



概率论与数理统计

主讲教师：朱丽娜 讲师

研究方向：智能交通，车联网与智能驾驶

电子邮件：lnzhu@xidian.edu.cn

个人主页：<http://web.xidian.edu.cn/lnzhu/>

第二章 随机变量及其分布

- ⇒ § 2.1 随机变量
- ⇒ § 2.2 离散型随机变量及其概率分布
- ⇒ § 2.3 随机变量的分布函数
- ⇒ § 2.4 连续型随机变量及其概率密度
- ⇒ § 2.5 随机变量的函数的分布

§ 2.5 随机变量的函数的分布

- 实际应用中，某些人们关心的随机变量 Y 往往不能直接测量得到，而它可能是某个能测量的随机变量的函数 X
 - 比如我们有时很关心圆柱轴截面的面积 S ，但无法直接获得，而我们能够获得圆柱轴截面的直径 D ，而随机变量 S 是随机变量 D 的函数，即 $S=(1/4)\pi D^2$
- 问题：
 - 怎样由已知的随机变量 X 的概率分布求它的函数 $Y=g(X)$ 的概率分布，其中 $g(\cdot)$ 是连续函数？
 - 一般的步骤是什么？

§ 2.5 随机变量的函数的分布

⇒ 例1: 设随机变量X具有以下的分布律, 试求 $Y=(X-1)^2$ 的分布律

•	X	-1	0	1	2
•	p_k	0.2	0.3	0.1	0.4

⇒ 解: 首先获得Y的所有可能取值: 0, 1, 4

• 求解各取值的概率

• $P\{Y=0\}=P\{(X-1)^2=0\}=P\{X=1\}=0.1$

• $P\{Y=1\}=P\{(X-1)^2=1\}=P\{X=0\}+P\{X=2\}=0.7$

• $P\{Y=4\}=P\{(X-1)^2=4\}=P\{X=-1\}=0.2$

• 即得Y的分布律为

•	Y	0	1	4
•	p_k	0.1	0.7	0.2

⇒ 对于离散型随机变量

• $P\{Y=y_k\}$ 等于所有满足 $y_k=g(X)$ 的X的取值的概率之和

§ 2.5 随机变量的函数的分布

例2: 设随机变量X具有概率密度 $f_X(x) = \begin{cases} x/8, & 0 < x < 4 \\ 0, & \text{其它} \end{cases}$

- 求随机变量Y=2X+8的概率密度

解: 分别记X,Y的分布函数为 $F_X(x)$, $F_Y(y)$, 先求Y的分布函数 $F_Y(y)$

- $F_Y(y) = P\{Y \leq y\}$ //由分布函数的含义

- 将Y=2X+8代入 $= P\{2X + 8 \leq y\}$

- 表示为关于X的概率 $= P\{X \leq (y - 8)/2\}$

- 即 $= F_X((y - 8)/2)$

- 将 $F_Y(y)$ 关于y求导得

- $f_Y(y) = dF_Y(y)/dy = dF_X((y - 8)/2)/dy = f_X((y - 8)/2) d((y - 8)/2)/dy$

- $= \begin{cases} (1/8)[(y - 8)/2](1/2), & 0 < (y - 8)/2 < 4 \\ 0, & \text{其它} \end{cases}$

- $= \begin{cases} (y - 8)/32, & 8 < y < 16 \\ 0, & \text{其它} \end{cases}$

- 注意: 代入时, f_X 的自变量及分段函数取值范围用 $g^{-1}(y)$ 来代

§ 2.5 随机变量的函数的分布

- ⇒ 求连续型随机变量 X 的函数 $Y=g(X)$ 的概率密度的一般步骤:
- ⇒ 已知 $f_X(x)$, $Y=g(X)$, 求 $f_Y(y)$
 - 1° 先写出 Y 的分布函数定义式: $F_Y(y)=P\{Y\leq y\}$
 - 由 $Y=g(X)$ 确定 Y 的值域, 当 y 不在值域范围内时单独讨论 $f_Y(y)$
 - 2° 将 $Y=g(X)$ 代入上式 $=P\{g(X)\leq y\}$ //在 y 的值域范围内讨论
 - 3° 由 $g(X)\leq y$ 求解 X 的范围 $=P\{X|g(X)\leq y\}$ 表示为 y 的形式
 - 4° 由 X 的分布函数 $F_X(x)$ 表示以上概率, 得到关于 y 的表达式 $F_Y(y)$, 其中 $F_X(x)$ 的自变量 x 用关于 y 的表达式来代。
 - 5° 求导得 $f_Y(y)=dF_Y(y)/dy$, 将 $f_Y(y)$ 的所有可能的情况合并
- ⇒ 掌握变上下限积分求导公式

