《数学中的小问题大定理》丛书（第五辑）

蒲丰投针问题

——从2009年清华大学的一道自主招生试题谈起

刘培杰数学工作室 编

- 一道自主招生试题
- 图形的格与蒲丰问题
- 几何概率问题
- 平面上的运动群和运动密度
- 平行四边形的弦长分布

哈尔滨工业大学出版社
HARB… LOGY PRESS

《数学中的小问题大定理》丛书（第五辑）

蒲丰投针问题

——从2009年清华大学的一道自主招生试题谈起

刘培杰数学工作室 编

◎一道自主招生试题

◎图形的格与蒲丰问题

◎几何概率问题

◎平面上的运动群和运动密

◎平行四边形的弦长分布

哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内容简介

本书从一道清华大学自主招生试题谈起，讲述了用概率计算圆周率的一个方法——蒲丰投针问题，介绍了随机方法在解决圆周率方面的一个应用。通过对这个著名问题的介绍，洞悉自主招生试题的深厚渊源。全书共分六章，分别为：一道自主招生试题、对 π 作统计估计的途径、图形的格与蒲丰问题、几何概率问题、平面上的运动群和运动密度、将蒲丰投针问题推广到 E_n。

本书适合于高中生、大学生以及数学爱好者参考阅读。

图书在版编目(CIP)数据

蒲丰投针问题:从2009年清华大学的一道自主招生试题谈起/刘培杰数学工作室编. — 哈尔滨:哈尔滨工业大学出版社,2014.1

ISBN 978-7-5603-4554-3

Ⅰ.①蒲… Ⅱ.①刘… Ⅲ.①中学数学课-高中-题解-升学参考资料 Ⅳ.①G634.605

中国版本图书馆CIP数据核字(2013)第309948号

策划编辑　刘培杰　张永芹
责任编辑　张永芹　刘家琳
封面设计　孙茵艾
出版发行　哈尔滨工业大学出版社
社　　址　哈尔滨市南岗区复华四道街10号　邮编150006
传　　真　0451-86414749
网　　址　http://hitpress.hit.edu.cn
印　　刷　哈尔滨工业大学印刷厂
开　　本　787mm×960mm　1/16　印张10.75　字数116千字
版　　次　2014年1月第1版　2014年1月第1次印刷
书　　号　ISBN 978-7-5603-4554-3
定　　价　38.00元

目录

第1章

一道自主招生试题

2009 年清华特色自主招生中曾考到过这样一题：

如图 1 所示，平面内间距为 d 的平行直线，任意放一长为 l 的针，求证：它与直线相交的概率为 $p=\frac{2l}{\pi d}$.

证明 令 M 表示针的中点；x 表示针投在平面上时，M 与最近一条平行线的距离；φ 表示针与最近一条平行线的交角. 显然

$$0\leqslant x\leqslant\frac{d}{2},0\leqslant\varphi\leqslant\pi$$

取直角坐标系，如图 2 所示，上式表示 φOx 坐标系中的一个矩形 R. 而 $x\leqslant$

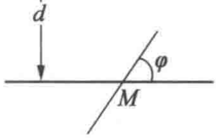

图 1

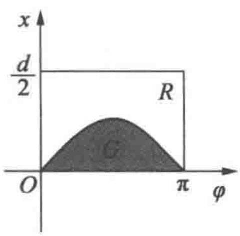

图 2

$\frac{l}{2}\sin\varphi$是使针与平行线(此线必为与点 M 最近的平行线)相交的充分必要条件,不等式 $x\leqslant\frac{l}{2}\sin\varphi$ 表示图 2 中的阴影部分,我们把抛掷针到平面上这件事理解为具有"均匀性". 因此,这个问题等价于向区域 R 中"均匀分布"地投掷点,求点落入阴影部分的概率 p. 由积分有关知识,阴影部分的面积为

$$\int_0^{\pi}\frac{l}{2}\sin\varphi\mathrm{d}\varphi=l$$

故

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

其实此题的背景为积分几何中的蒲丰(Buffon)投针问题.

