

随机回报的3阶矩模型研究*

孟开文, 方珂玮

(西南交通大学 经济管理学院, 成都 610031)

摘要:【目的】通过随机回报的3阶矩模型研究影响中国福利彩票销量的因素, 试图解释在奖池大量积累情况下彩票销量反而下降这种不正常现象。【方法】在彩民购买彩票不完全随机选择号码组合的情形下建立随机回报的3阶矩模型, 以研究彩票派奖规则对彩票销量的影响。【结果】奖池大量积累情况下彩票销量反而下降这种不正常现象, 在某种程度上是因为中国福利彩票最高奖的奖金有封顶。【结论】对现有福利彩票的玩法规则进行适当调整, 如提高奖金封顶额, 将很有可能提升彩票销量并减少奖池资金闲置, 进一步提高彩票的福利功能。

关键词: 双色球; 销售额; 奖池; 派奖规则; 时间序列

中图分类号: O212.2; F061.4

文献标志码: A

文章编号: 1672-6693(2018)06-0070-05

彩票发行是以筹集资金为目的, 以极小的中奖概率和极高的中奖金额的巨大诱惑刺激人们投资而获取收益的经营方式, 属于国民经济中第三产业的范畴。彩票筹集社会闲散资金, 可以缓解财政压力, 支持国家各种福利慈善事业的发展, 这些公益金通常被用于医疗、教育、救济贫困家庭、修建公共设施等领域。彩票业的发展也可带动相关行业的发展, 缓解就业压力, 提高税收收入等, 对维护社会稳定以及稳定宏观经济起到至关重要的作用。

在目前彩票销量不济的大背景下, 对销售额影响因素的研究是一个具有重要现实意义的研究课题。双色球作为中国彩票中最流行的玩法, 以双色球为研究对象具有代表意义。美国强力球彩票奖池累积高达15亿美元时, 一夜暴富成为亿万富翁的希望席卷美国, 彩票销售额暴增。而在中国, 近些年双色球的奖池长期高达10亿量级, 平均也在8亿人民币以上。但是即使当双色球奖池维持在较高水平时, 双色球销售额也并没有呈现井喷趋势。为了刺激双色球销量的增长, 对双色球曾进行了若干次的变动, 但是每次的调整并未达到预期的刺激作用。因此, 对彩票销售额影响因素的研究就显得尤为重要。

Walker 和 Young^[1]对一张彩票的期望收益 E_A , 获奖金额的离散度 V_A 以及可能获得奖金的分布偏度 S_K 等变量对彩票销量进行回归估计, 发现上述3个变量对销量的变化有很高的解释度, 具有显著的相关性。Matheson 等人^[2]和 Levi 等人^[3]对乐透类彩票销售数据进行分析, 验证了彩民购彩存在“光环效应”, 成因主要是彩民个体的从众心理和彩民原有的购买习惯。目前已有文献从彩票方案设计的角度对如何提高方案吸引力、增加彩票销售额进行了研究^[4-8]。张建明等人^[9]对彩票市场进行了研究。Denney 等人^[10]基于市场供需视角发现奖金分布的高偏度对彩民购彩具有吸引力。袁光辉等人^[11]分析影响各省市奖池动力不同的外部原因, 认为奖池大小会影响彩票销量。Jia^[12]则研究彩民购彩行为习惯和心理态度对销售额的影响。

纵观先前学者关于彩票销量的研究, 很少有学者从派奖规则与销量之间的关系角度开展研究工作。虽有少数学者提及封顶额、奖金分配等问题, 但对于封顶额仅仅是从政策角度进行分析和阐述, 并没有从理论分析和模型验证的角度进行深入研究, 对于奖金分配也只研究了个别因素, 缺乏系统性分析。本文基于前人的研究, 通过对双色球玩法的认真观察, 考虑了双色球的派奖规则设计变量, 建立模型, 研究了双色球等乐透型彩票派奖规则设计变量对销售额的影响, 以期发现影响双色球销售额的潜在因素, 进而提出针对性的指导建议。

本文所需要用到符号解释如下。 ω_i 表示第 i 等奖能够获得的中奖金额; π_i 表示第 i 等奖的理论中奖概率,

* 收稿日期: 2018-04-27 修回日期: 2018-09-10 网络出版时间: 2018-10-25 10:41

资助项目: 国家自然科学基金面上项目(No. 11671329)

