

# Introduction to Probability: Class-14

## Probability Distribution and Characteristic Function

YUTONG XIE<sup>1</sup>  
*ACM Class, 2016*  
2018 年 6 月 18 日

### 摘要

本节课中老师先是接着上节课的内容，补充了 Gamma 分布与 Cauchy 分布。然后便为了给下节课的中心极限定理做铺垫，着重对特征函数进行了讨论，并给出了常见概率分布的特征函数。其间还涉及到了母函数及一些有趣的数在概率问题中的应用此外囿于笔者本人浅薄的数学水平，笔记中难免会有说明错误或叙述不当之处，若有发现，还望不吝指出并赐教，感谢！

---

<sup>1</sup>如有任何问题，烦请联系 Email: [xxxxxyt@sjtu.edu.cn](mailto:xxxxxyt@sjtu.edu.cn)

# 目录

<b>1</b>	<b>常见概率分布</b>	<b>3</b>
1.1	Gamma 分布	3
1.1.1	习题: Gamma 分布与指数分布的关系	3
1.1.2	习题: Gamma 分布与 Poisson 分布的关系	4
1.2	Cauchy 分布	4
<b>2</b>	<b>概率分布的其他表示方法</b>	<b>4</b>
2.1	母函数	5
2.1.1	例题: 帽子问题	5
2.2	Laplace 变换	6
2.3	矩母函数	7
2.4	特征函数/Fourier 变换	7
2.4.1	习题: 独立随机变量的和的特征函数	7
2.4.2	定理: 随机变量的独立性与特征函数	8
<b>3</b>	<b>常见概率分布的特征函数</b>	<b>9</b>
3.1	Bernoulli 分布	9
3.2	二项分布	9
3.3	Poisson 分布	9
3.4	均匀分布	9
3.5	标准正态分布	10
3.6	指数分布	11
3.7	标准 Laplace 分布	11
<b>4</b>	<b>有趣的数</b>	<b>12</b>
4.1	37	12
4.1.1	投放炸弹问题	12
4.1.2	挑选驸马问题	13
4.2	$\pi$	13
4.2.1	Buffon 投针问题	13

# 1 常见概率分布

## 1.1 Gamma 分布

随机变量  $X$  服从 Gamma 分布时有

$$f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

记作  $X \sim \Gamma(\alpha, \beta) \equiv \text{Gamma}(\alpha, \beta)$

其中  $\alpha > 0$  为形状参数,  $\beta > 0$  为速率参数;  $\Gamma(\alpha)$  为 Gamma 函数, 满足

$$\begin{aligned} \Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi} \\ \Gamma(\alpha) &= (\alpha - 1)! \end{aligned}$$

### 1.1.1 习题: Gamma 分布与指数分布的关系

习题 随机变量  $X, Y$  分布服从 Gamma 分布与指数分布, 即

$$\begin{cases} X \sim \text{Gamma}(n, \lambda) \\ Y \sim \text{Exponential}(\lambda) \end{cases}$$

试证明

$$X + Y \sim \text{Gamma}(n + 1, \lambda)$$

解答

$$\begin{cases} P(X = x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}, & x \geq 0 \\ P(Y = y) = \lambda e^{-\lambda y}, & x \geq 0 \end{cases}$$

$$\begin{aligned} f(X + Y = m) &= \int_0^m P(X = x) \cdot P(Y = m - x) dx \\ &= \int_0^m \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \cdot \lambda e^{-\lambda(m-x)} dx \\ &= \frac{\lambda^n}{(n-1)!} \cdot e^{-\lambda m} \int_0^m x^{n-1} e^{-\lambda x} \cdot e^{\lambda x} dx \\ &= \frac{\lambda^n}{(n-1)!} \cdot e^{-\lambda m} \int_0^m x^{n-1} dx \\ &= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda m} \cdot \frac{1}{n} m^n \\ &= \frac{\lambda^{n+1}}{n!} m^n e^{-\lambda m} \end{aligned}$$

对比 Gamma 分布定义即得  $X + Y \sim \text{Gamma}(n + 1, \lambda)$

### 1.1.2 习题：Gamma 分布与 Poisson 分布的关系

习题 对于任意整数  $\alpha$ ，任意数  $\beta$ ，假设随机变量

$$\begin{cases} X \sim \text{Gamma}(\alpha, \beta) \\ Y \sim \text{Poisson}(x\beta) \end{cases}$$

