第 1 章　概率论基础

【引例】　现实中的统计。

X 商店位于 Y 市西郊，是一家以经营生鲜食品、日杂用品为主的中型百货商店。在 X 商店正式营业的第一年末，商店经理决定购入一批挂历进行销售，但购入挂历的数量成为困扰经理的一个难题——一方面如果购入的挂历数量不够，那么待挂历售尽便会出现缺货损失，从而只能眼睁睁地看着大笔生意被竞争对手抢走；另一方面如果购入的挂历数量过多，多余的存货积压必然会造成流动资金的短缺及存货费用的增加，因此只能做削价处理，这必将给商店带来经济上的损失。

为了使收益值的期望最大，经理请教了在高校任教的王老师来为商店确定一个合适的挂历进货量。已知商店每出售一件挂历可获得纯利润 7 元；但如果在春节以前不能售尽，则需要做削价处理，每件将亏损 3 元。

王老师调查了 Y 市与鼎文商店各种情况类似的十家商店，统计了每家商店在最近四年春节前挂历的销售情况。根据调查结果，得到了表 1-1 中的数据。

表 1-1　十家商店在最近四年的挂历销售量分布表

销售量	100	200	300	400	500	600	合计
次数	2	5	13	11	6	3	40

根据表 1-1 中的数据，王老师通过计算得到了另外两个表——挂历销售量的概率分布表和收益与收益期望分布状况表，并得出结论：当挂历的进货量为 400 件时，商店的期望收益最大，为 2075 元。按照经理的“收益值的期望最大”的要求，王老师向经理建议商店的进货量为 400 件。

上述引例中所涉及的概率、概率分布、期望等概念均属于概率论的范畴。概率论是研究随机现象规律性的数学分支，在科学研究和社会生产实践中有着十分广泛的应用，是统计研究的基础。本章将介绍一些概率论的基础理论，包括事件与概率、概率的基本性质、条件概率与事件独立性，以及随机变量及其分布。

1.1　事件与概率

事件与概率是概率论研究中的两个最基本的概念。围绕着这两个概念，本节将介绍三部分内容，包括随机试验与随机事件、事件的关系及运算、事件的概率等内容。

1.1.1　随机试验与随机事件

1. 随机试验与样本空间

在自然界和人类社会生产实践中，存在着两类现象。一类现象在一定条件下必然发生（或必然不发生）。例如，在标准大气压下，水的沸点是 100℃；又如向上抛掷一枚石子，由于受地心引力的作用，石子在上升到一定高度之后必然下落。由于这类现象具有确定的结果，故称为确定性现象。

然而，并不是所有的现象都具有确定性的结果。例如，抛掷一枚硬币，当硬币落地后，可能是正面朝上，也可能是反面朝上，而在硬币落地前不能预知究竟哪一面朝上。同样地，自动机床加工制造同一零件，加工出来的零件可能是合格品也可能是不合格品；同一门炮向同一目标发射多发同种炮弹，弹落点也不一样，等等。以上列举的现象均具有不确定性，即在基本条件不变的情况下，一系列试验或观察会得到不同的结果，并且在每次试验或观察之前不能预知会出现哪种结果，这类现象称为随机现象。概率论研究的对象就是随机现象。

【例 1–1】 生活中随机现象的例子。

① 抛掷一颗骰子，出现的点数；

② 一天内进入某超市的顾客数；

③ 某一生产线生产出的灯泡的寿命；

④ 某批产品的不合格率。

为了探索和研究随机现象的规律性，通过随机试验（简称试验）来对随机现象进行调查、观察或实验。具体来说，随机试验应满足如下条件：

- 试验可以在相同的条件下重复进行；
- 试验有多种可能的结果，并且事先可以明确所有可能出现的结果；
- 试验完成之前不能预知会出现哪一个结果。

一个随机试验的所有可能结果的集合称为样本空间，通常用 Ω 表示。样本空间的元素，即试验的每一个可能结果，称为这个试验的样本点，用 ω 表示。

【例 1–2】 试列出例 1–1 中随机现象的样本空间。

解：① 掷一颗骰子的样本空间为 $\Omega_1=\{\omega_1，\omega_2，\cdots，\omega_6\}$，其中 ω_i 表示出现的 i 点，$i=1，2，\cdots，6$。也即掷一颗骰子的样本空间为 $\Omega_1=\{1，2，\cdots，6\}$。

② 一天内进入某超市顾客数的样本空间为 $\Omega_2=\{0，1，2，\cdots\}$，其中“0”表示一天内无人光顾该超市。

③ 某生产线生产出的灯泡的寿命的样本空间为 $\Omega_3=\{t\mid t\geqslant 0\}$。

④ 产品的不合格率一定是介于 0 与 1 之间的一个实数，因此其样本空间为 $\Omega_4=\{y\mid 0\leqslant y\leqslant 1\}$。

在例 1–2 中，Ω_1 中的样本点的个数为有限个，是比较简单的样本空间；而 Ω_2、Ω_3 和 Ω_4 中的样本点的个数为无限个，但 Ω_2 中的样本点可以按照某种次序排列出来，即 Ω_2 中有可列个样本点。在概率研究中，将包含有限个或可列个样本点的样本空间称为离散样本空间，如例 1–2 中的 Ω_1 和 Ω_2；而将包含无限个或不可列个样本点的样本空间称为连续样本空间，如例 1–2 中的 Ω_3 和 Ω_4。

2. 随机事件

样本空间 Ω 的某个子集称为随机事件，简称事件，通常用大写字母 $A，B，C，\cdots$表示。

随机事件表示试验可能出现的结果，这个结果可以是仅由一个样本点组成的基本事件，也可以是由多个样本点组成的复合事件。

对于某一事件 A，当且仅当它所包含的某一样本点出现时，称事件 A 发生。

例如，在掷骰子试验中，投掷一颗均匀的骰子，其样本空间为 $\Omega=\{1，2，\cdots，6\}$。现在从不同的角度考察该实验的结果：记事件 A 为“出现 2 点”，事件 B 为“出现偶数点”，事件 C 为“出现的点数小于 7”，事件 D 为“出现的点数大于 6”。

其中，A 为基本事件，当且仅当掷出 2 点时，事件 A 发生，即 $A=\{2\}$；事件 B 发生当且仅当下列样本点之一发生：掷出 2 点、掷出 4 点和掷出 6 点，它由三个基本事件复合而成，即 $B=\{2，4，6\}$。

对于事件 C，在一次试验中，由于每次抛掷骰子出现的点数必然小于 7，因此事件 C 必然发生，即 $C=\{1, 2, \cdots, 6\}$。通常，把样本空间 Ω 本身称为必然事件，事件 C 就是一个必然事件。

同样地，对于事件 D，由于每次抛掷骰子出现的点数不可能大于 6，因此事件 D 不可能发生，即 $D=\{\}$。通常，把空集 $\varnothing$ 称为不可能事件，事件 D 即不可能事件。

严格来讲，必然事件和不可能事件反映了确定性现象，本质上不是随机事件，然而为了研究方便，还是把必然事件和不可能事件作为随机事件的两个极端情形来处理。

1.1.2　事件的关系及运算

在一个随机试验中，样本空间可以定义的随机事件显然不止一个，同时，事件与事件之间必然存在这样那样的联系。为了能够更好地理解及运用随机试验的结果，下面将借助文氏图分析事件的关系及运算。

1. 事件之间的关系

事件间的包含关系：若事件 A 发生必然导致事件 B 发生，则称 A 包含于 B，或 B 包含 A，记为 $A \subset B$ 或 $B \supset A$，如图 1–1 所示。

例如，在掷骰子试验中，若记事件 A 为“出现 2 点”，事件 B 为“出现偶数点”，则 $A \subset B$。显然，对于任一事件 A，必有 $\varnothing \subset A \subset \Omega$。

事件间的相等关系：若事件 A 发生必然导致事件 B 发生，同时事件 B 发生必然导致事件 A 发生，则称事件 A 与 B 相等，记为 $A=B$。相等的两事件在实质上是对同一事件的不同语言描述。

事件间的互不相容关系：若事件 A 与事件 B 不可能同时发生，则称事件 A 与 B 互不相容，如图 1–2 所示。

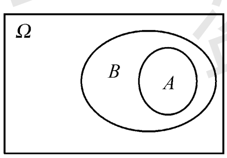


图 1–1　$A \subset B$

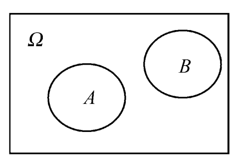


图 1–2　A 与 B 互不相容

同样以掷骰子试验为例，在试验中，“出现的点数小于 2”与“出现的点数大于 4”两个事件不可能同时发生，因而它们是互不相容事件。

2. 事件的运算

对于样本空间中的事件，可以通过以下四种事件的基本运算，得到新的事件。现定义两个事件 A 和 B，事件的运算有如下四种。

事件的并：由属于事件 A 或 B 的所有样本点构成的集合称为事件 A 与 B 的并，记为 $A \cup B$。特别地，对于互不相容的事件 A 和 B，称它们的并为和，记为 $A+B$。

事件 $A \cup B$ 表示事件 A 和 B 至少发生一个。例如，在掷骰子试验中，若记事件 A 为“出现的点数大于 1 小于 3”，事件 B 为“出现的点数大于 2 小于 4”，则事件 $A \cup B$ 表示“出现的点数大于 1 小于 4”。

事件的交：由同时属于事件 A 和 B 的所有样本点构成的集合称为事件 A 与 B 的交，记为 $A \cap B$ 或 AB。

事件 $A\cap B$ 表示事件 A 和 B 同时发生。例如，在掷骰子试验中，若记事件 A 为“出现的点数大于 1 小于 4”，事件 B 为“出现的点数大于 2 小于 5”，则事件 $A\cap B$ 表示“出现的点数大于 2 小于 4”，即出现 3 点。

事件的差：由属于事件 A，但不属于事件 B 的所有样本点构成的集合称为事件 A 与 B 的差，记为 $A-B$。

事件 $A-B$ 表示事件 A 发生而事件 B 不发生。例如，在掷骰子试验中，若记事件 A 为“出现的点数大于 1 小于 4”，事件 B 为“出现的点数大于 1 小于 3”，则事件 $A-B$ 表示“出现的点数大于等于 3 小于 4”，即出现 3 点。

事件的逆：由样本空间中不属于事件 A 的所有样本点构成的集合称为事件 A 的逆，记为 $\bar{A}$。

事件 $\bar{A}$ 是事件 A 的对立事件，表示事件 A 不发生。例如，在掷骰子试验中，若记事件 A 为“出现的点数小于 3”，则事件 $\bar{A}$ 表示“出现的点数大于等于 3”。

以上四种事件运算的文氏图如图 1–3 所示。

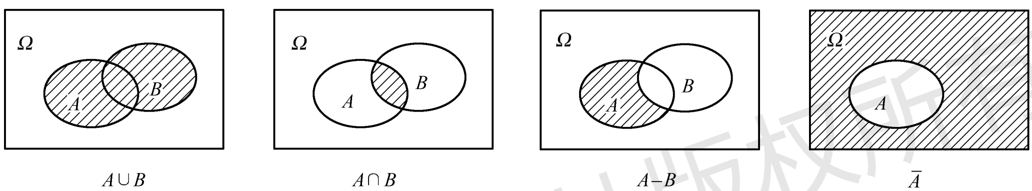


图 1–3　四种事件运算的文氏图

【例 1–3】 假设一批产品中有 3 件次品，且产品的外形没有任何差别。现随机地从这批产品中依次抽取 3 件，若以 A 记“第一次抽到次品”，以 B 记“第二次抽到次品”，以 C 记“第三次抽到次品”，试用 A、B 和 C 的关系表示下列各事件。