第2章 对 π 作统计估计的途径

§1 与 π 的统计估计有关的一个问题

1. 平行线网

蒲丰投针问题的解答，在历史上第一次开辟了对 π 作统计估计的途径. 由于蒲丰投针问题的解使 π 与蒲丰的概率 p 相联系，因而 π 的统计估计问题，实质上是蒲丰概率的统计估计问题.

考虑间隔为 1 的平行线网. 设 n 为投针次数，s 为小针实际与网相遇的次数，则

$$\hat{p} = sn^{-1} \tag{1}$$

是蒲丰概率 p(即小针与网相遇的概率)的一个无偏估计. 事实上，考虑以

$$\begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$$

为密度矩阵的随机变数 ξ. 投针 n 次，相当于对 ξ 进行 n 次独立观察，得一容量为 n 的子样$(\xi_1,\cdots,\xi_n)$. 由式(1)给出的

估计 $\hat{p}$ 实际上就是子样的平均值

$$\bar{\xi}=\frac{1}{n}\sum_{i=1}^{n}\xi_i$$

由于 $E\bar{\xi}=p=E\xi$,故 $\hat{p}$ 是 p 的无偏估计. 另外,不难看出,此估计的方差为

$$D\hat{p}=D\bar{\xi}=D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D\xi_i=\frac{1}{n}p(1-p) \quad (2)$$

下表是一个历史的记录:

试验者	针长	投针次数	触网次数	π 的估值
Wolf, 1850	0.8	5 000	2 532	3.159 6
Smith, 1855	0.6	3 204	1 218.5	3.155 3
De Morgan, c. 1860	1.0	600	382.5	3.137
Fox, 1884	0.75	1 030	489	3.159 5
Lazzerini, 1901	0.83	3 408	1 808	3.141 592 9
Reina, 1925	0.541 9	2 520	859	3.179 5
Gridgeman, c. 1960	0.785 7	2	1	3.143

2. 矩形网格,独立性条件

Schuster(1974)从试验设计的观点出发,提出如下的有趣的问题:一个试验者将长度为 l 的小针向布有间隔为 $2l$ 的平行线网的平面上投掷 200 次,记下小针与网相遇的次数;另一个试验者向布有正方形网格(以边长等于 $2l$ 的正方形作为基本区域)的平面上投掷小针 100 次,并分别记录小针与每组平行线网相遇的次数(注意,此正方形网格可看作是由两组互相正交的间隔为 $2l$ 的平行线网组成). 对于 π 的统计估计来说,这两种试验是否提供了同样的统计信息?

现在我们就矩形网格的情形作一般性的讨论. 设平面上有两组互相正交的平行线网,其中一组间隔为 b(不妨假定它平行于 Ox 轴),另一组间隔为 a(平行

于 Oy 轴). 设 $b \leqslant a$. 在随机投针的试验中,以 A 表示小针与平行于 Ox 轴的平行线网相遇事件,以 B 表示小针与平行于 Oy 轴的平行线网相遇事件. 我们先来探讨事件 A 与事件 B 是否独立的问题.

由于

$$P(AB) = P(A) + P(B) - P(A \cup B) \tag{3}$$

其中 $P(A \cup B)$ 为小针与矩形网格相遇的概率. 故根据两事件独立的定义,得到事件 A 与事件 B 互相独立的条件

$$P(A) \cdot P(B) = P(A) + P(B) - P(A \cup B) \tag{4}$$

对于矩形网格,当针长不超过基本区域较短边时(即 $l \leqslant b$),我们有

$$P(A) \cdot P(B) = \frac{2l}{\pi b} \cdot \frac{2l}{\pi a} = \frac{4l^2}{\pi^2 ab}$$

$$\begin{aligned} & P(A) + P(B) - P(A \cup B) \\ = & \frac{2l}{\pi b} + \frac{2l}{\pi a} - \frac{2l(a+b) - l^2}{\pi ab} \\ = & \frac{l^2}{\pi ab} \end{aligned}$$

显然此时条件(3)不成立.

