$$= \quad 3/4 - 1/8 + P(ABC).$$

根据 $P(AB) = 0$ 和 $ABC \subset AB$ 可知

$$0 \leqslant P(ABC) \leqslant P(AB) = 0$$

由此可知事件 A, B, C 中至少有一个事件发生的概率为 5/8.

1.3 古典概型与几何概型

本节介绍两种历史较为久远的经典概率模型: 古典概型与几何概型.

1.3.1 古典概型

首先研究一类简单的随机现象, 它是概率论早期最重要的研究对象, 其发展在概率论中具有重要的意义, 并在产品质量抽样检测等问题中具有广泛的应用.

定义 1.6 (古典概型) 如果试验 E 满足:

- 试验的结果只有有限种可能, 即样本空间 $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, 其中 w_i 为基本事件,
- 每种结果发生的可能性相同, 即 $P(\{\omega_i\}) = P(\{\omega_j\})$ $(i \neq j)$,

则称该类试验称为 **古典概型**, 又称 **等可能概型**.

根据上述定义以及 $P(\Omega) = 1$ 可知: 每个基本事件发生的概率为 $P(\{\omega_i\}) = 1/n$, 若事件 A 包含 k 个基本事件 $\{\omega_{i_1}, \omega_{i_1}, \ldots, \omega_{i_k}\}$, 则事件 A 发生的概率为

$$P(A) = k/n = |A|/|\Omega|,$$

这里 $|A|$ 表示事件 A 包含的事件的个数. 很显然古典概型的概率满足概率公理化体系的三条公理.

计算古典概率的本质是计数 (Counting), 计数是组合学研究的重要内容, 我们将在 1.4 节详细的介绍各种计数方法, 这里仅仅介绍一些基本原理和排列组合:

- **加法原理**: 若一项工作可以用两种不同的过程 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 完成, 且过程 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 分别有 n_1 和 n_2 种方法, 则完成该工作有 $n_1 + n_2$ 种方法.
- **乘法原理**: 若一项工作需要依次通过 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 两过程, 且过程 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 分别有 n_1 和 n_2 种方法, 则完成该工作有 $n_1 \times n_2$ 种方法.

上述两条原理可进一步推广到多个过程的情况.

排列: 从 n 个不同的元素中无放回地取出 r 个元素进行排列, 此时既要考虑取出的元素, 也要顾及其排列顺序, 则有 $(n)_r = n(n-1)\cdots(n-r+1)$ 种不同的排列. 若 $r = n$ 时称全排列, 有 $n!$ 种.

组合: 从 n 个不同的元素中无放回地取出 r 个元素, 取出的元素之间无顺序关系, 共有 $\binom{n}{r}$ 种不同的取法, 其中

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{(n)_r}{r!}, \quad \text{且记} \quad \binom{n}{0} = 1.$$

这里 $\binom{n}{r}$ 称为 **组合数** 或 **二项系数**, 它是二项展开式 $(a+b)^n = \sum_{r=0}^{n}\binom{n}{r}a^r b^{n-r}$ 中项 $a^r b^{n-r}$ 的系数.

很多经典的数学问题都可归纳为古典概型, 下面介绍一些典型的例子:

例 1.6 将 n 个不同的球随机放入 N $(N \geqslant n)$ 个不同的盒子中, 事件 A 表示恰有 n 个盒子且每盒一球; 事件 B 表示指定的 n 个盒子中各有一球; 事件 C 表示指定一盒子恰有 m 个球. 求事件 A, B, C 发生的概率. (盒子的容量不限, 放入同一个盒子内的球无顺序排列区别)

解 将 n 只不同的球随机放入 N 个不同的盒子中, 共有 N^n 种不同的放法. 而对事件 A, 有 $(N)_n = N!/n!$ 种不同的放法, 因此

$$P(A) = \frac{(N)_n}{N^n} = \frac{N!}{N^n n!}.$$

对事件 B, 有 $n!$ 种不同的放法, 因此

$$P(B) = \frac{n!}{N^n}.$$

对事件 C, 可分为两步: 第一步在指定的盒子内放入 m 个球, 有 $\binom{n}{m}$ 种不同的放法; 第二步将剩下的 $n-m$ 个球放入 $N-1$ 个盒子, 有 $(N-1)^{n-m}$ 种不同的放法. 因此

$$P(C) = \frac{\binom{n}{m}(N-1)^{n-m}}{N^n}.$$

生日问题是概率历史上有名的数学问题, 研究某次集会的 k 个人中至少有两人生日相同的概率, 或者有 k 人的班级中至少两人生日相同的概率.