试证明

$$P(X \leq x) = P(Y \geq \alpha)$$

提示 从上一习题中可看到 Gamma 分布与指数分布的关系，同时上节课中又提到了指数分布与 Poisson 分布的关系

## 1.2 Cauchy 分布

随机变量  $X$  服从 Cauchy 分布时有

$$f(x) = \frac{1}{\beta\pi} \frac{1}{1 + (x - \alpha)^2 / \beta^2}$$

记作  $X \sim \text{Cauchy}(\alpha, \beta)$

Cauchy 分布并不像 Gamma 分布那样自然，它是 Poisson 人为构造出的一个反例<sup>2</sup>

## 2 概率分布的其他表示方法

概率分布能够非常直观地描述事件空间的概率模型，但若只单纯从概率分布的朴素视角来分析概率模型，可能很难得到一些结论或性质的推导证明。而母函数、矩母函数、特征函数等概率分布的其他表示方法，虽然不能直观地体现出概率模型的特性，但在这些视角下却能方便地得到概率模型的很多性质

常见的概率分布的其他表示方法如下：

- 母函数 (Generating Function)
- Laplace 变换 (Laplace Transform)
- 矩母函数 (Moment Generating Function)
- 特征函数 (Characteristic Function) / Fourier 变换 (Fourier Transform)

之前提到的表示方法具有以下两个重要的性质，使得它们和原概率分布具有密切的联系：

---

<sup>2</sup>[Wikipedia: Cauchy Distribution - History](#)

- 唯一性：特征函数（或其他的表示有意义且收敛时）唯一地决定分布<sup>3</sup>，而将特征函数变回概率分布则需用到反演公式
- 独立性：设相互独立的随机变量的和  $X = \sum Y_j$ ，则  $h(X) = \prod h(Y_j)$ ，其中  $h(\cdot)$  为以上的某种表示<sup>4</sup>

## 2.1 母函数

定义 母函数是将  $(b_0, b_1, b_2, \dots)$  表示为

$$\sum_{i=0}^{\infty} b_i t^i$$

的一种表示方法。对于离散概率分布

$$(a_n)_{n=0}^{\infty}, \quad a_n = P(X = n)$$

定义其母函数为

$$g(z) := E[z^X] = \sum_{n=0}^{\infty} a_n z^n$$

关于母函数，老师在 Class15 中还讨论了它的一些重要性质，这里不再赘述，感兴趣的同学可以参见 Class15 的笔记

### 2.1.1 例题：帽子问题

**问题描述**  $n$  个人，每人各拥有 1 顶帽子。现将他们的帽子放在一起随机打乱，并再让他们随机拿取。试问没有 1 个人取回自己帽子的概率为？

**解答** 考虑如下数列

$$\begin{aligned} b_0 &= 1 \\ b_n &= P(X_n = 0), n \geq 1 \end{aligned}$$

其中  $X_n$  为在  $S_n$  中随机选取的置换  $\sigma$  中不动点的个数，即拿回自己帽子的人的人数，是为一随机变量

定义母函数

---

<sup>3</sup>证明请参考应坚刚. 何萍. 概率论. 定理 2.5.8; 同时概率分布也是唯一决定特征函数的，这是因为特征函数在定义中是由概率分布运算得到，而运算过程中无二义性，故其实特征函数与概率分布是一一对应的关系

<sup>4</sup> $X = \sum Y_j$  的概率密度的卷积关系可以在其他表示的视角下较为容易地看出

$$B(t) = \sum_{i=0}^{\infty} b_i t^i$$

注意到由全概率公式，对于人数为  $n$  的情况，有如下等式成立

$$1 = \sum_{k=0}^n P(X_n = k) = \frac{\binom{n}{k} (n-k)! b_{n-k}}{n!} = \frac{b_{n-k}}{k!} \quad (1)$$

等式1成立是因为使得  $n$  个人中恰好  $k$  个人拿回自己帽子的置换数，等于先确定哪  $k$  个人拿回帽子，再让剩下的  $n-k$  人都拿不回自己帽子的置换数。而先选定  $k$  人的方案数为  $\binom{n}{k}$ ，剩下的  $n-k$  人都没取回自己帽子的方案数为  $(n-k)! b_{n-k}$

