第一周　随机事件及其概率运算

1.1　随机试验与随机事件

同学们好！欢迎大家参加中国大学先修课程《概率论与数理统计》的学习。我是清华大学数学科学系的教师梁恒，很高兴在今后一段时间里与大家分享一些概率论与统计学中最基本、最重要也是最常用的经典成果。概率论与统计学集中对不确定性进行定量研究，建立了描述不确定性的有效数学模型和理论方法。随着现代科学技术的发展，深刻地理解不确定性有着越来广泛和紧要的需求，概率论与统计学已经成为科学研究、工程技术、经济管理，乃至人文社科等领域不可或缺的工具。这门课程的基本内容和方法不仅是提供了一些有效工具，更反映出独特的思维模式，很好地体现了数学理论和实际应用的联系。本课程面向已经有一些微积分基础的优秀中学生，强调抽象原理与现实应用的紧密结合，希望通过深入浅出的内容，引导同学们从传统的确定性思维模式逐步熟悉和掌握随机性思维模式，当然也希望能够激发同学们的学习兴趣，提升大家的科学素养，为同学们在大学后继课的学习奠定良好的数学基础并帮助同学们适应从初等数学到高等数学，学习观念上的转变，更好地适应即将到来的大学学习和生活。第一周我们主要学习随机事件及事件的概率运算。

****

偶然性与不确定性的概念几乎与人类文明本身一样的古老。人们不得不应付天气变化、传染病的侵袭、战争胜负，以及一次捕猎是否成功等等的不确定现象。很久以前，人们就对理解和运用不确定性的机理和特征产生了兴趣。早在公元前 **3500** 年左右，古埃及等地就已经出现了利用动物骨头制作的具有随机性质的游戏。

通常人们都认为近代的概率论，也就是概率的数学理论是由十七世纪法国数学家帕斯卡和费马共同开创的，他们成功地推导出一些赌博规则对应的实际概率，获得了一些有效的计算公式。从那时起，概率论得到了稳步的发展，被越来越多地应用到工程、科学、管理、医药等领域，成为与微积分、线性代数同等重要的最基础的数学工具之一。

****

随机试验与样本空间

如果一个试验事先能够明确地知道试验所有可能的基本结果，在每一次观察中，不能

事先准确地预言其中哪一个基本结果会发生，并且在相同条件下可以重复进行，则称此试验为随机试验。

随机试验的每种基本结果称为一个样本点ω，全体基本结果构成的集合称为样本空间，通常记为Ω。

****

例 1.1.1　考察下面几个随机试验的样本空间

试验 1：将一枚均匀硬币抛 3 次，观察出现正面的次数；　$\Omega_1=\{0,1,2,3\}$；

试验 2：同时掷两颗六面的色子，观察所得的点数和；　$\Omega_2=\{2,3,4,\cdots,12\}$

试验 3：某网站在某一段时间内被点击的次数；$\Omega_3=\{0,1,2,3,\cdots\}$

试验 4：在一批电子器件中任意取一只，测试其寿命。$\Omega_4=\{t\mid t\geq 0\}$

****

样本空间中具备某种属性的样本点的集合叫做一个随机事件，简称为事件，通常用大写字母 A、B、C 等表示。由一个样本点组成的单点集，称为基本事件。

对于 $A\subset\Omega$，$w\in A$ $\Leftrightarrow$ 事件 A 发生；　$w\notin A\Leftrightarrow$ 事件 A 未发生。