3. 有效性分析

以 ξ, η 表示下列随机变数:

$$\xi = \begin{cases} 1, \text{当小针与平行于 } Ox \text{ 轴之平行线网相遇} \\ 0, \text{当小针与平行于 } Ox \text{ 轴之平行线网不相遇} \end{cases};$$

$$\eta = \begin{cases} 1, \text{当小针与平行于 } Oy \text{ 轴之平行线网相遇} \\ 0, \text{当小针与平行于 } Oy \text{ 轴之平行线网不相遇} \end{cases}.$$

现在我们来考察

$$\hat{p} = \frac{1}{200} \sum_{i=1}^{100} (\xi_i + \eta_i) \tag{5}$$

的有效性. 我们有

$$D(\xi_i+\eta_i)$$
$$=D\xi_i+D\eta_i+2E\{(\xi_i-E\xi_i)(\eta_i-E\eta_i)\} \tag{6}$$

由于

$$D\xi_i=P(A)[1-P(A)] \tag{7}$$
$$D\eta_i=P(B)[1-P(B)] \tag{8}$$
$$E\{(\xi_i-E\xi_i)(\eta_i-E\eta_i)\}$$
$$=E(\xi_i\eta_i)-E\xi_iE\eta_i$$
$$=P(AB)-P(A)P(B) \tag{9}$$

从而有

$$D(\xi_i+\eta_i)=P(A)+P(B)+2P(AB)-[P(A)+P(B)]^2 \tag{10}$$

再利用式(3),得到

$$D(\xi_i+\eta_i)=3P(A)+3P(B)-[P(A)+P(B)]^2-2P(A\cup B) \tag{11}$$

在条件 $l\leqslant b\leqslant a$ 之下,有

$$D(\xi_i+\eta_i)=3\cdot\frac{2l}{\pi b}+3\cdot\frac{2l}{\pi a}-\left(\frac{2l}{\pi b}+\frac{2l}{\pi a}\right)^2-2\cdot\frac{2l(a+b)-l^2}{\pi ab}$$
$$=\frac{2}{\pi}\left[\frac{l}{b}+\frac{l}{a}+\frac{l^2}{ab}-\frac{2}{\pi}\left(\frac{l}{b}+\frac{l}{a}\right)^2\right] \tag{12}$$

由此得到式(5)所表示的 $\hat{p}$ 之方差

$$D\hat{p}=\frac{1}{200^2}\cdot 100\cdot D(\xi_i+\eta_i)$$
$$=\frac{1}{400}\cdot\frac{2}{\pi}\left[\frac{l}{b}+\frac{l}{a}+\frac{l^2}{ab}-\frac{2}{\pi}\left(\frac{l}{b}+\frac{l}{a}\right)^2\right] \tag{13}$$

另一方面,考虑与 Ox 轴平行的平行线网,投针 M 次. 置

$$\hat{p}_x=\frac{1}{M}\sum_{i=1}^{M}\xi_i \tag{14}$$

则

$$D\hat{p}_x = \frac{1}{M^2} \cdot M \cdot \frac{2l}{\pi b}\left(1 - \frac{2l}{\pi b}\right)$$
$$= \frac{1}{M} \cdot \frac{2l}{\pi b}\left(1 - \frac{2l}{\pi b}\right) \tag{15}$$

若要求 $D\hat{p}_x = D\hat{p}$，并记 $\frac{l}{b} = u$，$\frac{a}{b} = k$，则由式(13)和(15)两式有

$$M = \frac{400\left(1 - \frac{2}{\pi}u\right)}{1 + \frac{1}{k} + \left[\frac{1}{k} - \frac{2}{\pi}\left(1 + \frac{1}{k}\right)^2\right]u} \tag{16}$$

