infty} q^{k-1} p = pq^j \sum_{k=0}^{\infty} q^k = pq^j \frac{1}{1-q} = q^j \quad (q=1-p)$$

$$P(X>s+t \mid X>s)=\frac{P(X>s+t,\, X>s)}{P(X>s)}=\frac{P(X>s+t)}{P(X>s)}$$

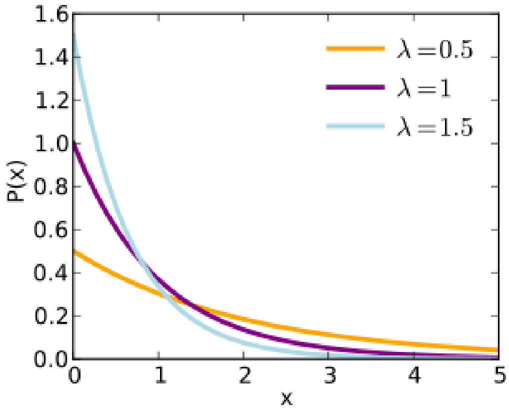
$$=\frac{q^{s+t}}{q^s}=q^t=P(X>t)$$

我们举一个例子说明几何分布的无记忆性。假设甲、乙两同学参加射击训练，每一轮射击两人同时发枪。假设两人每次射击命中的概率都是1/4，并且假设两人射击命中与否是互不影响、相互独立的。则甲、乙两人射中一次所需的射击次数均服从几何分布。如果刚刚结束的一轮射击甲命中，而乙已经连续8轮没有射中了。则在下一轮射击中，甲、乙命中的概率仍然相同；或更一般地描述，若在之后的 k 轮两人都没有命中，则在第k+1轮两人命中的概率仍然相同。这种条件概率大小与之前发生情况无关的性质，就是

几何分布的无记忆性。

****

指数分布



$X \sim Exp(\lambda),\ \lambda>0$

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

当 $x>0$ 时，$F(x)=P(X\le x)=\int_{-\infty}^{x} f(t)dx=\int_0^x \lambda e^{-\lambda t}dt=\int_0^x d\left(-e^{-\lambda t}\right)=-e^{-\lambda t}\Big|_0^x=1-e^{-\lambda x}$

$$F(x)=\begin{cases}0 & x\le 0\\ 1-e^{-\lambda x} & x>0\end{cases}$$

****

指数分布实例：假设一种电子元件的寿命 X 随机变量，对已使用了 t 小时的元件，在以后 Δt 小时内失效的概率为 $\lambda\Delta t+o(\Delta t)$，其中 λ 为不依赖 t 的常数，称为失效率，求该元件寿命的分布函数。

由题设有 $P(X\le t+\Delta t\mid X>t)=\lambda\Delta t+o(\Delta t)$， 记 $f(t)=P(X>t)$

$$f(t+\Delta t)=P(X>t+\Delta t)=P(X>t+\Delta t,X>t)=P(X>t)P(X>t+\Delta t\mid X>t)$$

$$=P(X>t)\left(1-P(X\le t+\Delta t\mid X>t)\right)=f(t)\left(1-\lambda\Delta t+o(\Delta t)\right)$$

$$\Rightarrow f(t+\Delta t)=f(t)[1-\lambda\Delta t+o(\Delta t)]\Rightarrow\frac{f(t+\Delta t)-f(t)}{\Delta t}=-\lambda f(t)+o(1)\Rightarrow\frac{df(t)}{dt}=-\lambda f(t)$$

