第 2 章教学要求：

1. 理解随机变量的概念，会计算与随机变量相联系的事件的概率.

2. 理解离散型随机变量及其分布律的概念，理解连续型随机变量及其概率密度的概念，掌握常见随机变量的分布及其应用.

3. 了解分布函数的概念及性质.

4. 会求简单随机变量的函数的分布.

第 2 章　随机变量及其分布

§2.1　随机变量的概念

在随机现象中，有很大一部分问题与数值有关，如：检验产品时出现的废品数、掷骰子出现的点数、某时段在银行等候办业务的顾客人数、桌面上的细菌数、某段时间内的话务量、人的身高、测量误差、气体分子运动的速度、降水量、洪峰值等.

还有些随机现象表面看起来与数值无关，但采用数值标识的方式，仍可以使其与数值关联.如，掷一次硬币的结果为“正面向上”或“反面向上”，我们可以用数 1 表示“正面向上”，用 0 表示“反面向上”，而 n 次抛掷中出现的正面数就是这个“1”出现的次数.又如，生产的产品是“优质品”记为 2，“次品”记为 1，“废品”记为 0.此外，新生儿性别登记、航班是否正点、系统工作是否正常等也都可以作数值标识.

在以上这些例子中，每一个试验结果（样本点）ω 都对应着一个实数 $X(\omega)$，即在样本空间 Ω 上定义了一个实值函数 $X=X(\omega)$. 反过来，一个实数 X 既是对具有“属性 X”的随机事件的标识.随机事件发生的概率既是标识随机事件的实数 X 发生的概率，因而 X 又具有随机性. 这种取值随机的变量称为**随机变量**.

随机变量常用英文大写字母 $X,Y,Z,\ldots$以及希腊字母 $\xi,\eta,\zeta,\ldots$等表示， 它的取值常用英文小写字母表示.

简单地说，将随机试验结果数量化即产生随机变量，随机变量取值即表示随机事件.

如， 在掷硬币试验中，“$X=1$”即表示出现正面的事件，“$X=0$”则表示出现反面的事件；在产品抽样中，随机地抽取 n 件产品，若以“k”标识其中有 k 件废品的结果，那么“$X=0$”即表示 n 件中没有废品的事件，“$X=1$”表示 n 件中有 1 件废品的事件，. . . ，“$X=n$”表示 n 件全是废品的事件，而“$X\le1$”则表示废品数不超过 2 的事件. 又如，如果测量误差的范围是[-0.5,0.5]， 那么“$X=0$”即表示测量误差为 0 的事件， “$-0.1<X\le0.2$”则表示测量误差范围在[-0.1, 0.2]间的事件.

从随机现象的可能结果看，定义的随机变量可分为两种类型：**离散型随机变量**和**非离散型随机变量**. 前者的特点是随机变量的可能取值为可列个，如前面提到的废品数、骰子点数和顾客人数等；后者的特点是随机变量的可能取值不止可列无穷多个，如前面提到的桌面上的细菌数、人的身高、测量误差等. 非离散型随机变量的范围广，情况复杂，其中重要而又常见的为**连续型随机变量**，如人的身高、降水量、洪峰值等.

随机变量
- 离散型随机变量
- 非离散型随机变量
  - 连续型随机变量
  - 非连续型随机变量

作为试验的目的，我们不仅想知道可能出现哪些结果，更重要的是想知道这些结果出现的可能性有多

大．这即是说，对于标识试验结果的随机变量，要知道它取哪些数值，以及取这些值的概率是多少.

把描述随机变量的全部可能取值以及它取到这些值所对应的概率关系叫做随机变量的概率分布律，简称**概率分布**．概率分布既是随机变量取值的概率分布规律．概率关系可以是函数、表或图等形式，分别称为随机变量的**概率函数**、**概率分布表**和**概率分布图**.

§2.2 离散型随机变量及其分布

本节讨论离散型随机变量的概率分布及性质、概率计算与常见分布.

2.2.1 离散型随机变量的分布律

离散型随机变量的概率分布又叫**分布律**，其概率函数的一般形式为

$$f(x_i)=P\{X=x_i\}=p_i,\quad i=1,2,\ldots. \tag{2.1}$$

这里$\{X=x_i\}$表示事件“$X=x_i$”，并规定：$f(x)=P\{X=x\}=0$，如果 x 不是 X 的取值.

概率分布表的一般形式为

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

因为它是数据表，所以也常用矩阵简单地表示为

$$X\sim\begin{pmatrix} x_1 & x_2\cdots x_i\cdots \\ p_1 & p_2\cdots p_i\cdots \end{pmatrix}.$$

为便于讨论，通常将 $x_i(i=1,2,\ldots)$ 按由小到大自左向右排序.

概率分布图则是一个散点图，它能直观地反映离散型随机变量的概率分布规律.