第一作者简介: 孟开文, 男, 副教授, 博士生导师, 博士, 研究方向为投资组合最优化、公交大数据分析 & 彩票设计, E-mail: mkwfly@126.com

网络出版地址: <http://kns.cnki.net/kcms/detail/50.1165.N.20181025.1041.016.html>

它是由玩法决定的固定值; n_i 表示第 i 等奖的中奖注数; W_g 表示用来派发固定奖奖金的总金额; Q 表示某一期双色球的销售金额; C 表示某一期奖池的累积金额(用作本期奖金的销售金额会按照一定规则分配到各等奖中,如果奖金经派发之后仍有剩余,则未被派发出去的奖金将会积累到下一期的奖池之中); U 表示一等奖奖金部分的封顶金额(现行规则中为 500 万); S 表示某一期出售的总的双色球彩票张数; r 表示销售额中取出用于返奖的比率(现行规则中为 50%,其中 1%为调节基金,文中不作考虑); l_1 表示可派发一等奖的总金额中第一部分包含的高等奖金总额的比例系数(决定奖池 C 对一等奖奖金的贡献量,现行规则中 $C \geq 1 \times 10^8$ 时为 55%, $C < 1 \times 10^8$ 时为 75%); l_2 表示可派发一等奖的总金额中第二部分包含的高等奖金总额的比例系数(现行规则中 $C \geq 1 \times 10^8$ 时为 20%, $C < 1 \times 10^8$ 时为 0); l 表示高等奖奖金中分配给一等奖的比率,现行规则下 $l = 0.75$,且 $l_1 + l_2 = l$; P_{ij} 表示在某一双色球开奖中,产生 j 注 i 等奖的概率; E_A 一张彩票随机回报的 1 阶矩,表示期望回报; V_A 为一张彩票随机回报的 2 阶矩,表示回报的不确定性; S_K 为一张彩票随机回报的 3 阶矩,表示回报分布的偏度。

1 双色球中奖注数分布与奖金分布

目前双色球采用 6/33+1/16 的模式,即从 1~33 个号码中选出不同的 6 个常规号码,再从 1~16 中选出一个特别号码,形成一个组合,根据相匹配的号码个数决定彩民的中奖等级。理论上每次中奖号码之间并不存在关联性。每一种号码组合出现的概率相等,号码组合的可能性是有限的,符合概率学中的“古典概型”,可以很容易计算出各等奖中奖概率。但人们在购买彩票时的各种难以预测的行为会给概率和统计问题带来诸多挑战。

其中一种行为,被称为“意识选择”。Cook 等人^[13]研究了美国彩票,Farrell 等人^[14]研究了英国彩票,Wang 和 Lin^[15]研究了台湾彩票,均发现“意识选择”的存在,并提出 2 个可能的解释:1) 彩民喜欢选择“幸运数字”;2) 倾向于选择容易记住的数字。Baker 和 Mc Hale^[16]认为人们购买彩票选择数字不是完全随机的,多为有意识的选择;泊松分布仅适用于随机选择,并建立复合模型分布拟合各等奖奖金概率分布情况。随后,Baker 和 Mc Hale^[17]又提出了拟合效果更好的理论模型,对英国和西班牙彩票进行研究。李珂^[18]通过研究分析福彩和体彩各省份的开奖数据,发现控制幻觉行为的存在,并且得出该种行为主要表现在彩民对于冷热号码的选择上。

若购买彩票时的择号行为是完全随机的,那么彩票中奖注数应服从泊松定理。因各等奖中奖注数受到销量的影响,理论上中奖注数 $n_i = \pi_i \times S = \pi_i \times \frac{Q}{2}$ (一张彩票价格为 2 元),需对各期中奖注数做标准化处理后再进行泊松分布验证,得到 $p < 0.05$,即各等奖中奖注数均不符合泊松分布。说明彩民购买彩票时择号行为是有意识的选择。因此需要对各等奖中奖数做拟合分布。