① 三次都抽到次品。

② 只有第一次抽到次品。

③ 三次都没有抽到次品。

④ 至少抽到一件次品。

⑤ 最多抽到一件次品。

⑥ 最多抽到两件次品。

解：

① ABC

② $A\bar{B}\bar{C}$

③ $\bar{A}\bar{B}\bar{C}$

④ $A\cup B\cup C$

⑤ 最多抽到一件次品，即 A，B，C 中只有一个发生或 A，B，C 全不发生，即

$$\bar{A}\bar{B}C\cup\bar{A}B\bar{C}\cup A\bar{B}\bar{C}\cup\bar{A}\bar{B}\bar{C}$$

⑥ 最多抽到两件次品，即是 A，B，C 全发生的对立事件，即

$$\overline{ABC}$$

（3）事件运算的性质

事件的运算与集合的运算一样，必须满足如下运算法则。

① 交换律：$A\cup B=B\cup A$，$AB=BA$。

② 结合律：$(A\cup B)\cup C=A\cup(B\cup C)$，$(AB)C=A(BC)$。

③ 分配律：$(A\cup B)\cap C=AC\cup BC$，$(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$。

④ 对偶律(德摩根公式)：$\overline{A\cup B}=\bar{A}\cap\bar{B}$，$\overline{A\cap B}=\bar{A}\cup\bar{B}$。

特别地,德摩根公式可以推广到对于 n 个事件或可列个事件求对偶的问题。

1.1.3　事件的概率

对于一次随机试验,在试验结束之前并不能确定某个事件是否会发生,这是由随机试验的基本性质所决定的。例如,在一次摸球试验中,假定袋子中有包括 9 只白球和 1 只黑球在内的 10 只小球,在实验结束之前,并不能确定会摸出黑球还是白球。显然摸出白球的可能性比摸出黑球的可能性大得多。

随机事件发生的可能性大小不仅可以比较,而且是可以量化的。对于一个随机事件 A,若可以用一个数 $P(A)$ 来表示其发生的可能性大小,这个数 $P(A)$ 就称为随机事件 A 的概率。

概率度量随机事件发生的可能性的大小,它由随机事件自身所决定,反映了随机现象的内在规律。那么,概率究竟应该如何量化呢?

1. 概率的统计定义

对于随机事件 A,如果它在 N 次试验中发生了 n 次,则称

$$F_N(A)=n/N \tag{1-1}$$

为随机事件 A 在 N 次试验中出现的频率。

显而易见，频率具有如下性质。

① 非负性：对于随机事件 A，必有 $F_N(A)\geqslant 0$。

② 规范性：对于必然事件 Ω，在 N 次试验中出现的次数应为 N，即 $F_N(\Omega)=1$。

③ 可加性：若 A 和 B 为互不相容事件，则 $F_N(A\cup B)=F_N(A)+F_N(B)$。

以上三条性质为频率的基本性质，根据这些性质，还可以得出许多其他的性质，例如，“任何随机事件在 N 次试验中出现的频率都不大于 1”，“不可能事件在 N 次试验中出现的频率为 0”，“对于有限个两两互不相容的事件的频率也具有可加性”等。

另外，对于多次重复试验，随机事件 A 的频率还具有另外一项重要的性质——频率稳定性。

人们经过长期的生产实践发现：在相同条件下进行的多次重复试验，随着试验重复次数 N 的增加，随机事件 A 的频率 $F_N(A)$ 会在某一固定的常数 a 附近摆动，频率的这个性质称为频率稳定性，这个固定的常数 a 就是概率。

下面，以抛硬币试验为例来说明频率的稳定性。

在抛一枚硬币时，既可能出现正面，也可能出现反面，预先做出判断是不可能的，但是假如硬币均匀，直观上出现正面与出现反面的机会应该相等，即在大量试验中出现正面的频率应该接近于 50%，为了验证这一点，历史上曾有不少人做过这个试验，结果如表 1-2 所示。

表 1-2　历史上抛硬币试验的若干结果

试　验　者	抛硬币次数	出现正面次数	出现正面频率
德摩根	2048	1061	0. 5181
蒲丰	4040	2048	0. 5069
费勒	10000	4979	0. 4979
皮尔逊	12000	6019	0. 5016
皮尔逊	24000	12012	0. 5005

从表 1-2 可以看出，出现正面的频率在 0.5 附近摆动，根据频率稳定性可知，出现正面的概率为 0.5。

由于概率与频率的密切关系，在实际应用中，常常需要根据历史数据，统计某一事件发生的频率，以估计其概率。例如，在北京奥运会前夕，中国气象局分析了北京市观象台 1951—2007 年历时 57 年的气象资料，得到北京地区在 2008 年 8 月 8 日奥运会开幕式当天降水的概率为 47%。

2. 概率的古典定义

在讨论概率的统计定义时曾提到过，抛一枚均匀的硬币，直观上出现正面与出现反面的概率是相等的，并且历史上大量的试验数据也验证了这一观点。类似于抛硬币的试验，在人类的生产实践中，存在着许多这类随机现象，诸如掷骰子、产品抽样检查等，对于这些随机现象进行深入分析之后，可以发现，它们之间存在以下两个基本的共同点：

① 试验具有有限个可能出现的结果；

② 试验的每个基本事件出现的可能性都是相等的。

具有以上两个基本特点的概率模型称为古典概型。古典概型在概率论的发展初期即被注意，它的内容简单，应用却很广泛，许多最初的概率论结果也是由它得出的。

在古典概型中，假定样本空间 $\Omega=\{\omega_1, \omega_2, \cdots, \omega_n\}$，则对于每个基本事件 ω_i 有

$$P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n)=\frac{1}{n}$$

进一步讲，对于古典概型，如果一个试验有 n 个基本事件，其中随机事件 A 包含的基本事件个数为 m，那么随机事件 A 的概率为

$$P(A)=\frac{\text{事件 } A \text{ 所包含的样本点的个数}}{\text{样本点总数}}=\frac{m}{n} \tag{1-2}$$

式中，由于事件 A 包含样本点的出现必然导致 A 的出现，因此又称这些样本点为 A 的“有利场合”。

在古典概型中，通过以上公式求解随机事件 A 的概率，首先要明确事件 A 中所包含的样本点数和样本空间中的样本点总数，计算时需要熟练地运用排列和组合的相关知识，具有一定的技巧性。

【例 1-4】 口袋中有 5 个白球、3 个黑球，从中任取两个，求取到的两个球颜色不同的概率。

解：从 8 个球中任取两个，共有 C_8^2 种不同的取法，将每一种取法作为该试验的一个样本点，可以得到取球试验的样本空间。由于是随机取球，任意两个小球被同时取出的概率是相等的，因此这个问题是古典概型。

记“取到的两个球颜色不同”为事件 A，则事件 A 包含的样本点数为 $C_5^1C_3^1$，因此取到两个不同颜色的球的概率为

$$P(A)=\frac{C_5^1C_3^1}{C_8^2}=\frac{15}{28}$$

摸球模型在实际问题中有很重要的应用。例如，如果把例 1-4 中的“白球”、“黑球”替换为“正品”、“次品”，就可以用来求解产品质量抽样检查问题。另外，还可以向口袋中加入其他颜色的球，使摸球模型更具有代表性，这时就可以描述具有更多等级的产品抽样问题，如将产品分为一等品、二等品、三等品和等外品的产品抽样检查问题。

3. 概率的几何定义

通过古典概型成功地解决了一类问题，这类问题有且只有 n 个基本事件，并且每个基本事

件的概率都为 $1/n$，如抛硬币、掷骰子及摸球等随机试验，以及由它们演化而来的一系列实际问题。然而，在实际问题中经常会遇到结果无限而又有某种等可能性的情况，这时就需要借助几何概型来求解。

一般地，设在空间上有一区域 Ω，随机地向 Ω 内投掷一点 M，则 M 落在区域 Ω 内的任意位置的可能性都是相等的。现规定区域 g 是包含在区域 Ω 内的任一区域，且区域 Ω 和区域 g 都是可以测度的，那么点 M 落在区域 Ω 的任何部分区域 g 内的概率只与 g 的测度（长度、面积、体积等）成正比，并且与 g 的位置和形状无关。具有这种性质的概率模型称为几何概型。

若以 A_g 记“向区域 Ω 中任意投掷一个点 M，点 M 落在 Ω 内的部分区域 g”这一事件，那么随机事件 A_g 的概率为

$$P(A_g)=\frac{g\text{ 的测度}}{\Omega\text{ 的测度}} \tag{1-3}$$

求解几何概型问题的关键是将样本空间 Ω 和所求事件 A_g 用几何图形描述清楚，然后计算出相关图形的测度。

【例 1-5】 甲、乙两人约定在 6 时到 7 时之间在某处会面，并约定先到者应等候另一人 20 分钟，过时即可离去，求两人能会面的概率。

解：如题意，甲乙两人均会在 6 时过后的第 0 ~ 60 分钟内到达，并且在任意时刻到达的可能性都相等，因此这是一个几何概型问题。

以甲到达的时刻为 x 轴，以乙到达的时刻为 y 轴，建立平面直角坐标系，如图 1-4 所示。因此，(x，y) 的所有可能结果为图中所示边长为 60 的正方形，由此得到样本空间 Ω 的测度为 $S_\Omega=60^2$。

如果两人能够会面，需要满足条件：

$$|x-y|\leqslant 20$$

即图 1-4 中的阴影部分，其面积为 $S_g=60^2-40^2$，故两人能会面的概率为

$$P(A_g)=\frac{60^2-40^2}{60^2}=\frac{5}{9}$$

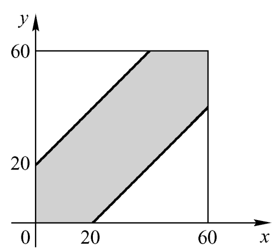


图 1-4　会面问题坐标图

在研究概率问题时，除了运用以上三种方法来确定随机事件的概率，对于一些不能重复的或不能大量重复的现象，人们通常根据个人的经验对随机事件发生的可能性进行估计，这样得出的概率称为主观概率。

主观概率在现实生活中的应用很多，例如，有些地方的气象预报有“降水概率”，根据播音员的预报，“今天夜间多云有阵雨，降水概率为 60%”，这是气象专家根据专业知识和最近的气象情况给出的主观概率。又如，一个外科医生根据自己多年来的临床经验和某患者的病情，认为“该患者手术成功的可能性为 90%”等。

1.2　概率的基本性质

在 1.1 节中，从三个不同的角度给出了概率的定义，它们各适合一类随机现象，有着各自确定概率的方法。那么，通过这些方法确定的概率是不是具有某些共性呢？下面来介绍概率的基本性质。

