

# 关于罐子模型一个极限分布的注记

努尔买买提·斯拉吉<sup>1</sup>, 杨纪龙<sup>2</sup>, 米 辉<sup>2</sup>

(1 新疆和田高等师范专科学校 数学系, 新疆 和田 848000 2 南京师范大学 数学与计算机科学学院, 江苏 南京 210097)

[摘要] 罐子模型在概率论的发展和实际应用中都具有十分重要的地位. 设一个罐子中装有  $b$  个黑球和  $r$  个红球, 从中随机地抽取一个球, 然后放回并同时加进  $c$  个与取出球颜色相同的球和  $d$  个与取出球颜色相反的球, 其中  $c, d$  为任意给定的整数. 如此反复进行下去. 当  $c > 0, d = 0$  时称为 Polya 罐子模型; 当  $c = 0, d > 0$  时称为 Friedman 罐子模型. 以  $S_n$  表示在前  $n$  次抽球中抽到黑球的次数, 证明了在 Polya 罐子模型中  $S_n/n$  依分布收敛于一个  $\beta$  分布随机变量, 在 Friedman 罐子模型中  $S_n/n$  依概率收敛于  $1/2$ .

[关键词] 罐子模型,  $\beta$  分布, 极限分布, 依概率收敛

[中图分类号] O212.2 [文献标识码] A [文章编号] 1672-1292(2007)02-0087-03

## Notes on a Limit Distribution of Urn Model

Nurmuhammat·Siraji<sup>1</sup>, Yang Jilong<sup>2</sup>, Mi Hui<sup>2</sup>

(1 Department of Mathematics Xinjiang Hotan Advanced Normal School, Hotan 848000, China)

2 School of Mathematics and Computer Science Nanjing Normal University, Nanjing 210097, China)

**Abstract** In this paper,  $S_n$  denotes the number of black balls chosen in the first  $n$  drawings and it is prove that  $S_n/n$  converges in probability as  $n \rightarrow \infty$  to 0.5 in Friedman Urn Model and  $S_n/n$  converges in distribution as  $n \rightarrow \infty$  to a random variable  $Z$  in Polya Urn Model which has a Beta distribution.

**Key words** Urn model,  $\beta$  distribution, limit distribution, converge in probability

## 0 引言

罐子模型在概率论的发展和实际应用中都具有十分重要的地位, 历史上许多著名学者都曾研究过这种模型. 设一个罐子中装有  $b$  个黑球和  $r$  个红球, 从中随机地抽取一个球, 然后放回并同时加进  $c$  个与取出球颜色相同的球和  $d$  个与取出球颜色相反的球, 其中  $c, d$  为任意给定的整数. 如此反复进行下去. 以  $S_n$  表示在前  $n$  次抽球中抽到黑球的次数, 本文研究抽到黑球的平均次数  $S_n/n$  的极限分布. 为方便计, 文中用  $a/b$  表示分数  $\frac{a}{b}$ .

罐子模型包括许多重要的分布模型.

(1) 二项分布模型. 取  $c = d = 0$  此时抽球为有放回抽球, 且罐子中的黑球和红球的个数始终保持不变, 因此  $S_n$  服从参数为  $n, b/(b+r)$  的二项分布. 由贝努里大数定律知,  $S_n/n$  依概率收敛于  $b/(b+r)$ , 即  $S_n/n$  的极限分布是一个退化的单点分布.

(2) 超几何分布模型. 取  $c = -1, d = 0$  此时抽球变为不放回抽球, 因此  $S_n$  服从分布列为  $P(S_n = k) = \frac{C_b^k C_r^{n-k}}{C_{b+r}^n}$ ,  $k = 0, 1, \dots, n (n \leq b+r)$  的超几何分布. 这里  $C_n^k$  表示组合数, 即  $C_n^k = \frac{n!}{k!(n-k)!}$ . 显然经过  $b+r$  次抽球后, 罐子中已无球可抽, 故不能讨论  $S_n/n$  的极限分布.