核密度估计是一种用于估计概率密度函数的非参数方法,原理是在对某一事物的概率分布未知的情况下,如果某一个数出现了,可以认为这个数的概率密度很大,和这个数比较近的数的概率密度也会比较大,而那些离这个数远的数的概率密度会比较小。针对观察中的第一个数可以用 $f(x - X_i)$ 去拟合出其远小近大率密度。针对每一个观察中出现的数拟合出多个概率密度分布函数之后,取平均。如果某些数是比较重要,某些数反之,则可以取加权平均。设 x_1, x_2, \dots, x_n 为独立同分布 F 的 n 个样本点,其概率密度函数为 f ,则核密度估计为

$$\hat{p}_{(x)} = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (1)$$

其中, n 为样本数据, X_i 为样本数据。 $K(\cdot)$ 为核函数, $h > 0$ 为一个平滑参数,称作带宽。通过核密度估计可以拟合概率密度函数和经验密度函数。因为核密度拟合的概率密度是连续分布的,而各等奖中奖注数是整数,属于离散分布,要求出各等奖中奖注数的每一种情况对应的发生概率,通过经验密度的区间差值可以求出 i 等奖中奖注数为 j 注的概率 $P_{ij} = p_i(x)(j+0.5) - p_i(x)(j-0.5)$,其中 $p_i(x)$ 表示 i 等奖中奖注数分布的经验概率函数。

根据基本假设和双色球的玩法、规则的可以构建以下公式:

$$W_g = 3\,000n_3 + 200n_4 + 10n_5 + 5n_6, \quad (1)$$

$$\omega_1 = \min\left\{U, \frac{c+l_1 \times (r \times Q - W_g)}{n_1}\right\} + \min\left\{U, \frac{(l-l_1) \times (r \times Q - W_g)}{n_1}\right\}, \quad (2)$$

$$\omega_2 = \max\left\{6\,000, \min\left\{U, \frac{(1-l) \times (r \times Q - W_g)}{n_2}\right\}\right\}, \quad (3)$$

根据已求得的经验概率分布,一张彩票随机回报的 3 个阶矩计算公式如下:

$$E_A = \sum_{i=1}^6 \sum_j P_{ij} \omega_i, \quad (4)$$

$$V_A = \sum_{i=1}^6 \sum_j P_{ij} \omega_i^2, \quad (5)$$

$$S_K = \sum_{i=1}^6 \sum_j P_{ij} \omega_i^3. \quad (6)$$

这里,因 E_A 值较小,在 V_A 和 S_K 公式中影响不大,为方便计算省去 E_A 。

2 奖金分布的 3 阶矩与销量的关系

彩票销量可以直接反映出彩票玩法对彩民的吸引力,但销量是多种因素相互作用的结果,无法直接测量。国外一些研究已经证明了奖金的 3 阶矩对销量的影响。例如 Walker^[2] 提出 E_A 和 S_K 与销量是正相关的,而 V_A 与销量是负相关的。

考虑到中国双色球彩票独特的封顶玩法设计及国情,拟采用中国双色球 2004 年(增设周二开奖期)到 2015 年网售叫停前(断点)的历史销售数据,对双色球销售额与 3 个阶距的关系进行研究。通过查阅资料,可以获得双色球的销售额、奖池、各等奖奖金额和各等奖中奖情况等数据,而 3 个阶距的数值需要利用已知数据计算获得,在已知各期开奖中奖人数时,代入(4)~(6)式。

将原始数据和计算得到的 3 个阶距的数据输入到 Eviews 软件中进行时间序列分析,对原始数据进行单位根检验,对处于非稳态的原始变量进行一阶差分处理再做最小二乘回归分析。得到的结果见表 1。

表 1 销量与 3 个阶距的最小二乘回归分析

Tab.1 Ordinary least square regression analysis between sales and the three moments of random return of a double-color lottery

变量	系数	标准误	T 统计量	p 值
C	-4 702 725	1 535 622	-3.062 425	0.002 2
E_A	5 176 306	1 978 939	2.615 698	0.009 0
V_A	-2.699 548	0.828 244	-3.259 362	0.001 1
S_K	3.60E-07	8.32E-08	4.320 922	0.000 0
T_7	14 758 755	1 246 624	11.838 98	0.000 0
T_2	-12 476 001	1 236 344	-10.091 04	0.000 0
D^{Δ_1}	-0.531 248	0.022 236	-23.891 29	0.000 0
D^{Δ_2}	-0.450 954	0.022 014	-20.484 77	0.000 0

表 1 中 D^{Δ} 表示销量 Q 的 1 阶差分,即销售额的增量。 T_2, T_7 为虚拟变量(双色球分别于周二、四、日开奖;如开奖期为周二则 T_2 为 1,否则为 0;如开奖期为周日则 T_7 为 1,否则为 0)。由模型输入可知,该模型具有良好的拟合效果。

$$D^{\Delta} = -4\,702\,725 + 5\,176\,306E_A - 2.699\,548V_A + (3.60e-07)S_K + 14\,758\,755T_7 - 12\,476\,001T_2 - 0.531\,248D^{\Delta_1} - 0.450\,954D^{\Delta_2} \quad (7)$$