§ 2.5 随机变量的函数的分布

例3: 设随机变量 X 具有概率密度 $f_X(x)$,

$-\infty < x < \infty$, 求 $Y=X^2$ 的概率密度

解: 1° $F_Y(y) = P\{Y \leq y\}$,

由 $Y=X^2 \geq 0$, 当 $y \leq 0$ 时, $F_Y(y) = 0$

2° 当 $y > 0$ 时, 代入 $Y=X^2$ 得

$$F_Y(y) = P\{Y \leq y\} = P\{X^2 \leq y\}$$

$$3^\circ \quad = P\{-\sqrt{y} \leq X \leq \sqrt{y}\}$$

$$4^\circ \text{ 由 } F_X(x) \text{ 得 } F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$5^\circ f_Y(y) = dF(y)/dy = \begin{cases} \frac{1}{2\sqrt{y}} [f_X(\sqrt{y}) + f_X(-\sqrt{y})], & y > 0 \\ 0, & y \leq 0 \end{cases}$$

§ 2.5 随机变量的函数的分布

- 在 $Y=g(X)$ ，且 $g(X)$ 为严格单调函数时的一般结果
- 定理：设随机变量 X 具有概率密度 $f_X(x)$ ， $-\infty < x < \infty$ ，又设函数 $g(x)$ 处处可导且恒有 $g'(x) > 0$ (或 $g'(x) < 0$)，则 $Y = g(X)$ 是连续型随机变量，其概率密度为

- $$f_Y(y) = \begin{cases} f_X[h(y)] |h'(y)|, & \alpha < y < \beta \\ 0, & \text{其它} \end{cases}$$

- 其中 $\alpha = \min\{g(-\infty), g(\infty)\}$,
- $\beta = \max\{g(-\infty), g(\infty)\}$,
- $h(y)$ 是 $g(x)$ 的反函数

§ 2.5 随机变量的函数的分布

⇒ 证: $g'(x) > 0$ 的情况

- $F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\}$

- 此时 $g(x)$ 为单调增函数, 所以 $h(y)$ 也是单调增函数, 且有 y 的取值范围为

- $\alpha = g(-\infty) \leq y \leq g(\infty) = \beta,$

- 因此当 $y \leq \alpha$ 时, $F_Y(y) = 0$; 当 $y \geq \beta$ 时, $F_Y(y) = 1$

- 当 $\alpha \leq y \leq \beta$ 时, $F_Y(y) = P\{Y \leq y\} = P\{g(X) \leq y\}$

- $h(y)$ 是单调增函数 $= P\{X \leq h(y)\}$

- $= F_X(h(y))$

- 求导 $f_Y(y) = dF_Y(y)/dy = dF_X(h(y))/dy = \begin{cases} f_X[h(y)]h'(y), & \alpha < y < \beta \\ 0, & \text{其它} \end{cases}$

- $g'(x) < 0$ 的情况, 此时 $h'(y) < 0$

- $P\{g(X) \leq y\} = P\{X \geq h(y)\} = 1 - F_X(h(y))$

- $f_Y(y) = dF(y)/dy = d[1 - F_X(h(y))]/dy = \begin{cases} f_X[h(y)][-h'(y)], & \alpha < y < \beta \\ 0, & \text{其它} \end{cases}$