例如：掷两颗色子的随机试验中，随机事件“点数和为 3 的倍数”对应集合 $A=\{3,6,9,12\}$。

****

例 1.1.2　考虑例 1.1.1 中的随机试验 3，即某网站在某一段时间内被点击的次数，试写出下列事件包含的样本点：

$A=\{$一小时内被点击次数在 10 到 20 次之间$\}$；　$A=\{10,11,12,\cdots,20\}$

$B=\{$一小时内被点击次数不多于 7 次$\}$；　$B=\{0,1,2,3,4,5,6,7\}$

$C=\{$一小时内被点击次数为偶数次$\}$；　$C=\{2,4,6,\cdots,2n,\cdots\}$

$D=\{$一小时内被点击次数至少为 0 次$\}$；　$D=\Omega=\{0,1,2,3,\cdots\}$

$E=\{$一小时内被点击次数少于 0 次$\}$.　$E=\Phi$

****

字幕(出镜)： 样本空间Ω包含所有的样本点，若事件集合等于Ω本身，则在每次试验中它总是发生的，故称为必然事件；而空集Φ中不包含任何样本点，它在每次试验中都不可能发生，故称为不可能事件。为讨论方便，以及概念的完整性，虽然必然事件与不可能事件并不具有不确定性，我们仍然常常将它们作为随机事件的特例，纳入到随机事件的范畴来统一考虑。

（备注 3： 话音结束后，再在此画面多停留 3-5 秒，然后切入下一页）

****

事件的示性函数

函数 $I_A(w)=\begin{cases}1, & w\in A\\ 0, & otherwise\end{cases}$，称为事件 A 的示性函数或标志函数（indicator）

****

1.2 古典概型

****

古典概型

如果每个基本事件都等可能出现，此时某一事件的概率为

$$P(A)=\frac{\text{事件}A\text{包含的基本事件数}}{\text{全部可能的基本事件数}} \quad \text{或} \quad P(A)=\frac{\text{事件}A\text{所占区域大小}}{\text{样本空间所占区域大小}}$$

称为古典概型，也叫等可能概型。

****

例 1.2.1（生日问题）$n(n\le 365)$ 个人的生日互不相同的概率 p_n。

解：我们只考虑一年 365 天的情况，

$$p_n=\frac{365\cdot(365-1)\cdots(365-(n-1))}{365^n}。$$

p_n 估值 $\quad p_n=\left(1-\frac{1}{365}\right)\cdot\left(1-\frac{2}{365}\right)\cdots\left(1-\frac{n-1}{365}\right)$

$\ln(1-x)\approx -x\ (x<<1)$,

$\ln p_n\approx-\frac{1+2+\cdots+(n-1)}{365}=-\frac{n(n-1)}{730},\qquad p_n\approx e^{-\frac{n(n-1)}{730}}$

n 越大则估计的精度越低，$\bar{p}_{10}=0.884, p_{10}=0.883$， $\bar{p}_{30}=0.3037, p_{30}=0.2937$

****

例 1.2.2（盒子模型） 设有 n 个球，每个球都等可能地被放到 $N\ (n\le N)$ 个不同的盒子中的任意一个，求指定的 n 个盒子中各有 1 球的概率 p

根据对球和盒子性质的不同假设，分为以下三种情况：

1）Maxwell-Boltzman 统计，n 个球不同（可分辨的），盒子容量不限； $p=\frac{n!}{N^n}$；

2）Fermin-Dirac 统计，每个盒子至多容纳 1 个球，小球不可分辨（小球完全相同）；

$p=1/C_N^n$；

****

例 1.2.2（盒子模型）设有 n 个球，每个球都等可能地被放到 $N\ (n\le N)$ 个不同的盒子中的任意一个，求指定的 n 个盒子中各有 1 球的概率 p