经过数据分析可知,中国双色球彩票的销售额与 3 个阶距的相关关系与国际上相关研究结论相符。很容易理解,当 E_A 增大,即购买彩票的预期回报增加,彩民购买彩票能获得更多的奖金自然更乐意购买彩票;当 S_K 增大,高等奖(主要是一等奖)奖金增多,根据前景理论人类具有强调小概率事件的倾向,彩民往往会放大中高等奖的极小概率事件,具有心存侥幸以小博大的心理,高等奖奖金的提高会增加彩票对彩民的吸引力,提升销量;当 V_A 增大,中奖的不确定性增加,根据前景理论人们面对获得时往往不愿意冒险,所以不确定性的奖金对彩民的吸引力比较小。

明确 3 阶矩与销量的关系后,研究彩票设计变量对销量的影响作用就可以通过奖金分布的 3 阶矩来实现。

3 派奖规则设计变量对销量的影响研究

运用蒙特卡罗方法,根据各等奖概率分布产生随机数(因核密度估计得到的概率密度分布是连续分布,产生的随机数多为小数,需要取整),模拟彩票开奖情况。本文模拟了 10 000 期开奖情况,得到每一期一至六等奖中奖注数,并根据彩票现行派奖规则,计算对应奖金额。根据 3 阶矩公式求出数值。运用(7)式可以观察变量对销量的影响,假设前两期销量不变($D^{\Delta_1}, D^{\Delta_2}$ 均为 0),开奖期为周四(T_7, T_2 均为 0)。

3.1 奖池金额 C 对 3 个阶距的影响

奖池金额虽不属于派奖设计变量,但目前双色球奖池资金积累水平非常高,如果高奖池对销量无促进作用,那么高额奖池的积累是对社会闲置资金的一种浪费,因此有必要研究高水平奖池对销量的影响。

保持除奖池金额 C 以外的因素均不变的情况下,改变奖池金额 C 进行蒙特卡罗仿真模拟,可以得出 3 个阶距的变化。研究发现在奖池处于较低水平时,奖池的增加对 3 个阶距均有促进作用,而随着奖池的增大,该促进作用逐渐减缓,奖金的 3 阶矩逐渐趋于平稳。这种变化是由于封顶额的存在,当奖池增大时,高等奖的单注奖金额逐渐增大,同时一张彩票的平均回报、奖金数额的不确定性和奖金的偏度都会增大,但当奖池大到一定程度使每一注高等奖的奖金都达到封顶额后,奖金的分布水平不会再变化,也就表现为 3 阶矩趋于平稳。近年的奖池累积早已超过 7 亿元,在现行规则下奖池的积累对销量的促进已无意义。

3.2 封顶额 U 对 3 个阶距的影响

保持除封顶额 U 以外的因素均不变的情况下,改变封顶额 U 进行蒙特卡罗仿真模拟,得出 3 个阶距的变化。随着封顶额的提高,3 个阶距均呈现明显上升趋势,封顶额的提升有助于销售额的提高。但当封顶额增大到一定程度,或者说当封顶额大于理论上分配给一二等奖的单注奖金额时,封顶额 U 对奖金已无封顶效果,可以看作奖金不封顶,在奖池和销售量都假设不变情况下,奖金分布情况是稳定的,所以 3 阶矩不再变化。但在奖金不封顶的情况下,如果结合奖池金额 C 的作用,随奖池积累 3 阶矩会一直增大。随着 U 的增大销量也是在不断增加,直到 U 非常大时,对销量不再有影响。

但是封顶额不能一直增大,目前双色球的一张彩票售价 2 元,如果 E_A 总是远大于 2 元,那么对于彩票发行方来讲是亏损的,奖池金额可能被掏空,甚至一直无法积累,用于返奖的金额甚至有可能不足以发给获奖者,所以按目前双色球的玩法规则下,不设封顶额不适用于双色球及其他彩票。

2009 年提高封顶额后,未再对封顶额进行调整。近年销售额进入瓶颈期,甚至有所下降,提高封顶额虽对 E_A 影响不大,但能提高奖金的不确定性 V_A 和偏度 S_R ,虽然前者对销量有抑制作用,后者对销量有促进作用,但促进作用是大于抑制作用的,也就是说提高封顶额可以增加高等奖奖金额,增加双色球对彩民的吸引力,从而拉动销量。