例 1.7 (生日问题) 有 k 个人 $(k < 365)$, 每个人的生日等可能地出现于 365 天中的任意一天, 求至少两人生日相同的概率.

解 用 A 表示至少有两人生日相同的事件, 其对立事件 $\bar{A}$ 表示任意两人生日均不相同的事件. k 个人的生日共有 365^k 种可能, 而 k 个人的生日两两互不相同的有 $(365)_k$ 种可能. 因此

$$P(A) = 1 - P(\bar{A}) = 1 - \frac{(365)_k}{365^k}.$$

易知当 $k = 30$ 时, $P(A) = 70.6\%$; 当 $k = 40$时, $P(A) = 89.1\%$; 当 $k = 50$时, $P(A) = 97\%$; 当 $k = 60$时, $P(A) = 99.4\%$; 当 $k = 100$时, $P(A) = 99.99\%$.

下面介绍古典概型计算中一类典型问题, 在产品质量检测等方面广泛应用.

例 1.8 设一批 N 件产品中有 M 件次品, 现从 N 件产品中不放回地任选 n 件, 求其中恰有 k 件次品的概率.

解 用 A 表示恰有 k 件次品的事件. 从 N 件产品中任选 n 件, 有 $\binom{N}{n}$ 种不同的选法; 在所选取的 n 件产品中, 有 k 件次品以及 $n-k$ 件正品, 即从 M 件次品中选出 k 件次品, 从 $N-M$ 件正品中选出 $n-k$ 件正品, 因此有 $\binom{M}{k}\binom{N-M}{n-k}$ 种不同的取法. 由此可得

$$P(A) = \binom{M}{k}\binom{N-M}{n-k}\bigg/\binom{N}{n}, \tag{1.2}$$

该概率 $P(A)$ 被称为 **超几何概率**.

在例 1.8 中若为有放回地任选 n 件, 则每次抽到一件非次品的概率为 $(N-M)/N$, 抽到一件次品的概率为 M/N, 因此 n 件中恰有 k 件次品的概率为

$$\binom{n}{k}\left(\frac{M}{N}\right)^k\left(\frac{N-M}{N}\right)^{n-k}.$$

抽签是人们引入随机性的一个简单例子, 广泛应用于各种体育赛事或日常生活中. 关于抽签的公平性, 即抽签结果虽然不同但出现这种结果的可能性相同, 需要通过计算概率来进行验证:

例 1.9 (抽签问题) 袋中有 a 个不同的白球, b 个不同的红球, 假设有 k 个人依次随机无放回地从袋中取一个球, 问第 i 个人 $(i \leqslant k)$ 取出红球的概率是多少?

解 用 A 表示第 i 个人取到红球的事件. 若 k 个人依次随机无放回地从袋中取一个球, 则有 $(a+b)_k$ 种不同的取法. 若事件 A 发生, 第 i 个人取到红球, 它可能是 b 个红球中的任意一个, 有 b 种取法; 其它剩余的 $k-1$ 个球可以从 $a+b-1$ 个球中取出, 有 $(a+b-1)_{k-1}$ 种不同的取法. 因此事件 A 的概率为

$$P(A) = \frac{b(a+b-1)_{k-1}}{(a+b)_k} = \frac{b}{a+b}.$$

由此可知第 i 个人取到红球的概率为 $b/(a+b)$, 与 i 的大小无关, 即抽签先后顺序对抽签的结果没有影响, 由此证明了抽签的公平性.

例 1.10 (匹配问题) 有 n 对夫妻参加一次聚会, 现将所有参会人员任意分成 n 组, 每组一男一女, 问至少有一对夫妻被分到同一组的概率是多少?