根据对球和盒子性质的不同假设，分为以下三种情况：

3）Bose-Einstein 统计，n 个球是不可分辨的（小球完全相同），盒子容量不限。

$N=3, n=2$ 时，第（1）种假设的样本空间为

$\boxed{1.(ab,-,-)}$ $\quad 2.(-,ab,-)\quad 3.(-,ab,-)$

$\boxed{4.(a,b,-)\quad 5.(b,a,-)}$ $\quad 6.(a,-,b)$

$7.(b,-,a)\quad 8.(-,a,b)\quad 9.(-,b,a)$

第（3）种假设的样本空间为

$1.(**,-,-)\quad 2.(-,**,-)\quad 3.(-,-,**)$

$4.(*,*,-)\quad 5.(*,-,*)\quad 6.(-,*,*)$

共 n+N-1 个可变位置

$$p=\frac{1}{C_{N+n-1}^{n}}=\frac{n!(N-1)!}{(N+n-1)!}$$

****

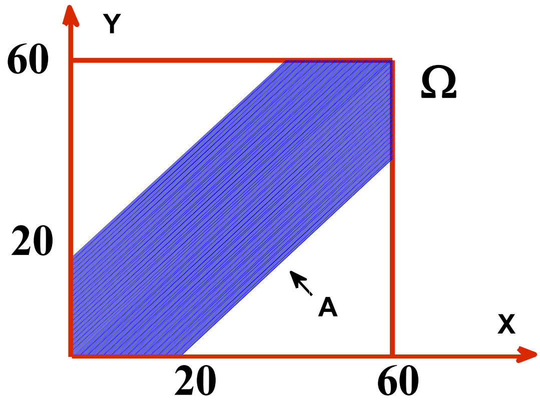
例 1.2.3 （会面问题） 甲、乙约定在下午 4 点至 5 点之间见面，并约定等候 20 分钟，过时即离去，求二人的会面概率。

解 设甲、乙二人分别在 4 点 x 分和 4 点 y 分到达，

则两人会面等价为下面不等式成立。

$$|x-y|\le 20,\ 0\le x,y\le 60$$

$$P=\frac{S_A}{S_\Omega}=1-\frac{2\times 40^2}{60^2}=\frac{5}{9}$$



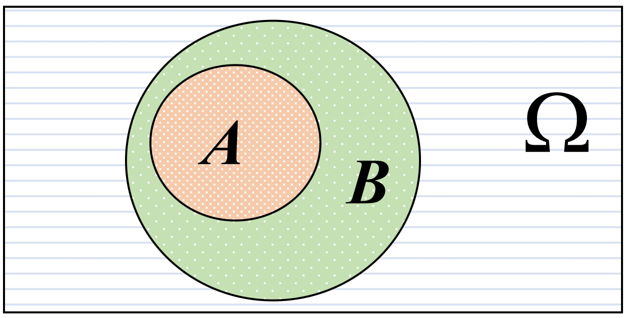
古典概型是做了等可能假设的非常特殊的样本空间下的一种概率模型，其功能是很有局限的。为了处理千变万化随机现象，还需要引入更一般的概率模型。下一节课我们学习一般样本空间中随机事件之间的关系和事件的概率运算。

****

1.3 事件间的关系与事件的运算

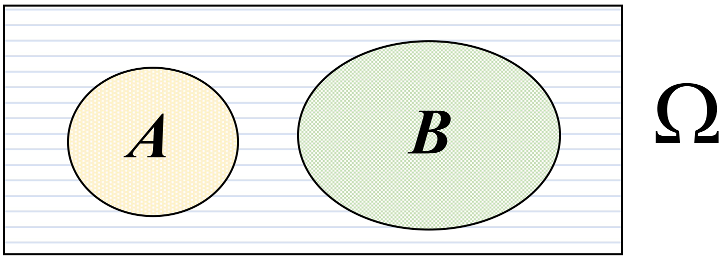
事件关系（包含，相等，互不相容，对立）

（1） 包含关系：若事件 A,B 满足 $A\subset B$，则称事件 B 包含事件 A，用示性函数表示为 $I_A(\omega)\le I_B(\omega)$.



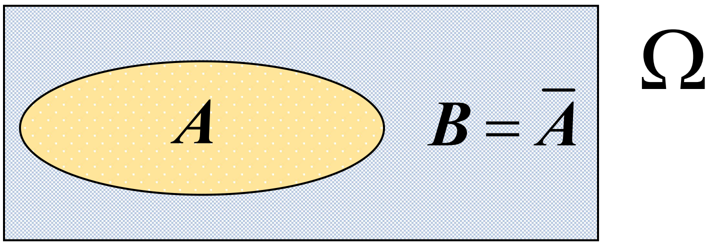
（2） 相等关系：若 $A \subset B$，且 $B \subset A$，即 $A = B$，则称事件 A 与事件 B 相等（或等价），为同一事件。用示性函数表示为 $I_A(\omega) = I_B(\omega)$.