(3) Polya 罐子模型. 取  $c > 0, d = 0$  在这种情况下, 每次抽球后, 增加了与取出球颜色相同的球的个数, 而与取出球颜色相反的球的个数不变. 这样在下次抽球中就增加了抽到与本次取出球颜色相同的球的

收稿日期: 2006-10-26

作者简介: 努尔买买提·斯拉吉 (1957-), 副教授, 主要从事概率统计的教学和研究. E-mail yangjilong@njjnu.edu.cn

概率. 该模型可作为“传染现象”的粗糙模型: 每一次传染后都增加了再传染的概率.  $S_n$  的分布称为 Polya 分布, 其分布列为:

$$P(S_n = k) = C_n^k \frac{b(b+c) \cdots (b+(k-1)c)r(r+c) \cdots (r+(n-k-1)c)}{(b+r)(b+r+c) \cdots (b+r+(n-1)c)}, k = 0, 1, \dots, n$$

(4) Friedman 罐子模型. 取  $c = 0, d > 0$  在这种情况下, 每次抽球后, 增加了与取出球颜色不同的球的个数, 而与取出球颜色相同的球的个数不变. 这样在下次抽球中就增加了抽到与本次取出球颜色不同的球的概率, 减少了抽到与本次取出球颜色相同的球的概率. 该模型可作为“安全运动”模型: 每当事故发生(即取到红球)时安全意识就抓紧, 于是发生事故的的概率就减少; 而没有发生事故时安全意识就放松, 于是发生事故的的概率就增大. 对于这个模型,  $S_n$  的分布十分复杂, 因此寻求  $S_n/n$  的极限分布较为困难.

本文主要研究在 Polya 罐子模型和 Friedman 罐子模型下  $S_n/n$  的极限分布.

## 1 Polya 罐子模型下 $S_n/n$ 的极限分布

以  $Y_n, Z_n$  分别表示在第  $n$  次抽球并放回(包括加入  $c$  个同色球)后罐中的黑球数及黑球比例数, 即:

$$Z_n = Y_n / (b + r + nc).$$

文[2]证明了  $Z_n \xrightarrow{d} Z \sim \beta(b_1, r_1) (n \rightarrow \infty)$ , 其中  $b_1 = b/c, r_1 = r/c, Z_n \xrightarrow{d} Z$  表示  $Z_n$  依分布收敛  $Z, Z \sim \beta(b_1, r_1)$  表示  $Z$  服从参数为  $b_1, r_1$  的  $\beta$  分布.

易见  $Y_n = b + S_n c$  于是:  $Z_n = (b + S_n c) / (b + r + nc)$  或  $S_n/n = a_n Z_n + b_n$ , 其中,  $a_n = 1 + (b+r)/(nc)$ ,  $b_n = -b/(nc)$ , 显然当  $n \rightarrow \infty$  时,  $a_n \rightarrow 1, b_n \rightarrow 0$  故由 Slutsky 定理知:  $S_n/n \xrightarrow{d} Z \sim \beta(b_1, r_1) (n \rightarrow \infty)$ . 于是证明了下述定理.

**定理 1** 对于 Polya 罐子模型, 以  $S_n$  表示在前  $n$  次抽球中抽到黑球的次数, 则  $S_n/n \xrightarrow{d} Z \sim \beta(b_1, r_1) (n \rightarrow \infty)$ . 其中,  $b_1 = b/c, r_1 = r/c$

文[2]介绍了利用黑球比例数  $Z_n$  产生  $\beta$  分布随机数的方法, 由定理 1 知, 利用黑球个数  $S_n$  也可产生  $\beta$  分布随机数.

以  $X_k = 1$  或  $0$  分别表示在第  $k$  次抽球中抽到黑球或红球, 则  $S_n = \sum_{k=1}^n X_k$ . 对于 Polya 罐子模型, 用数学归纳法可证  $X_k$  服从参数为  $p = b/(b+r)$  的  $0-1$  分布,  $k = 1, 2, \dots$ . 文[1]证明了  $\{X_n\}$  不服从大数定律, 而且  $\frac{\sum_{k=1}^n X_k}{n}$  的极限分布为  $\beta$  分布.

## 2 Friedman 罐子模型下 $S_n/n$ 的极限分布

记号  $X_n, S_n$  同前, 以  $Y_n$  表示在第  $n$  次抽样并放回(包括加入  $d$  个反色球)后罐子中的黑球数, 易见  $S_n$  与  $Y_n$  之间有下述关系式:  $Y_n = b + (n - S_n)d$ , 或  $S_n = n - (Y_n - b)/d$ .