题目要求的  $b_n$ ，可以通过等式1变形而得的

$$b_n = 1 - \sum_{k=0}^{n-1} \frac{b_{n-k}}{k!}$$

递推得到

此外，得到等式1后即可进一步得到

$$\frac{1}{1-t} = B(t)e^t \quad (2)$$

其中

$$\begin{aligned} \frac{1}{1-t} &= 1 + t + t^2 + \dots \\ e^t &= 1 + \frac{t}{1!} + \frac{t^2}{2!} + \dots \end{aligned}$$

对比展开后  $t^n$  系数可知需要满足

$$1 = \sum_{k=0}^n \frac{b_{n-k}}{k!}$$

而这一结果已在之前等式1中得到，故等式2成立

## 2.2 Laplace 变换

**定义** 对于  $\mathbb{R}$  上的随机变量，定义其 *Laplace* 变换结果为

$$L_X(\lambda) := E[e^{-\lambda X}] = \int e^{-\lambda x} P(x) dx$$

其中  $\lambda \in [0, \infty)$

注意对于任意概率分布，其 *Laplace* 变换的结果不一定有意义，即不一定收敛

## 2.3 矩母函数

定义 对于  $\mathbb{R}$  上的随机变量  $X$ ，定义其矩母函数为

$$M_X(t) := E[e^{tX}] = \int e^{tx} P(x) dx$$

其中  $t \in \mathbb{R}$

注意对于任意概率分布，其矩母函数同样不一定有意义

矩母函数在课本上和 Class15 中有较为详细的叙述，囿于篇幅，这里不再赘述，请感兴趣的同学参考教材内容或 Class15 的笔记

## 2.4 特征函数/Fourier 变换

定义 对于  $\mathbb{R}^n$  上的概率测度<sup>5</sup>  $\mu$ ，定义其 Fourier 变换结果为

$$\hat{\mu}(\vec{u}) := \int e^{i\langle \vec{u}, \vec{x} \rangle} \mu(d\vec{x})$$

其中  $\vec{u} \in \mathbb{R}^n$ ， $\langle \vec{u}, \vec{x} \rangle$  为  $\vec{u}, \vec{x}$  的内积

定义 对于  $\mathbb{R}^n$  上的随机变量  $\vec{X}$ （可理解为随机向量  $(X_1, X_2, \dots, X_n)$ ），定义其特征函数为

$$\varphi_X(\vec{u}) := E[e^{i\langle \vec{u}, \vec{X} \rangle}] = \int e^{i\langle \vec{u}, \vec{x} \rangle} P(d\vec{x}) = \int e^{i\langle \vec{u}, \vec{x} \rangle} P(\vec{x}) d\vec{x}$$

其中  $\vec{u} \in \mathbb{R}^n$  为随机变量  $\vec{X}$  的概率分布的特征向量，对其进行 Fourier 变换即可得到相应特征值  $\varphi_X(\vec{u}) \in \mathbb{C}$

可以看到矩母函数实际上是特征函数的一种特殊情况，实际上任何概率分布的特征函数都是存在的，所以经常会用特征函数来讨论概率分布的普遍性质

### 2.4.1 习题：独立随机变量的和的特征函数

习题 随机变量  $X = \sum_{j=0}^n Y_j$ ，其中  $Y_j$  为相互独立的随机变量。试证明

$$\varphi_X(u) = \prod_{j=0}^n \varphi_{Y_j}(u)$$

---

<sup>5</sup>Wikipedia: Measure 测度可粗略理解为某集合上定义的从子集到数的满足某些性质的映射，一个概率分布实质上就是该事件空间上的某种测度。这里应该是可以将“概率测度”直接替换成“概率分布”来理解的，但因老师上课时用到的描述是“测度”，故沿用

解答 根据题目条件有

$$\begin{aligned}\varphi_X(u) &= E[e^{iuX}] = E[e^{iu\sum Y_j}] \\&= E[\prod_{j=0}^n e^{iuY_j}] = \prod_{j=0}^n E[e^{iuY_j}] \quad \text{随机变量的独立性} \\&= \prod_{j=0}^n \varphi_{Y_j}(u) \quad \text{特征函数的定义}\end{aligned}$$

评注 请注意证明时不能直接进行形如  $E[e^{iuX}] = e^{iuE[X]}$  的变换, 因为  $E[g(X)] = g(E[X])$  仅在  $g(\cdot)$  为线性函数时成立<sup>6</sup>