同样，考虑平行于 Oy 轴的平行线网，投针 N 次，并置

$$\hat{p}_y = \frac{1}{N}\sum_{i=1}^{N} \eta_i \tag{17}$$

令 $D\hat{p}_y = D\hat{p}$，则有

$$N = \frac{400 \cdot \frac{1}{k}\left(1 - \frac{2}{\pi} \cdot \frac{1}{k}u\right)}{1 + \frac{1}{k} + \left[\frac{1}{k} - \frac{2}{\pi}\left(1 + \frac{1}{k}\right)^2\right]u} \tag{18}$$

例如，对 $k = 1$(即正方形网格)，有：

当 $u = 0$ 时，$M = N = 200$；

当 $u = \frac{1}{2}$ 时，$M = N = 222.273\,26$；

当 $u = 1$ 时，$M = N = 320.497\,01$.

又如，对于 $k = 2$(即 $a = 2b$)，有：

当 $u = 0$ 时，$M = 266.666\,67$，$N = 133.333\,33$；

当 $u = \frac{1}{2}$ 时，$M = 263.760\,22$，$N = 162.670\,31$；

当 $u = 1$ 时，$M = 256.079\,43$，$N = 240.198\,56$.

上述计算结果的意义可解释如下：以 $k=1, u=\frac{1}{2}$ 为例，计算的结果是 $M=N\approx 222$，它表明我们利用正方形网格作投针试验 100 次，大致相当于利用单一的平行线网作投针试验 222 次. 确切地说，利用正方形网格投针 100 次（针长 l 等于正方形边长之半），且由

$$\hat{p}=\frac{1}{200}\sum_{i=1}^{100}(\xi_i+\eta_i)$$

对蒲丰概率 p 作统计估计，其方差与利用单一的平行线网投针 222 次，并由

$$\hat{p}_x=\frac{1}{222}\sum_{i=1}^{222}\xi_i$$

对相应的蒲丰概率作统计估计的方差近似相等.

注意，在上述讨论中，无论是独立性条件的检验或是有效性分析，都是就 $l\leqslant b$ 的情况展开的. 其实对于 $b\leqslant l\leqslant a$ 及 $a\leqslant l\leqslant (a^2+b^2)^{\frac{1}{2}}$ 两款同样可以进行讨论. 因为上述讨论中关键之点在于利用了 $P(A\cup B)$，即小针与网格相遇的概率.

4. 平行四边形网格

我们已知关于平行四边形网格的蒲丰投针问题的完整的结果，因此前两段探讨的课题也可以就平行四边形网格情形展开讨论. 这里我们仅就独立性问题作一简短的讨论.

首先我们应当注意，在 2 中导出的独立性条件式(4)同样适用于现在的场合. 现在我们要问：怎样的平行四边形能使独立性条件式(4)成立？

根据各种类型的平行四边形域的 $P(A\cup B)$ 的表达式，并利用条件式(4)，不难回答这一问题. 此时

$$P(A)=\frac{2l}{\pi h_1},P(B)=\frac{2l}{\pi h_2}$$

$$P(A\cup B)=\frac{2l(a+b)-l^2\left[1+\left(\frac{\pi}{2}-\theta\right)\cot\theta\right]}{\pi ab\sin\theta}$$

$$=\frac{2l(h_1+h_2)-l^2\left[\sin\theta+\left(\frac{\pi}{2}-\theta\right)\cos\theta\right]}{\pi h_1h_2}$$

将这些表达式代入式(4)得

$$\sin\theta+\left(\frac{\pi}{2}-\theta\right)\cos\theta=\frac{4}{\pi},\theta\approx 0.766\ 05(\text{弧度})\tag{19}$$