（3）互不相容关系，也称互斥关系：对于事件 A、B，如果不可能同时发生，则 A、B 称为互不相容事件，此时 $AB = \Phi$。用示性函数表示为 $I_A(\omega) I_B(\omega) = 0$.



（4） 对立关系：如果两个事件 A、B 中，$B =$ “A 不发生”，则 A、B 称为具有对立关系（或互逆关系），又称 B 为 A 的对立事件，记为 $B = \overline{A}$。

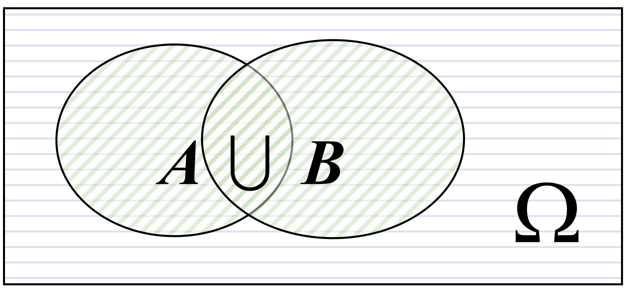
用示性函数表示为 $I_A(\omega) + I_B(\omega) = 1$.



****

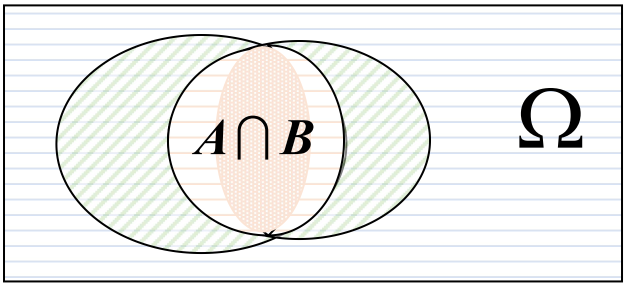
事件运算（和，积，差，交换律，结合律，分配律，结合律，对偶律）

（1） 事件的和：事件 A 与事件 B 的并集构成的事件称为事件 A 与事件 B 的和事件，记为 $A \cup B$ 或 $A + B$，即 $A \cup B = \{x \mid x \in A 或 x \in B\}$，如图所示的阴影部分. 显然，当且仅当事件 A 与事件 B 至少有一个发生时，事件 $A \cup B$ 才发生。



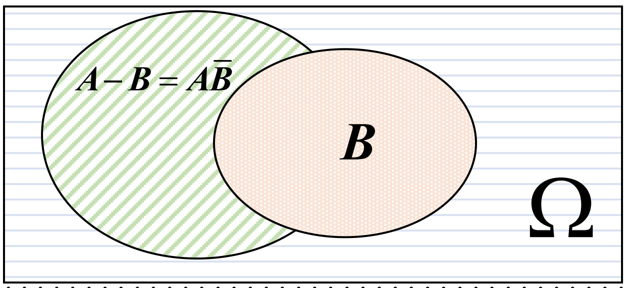
n个事件 $A_1,A_2,\cdots,A_n$ 的和事件，即为 n 个集合的并集 $\bigcup_{k=1}^{n} A_k$ 。

（2）事件的积（或交）：事件 A 与事件 B 的交集构成的事件称为事件 A 与事件 B 的积（或交）事件，事件 A 与事件 B 同时发生。 记为 $\boldsymbol{A \cap B}$ 或 AB 。



n个事件 $A_1,A_2,\cdots,A_n$ 的积事件，即为 n 个集合的交集 $\bigcap_{k=1}^{n} A_k$ 。

（3）事件的差：事件 A 与事件 B 的差集所构成的事件称为事件 A 与事件 B 的差事件，记为 $A-B$ 。$\boldsymbol{A-B=\{x \mid x \in A且x \notin B\}=A\overline{B}}$。当且仅当事件 A 发生但事件 B 不发生时，事件 $A-B$ 才发生.