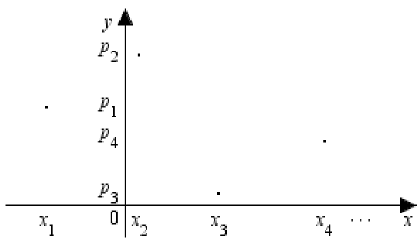


图 2.1 离散型随机变量的概率分布示意图

分布律(2.1)有如下性质：

(1) $p_i\geq 0,\ i=1,2,\ldots$;

(2) $\sum p_i=1$.

事件的概率计算：

(1) X 取单个值 x 的概率为

$$P\{X=x\}=p_i\ ，若\ x=x_i\ ；\ P\{X=x\}=0\ ，若\ x\neq x_i(i=1,2,\ldots). \tag{2.2}$$

(2) X 在区间 G 上取值的概率为

$$P\{X\in G\}=\sum_{x_i\in G}P\{X=x_i\}=\sum_{x_i\in G}p_i\ . \tag{2.3}$$

这里$\sum\limits_{x_i\in G}$表示对落在 G 中的 X 取值的项求和.

例 2. 1（P_{39}） 设离散型随机变量 X 的分布律为

X	-5	-3	-1	2	3	6
P	$0.05c$	$0.1c$	$0.15c$	$0.15c$	$0.25c$	$0.3c$

$G=(-3,2)\cup(4,7)$，求常数 c 及 $P\{X=-2\},P\{X<0\},P\{X\in G\}$.

解 由 $0.05c+0.1c+0.15c+0.15c+0.25c+0.3c=1$，得 c=1.

$P\{X=-2\}=0$，

$P\{X<0\}=P\{X=-5\}+P\{X=-3\}+P\{X=-1\}=0.3$，

$P\{X\in G\}=P\{-3<x<2\}+P\{4<x<7\}=P\{x=-1\}+P\{x=6\}=0.45$.

例 2. 2（P_{39}）

2. 2. 2 常见的离散型随机变量及其分布

1. 0–1 变量及其分布（P_{40}）

只有两个可能取值的随机变量叫**两点变量**，它所服从的分布叫**两点分布**.

两点变量中最常见的是取值为 0 和 1 的所谓 **0–1 变量**，其分布律

$$P\{X=k\}=p^k q^{1-k},\quad k=0,1\,(q=1-p,0<p<1) \tag{2.4}$$

称为参数为 p 的 **0–1 分布**.

0–1 变量常用来描述伯努利试验的结果：{X=1}表示事件 A，{X=0}表示事件 $\bar{A}$，故也称 0–1 分布为**伯努利分布**.

例如，掷一颗骰子，A 表示掷出数小于 3 的事件，如果用 0-1 变量来表示，则可令{X=1}表示事件 A，{X=0}表示掷出数大于等于 3 的事件 $\bar{A}$，事件{X=1}发生的概率即为 P{X=1}=P(A)=1/3，事件{X=0}发生的概率即为 P{X=0}=P($\bar{A}$)=2/3.

新生儿性别登记、抽样产品是否合格、投资中标与否、系统是否工作正常……都可由 0–1 变量及其分布来描述.

2. 二项变量及其分布（P_{40}）

表示 n 重伯努利试验中事件 A 发生次数的随机变量 X 的分布律

$$P\{X=k\}=C_n^k p^k q^{n-k},k=0,1,2,\cdots n(q=1-p) \tag{2.5}$$

称为参数为 n,p 的二项分布，记作 $X\sim B(n,p)$.

当 n=1 时，二项变量即为 0–1 变量.

二项变量在实际中经常遇到，有非常重要的应用.

例 2. 3（P_{41}） 某车间有 12 台车床，各台车床停开车是相互独立的，每台车床在任一时刻处于停车状态的概率均为 1/3，问：在任一时刻，车间

(1) 有 k(0≤k≤12)台车床处于停车状态的概率是多少？

(2) 有几台车床停车的可能性最大？

(3) 车床停车台数不超过一半的概率是多少？

解 将在任一时刻对一台车床的观察看作一次试验，由于观察结果只有停车或开车两种可能，且各台车床的停开车相互独立，各台车床在任一时刻处于停车状态的概率 p=1/3 固定不变，因而观察 12 台车床的试验是一个 12 重伯努利试验.

设 X 表示在某一时刻处于停车状态的车床台数，则 $X \sim B(12,1/3)$，于是

（1） $P\{X=k\}=C_{12}^{k}\left(\frac{1}{3}\right)^{k}\left(\frac{2}{3}\right)^{12-k},k=0,1,2,\cdots,12$.

详细结果见下表.

k	P	k	P	k	P
0	0.0077073	5	0.1907568	10	0.0004968
1	0.0462441	6	0.1112748	11	0.0000452
2	0.1271712	7	0.0476892	12	0.0000019
3	0.2119520	8	0.0149029		
4	0.2384660	9	0.0033118		

(2) 从上表中可见：有 4 台车床同时停车的可能性最大.

(3) $P\{X\le 6\}=\sum_{k=0}^{6}P\{X=k\}\approx 0.9336$.