这时

$$P(A)=\frac{2}{\pi}\arccos\frac{h_1}{l}+\frac{2}{\pi h_1}[l-(l^2-h_1^2)^{\frac{1}{2}}]$$

$$P(B)=\frac{2l}{\pi h_2}$$

$$P(A\cup B)=\frac{2ah_1\arccos\frac{h_1}{l}+2al-2a(l^2-h_1^2)^{\frac{1}{2}}+h_1^2}{\pi ab\sin\theta}$$

$$=\frac{2h_1h_2\arccos\frac{h_1}{l}+2h_2l-2h_2(l^2-h_1^2)^{\frac{1}{2}}+h_1^2\sin\theta}{\pi h_1h_2}$$

将这些表达式代入式(4),得独立性成立的条件

$$\sin\theta=\frac{1}{\pi h_1^2}\left[2\pi lh_1-4lh_1\arccos\frac{h_1}{l}-4l^2+4l(l^2-h_1^2)^{\frac{1}{2}}\right]\tag{20}$$

令$\frac{l}{h_1}=k$,则上式可改写为

$$\sin\theta=\frac{4k}{\pi}\left[\arcsin\frac{1}{k}-k+(k^2-1)^{\frac{1}{2}}\right]\tag{21}$$

例如,当 $k=2$ 时,$\theta\approx 0.708\ 909\ 4$(弧度)$\approx 40.6°$.

§2　平面上的带集

1. 带集密度

若平面上两条平行直线之间的距离为 a，则它们之间以及它们上面的点所构成的闭集叫作一个宽度为 a 的带.

我们用字母 B 代表带. 一个带的位置可以用它的平行中线[①]来确定. 设 p,ϕ 为这样的线的坐标，则具有固定宽度的带(的)集(合)(图 1)的密度是

$$\mathrm{d}B=\mathrm{d}p\wedge\mathrm{d}\phi \tag{1}$$

若要求密度在平面运动群下不变，则除一个常数因子外，这个密度是唯一的.

设 K 为有界凸集. 若 $B\cap K\neq\varnothing$，而 $K_{\frac{a}{2}}$ 为距 K 为 $\frac{a}{2}$ 的平行集，则 B 的平行中线同 $K_{\frac{a}{2}}$ 相交. 反过来，若 B 的平行中线和 $K_{\frac{a}{2}}$ 相交，则 B 和 K 相交. 因此，可得：

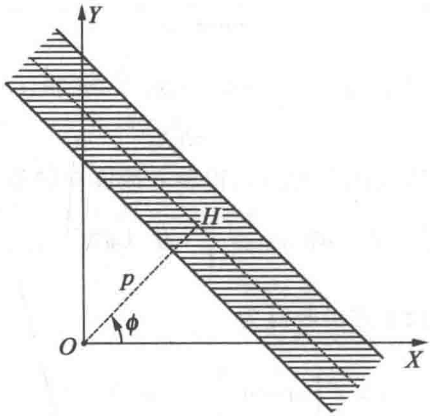

图 1

① 即同带的两界线平行而距离相等的直线.

同一个凸集 K 相交而宽度为 a 的带集的测度是

$$m(B;B\cap K\neq\varnothing)=\int_{B\cap K\neq\varnothing}\mathrm{d}B=L+\pi a\quad(2)$$

其中 L 为 K 的周长. 特殊地,有以下结果:

(a)含一个固定点 P 在内的宽度为 a 的一切带的测度是

$$m(B;P\in B)=\pi a\qquad(3)$$

(b)同一条长度为 s 的线段相交而宽度为 a 的一切带的测度是

$$m(B;B\cap S\neq\varnothing)=2s+\pi a\qquad(4)$$

(c)同一个连通但不一定凸的域相交而宽度为 a 的一切带的测度也用公式(2)确定,但这时 L 表示域的凸包的周长.

含一个已给点集在内的带的集合测度比较复杂,但若所给集 K 的直径 $D\leqslant a$,结果是简单的. 在此情况下,所求测度等于式(2)中的测度减去一切其边界同 K 相交的带的测度,而后一测度则是 $2L$. 故

$$m(B;K\subset B)=\pi a-L\qquad(5)$$

注意由于 $L\leqslant\pi D$,而 $D\leqslant a$,这个测度是非负的.