解 用 A 表示至少有一对夫妻被分到同一组的事件, 以及 A_i 表示第 i 对夫妻 $(i \in [n])$ 被分到同一组的事件, 于是有 $A = A_1 \cup A_2 \cup \cdots \cup A_n$. 根据容斥原理有

$$P(A) = P\left(\bigcup_{i=1}^n A_i\right) = \sum_{r=1}^n (-1)^{r+1} \sum_{i_1<\cdots<i_r} P(A_{i_1}\cdots A_{i_r}).$$

对任意 $r \in [n]$, 考虑事件 $A_{i_1}\cdots A_{i_r}$ 概率, 若参会人员任意分成 n 组且每组一男一女, 共有 $n!$ 种不同的分法, 若将第 $i_1, i_2, \cdots, i_r$ 对夫妻分别分组, 则有 $(n-r)!$ 种不同的分法. 根据等可能性原则有

$$P(A_{i_1}\cdots A_{i_r}) = \frac{(n-r)!}{n!}.$$

而和式 $\sum_{i_1<\cdots<i_r} P(A_{i_1}\cdots A_{i_r})$ 中共有 $\binom{n}{r}$ 项, 由此可得

$$\sum_{i_1<\cdots<i_r} P(A_{i_1}\cdots A_{i_r}) = \binom{n}{r}\frac{(n-r)!}{n!} = \frac{1}{r!},$$

于是事件 A 发生的概率

$$P(A) = 1 - \frac{1}{2!} + \frac{1}{3!} + \cdots + (-1)^{n+1}\frac{1}{n!}.$$

当 n 较大时, 利用泰勒展式 $e^x = 1 + x + x^2/2! + \cdots + x^n/n! + \cdots$ 以及令 $x=-1$ 有

$$e^{-1} = \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n\frac{1}{n!} + \cdots \approx \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n\frac{1}{n!},$$

由此近似有 $P(A) = 1 - 1/e = 0.632$.

在概率计算的过程中, 有时可适当利用概率的性质来简化计算, 例如,

例 1.11 从 $\{1,2,\ldots,9\}$ 数中有放回取 n 个, 试求取出 n 个数的乘积被 10 整除的概率.

解 令 $A=\{$取出 n 个整数的乘积能被 10 整除$\}$, $B=\{$取出的 n 个数中有偶数$\}$, $C=\{$取出的 n 个数中至少有一个 5$\}$, 于是有 $A=BC$. 直接计算事件 B 发生的概率较难, 我们因此考虑 B 的对立事件的概率

$$P(\bar{B}) = P\left(\{\text{取出的 } n \text{ 个数中无偶数}\}\right) = P\left(\{\text{取出的 } n \text{ 个数只包括} 1,3,5,7,9\}\right) = 5^n/9^n.$$

同理可得

$$P(\bar{C}) = 8^n/9^n \qquad \text{和} \qquad P(\bar{B}\bar{C}) = 4^n/9^n.$$

根据概率的性质有

$$P(A) = 1 - P(\overline{BC}) = 1 - P(\bar{B}\cup\bar{C}) = 1 - P(\bar{B}) - P(\bar{C}) + P(\bar{B}\bar{C}) = 1 - \frac{5^n}{9^n} - \frac{8^n}{9^n} + \frac{4^n}{9^n}.$$

1.3.2 几何概型

古典概型考虑有限的样本空间, 即有限个等可能的基本事件, 然而在很多实际应用中受到了限制. 本节研究可能有无限多种结果的随机现象, 具有如下两个特点:

- **样本空间无限可测** 样本空间包含无限不可列个样本点, 但可以用几何图形 (如一维线段、二位平面区域、或三维空间区域等) 来表示, 其相应的几何测度 (如长度、面积、体积等) 是一个非零有限的实数,

- **基本事件等可能性** 每个基本事件发生的可能性大小相等, 从而使得每个事件发生的概率与该事件的几何测度相关, 与具体位置无关,

称为 **几何概型**. 其形式化定义如下:

定义 1.7 在一个测度有限的区域 Ω 内等可能性投点, 落入 Ω 内的任意子区域 A 的可能性与 A 的测度成正比, 与 A 的位置与形状无关, 这样的概率模型称之为**几何概型**. 事件 A 发生的概率为

$$P(A) = \frac{A\text{ 的测度}}{\Omega\text{ 的测度}} = \frac{\mu(A)}{\mu(\Omega)}.$$

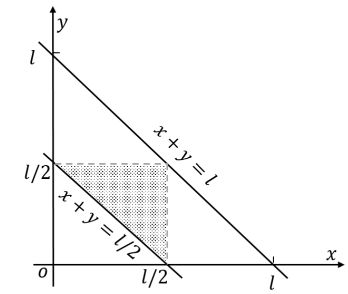
根据上述定义可验证几何概型的概率满足三条公理. 下面给出几何概型的案例.