下面先计算  $Y_n, S_n$  的均值. 以  $EY, DY$  分别表示  $Y$  的均值和方差, 以  $E(Y|X)$  表示  $Y$  关于  $X$  的条件均值.

$$\begin{aligned} EY_1 &= b \cdot \frac{b}{b+r} + (b+d) \frac{r}{b+r} = \frac{b^2 + br + rd}{b+r}, \\ E(Y_n | Y_{n-1}) &= Y_{n-1} \cdot \frac{Y_{n-1}}{b+r+(n-1)d} + (Y_{n-1} + d) \left( 1 - \frac{Y_{n-1}}{b+r+(n-1)d} \right) \\ &= d + \frac{b+r+(n-2)d}{b+r+(n-1)d} Y_{n-1}, \\ EY_n &= E(E(Y_n | Y_{n-1})) = d + \frac{b+r+(n-2)d}{b+r+(n-1)d} \cdot EY_{n-1} \\ &= d + \frac{b+r+(n-2)d}{b+r+(n-1)d} \cdot \left[ d + \frac{b+r+(n-3)d}{b+r+(n-2)d} \cdot EY_{n-2} \right] = \dots \end{aligned}$$

$$= \frac{\sum_{i=1}^{n-1} (b+r+id)d}{b+r+(n-1)d} + \frac{b+r}{b+r+(n-1)d} EY_1 = b + \frac{nd}{2} \cdot \frac{2r+(n-1)d}{b+r+(n-1)d} \quad (1)$$

$$ES_n = n - \frac{EY_n - b}{d} = \frac{n}{2} \cdot \frac{2b+(n-1)d}{b+r+(n-1)d} \quad (2)$$

$$E(S_n/n) = \frac{1}{2} \cdot \frac{2b+(n-1)d}{b+r+(n-1)d} \rightarrow \frac{1}{2} \quad (n \rightarrow \infty). \quad (3)$$

由式(2)可见, 当  $b > r$  时,  $ES_n > \frac{n}{2}$ , 当  $b = r$  时,  $ES_n = \frac{n}{2}$ , 当  $b < r$  时,  $ES_n < \frac{n}{2}$ .

**定理 2** 对于 Friedman 罐子模型, 以  $S_n$  表示在前  $n$  次抽球中抽到黑球的次数, 则  $S_n/n$  依概率收敛于  $1/2$  即  $S_n/n \xrightarrow{P} 1/2 \quad (n \rightarrow \infty)$ .

**证明** 不妨设  $b \leq r$ , 否则以  $S_n$  表示在前  $n$  次抽球中抽到红球的次数来证明本定理. 如前所述, 此时有

$ES_n \leq \frac{n}{2}$ . 以  $X_k = 1$  或  $0$  分别表示在第  $k$  次抽球中抽到黑球或红球, 则  $S_n = \sum_{k=1}^n X_k = S_{n-1} + X_n$ , 且  $EX_n^2 =$

$EX_n \leq 1$  于是:

$$\begin{aligned} ES_n^2 &= E(S_{n-1} + X_n)^2 = ES_{n-1}^2 + 2E[S_{n-1}E(X_n | S_{n-1})] + EX_n \\ &= ES_{n-1}^2 + 2E\left[S_{n-1} \cdot \frac{b+(n-1-S_{n-1})d}{b+r+(n-1)d}\right] + EX_n \\ &= \frac{b+r+(n-3)d}{b+r+(n-1)d} ES_{n-1}^2 + \frac{b+(n-1)d}{b+r+(n-1)d} \cdot 2ES_{n-1} + EX_n \\ &\leq \frac{b+r+(n-3)d}{b+r+(n-1)d} ES_{n-1}^2 + n \end{aligned}$$