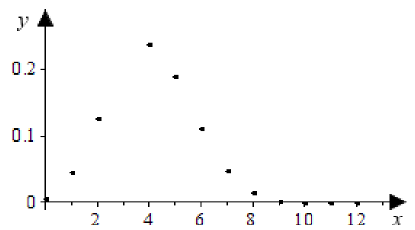


图 2. 2 例 2. 3 中的随机变量 *X* 的概率分布图

概率分布图 2. 2 显示，P{X=k}起初随 k 的增大而单调增大，直至达到最大值，随后又单调减少．这是二项分布的一般规律.

二项分布 B(n,p)具有这样的性质(证明留给读者)：P{X=k}在 k=[(n+1)p]达到最大值．当(n+1)p 为整数时，P{X=k}在 k=(n+1)p-1 与 k=(n+1)p 同时达到最大值．使概率 P{X=k}达到最大值的二项变量的取值称为**最可能数**.

例 2. 4 (P_{42})

例 2. 5 (P_{42})

例 2. 6 (P_{43})

3．泊松变量及其分布 (P_{43})

泊松分布的产生源于实际问题.当 n 很大时，给出二项变量分布律的具体结果很难，可是实际中又经常会遇到 n 很大的情形．如，牧草种子中的杂草种子数，物体表面的细菌数，放射性物质在某段时间内放射出的粒子数，5000 只纱绽在一分钟内的断头数，单位时间内电话交换台接到呼叫次数等．对于“n 很大但 p 或 q 很小”的二项变量，泊松（Poisson）给出了“二项分布的泊松逼近”定理.

定理（泊松定理）　在 n 重伯努利试验中，设一次试验中事件 A 发生的概率为 p_n(p_n 与试验次数 n 有关)．若 $\lim\limits_{n\to\infty} np_n=\lambda(\lambda>0, \lambda$ 为常数)，则对任意非负整数 k(≤n),有

$$\lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

显然 $P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!}>0,\quad k=0,1,\cdots$，且

$$P\{X=k\}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}\cdot e^{\lambda}=1.$$

因此

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},\quad k=0,1,\cdots(\lambda>0) \tag{2.6}$$

是分布律，称为参数为 λ 的**泊松分布**，记作 $X\sim P(\lambda)$．

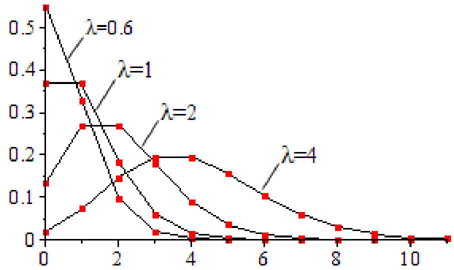


图 2. 5　泊松分布图

实际中常以泊松分布近似表示二项分布．一般当 n 很大、p 或 q 很小时，有如下近似式

$$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda},\quad k=0,1,\cdots. \tag{2.7}$$

其中 λ=np.

在实际应用中，当 n>50, p<0.1 且 np<5(或 nq<5)时，就用(2.7)式作近似计算.

泊松变量在实际应用中也很常见，前面提到的实际问题都可用泊松变量描述或近似描述．还有许多与人们日常生活密切相关的泊松变量，比如，某地一天内邮件的总数，某大医院一天内的就诊人数，某交通道口在某时段内的汽车流量，某汽车客运站在某时段的候车人数，某地区一年内发生的交通事故数等.

例 2. 7（P_{45}）　设每分钟通过某交叉路口的汽车流量 X 服从泊松分布，且已知在一分钟内无车辆通过与恰有一辆车通过的概率相同，求在一分钟内至少有两辆车通过的概率.

解　设 $X\sim P(\lambda)$.由题意知 $P\{X=0\}=P\{X=1\}$，即有

$$\frac{1}{0!}e^{-\lambda}=\frac{1}{1!}e^{-\lambda},$$

得 λ=1. 而一分钟内至少有两辆车通过的概率为

$$P\{X\ge 2\}=1-P\{X=0\}-P\{X=1\}=1-2e^{-1}\approx 0.2642.$$

例 2. 8（P_{45}）

4. 超几何变量及其分布（P_{46}）

超几何概型：有 N 件产品，其中 M 件为次品，从中抽取 n 件，则抽取出的次品数 X 的分布律

$$P\{X=k\}=\frac{C_M^k\cdot C_{N-M}^{n-k}}{C_N^n},\quad k=l_1,l_1+1,\cdots,l_2 \tag{2.8}$$

称为参数为 N,M,n 的**超几何分布**，记作 $X\sim H(N,M,n)$，其中 $l_1=\max\{0,n-(N-M)\},l_2=\min\{n,M\}$.

例 2. 9（P_{46}）

当很大，而 n 相对较小时，有

$$\frac{C_M^k\cdot C_{N-M}^{n-k}}{C_N^n}\approx C_n^k\left(\frac{M}{N}\right)^k\cdot\left(1-\frac{M}{N}\right)^{n-k}. \tag{2.9}$$

所以，当 N 很大而 n 相对较小时，不放回抽取可当作有放回抽取处理.