考虑 $f(0)=1$，有 $f(t)=e^{-\lambda t}$，$F(x)=P(X\le x)=1-P(X>x)=1-e^{-\lambda x}$。

****

指数分布的无记忆性：

$$P(X\le t+\Delta t\mid X>t)=\lambda\Delta t+o(\Delta t)$$

$X \sim Exp(\lambda),\ \lambda > 0$

对任意 $s > 0$ 和 $t > 0$，有 $P(Y > s + t \mid Y > s) = P(Y > t)$。

可以验证，指数分布也满足无记忆性的条件概率关系。而且指数分布是唯一具有无记忆性的连续型分布。同时，几何分布是唯一具有无记忆性的离散型分布。

****

4.4 正态分布

正态分布随机变量记为 $X \sim N(\mu, \sigma^2)\quad (\mu \in R,\ \sigma > 0)$

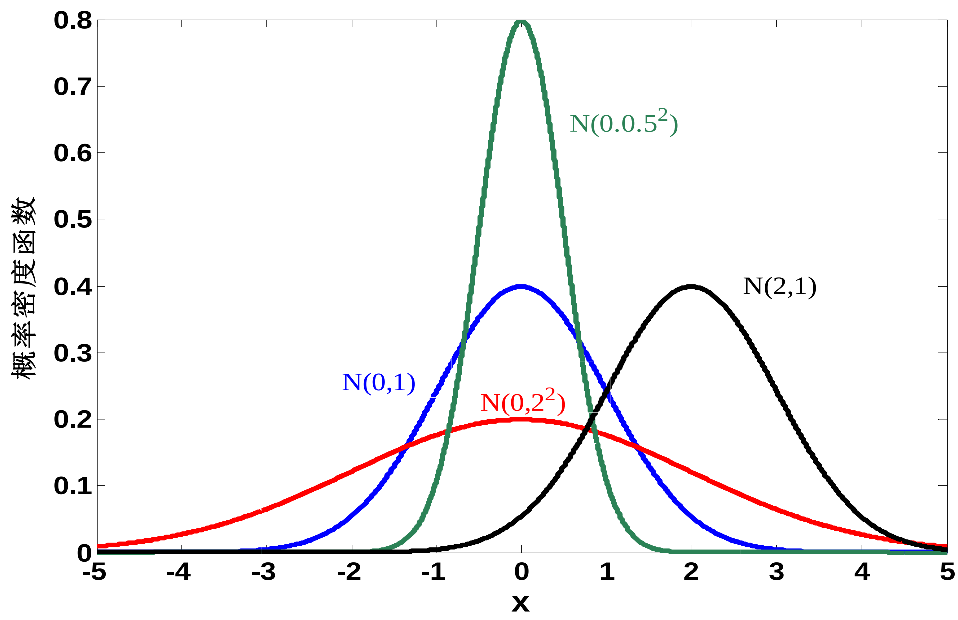
X 的密度函数为 $f(x) = \dfrac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\quad (x \in R)$.

$\mu = 0$，$\sigma^2 = 1$ 时的正态分布，即 $X \sim N(0,1)$，称为标准正态分布。

标准正态分布分布函数 $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\qquad x \in R$

标准正态分布分布函数 $\Phi(x) = \int_{-\infty}^{x} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$，$\Phi(x)$ 的值可查表得到。

若 $X \sim N(\mu, \sigma^2)$，则 $\dfrac{X-\mu}{\sigma} \sim N(0,1)$。

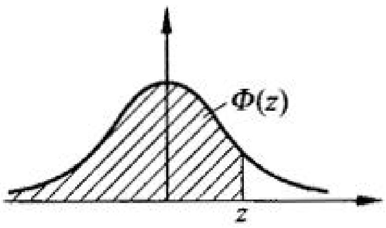


密度函数关于 $x=\mu$ 对称；当 $x=\mu$ 时，密度函数达到最大值；μ是分布的均值；

σ表示分散程度，σ越大则数据分布越散开，越小则数据分布越集中。

标准正态分布的分布函数表　$\Phi(-x)=1-\Phi(x)$，　$\Phi(0)=0.5$

$$\Phi(z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=P\{Z\leqslant z\}$$



z	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	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	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
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

续表

z	0	1	2	3	4	5	6	7	8	9
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	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	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	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

****

例 4.4.1 在现代的智力测验中，一种典型的规则是设定主体人口的平均智商为 100，而他们的智商服从正态分布 $N\left(100,15^2\right)$。求智商超出 140 和 168 的概率分别是多少？

解：设随机变量 $X\sim N\left(100,15^2\right)$，则 $\dfrac{X-100}{15}\sim N\left(0,1\right)$。 所以

$$P\left(X>140\right)=P\left(\frac{X-100}{15}>\frac{140-100}{15}\right)=P\left(\frac{X-100}{15}>2.67\right)$$

$$=1-P\left(\frac{X-100}{15}\le 2.67\right)=1-\Phi\left(2.67\right)\approx 1-0.9962=0.0038$$

$$P\left(X>168\right)=P\left(\frac{X-100}{15}>\frac{168-100}{15}\right)=P\left(\frac{X-100}{15}>4.53\right)$$

$$=1-P\left(\frac{X-100}{15}\le 4.53\right)=1-\Phi\left(4.53\right)\approx 2.95\times 10^{-6}$$

****

例 4.4.2 在现代典型的智力测验中，设定的主体人口智商服从正态分布 $X\sim N\left(\mu,\sigma^2\right)$，并将主体人口的平均智商设定为 100，即 $\mu=100$，同时还要求一半人口的智商，介于 90－110 之间，其中智商在 90－100 和 100－110 的人各占 25%。求满足要求的 σ^2 取值。

解：随机变量 $X\sim N\left(100,\sigma^2\right)$，则 $\dfrac{X-100}{\sigma}\sim N\left(0,1\right)$。

按要求，$P\left(100<X<110\right)=0.25 \;\Rightarrow\; P\left(100<X<110\right)=P\left(0<\dfrac{X-100}{\sigma}<\dfrac{10}{\sigma}\right)=0.25$

而 $P\left(0<\dfrac{X-100}{\sigma}<\dfrac{10}{\sigma}\right)=\Phi\left(\dfrac{10}{\sigma}\right)-\Phi\left(0.5\right)=\Phi\left(\dfrac{10}{\sigma}\right)-\dfrac{1}{2}=0.25 \;\Rightarrow\; \Phi\left(\dfrac{10}{\sigma}\right)=0.75$

查表得 $\Phi\left(0.675\right)\approx 0.75$，从而 $\dfrac{10}{\sigma}\approx 0.675 \;\Rightarrow\; \sigma\approx 14.8$。

实际上设定智商分布服从参数为 100，15 的平方的正态分布就是希望一半人集中在 90-110 之间。