性质 1（非负性）　对于任意事件 A，有

$$P(A)\geqslant 0 \tag{1-4}$$

性质 2（规范性）　必然事件 Ω 的概率为 1，即

$$P(\Omega)=1 \tag{1-5}$$

性质 3（可列可加性） 对于可列个两两互不相容事件 A_1，A_2，…，有

$$P(A_1\cup A_2\cup\cdots)=P(A_1)+P(A_2)+\cdots \tag{1-6}$$

性质 1，2，3 是苏联数学家柯尔莫哥洛夫提出的概率的公理化定义中所规定的概率必须满足的三条公理，由这三条公理可以推导出概率的其他性质。

性质 4 不可能事件$\varnothing$的概率为 0，即

$$P(\varnothing)=0 \tag{1-7}$$

证明：由于可列个不可能事件之和仍为不可能事件，所以

$$\Omega=\Omega\cup\varnothing\cup\varnothing\cup\cdots\cup\varnothing\cup\cdots$$

所以

$$P(\Omega)=P(\Omega)\cup P(\varnothing)\cup P(\varnothing)\cup\cdots\cup P(\varnothing)\cup\cdots$$

由性质 1 和性质 2 可得

$$P(\varnothing)=0$$

得证。

性质 5（有限可加性） 对于任意 n 个事件 A_1，A_2，…，A_n，若 $A_iA_j=\varnothing$（i，$j=1$，2，…，n；$i\neq j$），则

$$P(A_1\cup A_2\cup\cdots\cup A_n)=P(A_1)+P(A_2)+\cdots+P(A_n) \tag{1-8}$$

证明：令 $A_{n+1}=A_{n+2}=\cdots=\varnothing$，因此有 $A_iA_j=\varnothing(i$，$j=1$，2，…，$i\neq j)$，由性质 3 得

$$\begin{aligned}P(A_1\cup A_2\cup\cdots\cup A_n)&=P(A_1\cup A_2\cup\cdots\cup A_n\cup A_{n+1}\cup A_{n+2}\cdots)\\&=P(\bigcup_{k=1}^{\infty}A_k)=\sum_{k=1}^{\infty}P(A_k)=\sum_{k=1}^{n}P(A_k)+0\\&=P(A_1)+P(A_2)+\cdots+P(A_n)\end{aligned}$$

得证。

特别地，当 $n=2$ 时，对于任意两个互不相容事件 A 和 B，有 $P(A\cup B)=P(A)+P(B)$。

【例 1-6】 某工厂一个班组有男工 7 人、女工 4 人，现要选出 3 个代表，求 3 个代表中至少有一个女工的概率。

解：这是一个古典概型问题，样本空间包含的全部样本点数为 C_{11}^3。将“3 个代表中至少有一个女工”记为事件 A，则事件 A 是三个两两互不相容事件“3 个代表中有 i 个是女工”（$i=1$，2，3）的和。记“3 个代表中有 i 个是女工”（$i=1$，2，3）为 A_i，则 $A=A_1+A_2+A_3$，又

$$P(A_1)=\frac{C_4^1C_7^2}{C_{11}^3}=\frac{28}{55},\qquad P(A_2)=\frac{C_4^2C_7^1}{C_{11}^3}=\frac{14}{55},\qquad P(A_3)=\frac{C_4^3}{C_{11}^3}=\frac{4}{165}$$

故所求概率为

$$P(A)=P(A_1)+P(A_2)+P(A_3)=\frac{28}{55}+\frac{14}{55}+\frac{4}{165}=\frac{26}{33}$$

性质 6 对于任意事件 A，有

$$P(\bar{A})=1-P(A) \tag{1-9}$$

证明：因为 $A\cap\bar{A}=\varnothing$，根据性质 5，有

$$P(A\cup\bar{A})=P(A)+P(\bar{A})$$

又 $A\cup\bar{A}=\Omega$，故 $P(A\cup\bar{A})=1$，即 $P(A)+P(\bar{A})=1$，移项得 $P(\bar{A})=1-P(A)$，得证。

根据性质 6 可知，通过事件 A 的概率，可以很容易地得到它的逆事件的概率。同样地，事件 A

的概率可以借助于$\bar{A}$的概率来求解。在实际问题中，合理运用这一性质将有效地简化一些问题的求解步骤。

例如，例 1–6 中的问题也可以这样来求解：将“3 个代表中至少有一个女工”记为事件 A，则$\bar{A}$ = “3 个代表全部为男工”，而 $P(\bar{A})=\dfrac{C_7^3}{C_{11}^3}=\dfrac{7}{33}$，根据性质 6 可求得

$$P(A)=1-P(\bar{A})=1-\frac{7}{33}=\frac{26}{33}$$

性质 7　对于任意事件 A 和 B，若 $A\supset B$，则

$$P(A-B)=P(A)-P(B) \tag{1–10}$$

证明：由于 $A\supset B$，故有

$$A=B\cup(A-B)$$

又 $B\cap(A-B)=\varnothing$，根据性质 5，有 $P(A)=P(B\cup(A-B))=P(B)+P(A-B)$
移项可得

$$P(A-B)=P(A)-P(B)$$

得证。

特别地，当 $A\supset B$ 时，根据性质 1 有 $P(A-B)\geqslant 0$，因此必有 $P(A)\geqslant P(B)$。

性质 8（减法公式）　对于任意事件 A 和 B，有

$$P(A-B)=P(A)-P(AB) \tag{1–11}$$

证明：由于 $A-B=A-AB$，且 $AB\subset A$，根据性质 7 有

$$P(A-B)=P(A-AB)=P(A)-P(AB)$$

得证。

性质 9（加法公式）　对于任意事件 A 和 B，有

$$P(A\cup B)=P(A)+P(B)-P(AB) \tag{1–12}$$

证明：因 $A\cup B=A\cup(B-AB)$，又 $A\cap(B-AB)=\varnothing$，根据性质 5，有

$$P(A\cup B)=P(A)+P(B-AB)$$

又 $B\supset AB$，根据性质 5，$P(B-AB)=P(B)-P(AB)$，代入上式可得 $P(A\cup B)=P(A)+P(B)-P(AB)$，得证。

特别地，当 A 与 B 为互不相容事件时，$AB=\varnothing$，$P(A\cup B)=P(A)+P(B)$，同性质 5。

【例 1–7】　已知某学校向学生发行两种电子刊物 A 和 B，且该校学生中订阅刊物 A 的占 65%，订阅刊物 B 的占 50%，同时订阅刊物 A 和 B 的占 30%，试求：从该学校学生中随机地抽取一名，该学生订阅电子刊物的概率。

解：若以 A 记“学生订阅刊物 A”，以 B 记“学生订阅刊物 B”，则学生订阅电子刊物为事件 $A\cup B$。根据概率的加法公式，有

$$P(A\cup B)=P(A)+P(B)-P(AB)=0.65+0.5-0.3=0.85$$

概率的加法公式可以推广到具有多个事件的场合。

性质 10（一般加法公式）　对任意 n 个事件 A_1，A_2，$\cdots$，A_n，有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n) \tag{1–13}$$

特别地，

$$P(A_1\cup A_2\cup A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)-P(A_2A_3)-P(A_1A_3)+P(A_1A_2A_3)$$

【例 1-8】 某人写好 n 封信，又写好 n 个信封，然后在黑暗中把每封信放入一个信封中，试求至少有一封信与信封匹配的概率。

解： 若以 A_i 记第 i 封信与信封匹配，则所求事件为 $A_1 \cup A_2 \cup \cdots \cup A_n$，因此，根据一般加法公式，首先有

$$P(A_i)=\frac{(n-1)!}{n!}=\frac{1}{n}$$

$$P(A_iA_j)=\frac{(n-2)!}{n!}=\frac{1}{n(n-1)}$$

$$P(A_iA_jA_k)=\frac{(n-3)!}{n!}=\frac{1}{n(n-1)(n-2)}$$

$$\cdots$$

$$P(A_1A_2\cdots A_n)=\frac{1}{n!}$$

因此

$$\begin{aligned}P(A_1 \cup A_2 \cup \cdots \cup A_n) &= \mathrm{C}_n^1\frac{1}{n}-\mathrm{C}_n^2\frac{1}{n(n-1)}+\mathrm{C}_n^3\frac{1}{n(n-1)(n-2)}-\cdots+(-1)^{n-1}\frac{1}{n!}\\ &=1-\frac{1}{2!}+\frac{1}{3!}-\cdots+(-1)^{n-1}\frac{1}{n!}\end{aligned}$$

1.3　条件概率与事件独立性

前面两节介绍了概率论中两个最基本的概念——事件和概率，并讨论了概率的基本性质与运算法则。本节将更深入地介绍概率论中有关条件概率与事件独立性的知识，并利用这些知识来解决一些较为复杂的实际问题。

1.3.1　条件概率与乘法公式

1. 条件概率

前面从不同的角度讨论了概率的几种一般情形及概率的基本性质，这些讨论都基于一些固定的条件限制。然而，在处理实际问题时，经常需要在已知部分试验结果的基础上来求解概率，这就需要引入条件概率的概念。

假设一次考试的题型包括双选题，要求考生从 a，b，c，d 四个选项中依次选出两个正确答案，并且只有在两个答案全部选对的条件下才能够得分。在完成某道双选题时，如果考生没有掌握该题考查的内容，随机地选择两个答案，那么可能的选择有 {(a,b),(a,c),(a,d),(b,c),(b,d),(c,d)},并且每种选择的可能性都是相同的。若以 A 记事件“两个选项全部正确”，则考生得分（即事件 A 发生）的概率为 1/6。

但是，如果在四个选项中考生能够确定其中一个选项的正确性，即考生选择的答案中“至少第一个选项正确”，例如，考生可以确定 a 选项是正确的，那么事件 A 的概率便应是 1/3。

由于引入了条件“至少第一个选项正确”，事件 A 的概率发生了变化，若记“至少第一个选项正确”为 B，这时事件 A 的概率实际上是“在事件 B 发生的条件下，事件 A 的概率”，这就是我们所说的条件概率，记为 $P(A|B)$。

在上面的例子中，初始条件下样本点总数为 $n=6$，在规定了事件 B（仍假设考生选择了正确

选项 a）发生的前提后，样本空间也随之改变为 $\Omega_B=\{(a,b),(a,c),(a,d)\}$，样本点总数为 $m_B=3$，而有利场合（至少第一个选项正确，且两个选项全部正确）数为 $m_{AB}=1$，即

$$P(A\mid B)=\frac{1}{3}=\frac{m_{AB}}{m_B}=\frac{m_{AB}/n}{m_B/n}=\frac{P(AB)}{P(B)}$$

这个式子对于条件概率来说具有一般性，下面给出条件概率的定义。

对于任意两个事件 A，B，若 $P(B)>0$，则称

$$P(A\mid B)=\frac{P(AB)}{P(B)} \tag{1-14}$$

为在事件 B 发生的条件下事件 A 发生的条件概率，简称条件概率。

【例 1-9】 某公司共有 1200 名员工，其中男性 960 人，女性 240 人。在过去的三年里，有 324 名员工得到提升，其具体情况如表 1-3 所示。试根据条件概率的公式计算：