- 两种情况合并即可得证

§ 2.5 随机变量的函数的分布

推广

- 若 $f_X(x)$ 在有限区间 $[a,b]$ 以外等于0, 则只需假设在 $[a,b]$ 上恒有 $g'(x)>0$ (或 $g'(x)<0$), $\alpha = \min\{g(a), g(b)\}$, $\beta = \max\{g(a), g(b)\}$, 以上结论还成立
- 分段单调函数的情况下, 上述定理可进一步推广为下面形式
 - 假设 $g(x)$ 在下面两个区间 (a_1, b_1) 和 (a_2, b_2) 分别是单调的函数, 其中在 (a_1, b_1) 上的反函数为 $h_1(y)$, 在 (a_2, b_2) 上的反函数为 $h_2(y)$, 且 $f_X(x)$ 在其它有限区间等于0, 则有
 - $$f_Y(y) = \begin{cases} f_X[h_1(y)]|h_1'(y)| + f_X[h_2(y)]|h_2'(y)|, & \alpha < y < \beta \\ 0, & \text{其它} \end{cases}$$
 - 其中 y 的范围由 $g(x)$ 确定
 - 可参考教材中 $Y=X^2$ 的求解过程

§ 2.5 随机变量的函数的分布

⇒ 例4: 正态分布的随机变量的线性变换问题

- 已知随机变量 $X \sim N(\mu, \sigma^2)$, 试证明 X 的线性函数 $Y = aX + b (a \neq 0)$ 也服从正态分布。

⇒ 证: X 的概率密度为

- $$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

- 现在 $y = g(x) = ax + b$, 由这一式子解得 $x = h(y) = (y - b)/a$, 且有 $h'(y) = 1/a$

- 显然 $g(x)$ 是严格单调的, 由定理有

- $$f_Y(y) = \begin{cases} f_X[h(y)] |h'(y)|, & \alpha < y < \beta \\ 0, & \text{其它} \end{cases} = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right), \quad -\infty < y < \infty$$

- $$= \frac{1}{|a|} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\frac{y-b}{a} - \mu)^2}{2\sigma^2}} = \frac{1}{|a| \sigma \sqrt{2\pi}} e^{-\frac{[y - (b + a\mu)]^2}{2(a\sigma)^2}}, \quad -\infty < y < \infty$$

⇒ 所以有

- $Y = aX + b \sim N(a\mu + b, (a\sigma)^2)$
- 特别的当 $a = 1/\sigma$, $b = -\mu/\sigma$ 时, $Y = (X - \mu)/\sigma \sim N(0, 1)$
- 即上一节的引理的结果

思考

设 $f(x)$ 是连续函数, 若 X 是离散型随机变量, 则 $Y = f(X)$ 也是离散型随机变量吗? 若 X 是连续型的, 又怎样?

答 若 X 是离散型随机变量, 它的取值是有限个或可列无限多个, 因此 Y 的取值也是有限个或可列无限多个, 因此 Y 是离散型随机变量. 若 X 是连续型随机变量, 那末 Y 不一定是连续型随机变量.

本章小结 (一)

随机变量X的分布

分布函数 $F(x)$

- X落在区间 $(-\infty, x)$ 上的概率
- 不减函数、右连续性
- 求解分布函数, 概率密度要讨论区间 $(-\infty, \infty)$ 上的全部情况

数学期望

方差

离散型: 分布律	分布	数学期望	方差	应用
离散型: 分布律	(0-1)分布	p	$p(1-p)$	一次伯努利试验 $P(A)$
	二项分布 $X \sim b(n, p)$ 最值问题	np	$np(1-p)$	n 重伯努利试验, A发生的次数
	泊松分布 $X \sim \pi(\lambda)$ 问题	λ	λ	某确定时间段内A发生的次数
	超几何分布 $X \sim H(n, D, N)$	$n(M/N)$	$n(M/N)(1-M/N)[(N-n)/(N-1)]$	不放回抽样时A发生次数的概率, 而二项分布相当于放回
连续型: 概率密度	均匀分布 $X \sim U(a, b)$	$(a+b)/2$	$(b-a)^2/12$	
	指数分布, 无记忆性	θ	θ^2	事件发生的时间间隔
	正态分布 $X \sim N(\mu, \sigma^2)$	μ	σ^2	多种随机因素综合作用