例 2. 10（P_{47}）

5. 几何变量及其分布（P_{48}）

重复独立地做伯努利试验，直到事件 A (P(A)=p,0<p<1)首次发生为止所进行的试验次数 X 的分布律

$$P\{X=k\}=q^{k-1}p,\quad k=1,2,\cdots(q=1-p) \tag{2.10}$$

称为参数为 p 的**几何分布**，记作 $X\sim G(p)$.

例 2. 11（P_{48}） 某人射击的命中率为 0.45，求他在第偶数次射击中命中目标的概率.

解 射击一次为一次伯努利试验，射击若干次则为多重伯努利试验.若以 X 计此人首次击中目标时累计的射击次数，则 $X\sim G(0.45)$ ，其概率分布为

$$P\{X=k\}=0.55^{k-1}\cdot 0.45,\quad k=1,2,\cdots.$$

他在第偶数次射击中命中目标的概率则为

$$\sum_{k=1}^{\infty}P\{X=2k\}=0.45\sum_{k=1}^{\infty}0.55^{2k-1}=\frac{11}{31}.$$

例 2. 12（P_{48}） 十只同种电子元件，其中有两只废品.装配仪器时，从中任取一只，一旦发现是废品则扔掉再重取一只，直到取到正品为止.问在取到正品前已取出的废品件数 X 是几何变量吗？写出 X 的分布律.

解 X 不是几何变量. X 可取 0,1,2，其分布律为

$$P\{X=0\}=\frac{8}{10}=\frac{4}{5},\quad P\{X=1\}=\frac{2\times 8}{10\times 9}=\frac{8}{45},\quad P\{X=2\}=\frac{2\times 1\times 8}{10\times 9\times 8}=\frac{1}{45}.$$

作业（P_{49}-P_{50}）：1（2）. 2. 3. 8.–9. 12.–13. 16. 17. 19.–20.

4.–6. 7. 10.–11. 14.–15. 18.

§2. 3 连续型随机变量及其分布

本节讨论连续型随机变量的概率分布及性质、概率计算与常见分布.

2. 3. 1 连续型随机变量的概率密度函数

连续型随机变量的概率分布一般用概率函数表示，类似于刻画物体质量的“线密度”函数.

定义　如果存在实数域上的一个非负可积函数 f(x)，使得对于任意实数 a,b(a<b)，随机变量 X 的取值在区间(a,b]中的概率为

$$P\{a < X \le b\} = \int_a^b f(x)\,dx, \tag{2.11}$$

那么 X 是连续型随机变量. 这里的非负函数 f(x)即是描述连续型随机变量 X 取值规律的概率函数，称为 X 的概率密度函数，简称概率密度，记作 $X \sim f(x)$.

f(x)有时记作 $f_X(x)$，以有别于其它随机变量的密度函数.

从几何意义上来看，X 的取值在区间(a,b]内的概率等于以该区间为底边、以曲线 y=f(x)为高的曲边梯形的面积(图 2.6 中阴影部分).

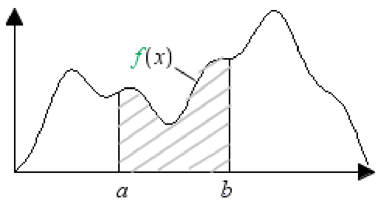


图 2.6 (2.11)式的几何解释

概率密度 f(x)有如下性质：

(1) f(x)≥0；

(2) $\int_{-\infty}^{+\infty} f(x)dx = 1$.

规定：f(x)=0，如果 x 不是 X 的取值.

如果一个函数 f(x)符合(1)、(2)，则可以将其视为一个连续型随机变量的概率密度函数.

事件的概率计算：

(1) X 取任意一个值 x 的概率为 0，即 P{X=x}=0.

(2) X 在区间 G 上取值的概率为

$$P\{X \in G\} = \int_G f(x)dx\,, \tag{2.12}$$

由于概率密度多是分段函数，以及 G 可能是多个区间的并集，所以计算概率的积分常为分段积分.

例 2.13(P_{51})　判断函数 $f(x) = \begin{cases} 0.5, & -1 < x < 0, \\ 0.25, & 0 \le x < 2, \\ 0, & \text{其他} \end{cases}$ 可否作为概率密度？

例 2.14(P_{51})

2.3.2 常见的连续型随机变量及其分布

1. 均匀变量及其分布(P_{52})

如果随机变量概率 X 只在区间(a,b)内取值，且在(a,b)内等长度的任意区间上取值的概率相同，则 X 是

区间(a,b)上的所谓均匀变量，其密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{其它},\end{cases} \tag{2.13}$$

称 X 服从参数为 a,b 的**均匀分布**，记作 $X \sim U(a,b)$.