① 若一个员工为男性，其得到提升的概率；

② 若一个员工为女性，其得到提升的概率。

表 1-3　某公司员工提升情况表

	男　性	女　性	合　计
升职人数	288	36	324
未升职人数	672	204	876
合计	960	240	1200

解：根据题意，分别以 M 记事件“某员工为男性”，以 W 记事件“某员工为女性”，以 A 记事件“某员工得到提升”。

根据表中数据，可以得到以下结果

$$P(M)=960/1200=0.80$$
$$P(W)=240/1200=0.20$$
$$P(A)=324/1200=0.27$$
$$P(AM)=288/1200=0.24$$
$$P(AW)=36/1200=0.03$$

① 根据条件概率的公式，若一个员工为男性，则其得到提升的概率为

$$P(A\mid M)=\frac{P(AM)}{P(M)}=\frac{0.24}{0.80}=0.3$$

② 若一个员工为女性，则其得到提升的概率为

$$P(A\mid W)=\frac{P(AW)}{P(W)}=\frac{0.03}{0.20}=0.15$$

2. 条件概率的性质

根据概率的性质，可以得到条件概率的一些类似的性质（以下均假定 $P(B)>0$）。

首先，条件概率也具有非负性、规范性、可列可加性三个基本性质，即

① 对于任意事件 A 和 B，有 $P(A\mid B)\geqslant 0$。

② 在事件 B 发生的条件下，必然事件 Ω 发生的概率为 1，即 $P(\Omega\mid B)=1$。

③ 对于可列个两两互不相容事件 A_1，A_2，…，以及任意事件 B，有

$$P(\bigcup_{i=1}^{\infty}A_i\mid B)=\sum_{i=1}^{\infty}P(A_i\mid B)$$

基于以上三个基本性质，可以推出条件概率的如下常用性质。

④ 对于任意事件B，有$P(\varnothing \mid B)=0$。

⑤ 对于任意n个两两互不相容事件A_1，A_2，…，A_n，有

$$P[(A_1 \cup A_2 \cup \cdots \cup A_n) \mid B]=P(A_1 \mid B)+P(A_2 \mid B)+\cdots+P(A_n \mid B)$$

⑥ 对于任意事件A和B，有$P(A \mid B)=1-P(\overline{A} \mid B)$。

⑦ 对于任意事件A_1，A_2和B，有$P[(A_1 \cup A_2) \mid B]=P(A_1 \mid B)+P(A_2 \mid B)-P(A_1A_2 \mid B)$。

特别地，当$B=\Omega$时，条件概率转化为无条件的一般概率。

3. 乘法公式

通过对条件概率公式移项，可以得到

$$P(AB)=P(B)P(A \mid B) \tag{1-15}$$

这个等式称为概率的乘法公式。

若$P(A)>0$，也可以定义$P(B \mid A)$，这时可以得到$P(AB)=P(A)P(B \mid A)$。

一般来说，可以把乘法公式推广到n个任意事件之交的场合，即当$P(A_1A_2\cdots A_{n-1})>0$时，有

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)\cdots P(A_n \mid A_1A_2\cdots A_{n-1})$$

【例1-10】 某商店出售一种零件，已知每箱装这种零件100件，且其中有4件是次品，商店采用“假一赔十”的营销方式，即顾客买一箱零件，如果随机地取1件发现是次品，商店立刻把10件合格品放入箱中，且次品不再放回。某顾客在一个箱子中先后取了3件进行测试，求这3件都不是合格品的概率。

解：根据题意，以A_i记“顾客在第i次测试时取到不合格品”（$i=1，2，3$），则有

$$P(A_1)=\frac{4}{100},\quad P(A_2 \mid A_1)=\frac{3}{99+10}=\frac{3}{109},\quad P(A_3 \mid A_1A_2)=\frac{2}{108+10}=\frac{2}{118}$$

根据乘法公式可知，顾客取出的3件都不是合格品的概率为

$$P(A_1A_2A_3)=P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2)=\frac{4}{100}\times\frac{3}{109}\times\frac{2}{118}=0.00002$$

1.3.2 事件独立性

在1.3.1节的引例中，由于引入了条件B“至少第一个选项正确后”，考生得分的概率从1/6变为了1/3，在这里，引入的条件对事件A“两个选项全部正确”的概率进行了“修正”。那么，是否存在这样一种情况，使事件B的发生与否对事件A的概率并没有影响呢？

假设袋中有a个白球、b个黑球，并且这些球除颜色外没有任何差别。现从袋中有放回地先后取两个球，若以A记“第一次取到黑球”，以B记“第二次取到黑球”，由于取球是有放回的，因此无论第一次取到什么颜色的球，第二次取到黑球的概率都会是$\frac{b}{a+b}$，即$P(B)=P(B \mid A)=P(B \mid \overline{A})=\frac{b}{a+b}$。在这种场合下，事件$A$的发生并不影响事件$B$的发生，它们之间具有某种“独立性”。

对于任意事件A和B，如果有

$$P(AB)=P(A)P(B) \tag{1-16}$$

则称事件A与B相互独立，简称A与B独立。

对于例1-10中的问题，根据已有结果$P(AM)=0.24$，而$P(M)\times P(A)=0.80\times 0.27=$

0.216，从二者的不相等性可以看出，员工的性别和员工是否升职是不独立的，即该公司在员工升职问题上存在性别歧视。

相互独立事件有如下两个性质：

性质 1　若事件 A 与 B 相互独立，则 $P(B|A)=P(B)$。

性质 2　若事件 A 与 B 相互独立，则$\bar{A}$与 B、A 与$\bar{B}$、$\bar{A}$与$\bar{B}$均独立。

【例 1-11】　甲、乙二人独立地向同一目标射击，其命中率分别为 0.6 和 0.7，试求目标被射中的概率。

解： 根据题意，以 A 记“甲射中目标”，以 B 记“乙射中目标”，以 C 记“目标被射中”，因此有 $C=A\cup B$。由于事件 A 和 B 是相互独立的，故目标被射中的概率为

$$\begin{aligned}P(C)&=P(A\cup B)\\&=P(A)+P(B)-P(AB)\\&=P(A)+P(B)-P(A)P(B)\\&=0.6+0.7-0.6\times0.7\\&=0.88\end{aligned}$$

也可以先考虑 C 的对立事件，显然有

$$\begin{aligned}P(C)&=1-P(\bar{C})=1-P(\overline{A\cup B})=1-P(\bar{A}\bar{B})=1-P(\bar{A})P(\bar{B})\\&=1-(1-0.6)\times(1-0.3)=0.88\end{aligned}$$

通过例 1-11 的求解，可以看到，在求解两个相互独立事件至少发生其一的概率问题时，可以通过求解它的逆事件，即两个相互独立事件都不发生的概率，由于相互独立事件的逆事件也一定是相互独立的，这样便简化了计算。

事件独立性的概念还可以推广到多个事件相互独立的场合，即 n 个事件相互独立，当且仅当它们中的任何 m（$2\leqslant m\leqslant n$）个事件也相互独立。

1.3.3　全概率公式

设事件 A_1，A_2，…，A_n 是样本空间 Ω 的一组事件，若这组事件满足

① $A_iA_j=\varnothing$（$i\neq j$；i，$j=1$，2，…，n）；

② $A_1\cup A_2\cup\cdots\cup A_n=\Omega$；

则称 A_1，A_2，…，A_n 是样本空间 Ω 的一个完备事件组。

假定 B 是样本空间 Ω 中的任意事件，显然有

$$(A_1\cup A_2\cup\cdots\cup A_n)\cap B=\Omega\cap B=B$$

根据事件运算的分配率，可以得到

$$B=\bigcup_{i=1}^{n}A_iB$$

显然，$A_iB(i=1,2,\cdots,n)$也是一个两两互不相容的事件组，因此，根据概率的有限可加性有

$$P(B)=P(\bigcup_{i=1}^{n}A_iB)=\sum_{i=1}^{n}P(A_iB)$$

结合式（1-15），当 $P(A_i)>0$ 时，有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B|A_i)\tag{1-17}$$

该等式称为全概率公式。

若将 A_1，A_2，…，A_n 理解为引起事件 B 发生的若干原因，那么全概率公式的意义在于，综合引起事件 B 发生的各个原因，求解事件 B 发生的概率。

在处理一些较为复杂的概率问题时，常常需要运用全概率公式，首先找到一个合适的完备事件组，将复杂事件分解为若干个互不相容的简单事件之和，然后分别计算这些简单事件的概率，再通过概率的可加性得到最终结果，即“化整为零，各个击破，积零为整”。

【例 1-12】 某车间有甲、乙、丙三条生产线共同加工一批零件，各生产线的产量分别占总产量的 40%、35% 和 25%，且在这三条生产线上加工该零件的次品率分别为 2%、4% 和 5%，求从这批零件中任意取出一个零件是次品的概率。

解： 根据题意，以 A_1 记“取出的零件来自甲生产线”，以 A_2 记“取出的零件来自乙生产线”，以 A_3 记“取出的零件来自丙生产线”，以 B 记“取出次品”，显然 A_1，A_2，A_3 构成这一随机取样试验样本空间 Ω 的完备事件组。

由已知条件可知

$$P(A_1)=40\%=0.40,\ P(A_2)=35\%=0.35,\ P(A_3)=25\%=0.25$$
$$P(B\mid A_1)=2\%=0.02,\ P(B\mid A_2)=4\%=0.04,\ P(B\mid A_3)=5\%=0.05$$

根据全概率公式，取出次品的概率为

$$\begin{aligned}P(B)&=P(A_1)P(B\mid A_1)+P(A_2)P(B\mid A_2)+P(A_3)P(B\mid A_3)\\&=0.40\times0.02+0.35\times0.04+0.25\times0.05=0.0345\end{aligned}$$

1.3.4　贝叶斯公式

在全概率公式的基础上，结合乘法公式，可以得到概率论中另一个非常重要的公式。

设 A_1，A_2，…，A_n 是样本空间 Ω 的一个完备事件组，且 $P(A_i)>0(i=1,2,\cdots,n)$，根据全概率公式，对于样本空间 Ω 中的任意事件 B 有

$$P(B)=\sum_{i=1}^{n}P(A_i)P(B\mid A_i)$$

而当 $P(B)>0$ 时，由条件概率的定义，有

$$P(A_i\mid B)=\frac{P(A_iB)}{P(B)}$$

根据乘法公式

$$P(A_iB)=P(A_i)P(B\mid A_i)$$

将以上三个式子结合，可以得到

$$P(A_i\mid B)=\frac{P(A_i)P(B\mid A_i)}{\sum_{j=1}^{n}P(A_j)P(B\mid A_j)}\tag{1-18}$$

这个等式称为贝叶斯公式。

贝叶斯公式反映了在复杂事件 B 已经发生的条件下，简单事件 A_i 发生的可能性大小，也即事件 B 的发生是由原因 A_i 引起的概率，因此 $P(A_i\mid B)$ 通常称为后验概率。相应地，$P(A_i)$ 表示的是引起事件 B 发生的各种原因发生的可能性大小，一般是根据经验事实得来的，并且在试验前已经知道，因此通常称为先验概率。

【例 1-13】 在例 1-12 中，如果取出的零件是次品，分别求这个零件是由甲、乙、丙生产线加工的概率。

解： 分别以 A_1，A_2，A_3 记取出的零件来自甲、乙、丙生产线，以 B 记“取出次品”，则根据例 1-12 的计算结果，有

$$P(B)=\sum_{i=1}^{3}P(A_i)P(B\mid A_i)=0.0345$$

根据贝叶斯公式，可以得到

$$P(A_1 \mid B) = \frac{P(A_1)P(B \mid A_1)}{\sum_{j=1}^{3} P(A_j)P(B \mid A_j)} = \frac{P(A_1)P(B \mid A_1)}{P(B)} = \frac{0.25 \times 0.05}{0.0345} = 0.3623$$