## 2.4.2 定理: 随机变量的独立性与特征函数

定理 对于定义在  $\mathbb{R}^n$  上的随机变量  $\vec{X} = (X_1, X_2, \dots, X_n)$ , 其每个维度  $(X_j)_{j=1}^n$  相互独立当且仅当

$$\varphi_X(u_1, u_2, \dots, u_n) = \prod_{j=1}^n \varphi_{X_j}(u_j)$$

其中

$$\varphi_X(u_1, u_2, \dots, u_n) := E[e^{i\sum u_j X_j}]$$

证明 由独立性推得等式可以用刚才习题的方法证明

$$\begin{aligned}E[e^{i\sum u_j X_j}] &= E[\prod_{j=1}^n e^{u_j X_j}] \\&= \prod_{j=1}^n E[e^{u_j X_j}] \quad \text{随机变量的独立性} \\&= \prod_{j=1}^n \varphi_{X_j}(u_j) \quad \text{特征函数的定义}\end{aligned}$$

再结合特征函数的唯一性定理<sup>7</sup>, 便能从等式推得随机变量的独立性

---

<sup>6</sup>最容易找出的反例即  $E[X^2] \neq (E[X])^2$

<sup>7</sup>假设  $(X_j)_{j=1}^n$  相互独立, 则  $\varphi(\vec{u}) = \prod \varphi_{X_j}(u_j)$ 。由于特征函数唯一地决定了概率分布, 则若  $(X_j)_{j=1}^n$  不相互独立的话  $\varphi(\vec{u}) \neq \prod \varphi_{X_j}(u_j)$ 。故在等式成立时,  $(X_j)_{j=1}^n$  必然相互独立, 即特征函数的关系和概率分布的关系一一对应



### 3 常见概率分布的特征函数

下面给出一些常见概率分布的特征函数及推导过程<sup>8</sup>

#### 3.1 Bernoulli 分布

$$X \sim \text{Bernoulli}(p)$$

$$\varphi_X(u) = E[e^{iux}] = e^{iu \cdot 0}(1-p) + e^{iu \cdot 1}p = pe^{iu} + 1 - p$$

其中  $X, u \in \mathbb{R}$ ,  $\varphi_X(u)$  在实轴上的任意点都有值, 而  $X$  仅离散地分布在  $0, 1$

#### 3.2 二项分布

$$X \sim \text{Binomial}(n, p)$$

$$\varphi_X(u) = (pe^{iu} + 1 - p)^n$$

该结果可利用 Bernoulli 分布的特征函数及随机变量的独立性直接得到

#### 3.3 Poisson 分布

$$X \sim \text{Poisson}(\lambda)$$

$$\begin{aligned}\varphi_X(u) &= \sum_{k=0}^{\infty} e^{iuk} \cdot e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{(\lambda e^{iu})^k}{k!} \\ &= e^{-\lambda} \cdot e^{\lambda e^{iu}} = e^{\lambda(e^{iu}-1)}\end{aligned}$$

#### 3.4 均匀分布

$$X \sim \text{Uniform}(-a, a)$$

$$\varphi_X(u) = \frac{1}{2a} \int_{-a}^a e^{iux} dx = \frac{e^{iua} - e^{-iua}}{2ai u} = \frac{\sin au}{au}$$

---

<sup>8</sup>常用概率分布及特征函数表

### 3.5 标准正态分布

$$X \sim \text{Normal}(0, 1)$$

$$\varphi_X(u) = e^{-\frac{u^2}{2}}$$

**推导 1** 朴素的推导方法是利用围道积分公式<sup>9</sup>

$$\begin{aligned}\varphi_X(u) &= \int e^{iux} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int e^{-\frac{x^2}{2} + iux} dx \\ &= \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} \int e^{-\frac{(x-iu)^2}{2}} dx \\ &= e^{-\frac{u^2}{2}}\end{aligned}$$

其中最后一个积分用到了围道积分即 **Cauchy** 积分原理,  $\int_{-N}^N \exp(x-iu)^2/2 dx$  可视为解析函数  $z \mapsto e^{-z^2/2}$  从  $-N-iu$  到  $N-iu$  的 (与路径无关) 积分

$$\int_{-N}^N e^{-\frac{(x-iu)^2}{2}} dx = \int_{-N-iu}^{N-iu} e^{-\frac{z^2}{2}} dz = \left( \int_{-N-iu}^{-N} + \int_{-N}^N + \int_N^{N-iu} \right) e^{-\frac{z^2}{2}} dz$$