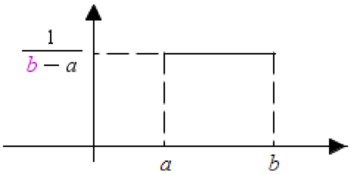


图 2. 7 均匀变量的概率密度函数图像

提问：若 $X \sim U(a,b)$ ，则对于 $\forall c,d(c \le d)$ ， $P\{c \le X \le d\}$=？

例 2. 15 (P_{52})　某公交汽车站在上午 7 时起每隔 15 分钟来一班车.有一乘客在 7:00-7:30 之间等可能地到达该站.问他候车时间不到 5 分钟的可能性有多大？

解　设乘客于 7 点过 X 分钟到达车站,依题意则有 $X \sim U(0,30)$ ，其密度函数为

$$f(x)=\begin{cases}\dfrac{1}{30}, & 0<x<30,\\ 0, & \text{其他}.\end{cases}$$

乘客候车时间不到 5 分钟，他必须在 7:10-7:15 之间或 7:25-7:30 之间到达车站.因此，乘客候车时间不到 5 分钟的可能性为

$$p=P\{10<X<15\}+P\{25<X<30\}=\frac{1}{30}\times(5+5)=\frac{1}{3}.$$

例 2. 16 (P_{53})

2. 指数变量及其分布 (P_{53})

如果随机变量 X 只在区间(0,+∞)内取值，密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\le 0.\end{cases} \quad \text{（其中 } \lambda>0 \text{ 为常数）} \tag{2.14}$$

则 X 为指数变量，称其服从参数为 λ 的**指数分布**，记作 $X \sim E(\lambda)$.

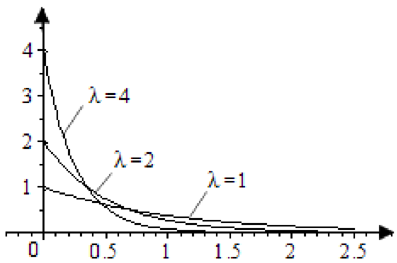


图 2. 8 指数变量的概率密度函数图像

提问：若 $X \sim E(\lambda)$ ，则对于 $\forall c,d(c \le d)$ ， $P\{c \le X \le d\}$=？

例 2. 17 (P_{54})　打一次电话所用的时间(单位：分钟)服从参数为 0.2 的指数分布，如果有人刚好在你前

走进公用电话亭，那么你等待时间超过 5 分钟的可能性与在 5 分钟与 10 分钟之间的可能性各有多大？

解 以 X 表示正在打电话的人所占用的时间，则 $X \sim E(0.2)$，其概率密度为

$$f(x)=\begin{cases}0.2e^{-0.2x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

等待时间超过 5 分钟，即事件{X>5}发生的可能性为

$$P\{X>5\}=e^{-0.2\times 5}=e^{-1}\approx 0.3679 .$$

等待时间在 5 分钟到 10 分钟之间的可能性为

$$P\{5<x<10\}=e^{-0.2\times 5}-e^{-0.2\times 10}=e^{-1}-e^{-2}\approx 0.2325 .$$

指数分布常叫“寿命”分布，在可靠性理论与排队论中有重要而广泛的应用.在可靠性问题中，电子元件的使用寿命常服从指数分布；随机服务系统中的服务时间也可以认为服从指数分布.

指数分布具有无记忆性(P_{54}).

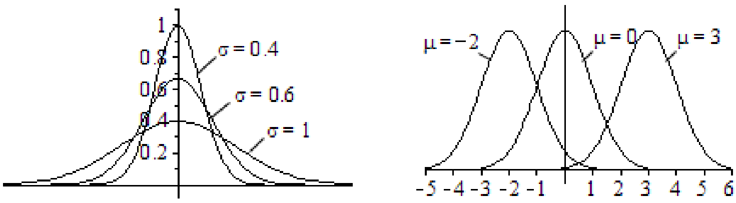
3．正态变量及其分布(P_{54})

如果随机变量 X 的密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}，（其中 \mu,\sigma(\sigma>0) 为常数） \qquad (2.15)$$

则 X 是正态变量，并称其服从参数为 μ,σ^2 的**正态分布**，记作 $X\sim N(\mu,\sigma^2)$.

当 $\mu=0,\sigma=1$ 时，X 称为**标准正态变量**.



(a) μ 相同但 σ 不同　　(b) σ 相同但 μ 不同

图 2. 9 正态变量的概率密度函数图像

正态变量的密度函数 f(x)有以下性质：

①密度函数曲线 f(x)关于直线 $x=\mu$ 对称，表明 $P\{X\ge\mu\}=P\{X\le\mu\}=0.5$；

②函数 y=f(x)在 $x=\mu$ 取得最大值 $\frac{1}{\sqrt{2\pi}\sigma}$，表明正态变量在 $x=\mu$ 附近取值的可能性最大；

③当 σ^2 大时，曲线平坦，表明正态变量的取值比较分散；当 σ^2 小时，曲线陡峭，表明正态变量的取值比较集中.