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_2)}{\sum_{j=1}^{3} P(A_j)P(B \mid A_j)} = \frac{P(A_2)P(B \mid A_2)}{P(B)} = \frac{0.35 \times 0.04}{0.0345} = 0.406$$

$$P(A_3 \mid B) = \frac{P(A_3)P(B \mid A_3)}{\sum_{j=1}^{3} P(A_j)P(B \mid A_j)} = \frac{P(A_3)P(B \mid A_3)}{P(B)} = \frac{0.4 \times 0.02}{0.0345} = 0.232$$

1.4　随机变量及其分布

概率论是研究随机现象数量规律的数学分支，为了能够更深入地研究这种规律，就需要对随机现象进行定量的数学处理，把随机现象的结果数量化，并掌握这些数量化的结果的取值规律，由此需要引入随机变量与分布函数的概念。

1.4.1　随机变量及其分布函数

1. 随机变量与分布函数

对于许多随机现象来说，其结果本身就是以数量的形式出现的，例如，掷一颗骰子可能出现的点数，一天内进入某超市的顾客数，某生产线生产的灯泡的寿命，产品抽样检查中的不合格率等。还有一些随机现象，其结果本身并不是数量的形式，如抛硬币试验，可能出现的结果为“正面朝上”或“反面朝上”，直观上它们与数值并没有直接的对应关系，但是如果将“正面朝上”指定为 1，“反面朝上”指定为 0，就可以实现结果的数量化了。简单地说，这种随机现象数量化的表现就是随机变量。

定义　设随机试验的样本空间为 Ω，若对于每个属于 Ω 的样本点 ω，总有一个实数 $X(\omega)$ 与其对应，则称实值函数 $X = X(\omega)$ 为随机变量，常用大写字母 X、Y、Z 等表示。

随机变量主要可以分为两种类型。对于一个随机变量 X，如果它的所有可能取值都能逐个列举出来，则称 X 为离散型随机变量；如果它的取值不能逐个列举，而是充满数轴上的某一区间，则称 X 为连续型随机变量。

若要全面地了解随机变量，仅仅知道它能取哪些值是不够的，更重要的是要知道它取这些值的规律，也就是说，需要掌握其概率分布。分布函数是用来刻画随机变量的概率分布的有效工具。

定义　设 X 是一个随机变量，对任意实数 x，称函数

$$F(x) = P(X \leqslant x) \tag{1-19}$$

为随机变量 X 的分布函数，记为 $X \sim F(x)$，读作 X 服从 $F(x)$。

通过分布函数，可以计算与随机变量 X 有关事件的概率。

以掷骰子试验为例，掷一颗骰子可能出现的点数 X 为一个随机变量，其可能的取值为 1，2，…，6。则事件 A“出现的点数小于等于 3”可以表示为 $A = \{X \leqslant 3\}$，通过分布函数求得 A 的概率为 $P(X \leqslant 3) = 1/2$。

2. 离散型随机变量及其分布

对于离散型随机变量，由于其所有可能取值可以一一列举出来，因此，对其概率分布定义如下。

定义　设 X 是一个离散型随机变量，并且它的所有可能取值为 x_1，x_2，…，x_n，…，则称 X 取 x_i 的概率

$$p_i = P(X = x_i),\quad i = 1, 2, \cdots, n, \cdots \tag{1-20}$$

为离散型随机变量 X 的概率分布，记为 $X \sim \{p_i\}$。

表 1-4　离散型随机变量的分布列

X	x_1	x_2	…	x_n	…
P	$p(x_1)$	$p(x_2)$	…	$p(x_n)$	…

离散型随机变量的概率分布也可以用表格形式来表示，称为分布列，如表 1-4 所示。

【例 1-14】　掷两颗骰子，若以 X 记出现的点数之和，试求 X 的分布列。

解：掷两颗骰子，可能出现的点数的组合为

$$\begin{array}{l}(1,1),(1,2),(1,3),(1,4),(1,5),(1,6)\\(2,1),(2,2),(2,3),(2,4),(2,5),(2,6)\\(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\\(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)\\(5,1),(5,2),(5,3),(5,4),(5,5),(5,6)\\(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\end{array}$$

计算可得 X 的分布列如表 1-5 所示。

表 1-5　两颗骰子点数之和的分布列

X	2	3	4	5	6	7	8	9	10	11	12
P	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

对于离散型随机变量的分布列，根据概率的非负性公理，首先一定有 $p_i \geqslant 0$。同时，由于 x_1，x_2，…，x_n，…构成样本空间的一个完备事件组，因此必有

$$\sum_{i=1}^{\infty} p_i = 1$$

如果已知离散型随机变量 X 的分布列，可以很容易地得到 X 的分布函数：

$$F(x) = P(X \leqslant x) = \sum_{x_i \leqslant x} p(x_i) \tag{1-21}$$

并且，对于任意实数 a，$b(a<b)$，有

$$P(a < x \leqslant b) = \sum_{a < x_i \leqslant b} p(x_i)$$

【例 1-15】　设随机变量 X 的分布列如表 1-6 所示。

试求 X 的分布函数。

解：根据分布列，得到 X 的分布函数如下：

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.2, & 1 \leqslant x < 2 \\ 0.2 + 0.3 = 0.5, & 2 \leqslant x < 3 \\ 0.2 + 0.3 + 0.5 = 1, & x \geqslant 3 \end{cases}$$

表 1-6　随机变量 X 的分布列

X	1	2	3
P	0.2	0.3	0.5

$F(x)$的图形呈一条阶梯状的曲线，且取值 1，2，3 处为跳跃点，其跳跃度分别为 0.2，0.3，0.5。如图 1-5 所示。

由于在求解离散型随机变量 X 的有关事件的概率时，用分布列比分布函数更方便，因此通常用分布列来描述其分布。

3. 连续性随机变量及其分布

与离散型随机变量不同，连续型随机变量的可能取值有无穷不可列个实数，这些实数覆盖数轴上的某一区间甚至整个数轴，因此不能像对离散型随机变量那样，通过分布列来描述其概率分布。在连续型随机变量的概率分布情况时，引入一个新的概念——概率密度函数。

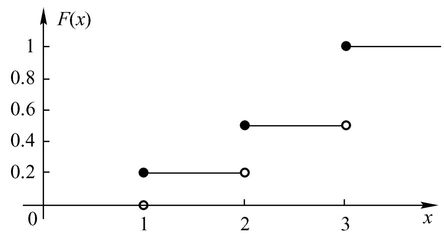


图 1-5　离散型随机变量的分布函数

定义　设随机变量 X 的分布函数是 $F(x)$，如果存在实数轴上的一个非负可积函数 $p(x)$，使得对任意实数 x，有

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} p(t)\,\mathrm{d}t \tag{1-22}$$

则称 X 为连续型随机变量，称 $p(x)$ 为 X 的概率密度函数，简称密度函数。

与离散型随机变量类似，对于连续型随机变量，其密度函数具有如下两个基本性质。

① 非负性：$p(x) \geqslant 0$。

② 正则性：$\int_{-\infty}^{+\infty} p(x)\,\mathrm{d}x = 1$。

这两条基本性质作为判别某个函数是否为密度函数的充要条件，如果连续型随机变量 X 的密度函数存在，则对于任意实数 a，b（$a<b$），有

$$P(a < x \leqslant b) = \int_{a}^{b} p(x)\,\mathrm{d}x$$

结合积分的几何意义，则 X 落在区间（a，b］上的概率等于曲线 $y=p(x)$ 在区间 $(a,b]$ 上与 x 轴构成的曲边梯形的面积，如图 1-6 所示。

不难得出，离散型随机变量 X 仅取一点时 $y=p(x)$ 与 x 轴所积面积为 0，即此时的概率恒为 0，因此在计算 X 落在某一区间上的概率时可以不用计较区间的开闭。

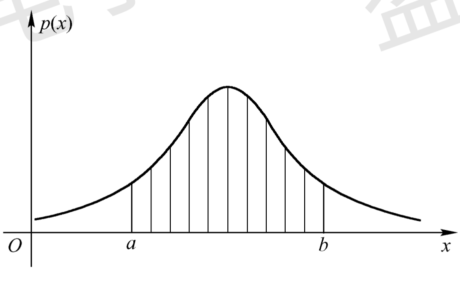


图 1-6　连续型随机变量落在区间 $(a,b]$ 上的概率

【例 1-16】　设连续型随机变量 X 的密度函数为

$$p(x) = \begin{cases} 1+x, & -1 \leqslant x < 0 \\ 1-x, & 0 \leqslant x < 1 \\ 0, & \text{其他} \end{cases}$$

试求 X 的分布函数 $F(x)$。

解：由分布函数的定义可得

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} p(t)\,\mathrm{d}t$$

当 $x < -1$ 时，$p(x)=0$，所以

$$F(x) = \int_{-\infty}^{x} p(x)\,\mathrm{d}x = 0$$

当 $-1 \leqslant x < 0$ 时，

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

当 $0 \leqslant x < 1$ 时，

$$F(x) = \int_{-\infty}^{x} p(x)\,\mathrm{d}x = \int_{-1}^{0} (1+x)\,\mathrm{d}x + \int_{0}^{x} (1-x)\,\mathrm{d}x = -\frac{x^2}{2} + x + \frac{1}{2}$$

当 $x \geqslant 1$ 时，

$$F(x)=\int_{-\infty}^{x}p(x)\,\mathrm{d}x=\int_{-1}^{0}(1+x)\,\mathrm{d}x+\int_{0}^{1}(1-x)\,\mathrm{d}x=1$$

综上所述，X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1\\ \dfrac{x^2}{2}+x+\dfrac{1}{2}, & -1\leqslant x<0\\ -\dfrac{x^2}{2}+x+1, & 0\leqslant x<1\\ 1, & x\geqslant 1\end{cases}$$

1.4.2　随机变量的数字特征

随机变量的概率分布能够完整地描述随机变量的统计特征，并且据此可以求得与随机变量有关事件的概率。然而，在一些场合中，并不需要了解随机变量的全面情况，而只须从某个侧面考察随机变量的特征。例如，假设某地区成年男子的身高为随机变量 X，在统计该地区男子的身高情况时，只须注意男子的平均高度，以及个体的身高与平均身高的偏离程度。这种用数字表示的随机变量的特征称为随机变量的数字特征。

本节将主要介绍随机变量的常用数字特征：数学期望、方差和标准差。

1. 数学期望

数学期望表示随机变量所有可能取值的平均水平，记为 $E(X)$ 或 μ。下面，对于离散型随机变量和连续型随机变量，分别给出数学期望的定义和性质。

（1）离散型随机变量的数学期望

定义　设离散型随机变量 X 的所有可能取值为 x_1，x_2，…，x_n，…，且 $X\sim\{p_i\}$，如果

$$\sum_{i=1}^{+\infty}x_ip_i$$

绝对收敛，则称

$$E(X)=\sum_{i=1}^{+\infty}x_ip_i \tag{1-23}$$

为随机变量 X 的数学期望，简称期望；否则，称 X 的数学期望不存在。

从定义可以看出，求解离散型随机变量 X 的数学期望，也就是求解 X 的所有可能取值为 x_1，x_2，…，x_n，…关于权 p_1，p_2，…，p_n，…的加权平均值。