例 1.12 将一根长度为 l 的木棍随意折成三段, 这三段能构成平面三角形的概率是多少?

解 在此例中将一根木棍折成三段有无穷种可能, 根据其随意性任何一种折法的可能性大小相等, 且木棍的长度可度量, 由此采用几何概型. 用 x, y 分别表示第一段、第二段木棍的长度, 第三段的长度为 $l-x-y$, 由此可得样本空间

$$\Omega = \{(x,y)\colon x>0, y>0, l-x-y>0\}.$$

用 A 表示折成的三段能构成平面三角形的事件, 而构成平面三角形的条件是任意两边之和大于第三边, 由此可得



$$\begin{aligned} A &= \{(x,y)\colon x+y>l-x-y, 1-y>y, l-x>x\} \\ &= \{(x,y)\colon x+y>l/2, y<l/2, x<l/2\}. \end{aligned}$$

如右图所示, 计算事件 A 发生的概率为

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{(l/2)^2/2}{l^2/2} = \frac{1}{4}.$$

例 1.13 假设一乘客到达汽车站的时间是任意的, 客车间隔一段时间发班, 请规划最长的间隔发车时间, 才能确保乘客候车等待时间不超过 20 分钟的概率大于 80%.

解 设客车的间隔时间为 $l\ (l>20)$, 选择特定的连续的 l 分钟为样本空间, 则乘客到达时间的样本空间为 $\Omega=\{x\colon 0<x\leqslant l\}$. 用 B 表示乘客的等待时间超过 20 分钟的事件, 而事件 B 发生则可知乘客到达车站的时间在 0 与 $l-20$ 之间, 即

$$B=\{x\colon 0<x<l-20\}.$$

可知事件 B 发生的概率小于或等于 20%, 即

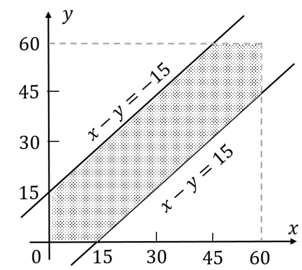
$$P(B)=\frac{l-20}{l}\leqslant 0.2,$$

求解可得 $l\leqslant 25$.

例 1.14 (会面问题) 两银行经理约定中午 12 : 00 − 13 : 00 到某地会面, 两人到达时间随机, 先到者等另一人 15 分钟后离开, 求两人见面的概率.

解 用 x, y 分别表示两人的到达时间 (分钟), 则样本空间 $\Omega = \{(x,y)\big|0 \leqslant x, y \leqslant 60\}$. 用 A 表示两人见面的事件, 则

$$A = \{(x,y)\big||x-y| \leqslant 15\} = \{(x,y)\big|x-y \leqslant 15 \text{ 且 } x-y \geqslant -15\}.$$



根据右图计算事件 A 发生的概率

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{60^2 - 45^2}{60^2} = \frac{7}{16}.$$

进一步思考: 若两银行经理非常聪明且都非常希望能促成此次见面, 但没有通讯方式进行联系, 能否找出一些策略来解决会面问题?

很多几何概型的概率可通过计算机模拟仿真来近似计算, 即 **统计模拟法** 或 **蒙特卡洛 (Monte Carlo) 法**. 先构造相应的概率模型, 再进行计算机模拟试验, 用统计的方法计算其估计值, 作为所求问题的近似值. 例如, 可利用蒙特卡洛法来近似计算例 1.14 的概率, 伪代码如下:

```
输入参数: 试验总次数 N.                    %% 取较大正整数 N, 更能精确计算两人见面的概率
初始化: 事件 A 最初始发生的次数 n_A ← 0.      %% 此处事件 A 表示两人见面的事件
For i = 1 : N
    x ← Random(0, 60), y ← Random(0, 60).   %% 在区间 (0, 60) 以随机任意选取两个数
    If |x − y| ⩽ 15 then
        n_A ← n_A + 1.                      %% 若两人见面则频数+1
    End
End
输出概率: n_A/N.
```

接下来介绍几何概型的一个经典问题, 由法国科学家蒲丰于 1777 年提出.