由上述递推公式及  $ES_2^2 \leq 4$  得:

$$\begin{aligned} ES_n^2 &\leq n + \frac{b+r+(n-3)d}{b+r+(n-1)d} \cdot \left[ (n-1) + \frac{b+r+(n-4)d}{b+r+(n-2)d} ES_{n-2}^2 \right] \leq \dots \\ &\leq n + (n-1) \frac{b+r+(n-3)d}{b+r+(n-1)d} + (n-2) \frac{b+r+(n-3)d}{b+r+(n-1)d} \cdot \frac{b+r+(n-4)d}{b+r+(n-2)d} + \dots \\ &\quad + 3 \cdot \frac{b+r+2d}{b+r+(n-1)d} \cdot \frac{b+r+d}{b+r+(n-2)d} + \frac{b+r+d}{b+r+(n-1)d} \cdot \frac{b+r}{b+r+(n-2)d} ES_2^2 \\ &\leq \frac{1}{b+r+(n-1)d} \cdot \frac{1}{b+r+(n-2)d} \left[ \frac{1}{2} n(n+1)(b+r)^2 + 2n^3(b+r)d \right. \\ &\quad \left. + \frac{1}{4} (n-2)(n-1)n(n+1) + 4(b+r+d)(b+r) \right]. \end{aligned}$$

由上式及式(3)有:

$$\begin{aligned} \frac{1}{n^2} D(S_n) &= \frac{1}{n^2} [ES_n^2 - (ES_n)^2] = \frac{1}{n^2} ES_n^2 - [E(S_n/n)]^2 \\ &\leq \frac{1}{n^2(b+r+(n-1)d)(b+r+(n-2)d)} \cdot \left[ \frac{1}{2} n(n+1)(b+r)^2 + 2n^3(b+r)d \right. \\ &\quad \left. + \frac{1}{4} (n-2)(n-1)n(n+1) + 4(b+r+d)(b+r) \right] - [E(S_n/n)]^2 \\ &\rightarrow \frac{1}{4} - \frac{1}{4} = 0 \quad (n \rightarrow \infty). \end{aligned}$$

由于对一切  $n$  有  $D(S_n) \leq n^2 < \infty$ , 故由马尔可夫大数定律知  $\{X_n\}$  服从大数定律, 即  $S_n/n - E(S_n/n) \xrightarrow{P} 0$  于是  $S_n/n \xrightarrow{P} 1/2$  证毕.

定理 2 表明  $S_n/n$  的极限分布是取值  $1/2$  的退化分布, 它与常数  $b, r, d$  无关. 因此如果采用这种“安全运动”模型, 不管开始时“发生事故”的概率  $r/(b+r)$  是多么小, 但长久下去 ( $n \rightarrow \infty$ ) “发生事故”的概率都会接近  $1/2$

(下转第 92 页)