【例 1-17】　试求例 1-14 中随机变量 X 的数学期望。

解：根据表 1-5 中的计算结果，计算得到 X 的数学期望为

$$E(X)=2\times\frac{1}{36}+3\times\frac{1}{18}+4\times\frac{1}{12}+5\times\frac{1}{9}+6\times\frac{5}{36}+7\times\frac{1}{6}+8\times\frac{5}{36}$$
$$+9\times\frac{1}{9}+10\times\frac{1}{12}+11\times\frac{1}{18}+12\times\frac{1}{16}=7$$

（2）连续型随机变量的数学期望

定义　设连续型随机变量 X 的概率密度函数为 $p(x)$，如果

$$\int_{-\infty}^{+\infty}xp(x)\,\mathrm{d}x$$

绝对收敛，则称

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\,\mathrm{d}x \tag{1-24}$$

为随机变量 X 的数学期望，简称期望。同样地，如果级数的收敛条件不成立，则称 X 的数学期望不存在。

【例 1-18】 试求例 1-16 中随机变量 X 的数学期望。

解： 已知 X 的密度函数为

$$p(x)=\begin{cases}1+x, & -1\leqslant x<0\\ 1-x, & 0\leqslant x<1\\ 0, & \text{其他}\end{cases}$$

因此 X 的数学期望为

$$E(X)=\int_{-\infty}^{+\infty}xp(x)\mathrm{d}x=\int_{-1}^{0}x(1+x)\mathrm{d}x+\int_{0}^{1}x(1-x)\mathrm{d}x=0$$

(3) 数学期望的性质

假定以下所涉及随机变量的数学期望均存在，根据数学期望的定义可以得出下列性质。

性质 1　对于任意常数 c，有

$$E(c)=c \tag{1-25}$$

性质 2　对于任意随机变量 X 和常数 a，b，有

$$E(aX)=aE(X) \tag{1-26}$$

$$E(X+b)=E(X)+b \tag{1-27}$$

性质 3　对于任意随机变量 X 和 Y，有

$$E(X\pm Y)=E(X)\pm E(Y) \tag{1-28}$$

性质 4　对于任意随机变量 X 和 Y，若 X 与 Y 相互独立，有

$$E(XY)=E(X)E(Y) \tag{1-29}$$

2. 方差与标准差

随机变量 X 的数学期望在一定程度上反映了随机变量的集中趋势，它反映了 X 的取值总在 $E(X)$ 周围波动，但是却不能反映出这种波动的大小，即 X 的取值与 $E(X)$ 的偏离程度。例如，在统计某地区成年男子的身高情况时，不仅要注意男子的平均高度，还要观察个体的身高与平均身高的偏离程度。为了度量这种偏离程度，下面引入方差和标准差的概念。

定义　设 X 是一个随机变量，若 $E[X-E(X)]^2$ 存在，则称

$$D(X)=E[X-E(X)]^2 \tag{1-30}$$

为随机变量 X 的方差，称 $\sigma(X)=\sqrt{D(X)}$ 为随机变量 X 的标准差。

根据定义可知，方差实际上就是随机变量 X 的取值相对于均值 $E(X)$ 的偏差平方的数学期望，这是由于偏差 $X-E(X)$ 的值有正有负，直接相加则会出现正负抵消的现象，因此利用偏差的平方来计算随机变量 X 的方差，然后对得到的方差开平方，就得到了与数学期望的量纲相同的标准差。

以上是方差的一般定义，结合数学期望的计算公式，针对不同类型的随机变量，有如下结论。

① 对于离散型随机变量 X，如果 X 的所有可能取值为 x_1，x_2，…，x_n，…，且 $X\sim\{p_i\}$，则有

$$D(X)=\sum_{i=1}^{+\infty}[x_i-E(X)]^2p_i \tag{1-31}$$

在例 1-17 中，

$$
\begin{aligned}
D(X)=&(2-7)^2\times\frac{1}{36}+(3-7)^2\times\frac{1}{18}+(4-7)^2\times\frac{1}{12}+(5-7)^2\times\frac{1}{9}+\\
&(6-7)^2\times\frac{5}{36}+(7-7)^2\times\frac{1}{6}+(8-7)^2\times\frac{5}{36}+(9-7)^2\times\frac{1}{9}+(10-7)^2\times\\
&\frac{1}{12}+(11-7)^2\times\frac{1}{18}+(12-7)^2\times\frac{1}{36}\\
=&5.83
\end{aligned}
$$

$$\sigma(X)=\sqrt{D(X)}=\sqrt{5.83}=2.42$$

② 对于连续型随机变量 X，如果 X 的概率密度函数为 $p(x)$，则有

$$D(X)=\int_{-\infty}^{+\infty}[x-E(x)]^2p(x)\mathrm{d}x \tag{1-32}$$

在例1-18中，

$$D(X)=\int_{-\infty}^{+\infty}[x-E(x)]^2p(x)\mathrm{d}x=\int_{-1}^{0}(x-0)^2(1+x)\mathrm{d}x+\int_{0}^{1}(x-0)^2(1-x)\mathrm{d}x=0$$

$$\sigma(X)=\sqrt{D(X)}=0$$

③ 方差的性质。若以下随机变量的数学期望和方差均存在，根据方差的定义容易得出下列性质。

性质1　对于任意常数 c，有

$$D(c)=0 \tag{1-33}$$

性质2　对于任意随机变量 X 和常数 a，b，有

$$D(aX)=a^2D(X) \tag{1-34}$$

$$D(X+b)=D(X) \tag{1-35}$$

性质3　对于任意随机变量 X 和 Y，若 X 与 Y 相互独立，有

$$D(X+Y)=D(X)+D(Y) \tag{1-36}$$

性质4　对于任意随机变量 X，有

$$D(X)=E(X^2)+[E(X)]^2 \tag{1-37}$$

1.4.3　常用的离散型分布

在实际问题中，常常会遇到许多不同类型的离散型随机变量，下面就来介绍四种常见的离散型分布。

1. 0-1分布

若离散型随机变量 X 的概率分布为

$$P(X=1)=p,\ P(X=0)=1-p,\qquad (0<p<1) \tag{1-38}$$

则称随机变量 X 服从参数为 p 的0-1分布，记为 $X\sim B(1,p)$。0-1分布也称为两点分布或伯努利分布。

例如，从一批产品中任取一个做测试，以 $X=1$ 记产品是好品，以 $X=0$ 记产品是废品，若这批产品的合格率为90%，则 X 服从参数为0.9的0-1分布，记为 $X\sim B(1,0.9)$，其分布列如表1-7所示。

表1-7　0-1分布的分布列

X	0	1
P	0.1	0.9

0-1分布的数学期望、方差和标准差分别为

$$E(X)=1\times p+0\times(1-p)=p$$

$$D(X)=(1-p)^2\times p+(0-p)^2\times(1-p)=p(1-p)$$
$$\sigma(X)=\sqrt{D(X)}=\sqrt{p(1-p)}$$

2. 二项分布

在处理实际问题时，常常会遇到只有两种可能结果的试验。例如，在产品抽样调查中，随机抽取的某个产品可能是合格品，也可能是废品；在调查新出台的政策是否符合民意时，对参与调查的某个公民来说可能是支持这项政策，也可能是反对这项政策等。在这些问题中，每次试验都只有两种可能出现的结果：事件 A 发生或事件 $\overline{A}$ 发生，并且在每次试验中事件 A 出现的概率都相同，记为 $p(0<p<1)$，如果将这种只有两种结果的试验在相同条件下重复独立地进行 n 次，那么这些试验便构成了一个新的试验，称为 n 重伯努利试验。

若以 X 记 n 重伯努利试验事件 A 出现的次数，X 的可能取值为 1，2，…，n，则 X 的概率分布为

$$P(X=k)=C_n^k p^k(1-p)^{n-k},\ k=0,1,2,\cdots,n \tag{1-39}$$

在 n 重伯努利试验中，称事件 A 出现的次数 X 服从以 n，p 为参数的二项分布，记为 $X\sim B(n,p)$。

二项分布是一种常用的离散分布，特别是当 $n=1$ 时，k 只能取 0 和 1，且 $P(X=1)=p$，$P(X=0)=1-p$，也就是前面所说的 0 – 1 分布。

二项分布的数学期望、方差和标准差分别为

$$E(X)=np$$
$$D(X)=np(1-p)$$
$$\sigma(X)=\sqrt{np(1-p)}$$

【例 1–19】 在 10 件产品中混入了 2 件次品，现有放回地先后取出 3 件产品，用随机变量 X 表示次品数，试求 X 的分布列、$E(X)$ 和 $D(X)$。

解： 由于抽样是有放回的，因此每次取出次品的概率都相同，这是一个 n 重伯努利试验。随机变量 X 的可能取值为 0，1，2，3，且

$$P(X=0)=C_3^0\left(\frac{2}{10}\right)^0\left(1-\frac{2}{10}\right)^3=\frac{64}{125}$$

$$P(X=1)=C_3^1\left(\frac{2}{10}\right)^1\left(1-\frac{2}{10}\right)^2=\frac{48}{125}$$

$$P(X=2)=C_3^2\left(\frac{2}{10}\right)^2\left(1-\frac{2}{10}\right)^1=\frac{12}{125}$$

$$P(X=3)=C_3^3\left(\frac{2}{10}\right)^3\left(1-\frac{2}{10}\right)^0=\frac{1}{125}$$

故 X 的分布列如表 1–8 所示。且有

$$E(X)=3\times\frac{2}{10}=\frac{3}{5}$$

$$D(X)=3\times\frac{2}{10}\times\left(1-\frac{2}{10}\right)=\frac{12}{25}$$

表 1–8　产品质量抽样检查中次品数的分布列

X	0	1	2	3
P	64/125	48/125	12/125	1/125

3. 泊松分布

（1）泊松分布的概率分布

若随机变量 X 的概率分布为

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},\quad \lambda>0,\ k=0,1,2,\cdots \tag{1-40}$$

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

泊松分布是由法国数学家泊松于 1937 年引入的，其应用十分广泛。在现实问题中，许多随机现象都服从泊松分布，如在单位时间内电话交换台接到的用户呼叫数，1m^2 内玻璃上的气泡数，单位时间内公共汽车站来到的乘客数等。

泊松分布的数学期望、方差和标准差分别为

$$E(X)=\lambda$$
$$D(X)=\lambda$$
$$\sigma(X)=\sqrt{\lambda}$$

【例 1-20】 假定某航空公司预订票处十分钟内接到订票电话的次数服从参数为 7 的泊松分布，试求订票处在十分钟内恰好接到 6 次电话的概率。

解： 以随机变量 X 表示订票处在 10 分钟内接到订票话的次数，则 $X\sim P(7)$，故

$$P(X=6)=\frac{7^6}{6!}e^{-7}$$

对这个式子直接计算会比较麻烦，可以利用泊松分布表来求解，当 $k=6$，$\lambda=7$ 时

$$P(X=6)=\frac{7^6}{6!}e^{-7}=0.149$$

（2）二项分布的泊松近似

泊松分布常常被看成二项分布的近似：在 n 重伯努利试验中，当试验次数 n 很大，而事件 A “成功”发生的概率 p 很小时，二项分布可以用 $\lambda=np$ 的泊松分布来近似。

【例 1-21】 已知某种疾病的发病率为 0.001，某地区共有 5000 居民，现有一医疗团队为该地区居民义务会诊，试求该地区患有这种疾病的人数不超过 5 人的概率。