****

例 1.3.1 某同学在篮球场上进行了连续 3 次投篮练习，记 $A_i=\{$第 i 次投中篮筐$\}$，试用 $A_i\ (i=1,2,3)$ 表示事件：

（1） $B_j=\{$连续 3 次投篮中恰好有 j 次投中篮筐$\}\ (j=0,1,2,3)$；

（2） $C_k=\{$连续 3 次投篮中至少有 k 次投中篮筐$\}\ (k=0,1,2,3)$.

解：（1） $B_0=\overline{A_1}\,\overline{A_2}\,\overline{A_3}$；　　$B_1=A_1\overline{A_2}\,\overline{A_3}\cup\overline{A_1}A_2\overline{A_3}\cup\overline{A_1}\,\overline{A_2}A_3$；

$$B_2 = A_1A_2\overline{A_3} \cup A_1\overline{A_2}A_3 \cup \overline{A_1}A_2A_3 \quad ; \qquad B_3 = A_1A_2A_3 .$$

（2）$C_0 = \Omega = B_0 \cup B_1 \cup B_2 \cup B_3$； $\qquad C_1 = A_1 \cup A_2 \cup A_3 = B_1 \cup B_2 \cup B_3$；

$$C_2 = A_1A_2 \cup A_2A_3 \cup A_3A_1 = B_2 \cup B_3 ; \quad C_3 = A_1A_2A_3 = B_3$$

****

事件的运算律

交换律： $A \cup B = B \cup A$， $AB = BA$ （即 $A \cap B = B \cap A$）

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

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

对偶律： $\overline{A \cup B} = \overline{A} \cap \overline{B}$， $\overline{A \cap B} = \overline{A} \cup \overline{B}$

****

事件概率的几条基本性质：

$P(\Phi) = 0$， $P(\Omega) = 1$， $P(\overline{A}) = 1 - P(A)$，

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

****

例 1.3.2 袋中有编号为 $1,2,\cdots,n$ 的 n 个球，从中有放回地任取 m 次，求取出的 m 个号码中最大编号恰好是 k 的概率。

分析：最大编号不超过 k 的基本事件数为 $\dfrac{k^m}{n^m}$

解：设事件 A_k 为最大号码恰为 k， B_k 表示最大号码不超过 k，

则 $A_k = B_k - B_{k-1}$，且 $B_{k-1} \subset B_k$， 所以 $P(A_k) = P(B_k) - P(B_{k-1})$

又知 $P(B_k)=\dfrac{k^m}{n^m}\quad(k=1,2,\cdots,n)$，

得 $P(A_k)=\dfrac{k^m-(k-1)^m}{n^m}\quad(k=1,2,\cdots,n)$。

****

例 1.3.3 （匹配问题）n 封写给不同人的信随机放入 n 个写好收信人姓名的信封，求所有信件都装错了信封的概率。

解：将 n 封不同的信分别编号 $1,2,\cdots,n$， n 个对应的信封同样编号 $1,2,\cdots,n$，

设事件 A_i 表示编号为 i 的信恰好装入了编号为 i 的信封，则所求概率

$$p_0=1-P(A_1\cup A_2\cup\cdots\cup A_n)$$

概率的加法公式：

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

$$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$$

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\le i_1<i_2\le n}P(A_{i_1}A_{i_2})+\sum_{1\le i_1<i_2<i_3\le n}P(A_{i_1}A_{i_2}A_{i_3})-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)$$

$$p_0=1-P(A_1\cup A_2\cup\cdots\cup A_n)$$

$$=1-\left[\sum_{i=1}^{n}P(A_i)-\sum_{1\le i_1<i_2\le n}P(A_{i_1}A_{i_2})+\sum_{1\le i_1<i_2<i_3\le n}P(A_{i_1}A_{i_2}A_{i_3})-\cdots+(-1)^{n-1}P(A_1\cdots A_n)\right]$$

$$P(A_{i_1})=\frac{1}{n},\quad P(A_{i_1}A_{i_2}\cdots A_{i_k})=\frac{1}{n(n-1)\cdots(n-k+1)}\quad(\forall 1\le i_1<i_2<\cdots<i_k\le n)$$

$$P(A_1\cup A_2\cup\cdots\cup A_n)=1-\frac{1}{2!}+\frac{1}{3!}-\frac{1}{4!}+\cdots+(-1)^{n-1}\frac{1}{n!}$$