令  $N \rightarrow \infty$ , 则算式的第一个与第三个积分趋近于零<sup>10</sup>, 而第二个积分是  $\sqrt{2\pi}$

**推导 2** 课上还提供了一种具有技巧性的推导方法

$$\begin{aligned}\varphi_X(u) &= \int e^{iux} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \\ &= \int \frac{\cos ux}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx + i \int \frac{\sin ux}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int \cos ux e^{-\frac{x^2}{2}} dx\end{aligned}$$

此时, 对  $\varphi_X(u)$  求导<sup>11</sup>有

<sup>9</sup>此处证明参考了应坚刚. 何萍. 概率论. 例 2.5.9

<sup>10</sup>这里需要用到复积分的知识

<sup>11</sup>积分号下求导定理, 囿于篇幅, 请参见陈纪修. 数学分析. 第二版. 下册. 定理 15.2.7

$$\begin{aligned}
\varphi_X(u) &= \frac{1}{\sqrt{2\pi}} \int -x(\sin ux)e^{-\frac{x^2}{2}} dx \\
&= \frac{-1}{\sqrt{2\pi}} \int u(\cos x)e^{-\frac{x^2}{2}} dx \quad \text{分部积分} \\
&= -u\varphi_X(u)
\end{aligned}$$

$$\Rightarrow \frac{\varphi_X(u)}{\varphi_X(u)} = -u \Rightarrow \ln |\varphi_X(u)| = -\frac{u^2}{2} + C$$

再由边界条件

$$\begin{aligned}
\varphi_X(0) &= \int e^{i \cdot 0 \cdot x} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx = \int \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx = 1 \\
&\Rightarrow \varphi_X(u) = e^{-\frac{u^2}{2}}
\end{aligned}$$

有兴趣的同学还可以类似地对高维正态分布的特征函数进行求解

### 3.6 指数分布

$$X \sim \text{Exponential}(\lambda)$$

$$\begin{aligned}
\varphi_X(u) &= \int_0^\infty e^{iux} \cdot \lambda e^{-\lambda x} dx \\
&= \lambda \int_0^\infty (e^{iu-\lambda})^x dx \\
&= \frac{\lambda}{iu - \lambda} \int_0^\infty d(e^{iu-\lambda})^x \\
&= \frac{\lambda}{\lambda - iu}
\end{aligned}$$

其中  $\lambda > 0 \Rightarrow |e^{iu-\lambda}| < 1$

### 3.7 标准 Laplace 分布

$$X \sim \text{Laplace}(0, 1), \quad P(X = x) = \frac{1}{2}e^{-|x|}$$

$$\begin{aligned}
\varphi_X(u) &= \int e^{iux} \cdot \frac{1}{2} e^{-|x|} dx \\
&= \frac{1}{2} \left( \int_{-\infty}^0 e^{(iu+1)x} dx + \int_0^{\infty} e^{(iu-1)x} dx \right) \\
&= \frac{1}{2} \left( \frac{1}{iu+1} \int_{-\infty}^0 de^{(iu+1)x} + \frac{1}{iu-1} \int_0^{\infty} de^{(iu-1)x} \right) \\
&= \frac{1}{2} \left( \frac{1}{iu+1} - \frac{1}{iu-1} \right) = \frac{1}{1+u^2}
\end{aligned}$$

Laplace 分布又称双指数分布，这是因为 Laplace 分布可由两个对称的指数分布相加得到

$$\begin{aligned}
Y &\sim \text{Exponential}(1) \\
P(Y = y) &= e^{-y} \\
P(X = x) &= \frac{1}{2} (P(Y = x) + P(Y = -x)) = \frac{1}{2} e^{-|x|}
\end{aligned}$$