解： 以随机变量 X 记该地区患有这种疾病的人数，则 $X\sim B$（5000，0.001），所以有

$$P(X\leqslant 5)=\sum_{k=0}^{5}C_{5000}^{k}0.001^k\,0.999^{5000-k}$$

通过二项分布来求解这个问题计算量是很大的，由于 n 很大，而 p 很小，这时可以利用泊松分布来求解

$$\lambda=np=5000\times 0.001=5$$
$$P(X\leqslant 5)=\sum_{k=0}^{5}\frac{5^k}{k!}e^{-5}=0.616$$

4．超几何分布

在产品检验问题中，常常会遇到采取不放回抽样的情况，例如，对电灯泡寿命的检验和棉纱强度的检验，由此需要引入超几何分布的概念。

设一批产品共有 N 件，其中有 M 件不合格品，现从这 N 件产品中不放回地先后抽取 n 件，则其中含有的不合格品的个数 X 服从超几何分布，记为 $X\sim H(n,N,M)$。超几何分布的概率分布列为

$$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n},\quad k=0,1,2,\cdots,m \tag{1-41}$$

式中，$m=\min\{M,n\}$，且 $M\leqslant N$，$n\leqslant N$，n，N，M 均为正整数。

超几何分布的数学期望、方差和标准差分别为

$$E(X)=n\frac{M}{N}$$

$$D(X)=\frac{nM(N-M)(N-n)}{N^2(N-1)}$$

$$\sigma(X)=\sqrt{\frac{nM(N-M)(N-n)}{N^2(N-1)}}$$

特别地，在实际问题中，当抽样的个数远远小于产品的总数时，每次抽样之后总体中的不合格率 $p=M/N$ 改变甚微，这时，不放回抽样可以近似地看成有放回抽样，因此可以计算二项分布作为近似值。

1.4.4　常用的连续型分布

1. 均匀分布

若随机变量 X 的密度函数为

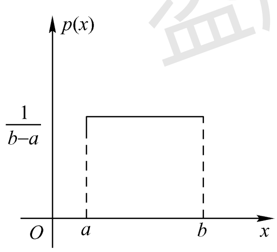
$$p(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其他}\end{cases} \tag{1-42}$$

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

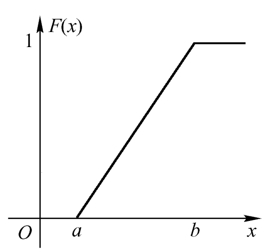
相应地，X 的分布函数为

$$F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\leqslant x<b\\ 1, & x\geqslant b\end{cases} \tag{1-43}$$

区间（a，b）上的均匀分布如图 1-7 所示。



（a）密度函数$p(x)$



（b）分布函数$F(x)$

图 1-7　区间(a，b)上的均匀分布

若随机变量 X 服从区间(a，b)上的均匀分布，则 X 在(a，b)中取值落在某一区域内的概率与这个区域的测度成正比。

均匀分布的数学期望、方差和标准差分别为

$$E(X)=\frac{a+b}{2}$$

$$D(X)=\frac{(b-a)^2}{12}$$

$$\sigma(X)=\sqrt{\frac{(b-a)^2}{12}}$$

【例 1-22】 设随机变量 X 服从(0，10)上的均匀分布，试求 $P(3<x<7)$ 与 $P(5<x\leqslant 12)$。

解： 由 X 服从(0，10)上的均匀分布可知

$$p(x)=\begin{cases}0.1, & 0<x<10\\ 0, & 其他\end{cases}$$

因此

$$P(3<x<7)=\int_3^7 0.1\mathrm{d}x=0.4$$

$$P(5<x\leqslant 12)=\int_5^{12} p(x)\mathrm{d}x=\int_5^{10} 0.1\mathrm{d}x=0.5$$

2. 指数分布

若随机变量 X 的密度函数为

$$p(x)=\begin{cases}\lambda \mathrm{e}^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases} \tag{1-44}$$

则称 X 服从参数为 λ 的指数分布，记为 $X\sim \mathrm{Exp}(\lambda)$，其中 $\lambda>0$。

相应地，X 的分布函数为

$$F(x)=\begin{cases}1-\mathrm{e}^{-\lambda x}, & x\geqslant 0\\ 0, & x<0\end{cases} \tag{1-45}$$

指数分布常常用于表示各种“寿命”分布，如无线电元件的生命周期、动物的寿命、电话的通话时间、随机服务系统中的服务时间等，都可假定服从指数分布。

指数分布的数学期望、方差和标准差分别为

$$E(X)=\frac{1}{\lambda}$$

$$D(X)=\frac{1}{\lambda^2}$$

$$\sigma(X)=\sqrt{\frac{1}{\lambda^2}}=\frac{1}{\lambda}$$

【例 1-23】 假设某商店从早晨开始营业起直到第一个顾客到达的等待时间（分钟）服从参数 $\lambda=0.4$ 的指数分布，试求等待时间不超过 3 分钟的概率。

解： 根据题意，可知等待时间 X 的分布函数为

$$F(x)=\begin{cases}1-\mathrm{e}^{-0.4x}, & x\geqslant 0\\ 0, & x<0\end{cases}$$

因此，等待时间不超过 3 分钟的概率为

$$P(X\leqslant 3)=F(3)=1-\mathrm{e}^{-0.4\times 3}=1-\mathrm{e}^{-1.2}=0.699$$

3. 正态分布

正态分布是连续型随机变量的一个最重要的分布，它对于统计研究具有十分重要的意义。在自然界和社会经济问题中，许多随机现象都可以用正态分布来描述或近似描述，如测量的误差、炮弹落地点的分布、人的身高和体重、农作物的收获量、年降雨量等都近似服从正态分布。

（1）正态分布的密度函数和分布函数

若随机变量 X 的密度函数为

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty<x<+\infty \tag{1-46}$$

则称 X 服从正态分布，记为 $X\sim N(\mu,\sigma^2)$，其中参数 $-\infty<\mu<+\infty,\sigma>0$。

$y=p(x)$ 的图形关于 $x=\mu$ 对称，且在 $(-\infty,\mu)$ 单调递增，在 $(\mu,+\infty)$ 内单调递减，在 $x=\mu$ 时达到最大值，如图 1–8 所示。

相应地，X 的分布函数为

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{1}{2\sigma^2}(t-\mu)^2}dt \tag{1-47}$$

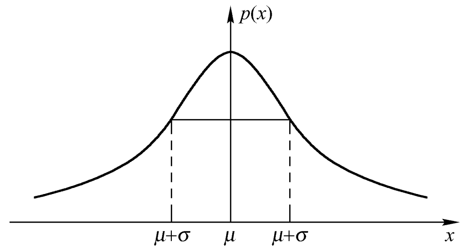
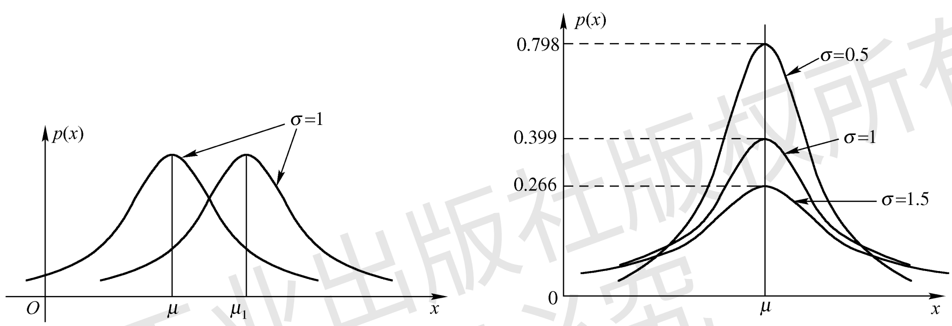


图 1–8　正态分布的密度函数

正态分布的密度函数具有如下性质：

- 如果保持 σ 不变，改变 μ 的值，则曲线沿 x 轴平移，而曲线的形状不改变。也就是说正态密度函数在平面直角坐标系的位置是由参数 μ 确定的。
- 如果保持 μ 不变，改变 σ 的值，则曲线随着 σ 值的增加而变得平缓，或随着 σ 值的减小而变得陡峭，而曲线的中心位置保持不变。也就是说正态密度函数的尺度由参数 σ 所确定。

正态密度函数图形与参数的关系如图 1–9 所示。



（a）保持σ不变，改变μ值　　（b）保持μ不变，改变σ值

图 1–9　正态密度函数图形与参数的关系

正态分布的数学期望、方差和标准差分别为

$$E(X)=\mu$$
$$D(X)=\sigma^2$$
$$\sigma(X)=\sigma$$

故正态分布 $X\sim N(\mu,\sigma^2)$ 通常读为随机变量 X 服从均值为 μ，方差为 σ^2 的正态分布。

（2）标准正态分布

当参数 $\mu=0$，$\sigma=1$ 时，正态分布 $N(0,1)$ 称为标准正态分布。对于标准正态分布，通常用 $\varphi(x)$ 表示密度函数，用 $\Phi(x)$ 表示分布函数，故

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \tag{1-48}$$

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt \tag{1-49}$$

对于服从标准正态分布的随机变量，可以通过“正态分布表”查得 $\Phi(x)$，然后通过一定的换算得到所要求解的概率，主要的换算法则有

- $\Phi(-x)=1-\Phi(x)$
- $P(X>x)=1-\Phi(x)$
- $P(a<X<b)=\Phi(b)-\Phi(a)$
- $P(|X|<c)=2\Phi(c)-1$

实际上，恰好服从标准正态分布的随机变量很少，但可以通过一定的线性变换将一般正态分布转化为标准正态分布。对于一般正态分布，进行变换

$$Z=\frac{X-\mu}{\sigma} \tag{1-50}$$

便可将其转换为标准正态分布 $Z \sim N(0,1)$，这时，就可以借助于正态分布表来求解一般正态分布问题了。

【例 1-24】 王某家住市区西郊，工作单位位于东郊。王某的上班的路线可以有两种选择：一是横穿市区，这条路线路程较短，但交通堵塞严重，所需时间 $X \sim N(30,100)$；二是选择环城公路，这条路线路程较远，但堵塞少，所需时间 $Y \sim N(40,16)$。

① 若距上班时间还有 50 分钟，应选择哪条路线？

② 若距上班时间还有 45 分钟，又应选择哪条路线？

解：根据题意，设王某选择第一条路线需要花费的时间为 x，选择第二条路线需要花费的时间为 y。

① 若距离上班时间还有 50 分钟，则对于两条路线，王某准时上班的概率分别为

$$P(x \leqslant 50)=P\left(\frac{x-30}{10} \leqslant \frac{50-30}{10}\right)=\Phi(2)=0.9772$$

$$P(y \leqslant 50)=P\left(\frac{y-40}{4} \leqslant \frac{50-40}{4}\right)=\Phi(2.5)=0.9938$$

此时选择第二条路线时准时上班的概率大于选择第一条路线，因此应选第二条路线。

② 若距离上班时间还有 45 分钟，则对于两条路线，王某准时上班的概率分别为

$$P(x \leqslant 45)=P\left(\frac{x-30}{10} \leqslant \frac{45-30}{10}\right)=\Phi(1.5)=0.9332$$

$$P(y \leqslant 45)=P\left(\frac{y-40}{4} \leqslant \frac{45-40}{4}\right)=\Phi(1)=0.8944$$