以上结果可用于几何概率如下:

(a)设 K_1 为含于凸集 K 内的凸集. 一个宽度为 a,而同 K 相交的随机带也同 K_1 相交的概率是

$$p=\frac{L_1+\pi a}{L+\pi a}\qquad(6)$$

其中 L_1 和 L 依次为 K_1 和 K 的周长.

若 K_1 的直径不超过 K 的直径,则带 B 含 K_1 在内的概率是

$$p=\frac{\pi a-L_1}{\pi a+L} \tag{7}$$

若 K_1 缩成一点,则只需在式(7)中令 $L_1=0$,该式就适用.

(b)考虑在有界凸集 K 内 N 个凸集 $K_i(i=1,2,\cdots,N)$(图2).设 L 为 K 的周长,L_i 为 K_i 的周长.若 n 为和带 B 相交的集 K_i 的个数(在图2里,$n=3$),则

$$\int_{B\cap K\neq\varnothing} n\mathrm{d}B=\sum_{i=1}^{N} m(B;B\cap K_i\neq\varnothing)=\sum_{i=1}^{N} L_i+\pi Na \tag{8}$$

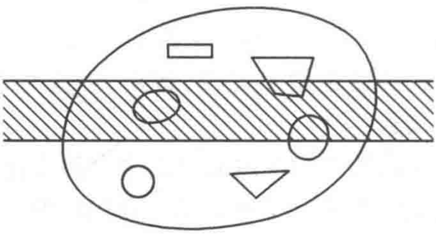

图2

若一切 K_i 的周长都不超过 a,而 n_i 为含于带 B 内的 K_i 的个数,则从式(5)可得

$$\int_{B\cap K\neq\varnothing} n_i\mathrm{d}B=\pi Na-\sum_{i=1}^{N} L_i \tag{9}$$

由式(5),(8)和(9),得:

设 $K_i(i=1,2,\cdots,N)$ 为含于有界凸集 K 内的 N 个凸集,而 B 为一个随机地同 K 相交而宽度为 a 的带.则同 B 相交的 K_i 的个数的平均值是

$$E(n)=\frac{\sum_{i=1}^{N} L_i+\pi Na}{L+\pi a} \tag{10}$$

若一切 K_i 的直径都不超过 a，则含于带内的 K_i 的个数的平均值是

$$E(n_i)=\frac{\pi Na-\sum_{i=1}^{N}L_i}{\pi a+L} \tag{11}$$

2. 蒲丰投针问题

假设 K 为幅度等于 D 的凸集，而 K_1 为含于 K 内的任意凸集. 我们曾经指出，幅度 $D_1\leqslant D$ 的任意凸集都可以含在 K 内. 一个和 K 相交而宽度为 a 的带 B 同时和 K_1 相交的概率为式(6)所确定. 我们原来假定 K 固定，而带 B 则是随机位置的，现在反过来，设想在整个平面上画上平行的带 B，其间的间隔是 D，然后把 K 和 K_1 一起随机地放上去(图 3). 这样 K 肯定要和唯一的一个带相交(除非 K 同带相切，但这样位置的 K 测度是零). 而 K_1 和一个带相交的概率为式(6)所确定；即，若令 $L=\pi D$，就有

$$p=\frac{L_1+\pi a}{\pi(a+D)} \tag{12}$$

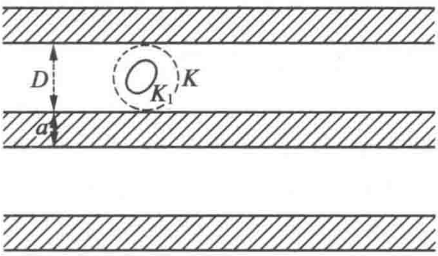

图 3

显然，不必要假定 K 集存在，因此，可以说，若一个幅度为 $D_1\leqslant D$，周长等于 L_1 的凸域 K_1 随机地放在平面上，则它和一个带相交的概率为式(12)所确定.