正态变量的概率计算一般无法通过积分直接得到结果. 这是因为

若 $X\sim N(\mu,\sigma^2)$，则对 $\forall a,b(a<b)$，

$$P\{a<X\le b\}=\int_a^b\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx，$$

上式右端积分积不出来.

作变量代换：$t=\dfrac{x-\mu}{\sigma}$，那么

$$\int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt . \tag{2.16}$$

(2.16)式右端的被积函数 $\dfrac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$ 是标准正态变量的密度函数，所以右端的积分表示的是标准正态变量在 $(a-\mu)/\sigma$ 与 $(b-\mu)/\sigma$ 之间取值的概率，这个积分仍然积不出来. (2.16)式表明：一般正态变量事件的概率计算问题归结为标准正态变量的事件的概率计算问题.

为得出(2.16)式右端的积分结果，设函数

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

Φ 在 x 处的函数值即是标准正态变量在(-∞,x]内取值的概率. Φ(x)的函数值已制成**标准正态分布表**(见附录3). 于

$$\int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \int_{-\infty}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt - \int_{-\infty}^{\frac{a-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right).$$

所以，若 $X\sim N(\mu,\sigma^2)$，则

$$P\{a<X\le b\}=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right). \tag{2.17}$$

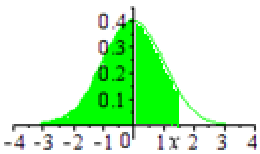


图 2.10　Φ(x)的几何解释

标准正态分布的概率密度函数一般用 $\varphi(x)$ 表示，即

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

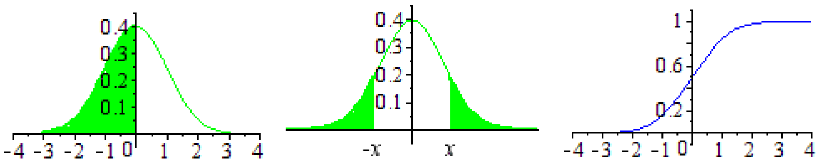
根据 $y=\varphi(x)$ 的特点以及 Φ(x)的定义，易知 Φ(x)有以下性质：

（1） Φ(0)=0.5；

（2） Φ(-x)=1-Φ(x)；

（3） Φ(x)是 x 的单调增函数；

（4） Φ(-∞)=0, Φ(+∞)=1.



(a) 性质(1)的几何解释　　(b) 性质(2)的几何解释　　(c) 性质(3)、(4)的几何解释

图 2. 11　Φ(x)性质的几何解释

例 2. 18（P_{57}）　设 $X\sim N(0,1)$，计算 $P\{X>0\},P\{-2.31<X<-1.25\}$，$P\{|X|<1.54\}$，$P\{1.23<X<2.15\}$.

解　$P\{X>0\}=0.5$，

$P\{-2.31<X<-1.25\}=\Phi(-1.25)-\Phi(-2.31)=\Phi(2.31)-\Phi(1.25)=0.9896-0.8944=0.0952$，

$P\{|X|<1.54\}=\Phi(1.54)-\Phi(-1.54)=2\Phi(1.54)-1=2\times 0.9382-1=0.8764$，

$P\{1.23<X<2.15\}=\Phi(2.15)-\Phi(1.23)=0.9842-0.8907=0.0935$.

例 2. 19（P_{57}）

例 2. 20（P_{58}）

解　(1) $P\{|X-\mu|<a\}=P\left\{\frac{|X-\mu|}{\sigma}<\frac{a}{\sigma}\right\}=2\Phi\left(\frac{a}{\sigma}\right)-1$.

(2) 当 $a=k\sigma$ 时，$P\{|X-\mu|<k\sigma\}=2\Phi(k)-1$.

且当 $k=1,2,3$ 时，有

$$P\{|X-\mu|<\sigma\}=2\Phi(1)-1=0.6836,$$

$$P\{|X-\mu|<2\sigma\}=2\Phi(2)-1=0.9544,$$

$$P\{|X-\mu|<3\sigma\}=2\Phi(3)-1=0.9974,$$

当 $k\geq 4$ 时，$P\{|X-\mu|<k\sigma\}=2\Phi(k)-1\approx 1$.

该结果表明：正态变量 X 的取值几乎都在区间(μ-3σ,μ+3σ)内，这称为正态变量的 **3σ 原则**.

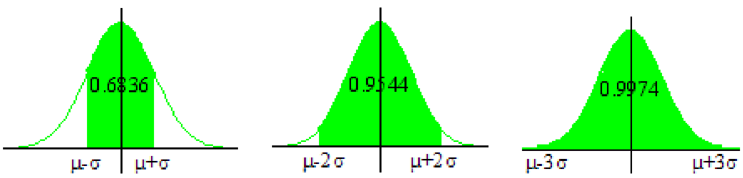


图　正态变量 3σ 原则的几何解释

设 $X\sim N(0,1)$，对任意给定的 $\alpha(0<\alpha<1)$，称使

$$P\{X>z_\alpha\}=\alpha$$

成立的 z_α 为标准正态分布 N(0,1)的**上 α 分位数**.