例 1.15 (投针问题) 平面上有两条平行线, 相距为 a, 向此平面任投一长度为 l $(l < a)$ 的针, 求此针与任一平行线相交的概率.

解 用 x 表示针的中点到最近的一条平行线的距离, 用 θ 表示针与平行线的夹角. 针与平行线的位置关系图 1.3 所示.

容易知道 $x \in [0, a/2]$ 和 $\theta \in [0, \pi]$, 以 Ω 表示边长分别为 $a/2$ 和 π 的长方形, 用 A 表示针与平行线相交的事件, 若事件 A 发生则必有 $x \leqslant l\sin(\theta)/2$ 成立, 由此得到事件 A 的发生如图 1.3 中阴影部分所示. 求解可得

$$P(A) = \frac{A \text{ 的面积}}{\Omega \text{ 的面积}} = \frac{\int_0^\pi l\sin(\theta)/2d\theta}{a\pi/2} = \frac{2l}{a\pi}\ . \tag{1.3}$$

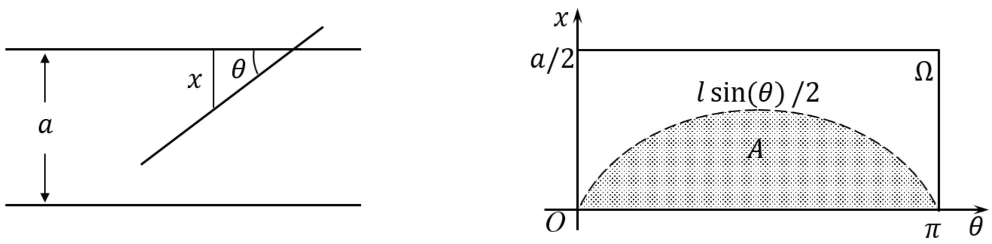


图 1.3 蒲丰投针问题

在上例中, 事件 A 发生的概率包含圆周率 π, 由此蒲丰设想出计算 π 的概率近似方法, 通过频率来近似计算事件 A 发生的概率, 再根据 (1.3) 计算圆周率 π, 即

$$\frac{n_A}{n} \approx P(A) = \frac{2l}{a\pi} \qquad \Rightarrow \qquad \pi \approx \frac{2ln}{an_A} \ .$$

这里 n 表示试验的总次数, 而 n_A 表示事件 A 发生的次数. 历史上有多人根据蒲丰的设想还真做了试验来近似计算圆周率 π, 有兴趣的读者可查找相关文献.

在二十世纪之前, 很多人都相信只要找到合适的等可能性描述, 概率是可以被唯一定义的. 然而贝特朗 (Bertrand) 对这种观点提出了质疑, 他通过下面的一个具体例子说明: 几何概型的等可能性概率存在多种看似合理但相互矛盾的结果.

例 1.16 (贝特朗奇论) 在半径为 1 的圆内随机地取一条弦, 求其弦长超过该圆内接等边三角形边长 $\sqrt{3}$ 的概率.

解 对于"等可能性"或"随机性"含义的不同解释, 这个问题存在着多种不同答案的解决方法, 下面给出三种不同的方法:

(1) 任何与圆相交的弦一般有两个交点, 不妨在圆上先固定其中一点, 以此点为顶点作一个等边三角形, 只有落入此三角形内的弦才满足弦长超过 $\sqrt{3}$. 这种弦的另一端跑过的弧长为整个圆周的 1/3, 故所求概率等于 1/3 (如图 1.4-**a**).

(2) 弦长只跟到圆心的距离有关, 与方向无关, 因此可以假定它垂直于某一条直径. 当且仅当它与圆心的距离小于 1/2 时, 其弦长才会大于 1/3, 因此此时所求的概率为 1/2 (如图 1.4-**b**).