注意到 $1-e^{-1}=1-\dfrac{1}{2!}+\dfrac{1}{3!}-\dfrac{1}{4!}+\cdots+(-1)^{n-1}\dfrac{1}{n!}+\cdots=\displaystyle\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k!}$

可得 $p_0 \approx e^{-1} \approx 0.37$，当$n \geq 4$时，这一估计值的误差不超过小于 **1%**。

****

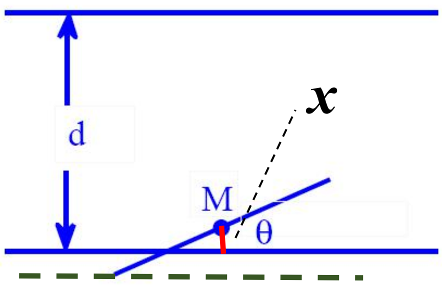
1.4　两个著名的例子

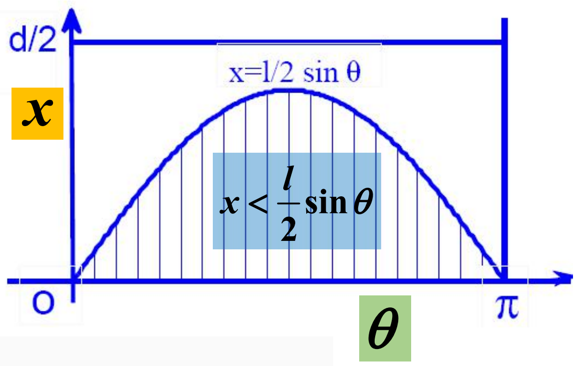
布丰(Buffon，1707-1788)投针问题

18 世纪时，法国学者布丰提出了一个“投针问题”，记载于布丰 1777 年出版的著作中。问题是这样的 “在平面上画有一组间距为 d 的平行线，将一根长度为 L 的的针任意掷在这个平面上，求此针与平行线中任一条相交的概率。

在平面上画有一组间距为 d 的平行线，将一根长度为 $l\left(l \leq d\right)$ 的针任意掷在这个平面上，求此针与平行线中任一条相交的概率。

布丰证明了，这个概率是 $p = \dfrac{\int_0^{\pi} \frac{l}{2}\sin\theta d\theta}{\frac{d}{2}\pi} = \dfrac{2l}{\pi d}$，　　$\dfrac{m}{n} \approx \dfrac{2l}{\pi d} \Rightarrow \pi \approx \dfrac{2nl}{md}$





利用布丰投针试验估计圆周率的一些历史资料

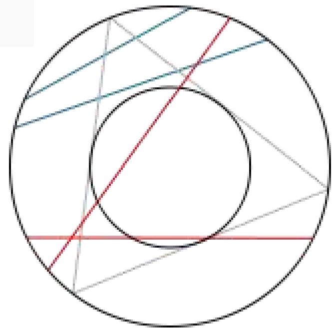
Experimenter	Length of needle	Number of casts	Number of crossings	Estimate for π
Wolf, 1850	.8	5000	2532	3.1596
Smith, 1855	.6	3204	1218.5	3.1553
De Morgan, c.1860	1.0	600	382.5	3.137
Fox, 1864	.75	1030	489	3.1595
Lazzerini, 1901	.83	3408	1808	3.1415929
Reina, 1925	.5419	2520	869	3.1795

****

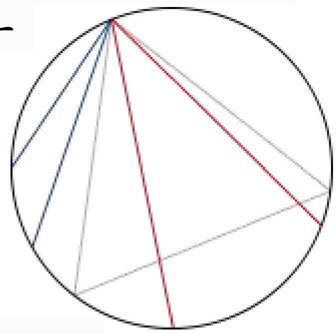
贝特朗奇论(Bertrand's paradox)