随机变量X的函数的分布一般步骤

已知 $f_X(x)$, $Y=g(X)$, 求 $f_Y(y)$

1° 先写出Y的分布函数定义式: $F_Y(y) = P\{Y \leq y\}$

由 $Y=g(X)$ 确定Y的值域, 当y不在值域范围内时单独讨论

2° 将 $Y=g(X)$ 代入上式 $= P\{g(X) \leq y\}$

3° 由 $g(X) \leq y$ 求解X的范围 $= P\{X | g(X) \leq y\}$ 表示为y的形式

4° 由X的分布函数 $F_X(x)$ 表示以上概率, 得到关于y的表达式 $F_Y(y)$, 其原自变量x用关于y的表达式来代

5° 求导得 $f_Y(y) = dF_Y(y)/dy$

图象特性

表达式

标准正态分布

1° $X \sim N(0, 1)$, $\phi(x)$ 概率密度函数

2° 分布函数 $F(x) = 1 - F(-x)$

3° $Z = (X - \mu) / \sigma \sim N(0, 1)$

4° 3σ 准则

5° 上 α 分位点

其它需了解的分布

对数正态分布

Γ 分布

β 分布

Weibull分布

本章作业

- P_{55} : 2, 4, 5, 8
- P_{56} : 10, 12, 13, 17, 21, 24, 26, 29
- P_{59} : 32, 33, 35, 36, 37

基于MATLAB的概率统计数值实验一

古典概型

MATLAB常用的及与随机数产生相关的函数

实验1：计算超几何分布

实验2：频率稳定性实验

实验3：利用频率估计自然对数底e

实验4：蒲丰投针实验，利用频率估计圆周率 π

实验5：生日悖论实验

一、古典概型

- 利用MATLAB 软件的图形可视功能将概率统计的内容用图形表示出来，以加深对概率的理解
- MATLAB常用的及与随机数产生相关的函数
 - **factorial(n)**：阶乘， $n!$ ，可通过阶乘来计算排列组合数
 - **1.rand(m,n)**：生成 $m \times n$ 的随机矩阵，每个元素都在 $(0,1)$ 间，生成方式为均匀分布。
 - **2.randn(m,n)**：生成 $m \times n$ 的随机矩阵，每个元素都在 $(0,1)$ 间，生成方式为正态分布
 - **3.randperm(m)**：生成一个 $1 \sim m$ 的随机整数排列
 - **4.perms(1:n)**：生成一个 $1 \sim n$ 的全排列，共 $n!$ 个
 - 5.取整函数系列：
 - (1) **fix(x)**：截尾法取整；
 - (2) **floor(x)**：退一法取整（不超过 x 的最大整数）；向负方向舍入
 - (3) **ceil(x)**：进一法取整（ $= \text{floor}(x)+1$ ）；向正方向舍入
 - (4) **round(x)**：四舍五入法取整。
 - **6.unique(a)**：合并 a 中相同的项
 - **7.prod(x)**：向量 x 的所有分量元素的积

一、古典概型

⇒ 示例:

```
>> rand(1) %生成一个(0,1)间的随机数
```

```
ans = 0.8147
```

```
>> rand(2,2) %生成一个2×2阶(0,1)间的随机数矩阵
```

```
ans = 0.9134 0.0975
```

```
0.6324 0.2785
```

```
>> randperm(5) %生成一个1~5的随机整数排列
```

```
ans = 4 1 5 2 3
```

```
>> a=[1 2 4 2 3 3 2];
```

```
unique(a)
```

```
ans = 1 2 3 4
```

⇒ 实验 1: 计算超几何分布的结果

- 设有N件产品，其中D件次品，今从中任取n件
- 问其中恰有k(k≤D)件次品的概率是多少?
- (令N=10, D=3, n=4, k=2)

$$\frac{C_D^k \times C_{N-D}^{n-k}}{C_N^n}$$

⇒ 解: 编辑组合函数zuhe.m文件

- `function y=Com(n,r)`
- `y=factorial(n)/(factorial(r)*factorial(n-r))`

⇒ 计算如下:

- `>> N=10; D=3; n=4; k=2;`
- `p=Com(3,2)*Com(10-3,4-2)/Com(10,4)=0.3`

⇒ 实验2 频率稳定性实验

- 随机投掷均匀硬币，观察国徽朝上与国徽朝下的频率

⇒ 解

- `>> n= 3000~100000000;m=0;`
- `for i=1:n`
- `t=randperm(2); %生成一个1~2的随机整数排列`
- `x=t-1; %生成一个0~1的随机整数排列`
- `y=x(1); %取x排列的第一个值`
- `if y==0;`
- `m=m+1;`
- `end`
- `end`
- `p1=m/n`
- `p2=1-p1`

试验次数n	3000	5000	1万	2万	3万
国徽朝上 频率	0.5040	0.5006	0.4879	0.4999	0.5046
国徽朝下 频率	0.4960	0.4994	0.5121	0.5001	0.4954

试验次数n	5万	10万	100万	100万	1亿
国徽朝上 频率	0.5021	0.4999	0.4999	0.5001	0.5000
国徽朝下 频率	0.4979	0.5001	0.5001	0.4999	0.5000

可见当 $n \rightarrow \infty$ 时, $f_n(A) = P(A)$

实验3 用频率估计自然对数 e

- 某班有 n 个人，每人各有一支枪，这些枪外形一样。某次夜间紧急集合，若每人随机地取走一支枪，求没有一个人拿到自己枪的概率？

解：记事件 A_i 为第 i 个人拿到自己枪，事件 \bar{A}_i 为第 i 个人没拿到自己枪，易知：

- $P(A_i) = \frac{1}{n}$; $P(\bar{A}_i) = \frac{n-1}{n}$, $(i = 1, 2, \dots, n)$

- 又记 p_0 为没有一个人拿到自己枪的概率。

$$p_0 = P(\bar{A}_1 \bar{A}_2 \cdots \bar{A}_n) = 1 - P\left(\bigcup_{i=1}^n A_i\right)$$

- 由乘法公式可知

$$P(A_i A_j) = P(A_i) \times P(A_j | A_i) = \frac{1}{n(n-1)} \quad (1 \leq i < j \leq n)$$

- $P(A_i A_j A_k) = P(A_i A_j) \times P(A_k | A_j A_i) = \frac{1}{n(n-1)(n-2)} \quad (1 \leq i < j < k \leq n)$
... ..

$$P(A_1 A_2 A_3 \cdots A_n) = \frac{1}{n!}$$

于是
$$\sum_{i=1}^n P(A_i) = 1, \quad \sum_{1 \leq i < j \leq n} P(A_i A_j) = \frac{C_n^2}{n(n-1)}$$

$$\sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) = \frac{C_n^3}{n(n-1)(n-2)}, \dots$$

$$P(A_1 A_2 A_3 \dots A_n) = \frac{1}{n!}$$

所以
$$p_0 = 1 - P\left(\bigcup_{i=1}^n A_i\right) = 1 - \left[1 - \frac{C_n^2}{n(n-1)} + \frac{C_n^3}{n(n-1)(n-2)} + \dots + \frac{(-1)^{n+1}}{n!} \right] = \sum_{k=0}^n \frac{(-1)^k}{k!}$$

特别地，当 n 较大时， $p_0 \approx e^{-1}$ 。

因此，可随机模拟出没有人拿到自己枪的频率，根据频率的稳定性，近似当做概率，然后去估计自然对数 e 。并考虑估计精度与人数是否有关系，为什么。算法如下：

- ⇒ 1、产生 n 个随机数的随机序列；
- ⇒ 2、检验随机列与自然列是否至少有一个配对；
- ⇒ 3、对没有一个配对的序列进行累积 t ；
- ⇒ 4、重复1、2、3步 m 次；
- ⇒ 5、估计 $e = \frac{m}{t}$

⇒ 具体程序及相关结果如下页图

- 注：自然常数 $e \approx 2.7183$

```

>> m=40000;
n=50;
t=0;
for j=1:m
    k=0;
    sui=randperm(n);
    for i=1:n
        if sui(i)==i
            k=k+1;
        else
            k=k;
        end
    end
end
if k==0
    t=t+1;
else
    t=t;
end
end
e=m/t
e = 2.7313