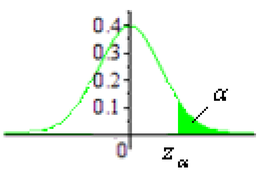


图 2. 12　标准正态分布的上 α 分位数图示

易见，z_α 满足

$$\Phi(z_\alpha)=1-\alpha\,. \tag{2.20}$$

z_α 的值可从标准正态分布表中查到．下表列出的是几个常用的 z_α 值.

α	0.001	0.005	0.010	0.025	0.050	0.100
z_α	3.090	2.576	2.327	1.960	1.645	1.282

作业 (P_{59})：1.　4.−6.　9.−12.　15.　17.　　3.　7.−8.　13.−14.　16.

§2.4　随机变量的分布函数

本节以统一的方式研究各种类型的随机变量，分布函数是这一方式的工具.

一个与随机变量 X 的类型无关，只与 X 所表达的特定事件的概率有关的一元函数

$$F(x)=P\{X\le x\} \tag{2.21}$$

称为随机变量 X 的**分布函数**. 这里{X≤x}表示 X 的取值不超过 x 的事件，函数 F 在 x 处的函数值即是事件{X≤x}发生的概率.

X 的分布函数有时记为 $F_X(x)$，以有别于其它随机变量的分布函数.

分布函数 F(x)的计算公式：

(1) 若 X 是离散型随机变量，概率分布为(2.1)式，则

$$F(x)=\sum_{x_i\le x}P\{X=x_i\}=\sum_{x_i\le x}p_i=\begin{cases}0, & x<x_1\\ p_1, & x_1\le x<x_2\\ p_1+p_2, & x_2\le x<x_3\\ \vdots & \end{cases}. \tag{2.22}$$

反之，对于任意实数 x ，有

$$P\{X=x\}=F(x)-F(x^-). \tag{2.23}$$

(2)若 X 是连续型随机变量，概率密度函数为 f(x)，则

$$F(x)=\int_{-\infty}^{x}f(x)dx. \tag{2.24}$$

反之，在密度函数 f(x)的任意连续点 x 处，有

$$f(x)=\frac{dF(x)}{dx}. \tag{2.25}$$

根据(2.22)式与(2.24)式可知，离散型随机变量的分布函数是右连续的、单调不减的阶跃函数；连续型随机变量的分布函数是连续的单调不减函数.

例 2.21 (P_{60})　一只袋中有 6 个球，一个标有-2，三个标有 1，两个标有 3.今从口袋中任取一个球，并以 X 记球上标有的数字，求 X 的分布函数.

解　X 的分布律为 $P\{X=-2\}=1/6, P\{X=1\}=1/2, P\{X=3\}=1/3$. 按公式(2.22)，X 的分布函数为

$$F(x)=\begin{cases}0, & x<-2,\\ \dfrac{1}{6}, & -2\le x<1,\\ \dfrac{2}{3}, & 1\le x<3,\\ 1, & x\ge 3.\end{cases}$$

其图像如下图 2.13 所示.

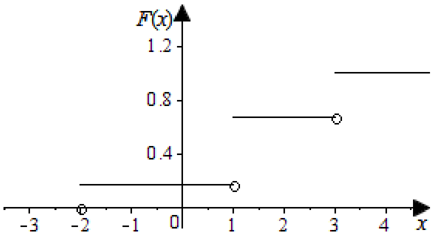


图 2.13　例 2.21 中的随机变量 X 的分布函数图像

例 2.22（P_{61}）求常见的连续型随机变量的分布函数.

解 (1) 均匀分布的分布函数为

$$F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\le x<b,\\ 1, & x\ge b.\end{cases}$$

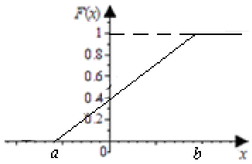


图 2.14　均匀变量的分布函数图像

(2) 指数分布的分布函数为

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

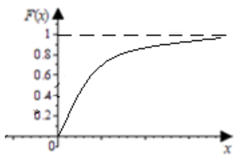


图 2.15　指数变量的分布函数图像

(3) 正态分布的分布函数为

$$F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right).$$

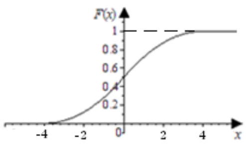


图 2.16　正态变量的分布函数图像

分布函数 F(x)的性质：

(1) $0\le F(x)\le 1$；

(2) 对 $\forall\, x_1\le x_2$，有 $F(x_1)\le F(x_2)$，即 F(x)是 x 的单调不减函数；

(3) $F(-\infty)=0, F(+\infty)=1$；

(4) $F(x^+)=F(x)$，即 F(x)是右连续函数.（当 X 是连续型随机变量时，F(x)是连续函数）.

如果一个函数符合(1)-(4)，则可以将其视为随机变量的分布函数.