- Li Buhong Xie Shusen Lu Zukang Spectral properties of new photosensitizers for photodynamic diagnosis and therapy[J]. Spectroscopy and Spectral Analysis, 2002, 22( 6): 902– 904 ( in Chinese)
- [ 2] 许德余. 光动力制癌药物的历史、现状、进展、问题和前景 [ J]. 中国激光医学杂志, 2001, 10( 1): 44– 47.  
Xu Deyu The history, current status, progress, problems and prospects of the photodynamic cancer drugs[ J]. Chinese Journal of Laser Medicine and Surgery, 2001, 10( 1): 44– 47 ( in Chinese)
- [ 3] 章申峰, 龚兴国. 光动力疗法对肿瘤的作用机制及其影响因素 [ J]. 细胞生物学杂志, 2005, 27( 4): 395– 399.  
Zhang Shenfeng, Gong Xingguo The anti-cancer mechanism and influence factors of photodynamic therapy[ J]. Chinese Journal of Cell Biology, 2005, 27( 4): 395– 399. ( in Chinese)
- [ 4] Wei Shaohua Zhou Jiahong Huang Deyin et al Synthesis and Type I/Type II photosensitizing properties of a novel amphiphilic zinc phthalocyanine[ J]. Dyes and Pigments, 2006, 71( 1): 61– 67
- [ 5] Gao Lindong Qian Xuhong Zhang Li Tetra-trifluoromethoxyl zinc phthalocyanine potential photosensitizer for use in the photodynamic therapy of cancer[ J]. Journal of Photochemistry and Photobiology B: Biology, 2001, 65: 35– 38
- [ 6] 黄金陵, 黄剑东, 刘尔生, 等. 酞菁配合物的结构与其光动力抗癌活性 [ J]. 物理化学学报, 2001, 17( 7): 662– 671.  
Huang Jinling Huang Jiandong Liu Ersheng, et al Some relationships between structures and photodynamic anti-cancer activities of phthalocyanines[ J]. Acta Physico-chimica Sinica, 2001, 17( 7): 662– 671. ( in Chinese)
- [ 7] 李永新, 赵丹华, 卓淑娟, 等. 鲁米诺 – 四磺基锰酞菁 – 过氧化氢化学发光体系在蛋白质测定中的应用 [ J]. 分析化学, 2003, 31( 5): 638– 641.  
Li Yongxin Zhao Danhua Zhuo Shujuan, et al Application of the chemiluminescence system of luminol-MnTSPc-H<sub>2</sub>O<sub>2</sub> in the mensuration of protein[ J]. Chinese Journal of Analytical Chemistry, 2003, 31( 5): 638– 641. ( in Chinese)
- [ 8] 杨功俊, 徐静娟, 陈洪渊. 儿茶酚胺衍生物与 DNA 之间相互作用的光谱和电化学法研究 [ J]. 高等学校化学学报, 2004, 25( 7): 1 235– 1 239.  
Yang Gongjun Xu Junjuan Chen Hongyuan Studies on the interaction between catecholamine derivatives and DNA by means of spectroscopic and voltammetric methods[ J]. Chemical Research in Chinese Universities, 2004, 25( 7): 1 235– 1 239. ( in Chinese)
- [ 9] 黄剑东, 刘尔生, 杨素苓, 等. 抗癌光敏剂 ZnPcSP 在溶液中的存在状态及其对活性的影响 [ J]. 高等学校化学学报, 2002, 23( 12): 2 287– 2 291  
Huang Jiandong Liu Ersheng Yang Suling et al Existence states and activities of ZnPcSP as photosensitizer against cancer in solutions[ J]. Chemical Research in Chinese Universities, 2002, 23( 12): 2 287– 2 291. ( in Chinese)
- [ 10] 吕琳, 吴星, 袁诗海. 磺化酞菁镓、钒、铝、锌在水、醇、胶束中的聚合行为研究 [ J]. 光谱学与光谱分析, 1999, 19( 5): 751– 754.  
L Lin Wu Xing Yuan Shihai Dimerization of Aluminum, Zinc, Vanadium and Gallium phthalocyanine sulfonates in water, aqueous alcoholic solutions and micelles[ J]. Spectroscopy and Spetraanalysis, 1999, 19( 5): 751– 754. ( in Chinese)

[ 责任编辑: 丁蓉 ]

( 上接第 89 页 )

## [参考文献] (References)

- [ 1] Feller W. Probability Theory and Its Applications[M]. 2nd ed. New York: John Wiley and Sons, 1957.
- [ 2] 杨纪龙, 叶尔骅. Polya 罐子模型的一个极限分布推广及其应用 [ J]. 南京航空学院学报, 1988, 21( 4): 112– 118  
Yang Jilong Ye Erhua Generalization of a limit distribution of Polya urn model[ J]. Journal of Nanjing University of Aeronautics and Astronautics, 1988, 21( 4): 112– 118 ( in Chinese)
- [ 3] Athreya K B. On a characteristic property of Polya urn[ J]. Stud Sci Math Hung, 1969( 4): 31– 35
- [ 4] Consul P C, Mittal S P. A new urn model with predetermined strategy[ J]. Biom Z, 1975, 17: 67– 75.
- [ 5] Dyczka W. Polya distribution connected with the problem of Bayes [ J]. Demonstr Math, 1972( 4): 145– 165.
- [ 6] Johnson N L, Kotz S. Two variants of Polya’ s urn models [ J]. Am Stat, 1976, 30( 4): 186– 188
- [ 7] Karlin S, Taylor H M. A First Course in Stochastic Processes [M]. New York: Academic Press, 1975.
- [ 8] Karlin S. Central limit theorems for certain infinite urn schemes [ J]. J Math Mech, 1967, 17( 4): 373– 401.
- [ 9] Kolchin V E. Uniform local limit theorems in the classical ball problem for a case with varying lattices[ J]. Theory Prob Appl, 1967( 12): 57– 67.

[ 责任编辑: 严海琳 ]