```

模拟次数			
m	4000	40000	400000
人数			
n	50	50	50
e	2.7379	2.7313	2.7194

模拟次数			
m	40000	40000	40000
人数			
n	1000	2000	5000
e	2.7155	2.7082	2.7202

实验4: 蒲丰(Buffon)投针实验, 用频率估计 π 值

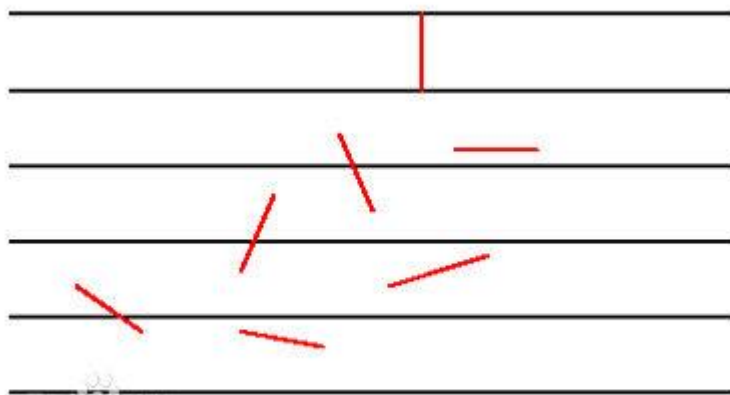
18世纪, 法国数学家布丰提出的“投针问题”, 记载于布丰1777年出版的著作中: “在平面上画有一组间距为 a 的平行线, 将一根长度为 l ($l \leq a$) 的针任意掷在这个平面上, 求此针与平行线中任一条相交的概率。”

布丰本人证明了, 这个概率是:

$$p = \frac{2l}{\pi a}$$

由于它与 π 有关, 于是人们想到利用投针试验来估计圆周率的值。

布丰惊奇地发现: 有利的扔出与不利的扔出两者次数的比, 是一个包含 π 的表示式. 如果针的长度等于 $a/2$, 那么有利扔出的概率为 $1/\pi$. 扔的次数越多, 由此能求出越为精确的 π 的值。



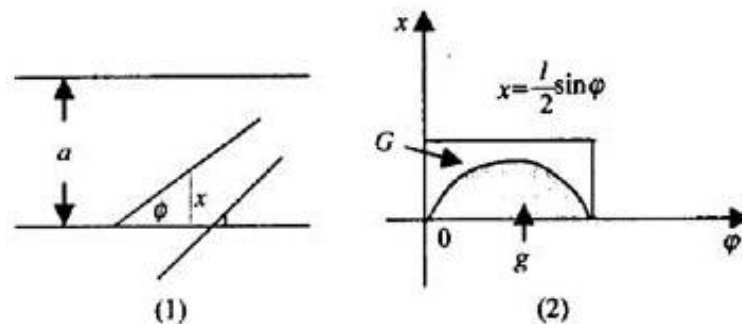
实验4: 蒲丰(Buffon)投针实验, 用频率估计 π 值

- 在画有许多间距为 d 的等距平行线的白纸上, 随机投掷一根长为 $l(l \leq d)$ 的均匀直针, 求针与平行线相交的概率, 并计算 π 的近似值

- 解: 设针与平行线的夹角为 $\alpha(0 \leq \alpha \leq \pi)$, 针的中心与最近直线的距离为 $x(0 \leq x \leq d/2)$ 。针与平行线相交的充要条件是 $x \leq (l/2)\sin \alpha$, 这里 $x(0 \leq x \leq d/2)$ 并且 $0 \leq \alpha \leq \pi$ 。建立直角坐标系, 上述条件在坐标系下将是曲线所围成的曲边梯形区域, 总的区域即 x 和 α 所有可能取值构成的矩形区域, 且所有可能取值是机会均等的, 符合几何概型, 则所求概率为

$$p = \frac{g \text{ 的面积}}{G \text{ 的面积}} = \frac{\int_0^\pi \frac{l}{2} \sin \alpha d\alpha}{\pi \frac{d}{2}} = \frac{2l}{\pi d} \approx \frac{m}{n}$$