(3) 弦可被弦的中心点唯一确定, 当且仅当弦的中心点位于半径为 1/2 的同心圆内时, 弦长才会大于 1/3. 半径为 1/2 的小圆面积为 1/4, 大圆面积为 1, 故所求概率等于 1/4 (如图 1.4-**c**).

同一问题却有三种不同的答案, 其根本原因在于取弦时采用不同的等可能性假定. 第一种方法假定端点在圆周上均匀分布, 第二种方法假定弦的中点在直径上均匀分布, 第三种方法假定弦的中心点在小圆内均匀分布. 这三种方法采用三种不同的随机试验, 对于各自的随机试验而言, 它们都是正确的, 因此在概率的计算中一定要明确具体的试验.

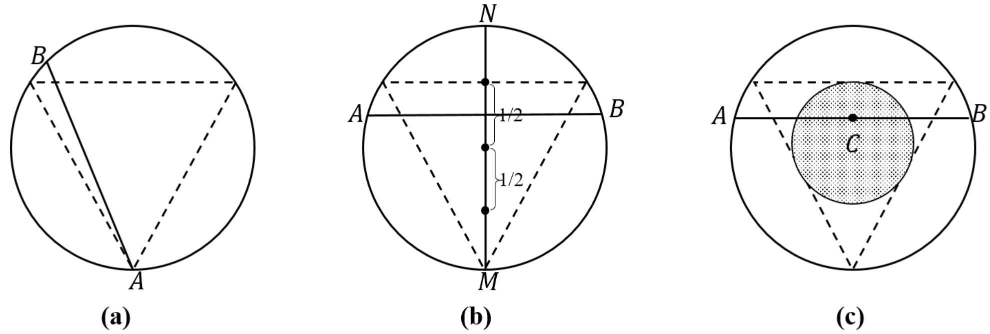


图 1.4 贝特朗奇论

贝特朗奇论在现代概率的发展中起到过重要作用, 上述例子由贝特朗于 1899 年提出, 以此反驳了"等可能性可完全定义概率"的观点. 从此概率论开始向公理化方面发展, 从应满足的基本性质来定义概率, 而不是某些具体事件的概率. 正因为如此, 希尔伯特于 1900 年在巴黎举行的第二届数学家大会上提出了著名的 23 个数学问题, 其中第六个问题就是概率公理化.

与古典概型一样, 几何概型的研究有助于发现概率的一些基本性质, 有助于对某些概率问题的直观理解和具体计算.

1.4 组合计数*

组合计数研究满足一定条件的计数对象的数目, 概率论中的很多问题都可以通过计算一个事件发生的数目来解决, 如古典概型. 此外, 组合计数在人工智能、计算机等领域具有广泛的应用, 本节将介绍经典的组合计数: 十二重计数 (The twelvefold way).

十二重计数由著名的组合学家 G.-C. Rota (1932-1999) 提出, 最初的问题表述为研究一定条件下两个集合之间映射的数目. 为了问题的可理解性, 我们采用《计算机程序设计艺术》中的表述: 将 n 支不同或相同的球, 放入 m 个不同或相同的箱子, 在无任何限制、或每个箱子至多或至少放一个球的条件, 研究在这十二种情形下分别有多少种不同的方法数. 我们首先给出十二重计数的结果, 如表 1.3 所示, 相关知识和符号说明将在后续小节逐一介绍.

表 1.3 十二重计数.

n 个球	m 个箱子	无任何限制	每个箱子至多放一球	每个箱子至少放一球
不同	不同	m^n	$(m)_n$	$m!S(n,m)$
相同	不同	$\binom{n+m-1}{n}$	$\binom{m}{n}$	$\binom{n-1}{m-1}$
不同	相同	$\sum_{k=1}^{m} S(n,k)$	$\begin{cases} 1 & n \leqslant m \\ 0 & n > m \end{cases}$	$S(n,m)$
相同	相同	$\sum_{k=1}^{m} p(n,k)$	$\begin{cases} 1 & n \leqslant m \\ 0 & n > m \end{cases}$	$p(n,m)$