分布函数还有一个很重要的应用，可以用来计算随机变量在左开、右闭区间(a,b]内取值的概率：

$$P\{a<X\le b\}=F(b)-F(a). \tag{2.28}$$

例 2.24（P_{64}）　已知随机变量 X 的分布函数为

$$F(x)=\begin{cases} a, & x<0, \\ b\sin x, & 0\le x\le \pi/2, \\ c, & x>\pi/2. \end{cases}$$

求 $P\{|x|<\pi/6\}$.

解 由 F(-∞)=0,F(+∞)=1 及 F(x)在 $x=\pi/2$ 处的右连续性，有

$$\begin{cases} a=0, \\ 1=b\sin\dfrac{\pi}{2}, \\ c=1. \end{cases}$$

得 $a=0,b=1,c=1$. 于是

$$P\{|X|<\frac{\pi}{6}\}=P\{-\frac{\pi}{6}<X<\frac{\pi}{6}\}=F\left(\left(\frac{\pi}{6}\right)^{-}\right)-F\left(-\frac{\pi}{6}\right)=\sin\frac{\pi}{6}-0=\frac{1}{2}.$$

例 2. 25（P_{64}）

作业（P_{65}）：2. 4. 6. 8.－10. 5. 11.－14.

§2. 5 随机变量的函数的分布

通过已知的概率分布得出随机变量的函数（也是随机变量）的概率分布，就是随机变量的函数的分布问题. 如：在分子运动中，希望由分子运动速度的绝对值变量 X 的概率分布推导出分子运动动能 $W=mX^2/2$ 的概率分布；希望由风速 V 的概率分布推导出飞机机翼所受压力 $F=kV^2$（k>0 是常数）的概率分布；希望由圆半径 R 的概率分布推导出圆面积 $A=\pi R^2$ 的概率分布， 等等.

随机变量函数的分布问题一般提法：已知随机变量 X 的概率分布，求随机变量 Y=g(X)（g 是函数）的概率分布.

2. 5. 1 g 为连续函数的分布

1. X 为离散型随机变量

例 2. 26（P_{66}）

2. X 为连续型随机变量

如果 X 是连续型随机变量，那么由连续函数 g 所产生的随机变量 Y=g(X)也是连续型随机变量.

设随机变量 X 的概率密度为 $f_X(X)$，则求 Y=g(X)的概率分布的步骤如下：

(1) 求 Y 的分布函数

$F_Y(y)=P\{Y\le y\}$ （形式上写出 Y 的分布函数）

$=P\{g(X)\le y\}$ （作变量代换，转化为 X 的事件）

$=\int_{g(x)\le y} f_X(x)dx$ （计算关于 X 的事件的概率）

其中 $\int_{g(x)\le y}$ 表示在由 g(x)≤y 所决定的 x 的取值区域上积分.

(2) 求 Y 的密度函数

$$f_Y(y)=\frac{dF_Y(y)}{dy}.$$

如果 X 的密度函数为 $f_X(X)$，且 X 的全部可能取值为区间(a,b)(也可以是[a,b],(a,b],[a,b))内的值，即 $f_X(X)$ 只在区间(a,b)内不为零，而函数 y=g(x)的一阶导数 $g'(x)$ 在(a,b)内几乎处处大于零(或小于零)，那么按照上面的步骤可以推导出 Y=g(X)的密度函数为

$$f_Y(y)=\begin{cases} f_X(h(y))\cdot|h'(y)|, & c<y<d, \\ 0, & 其他. \end{cases} \tag{2.29}$$

其中 x=h(y)是 y=g(x)的反函数，c=min{g(a), g(b)}, d=max{g(a), g(b)}.

例 见 P_{71}

解

例 2.27 (P_{68})

结果表明：正态变量的线性函数仍然是正态变量，即 $Y=aX+b(a\neq 0)\sim N(a\mu+b,(a\sigma)^2)$. 特别地，$Y=\frac{X-\mu}{\sigma}\sim N(0,1)$.

例 2.28 (P_{69}) 设 $X\sim f(x)=0.5e^{-|x|}$，求 $Y=X^2$ 的密度函数.

解

例 (P_{71}) 设 $X\sim U(x_1,x_2)$，Y=cX+d，证明：$Y\sim U(y_1,y_2)$，其中 $y_1=\min\{cx_1+d, cx_2+d\}$，$y_2=\max\{cx_1+d, cx_2+d\}$.

证

2.5.2 g 为非连续函数的分布

下面举例说明，如果 g 是不连续函数，如何求 Y=g(X)的概率分布.

例 2.29 (P_{69}) 设随机变量 $X\sim P(\lambda)$，随机变量 $Y=\begin{cases}0, & X\le 1,\\ 1, & X>1\end{cases}$，求 Y 的概率分布.

解

例 2.30 (P_{69}) 设 $X\sim N(0,1)$，随机变量 $Y=\begin{cases}-1, & X\le 0,\\ 1, & X>0\end{cases}$，求 Y 的概率分布.

解

作业 (P_{71})：1.−2. 3. 4.−7.（各题任意选做一小题） 8. 10. 9. 11^*.