- 故可得 π 的近似计算公式 $\pi \approx \frac{2nl}{md}$, 其中 n 为随机试验次数, m 为针与平行线相交的次数。



➤ 解

- `>> clear,clf`
- `n=10000000;l=0.5;m=0;d=1;`
- `for i=1:n`
- `x=(l/2)*sin(rand(1)*pi); y=rand(1)*d/2;`
- `if x>=y`
- `m=m+1;`
- `end`
- `end`
- `p1=m/n`
- `pai=2*n*l/(m*d)`

确定针与平行线的角度

确定针中心点的位置

试验次数n	5千	1万	10万	100万	1000万
针长l/平行间距d	3/10	3/10	3/10	3/10	3/10
相交频率	0.1836	0.1971	0.1887	0.1905	0.1912
π 的近似值	3.2680	3.0441	3.1798	3.1498	3.1387
试验次数n	5千	1万	10万	100万	1000万
针长l/平行间距d	2/5	2/5	2/5	2/5	2/5
相交频率	0.2496	0.2562	0.2549	0.2544	0.2543
π 的近似值	3.2051	3.1226	3.1386	3.1451	3.1433

试验次数n	5千	1万	10万	100万	1000万
针长l/平行间距d	1/2	1/2	1/2	1/2	1/2
相交频率	0.3254	0.3148	0.3158	0.3178	0.3183
π 的近似值	3.0731	3.1766	3.1667	3.1470	3.1417
试验次数n	5千	1万	10万	100万	1000万
针长l/平行间距d	4/5	4/5	4/5	4/5	4/5
相交频率	0.5142	0.5134	0.5086	0.5093	0.5093
π 的近似值	3.1116	3.1165	3.1460	3.1418	3.1418

试验次数n	5千	1万	10万	100万	1000万
针长l/平行间距d	17/20	17/20	17/20	17/20	17/20
相交频率	0.5432	0.5452	0.5420	0.5412	0.5410
π 的近似值	3.1296	3.1181	3.1366	3.1413	3.1426

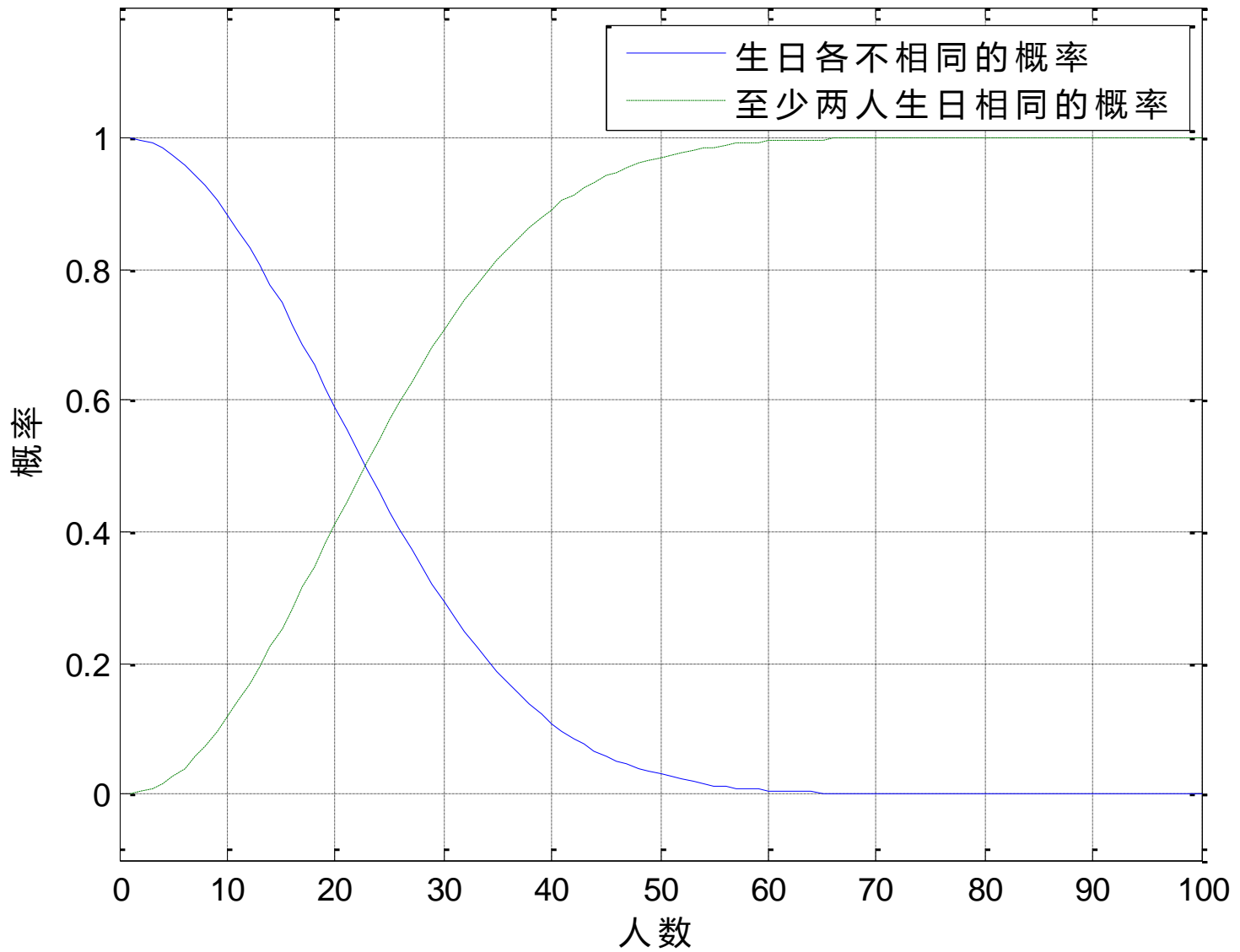
试验次数n	5千	1万	10万	100万	1000万
针长l/平行间距d	9/10	9/10	9/10	9/10	9/10
相交频率	0.5860	0.5700	0.5756	0.5733	0.5731
π 的近似值	3.0717	3.1579	3.1272	3.1395	3.1410