此时选择第一条路线时准时上班的概率大于选择第二条路线，因此应选第一条路线。

(3) 3σ 原则

如果随机变量 $X \sim N(0,1)$，则

$$P(|X-\mu| \leqslant \sigma)=0.6826$$

$$P(|X-\mu| \leqslant 2\sigma)=0.9545$$

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

显然，尽管正态变量 X 的取值范围是 $(-\infty,+\infty)$，但是它的取值有 99.73% 落在区域 $(\mu-3\sigma,\ \mu+3\sigma)$ 内，因此常把在此范围之外的随机变量取值忽略不计，这一性质在统计学上称为 3σ 原则。3σ 原则在实际问题中有许多应用，例如，在工业生产上，一些产品质量指数就是根据 3σ 原则制定的。

4. χ^2 分布，t 分布，F 分布

下面来介绍三种特殊的连续型分布，它们是由若干正态随机变量构成的特殊函数，在参数估计、假设检验等方面具有重要应用。

① 设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立且都服从 $N(0,1)$，则随机变量 $Y=\sum\limits_{k=1}^{n} X_k^2$ 服从自由度为 n 的 χ^2 分布，记为 $Y \sim \chi^2(n)$。

② 设随机变量 $X \sim N(0,1)$，$Y \sim \chi^2(n)$，且它们相互独立，则随机变量 $Z=\dfrac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布，记为 $Z \sim t(n)$。

③ 设随机变量 X 和 Y 相互独立，且 $X \sim \chi^2(m), Y \sim \chi^2(n)$，则随机变量 $Z = \dfrac{X/m}{Y/n}$ 服从第一自由度为 m，第二自由度为 n 的 F 分布，记为 $Z \sim F(m,n)$。

1.5　案　　例

案例 1.1　概率论在可靠性检验中的应用

随着社会生产力的迅速发展，科学技术日新月异，在科学研究和生产实践中出现了许多大型产品和复杂的系统工程，如美国曼哈顿计划研制的原子弹、海军研制的“北极星导弹潜艇”、火箭发射、人造卫星、阿波罗宇宙飞船等。对于这些大型产品和系统工程来说，对其可靠性提出的要求是空前的。以宇航工业产品为例，宇航工业产品的可靠性要求达到 99.9999%，即这项极为复杂的系统工程在 100 万次动作中，只允许有一次失灵。它们所用的电子元件、器件、机械零件等，持续安全运转工作时间要在 1 亿小时，甚至 10 亿小时。那么，可靠性应当如何定义呢？

对于一个元件，将它能够正常工作的概率称为它的可靠性。若由若干个元件组成一个系统，则这个系统正常工作的概率称为该系统的可靠性。如何从组成元件的可靠性寻求系统的可靠性，或者如何从系统可靠性的要求反过来求出元件的可靠性，在实际问题中具有十分现实的意义。

元件的连接有两种基本的方式：串联和并联。

1. 串联

把几个元件首尾相连组成一条通路，这样的系统称为串联系统。在串联系统中，只有当所有元件都能够正常工作时，系统才能够正常工作。

2. 并联

把几个元件首首相连、尾尾相连组成一条通路，这样的系统称为并联系统。在并联系统中，只要有一个以上的元件能够正常工作，系统就能正常工作。

已知有 $2n$ 个完全相同的元件，每个元件的可靠性都是 p。现将这 $2n$ 个元件以 4 种方式连接成为一个系统，如图 1-10 所示。

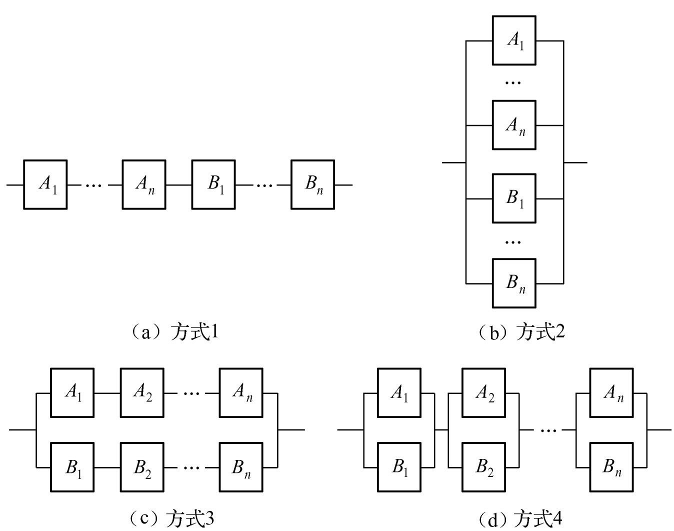


图 1-10　$2n$ 个元件的 4 种连接方式

思考题：

① 试求每个系统的可靠性；

② 比较各个系统的可靠性，得出可靠性大小的顺序。

案例 1.2　概率论在民事纠纷中的应用

在民事纠纷案件中，受害人将案件提交法院诉讼，除了要考虑胜诉的可能性外，还应考虑到诉讼费用的负担。理性的当事人往往通过私下协商赔偿费用而趋于和解，免于起诉。

在一次交通事故案件中，司机（致害人）开车撞伤了受害人，使受害人遭受了 10 万元的经济损失。假若将案件提交诉讼，则诉讼费用共需要 0.4 万元，并按所负责任的比例由双方承担。从事故发生的情形分析，法院对事故判决可能有三种情况：

（1）致害人应承担 100% 的责任，要向受害人赔偿 10 万元的损失费用，并支付全部 0.4 万元的诉讼费；

（2）致害人应承担 70% 的责任，要向受害人赔偿 7 万元的损失费用，并支付 0.4 万元诉讼费的 70%，诉讼费另外的 30% 由受害人支付；

（3）致害人应承担 50% 的责任，要向受害人赔偿 5 万元的损失费用，0.4 万元的诉讼费由双方各负担一半。

受害人估计情况（1）、（2）、（3）发生的概率分别为 0.2、0.6 和 0.2。

（资料来源：陈卫东．离散型随机变量的数学期望在法律、医学和经济问题中的应用［J］．广东广播大学学报，2005，14（14）：103 ~ 104.）

思考题：

（1）如果受害人将案件提交诉讼，列出其获得收益的分布列。

（2）如果致害人希望私下和解而免于起诉，他应至少给受害人多少数额的赔偿费，才会使受害人从经济收益上考虑而趋于和解？

习　题　1

1. 试针对样本空间中样本点的不同类型，列举出几个日常生活中的随机现象。

2. 设有甲、乙两种产品，现分别从这两种产品中取出一件，若以 A 记从甲产品中取出次品，以 B 记从以产品中取出次品，试表示如下事件：

① 两件产品都是次品；

② 至少有一件产品是次品；

③ 恰好有一件产品是次品。

3. 设有 n 个球，每个球都等可能地被放到 N 个不同的盒子中的任一个，每个盒子所放球数不限。试求：

① 指定的 n（$n \leqslant N$）个盒子中各有一球的概率 p_1；

② 恰好有 n（$n \leqslant N$）个盒子中各有一球的概率 p_2。

4. 从 1，2，…，10 这十个数中任取一个，假定各个数都以同样的概率被取中，取后还原，先后取 7 个数，试求：

① 7 个数全不相同的概率；

② 7 个数中不含 9 和 2 的概率；

③ 8 恰好出现三次的概率；

④ 5 至少出现两次的概率；

⑤ 取到的最大数为 6 的概率。

5. 某码头只能容纳一只船，现预知某日将独立来两只船，且在 24 小时内各时刻来到的可能性都相等，如果它们需要停靠的时间分别为 3 小时和 4 小时，试求有一只船要在江中等候的概率。

6. 口袋中装有 10 个球，其中有 3 个黑球，7 个白球，先后两次从袋中各取一球（不放回）。

① 已知第一次取出的是黑球，求第二次取出的仍是黑球的概率；

② 已知第二次取出的是黑球，求第一次取出的也是黑球的概率。

7. 设某光学仪器厂制造的透镜，第一次落下时被打破的概率为 1/2，若第一次落下未打破，第二次落下打破的概率为 7/10，若前两次落下未打破，第三次落下打破的概率为 9/10。试求透镜落下三次而未打破的概率。

8. 三人独立地去破译一份密码，已知他们能破译该密码的概率分别为 1/5，1/3，1/4，试求该密码被破译的概率。

9. 两台车床加工同样的零件，第一台出现不合格品的概率是 0.03，第二台出现不合格品的概率是 0.06，加工出来的零件放在一起，并且已知第一台加工的零件比第二台加工的零件多一倍。

① 求任取一个零件是合格品的概率；

② 如果取出的零件是不合格品，求它是由第二台车床加工的概率。

10. 一批产品共有 100 件，其中 10 件是不合格品。根据验收规则，从中任取 5 件产品进行质量检验，假如 5 件产品中无不合格品，则这批产品被接收，否则就要重新对这批产品逐个检验。

① 试求 5 件产品中不合格品数 X 的分布列；

② 需要对这批产品进行逐个检验的概率是多少？

11. 一种电子管的使用寿命 X（单位：小时）的概率密度函数为

$$p(x)=\begin{cases}\dfrac{1000}{x^2}, & x\geqslant 1000\\ 0, & x<1000\end{cases}$$

设某种仪器中装有 5 个这种工作相互独立的电子管，试求：

① 使用最初 1500 小时没有一个电子管损坏的概率；

② 这段时间内至少有两个电子管损坏的概率。

12. 设随机变量 X 的概率分布为

X	0	1	2	3
概　率	1/8	1/4	3/8	1/4

求 $E(X)$ 和 $D(X)$。

13. 设随机变量 X 的概率密度函数为

$$p(x)=\begin{cases}ax+\dfrac{1}{2}, & 0\leqslant x\leqslant 1\\ 0, & \text{其他}\end{cases}$$

已知 $E(X)=\dfrac{7}{12}$，试确定 a 的值，并求分布函数 $F(x)$。

14. 设事件 A 在每次独立试验中发生的概率为 p，当 A 发生不少于 3 次时，指示灯发出信号。若

① 进行了 5 次独立试验，试求指示灯发出信号的概率；

② 进行了 7 次独立试验，试求指示灯发出信号的概率。

15. 一批产品的不合格率为 0.03，现从中任取 40 只进行检查，若发现两只或两只以上不合格品就拒收这批产品。分别用以下方法求拒收的概率：

① 用二项分布做精确计算；

② 用泊松分布做近似计算。

16. 某公共汽车站每隔 10 分钟有一辆车通过，一个乘客随机到达此车站候车，候车时间 X 服从［0，10］上的均匀分布。试求，这个人至少等候 6 分钟的概率。

17. 设顾客在某银行的窗口等待服务的时间（单位：分钟）X 服从指数分布，其概率密度函数为

$$p(x)=\begin{cases}\dfrac{1}{5}\mathrm{e}^{-x/5}, & x>0\\ 0, & 其他\end{cases}$$

某顾客在窗口等待服务，若超过 10 分钟，他就离开。试求该顾客等到服务的概率。

18. 公共汽车车门的高度是按男子与车门顶碰头概率在 0.01 以下来设计的，设男子身高 $X\sim N(170,6^2)$，问车门高度应如何确定？

19. 设某幢建筑物使用寿命（单位：年）X 服从正态分布 $N(50,100)$，已知这幢建筑物已经使用了 30 年，试求它还可以再使用 30 年的概率。

20. 试根据引例中的数据，计算得出年历销售量的概率分布表和收益与收益期望分布状况表，并给出计算过程。