同时标准 Laplace 分布与标准指数分布的特征函数满足如下关系<sup>12</sup>

$$\frac{1}{1+u^2} = \frac{1}{1-iu} + \frac{1}{1+iu}$$

## 4 有趣的数

### 4.1 37

#### 4.1.1 投放炸弹问题

**问题** 向  $n$  个区域等概率地随机投放  $n$  个炸弹，试求某个区域不被炸到的概率

**解答** 某个区域在  $n$  次炸弹投放中始终不被炸到的概率为  $(1 - \frac{1}{n})^n$ ，取极限得到

$$\lim_{n \rightarrow \infty} (1 - \frac{1}{n})^n = \frac{1}{e} \approx 0.37$$

**其他** 机器学习中自助采样法 (booststrapping) 的数学模型与该问题相同

---

<sup>12</sup>可从 Laplace 分布特征函数的推导中看出，该性质来源于上面叙述的，Laplace 分布可由两个对称的指数分布相加得到

### 4.1.2 挑选驸马问题

**问题** 波斯公主到了适婚年龄，要选驸马。候选男子 100 名，都是公主没有见过的。百人以随机顺序，从公主面前逐一经过。每当一位男子在公主面前经过时，公主要么选他为驸马，要么不选。如果选他，其余那些还没有登场的男子就都遣散回家，选驸马的活动也结束了。如果不选，当下这名男子就离开，也就是淘汰此人，下一人登场。被淘汰的，公主不可以反悔再从选。规则是，公主必须在这百人中选出一人做驸马，也就是说，如果前 99 人公主都看不中的话，她必须选择第 100 名男子为驸马，不管他有多么丑陋。试问用何种策略，能让公主有最高概率选到百人中最英俊的男子作为驸马？

**解答** 现只考虑波斯公主的选择策略为一个单纯的数学模型，即先拒绝掉  $k$  人，在后  $n - k$  人中，一旦发现有人比之前的人都英俊，就选择他作为驸马。那么现在问题就变为了在  $n = 100$  的情况下， $k$  为多少时能使得公主有最高概率选到最英俊的男子

对于某个固定的  $k$ ，假设最英俊的男子所处的位置为  $i (k < i \leq n)$ ，那么他能被选中当且仅当前  $i - 1$  人中最英俊的人在前  $k$  人里，概率为

$$P_{k,i} = \frac{k}{i-1}$$

考虑所有的  $i$ ，利用全概率公式，便能得到先看前  $k$  个人时能选中最英俊男子的概率

$$P_k = \sum_{i=k+1}^n \frac{1}{n} P_{k,i} = \sum_{i=k+1}^n \frac{1}{n} \frac{k}{i-1} = \frac{k}{n} \sum_{i=k+1}^n \frac{1}{i-1}$$

不妨让  $n \rightarrow \infty$ ，令  $\frac{k}{n} = x$ ， $\frac{i}{n} = t$ ，此时上式就可以写成

$$P_k = x \int_x^1 \frac{1}{t} dt = -x \ln x$$

求导得最优的  $x$  取值为  $\frac{1}{e}$ ，即若共有  $n$  个求爱者，最佳的选择策略是先拒绝掉前  $\frac{n}{e}$  个人，而当  $n = 100$  时，这个数字大约是 37

**评注** 公主选择的难处就在于她不知道这百人的英俊程度是如何分布的，所以她最佳的策略是，淘汰最初 37 位男子，并把他们看成一个有代表性的样本集，从而了解这百人相貌的大致分布。然后在这个认知的基础上进行选择

## 4.2 $\pi$

### 4.2.1 Buffon 投针问题

**问题** 现平面上有无数条平行且间距为  $d$  的细线，随机地向平面上投放一根长度为  $\ell$  的针，试问针碰到细线的概率是多少？

**解答** 定义  $P_i(\ell)$  为长度为  $\ell$  的针与细线有  $i$  个交点的概率，则针碰到细线的概率为

$$P(\ell) = \sum_{i=1}^{\infty} P_i(\ell)$$

直接对概率进行计算往往是复杂的，而通常随机变量的期望却更加容易计算，所以这里我们考虑针与细线交点数的期望

$$E(\ell) = \sum_{i=1}^{\infty} iP_i(\ell)$$

由期望的线性可加性<sup>13</sup>可得

$$E(\ell) = c \cdot \ell$$

同时，针的形状也不会影响交点数的期望，这是因为一根长相奇怪的针可以被视作是由无数小针线段拼接而成。特别地，我们考虑直径为  $d$  的圆形的针，可以发现，不管它被投放到哪里，其与细线的交点数都是 2，即

$$E(d\pi) = 2$$

当  $\ell < d$  时，有

$$P(\ell) = P_1(\ell) = E(\ell) = \frac{2\ell}{d\pi}$$

---

<sup>13</sup>一根针可看作是由 A,B 两部分拼接构成，可以对这两部分分别计算期望。虽然由于总的针的形状固定，A,B 两部分的投放结果并非独立事件，但期望的线性可加性与独立性无关，故仍然可用两部分的期望的和来作为总的针的期望。同时再假设单位长度针的期望为  $c$ ，即可获得上述表达