在半径为 1 的圆内随机地取一条弦，问其长度超过该圆内接等边三角形的边长 $\sqrt{3}$ 的概率等于多少？

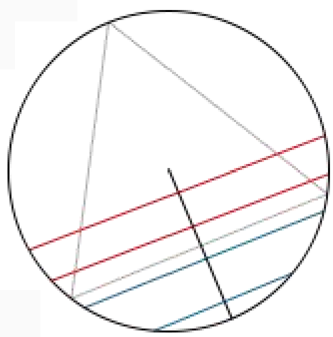
分析一：弦长被其中心唯一确定，当且仅当其中点属于半径为 $\frac{1}{2}$ 的同心圆内时，弦长大于 $\sqrt{3}$，此小于圆的面积为大圆面积的 $\frac{1}{4}$，因此所求的概率为 $\frac{1}{4}$



分析二：任何弦交圆两点，不失一般性，先在圆周上固定其中一点，则弦的长度由另一个端点的位置所决定。以固定点为顶点，过该顶点的直径为对称轴做一等边三角形，显然只有落如此三角形内的弦才满足要求，这种弦的另一端点对应弧长为整个圆周的 $\frac{1}{3}$，故所求概率为 $\frac{1}{3}$



分析三：弦长只跟它与圆心的距离有关，而与方向无关，因此可以假定它与某一直径垂直，当且仅当它与圆心的距离小于 $\frac{1}{2}$ 时，其长度才大于 $\sqrt{3}$，因此所求概率为 $\frac{1}{2}$



同一问题有三种不同的答案，细究原因，发现是在取弦时采用了不同的等可能性假设。在第一种分析中，假定弦的中点在圆内均匀分布；第二种分析中则假定弦的端点在圆周上均匀的分布，而第三种解法中又假定弦的中心在直径上均匀分布。这三种答案针对的是三种不同的随机试验，从而有着不同的样本空间和样本点。三种方法出发点不同，所以得到不同的结果也就不足为奇了。再进一步讨论究竟哪一个更合理也没有太多意义。好比，我们问一位同学“从家里到学校需要多长时间？”若没有指明走着去，还是坐车去，还骑自行车去等，这个问题的提法就是有缺陷的，无需过多地讨论。因此，笼统地使用“随机”、“等可能”、“均匀分布”等术语，其含义可能是模糊的，可能会产生不同的理解。作为一门数学理论，受到自然语言所产生歧义的影响是不能令

人满意的。1899年贝特朗在巴黎出版《概率论》，书中对几何概率提出了批评，并以生动的实例引起大家的注意。这种善意的批评，推动了概率论的发展。它促使人们思考到底什么是随机，应该如何给出更严格的定义。

勒贝格、波雷尔等数学家在19世纪末、20世纪初建立了测度的理论，为概率论的公理化做了坚实的准备。20世纪30年代，苏联数学家科尔莫格罗夫完善了概率论的公理体系，使概率论的发展有了严格的数学基础。

****

概率的公理化定义：

Ω 是一个样本空间，F 为 Ω 的某些子集组成的集合，称为一个事件域。

如果对任一事件 $A\in F$，定义在 F 上的一个实值函数 $P(A)$ 满足

（1）非负性公理　　若 $A\in F$，则 $P(A)\geq 0$；

（2）正则性公理　　$P(\Omega)=1$；

（3）可列可加性公理　若 $A_1,A_2,\cdots,A_n,\cdots$ 互不相容，有 $P\left(\bigcup_{i=1}^{+\infty}A_i\right)=\sum_{i=1}^{+\infty}P(A_i)$，

则称 $P(A)$ 为事件 A 的概率，称三元素 (Ω,F,P) 为概率空间。

只要满足这样几条简单的假设，就可以清晰地定义和研究各种随机试验的概率模型，推演出千变万化的概率结果。公理化叙述中的每个细节都是必不可少的，有其丰富的内涵，充分地理解这些内容需要测度论的知识，已经超出了本课程的范围，这里我们只是稍稍了解一下就可以了。