实验5 生日悖论实验

- 在100个人的团体中，不考虑年龄差异，研究是否有两个以上的人生日相同。假设每人的生日在一年365天中的任意一天是等可能的，那么随机找 n 个人（不超过365人）。
- (1) 求这 n 个人生日各不相同的概率是多少？从而求这 n 个人中至少有两个人生日相同这一随机事件发生的概率是多少？
- (2) 近似计算在30名学生的一个班中至少有两个人生日相同的概率是多少

➤ 解: (1)

- `>> clear,clf`
- `for n=1:100`
- `p0(n)=prod(365:-1:365-n+1)/365^n;`
- `p1(n)=1-p0(n);`
- `end`
- `p1=ones(1,100)-p0;`
- `n=1:100;`
- `plot(n,p0,n,p1,'--')`
- `xlabel('人数'),ylabel('概率')`
- `legend('生日各不相同的概率','至少两人生日相同的概率')`
- `axis([0 100 -0.1 1.199]),grid on`



⇒ $p_1(30)=0.7063$, $p_1(60)= 0.9941$

- (2) 在30名学生中至少两人生日相同的概率为70.63%。
- 下面进行计算机仿真。
 - 随机产生30个正整数，代表一个班30名学生的生日，然后观察是否有两人以上生日相同。当30个人中有两人生日相同时，输出“1”，否则输出“0”。如此重复观察100次，计算出这一事件发生的频率 f_{100}

```
>> clear,clf
n=0;
for m=1:100                %做100次随机试验
    y=0;
    x=1+fix(365*rand(1,30)); %产生30个随机数
    for i=1:29 %用二重循环寻找30个随机数 中是否有相同数
        for j=i+1:30
            if x(i)==x(j)
                y=1;break;
            end
        end
    end
    n=n+y;                %累计有两人生日相同的试验次数
end
f=n/m                    %计算频率
```

⇒ 作业

- 1. 利用实验3求自然对数底 e ，精确到2.7183，并给出得到该精确值时的实验次数。将该实验做三次，比较一下所得结果。
- 2. 求50个人中至少两人生日相同的概率