4 结论

每一种彩票从发行以来,都会经历一个完整的生命周期。近两年双色球的疲软表现就表明双色球已经处于平稳期,为了防止双色球的销售额进一步下滑,就必须采取一些有效的措施。

1) 提高封顶额。当封顶额提升时,双色球对彩民的吸引力是明显提高的。虽然方差 V_A 也变大了,但是人们在面对高额回报时,往往会忽视风险因素。并且双色球的购买成本很低,所以购买双色球的风险损失本身也是不大的。考虑到实际情况,不建议彻底取消封顶制度,而是应该根据目前国民人均收入和 GDP 来适时调整封顶额。

2) 调控一等奖派奖方式。受到一等奖奖金中两部分的高等奖奖金分配比例的启发,可以使一等奖第一部分的奖池金额更多地派发出去。具体做法是尽可能的降低 l_1 , 提高 l_2 , 这样既有利于 3 个阶距的上升,又有利于减弱奖池的疯狂积累,从而增强资金的流动性,减弱资金闲置的负面影响。同时,还有利于双色球销售额的增长,最终充分发挥双色球的福利性本质,为社会创造更多福利。但是,奖池过低也会对每张彩票的中奖金额产生影响。过低的奖池会导致头等奖奖金额的降低,减少每张彩票的期望收益。

3) 提高高等奖级金额中一等奖的分配率。当一等奖的分配率提高,一等奖奖金增大,即使一张彩票的平均回报降低。但是由于 E_A 本身的值就比较低,双色球定价低,所以这极低的亏损额在巨额奖金前往往被忽略,鉴于人们好奇、以小博大的冒险心理,高等奖奖金的提高会增加彩票对彩民的吸引力。

参考文献:

- [1] WALKER I, YOUNG J. An economist's guide to lottery design[J]. The Economic Journal, 2001, 111(11): 700-722.
- [2] GROTEA K R, MATHESON B V A. Examining the "Halo Effect" in lottery games[J]. Applied Economics Letters, 2007, 14(4): 307-310.
- [3] DAVID F, LEVI P. Just like the lottery? Player behaviour and anomalies in the market for football pools[J]. Journal of Gambling Studies, 2015, 31(2): 471-482.

- [4] 王金山,王伦夫,黎明曦.彩票发行方案的合理性评价[J].
运筹与管理,2015,14(4):125-129.
WANG J S,WANG L F,LI M X. Lottery project rationality evaluation[J]. Operations Research and Management Science, 2015,14(4):125-129.
- [5] MAEDA A. Optimal lottery design for public financing[J].
Economic Journal,2008,118:1698-1718.
- [6] PENG T,RONG P,LIU J Y. Influencing factors of Sports Lottery marketing: empirical study based on data from Jiangsu[J]. Innovative Computing and Information, 2010, 232:1-9.
- [7] CHEN X, DIAKONIKOLAS I, ORFANO A, et al. On the complexity of optimal lottery pricing and randomized mechanisms[C]//IEEE. Annual Symposium on Foundations of Computer Science (FOCS). Berkeley, California: [s. n.], 2015:1464-1479.
- [8] 刘圣文. 不确定性背景下的中国体育彩票奖金分布策略研究[D]. 济南:山东大学,2017.
LIU S W. Research on prize distribution strategy of China Sports Lottery under uncertainty[D]. Jinan:Shandong University,2017.
- [9] 张建明,史安娜,陈理飞. 基于 Logistic 模型的中国彩票市场实证分析[J]. 统计与决策,2010,26(21):82-85.
ZHANG J M, SHI A N, CHEN L F. Empirical study of lottery market in China based on logistic model[J]. Statistics and Decision,2010,26(21):82-85.
- [10] DENNERY C,DIRER A. Optimal lottery [J]. Journal of Mathematical Economics,2014,55:15-23.
- [11] 袁光辉,杨飞. 奖池动力的省域聚类实证分析[J]. 中国商论,2014,18:210-211.
YUAN G H,YANG F. Empirical study and clustering analysis of prize pool dynamics for different provinces[J]. China Business & Trade,2014,18:210-211.
- [12] JIA Y. Examining the gambling behaviors of Chinese online lottery gamblers: are they rational? [J]. Journal of Gambling Studies,2015,31(2):573-584.
- [13] COOK P J,CLOTFELTER C T. The peculiar scale economies of Lottery[J]. Am Econ Rev,1993,83(3):634-643.
- [14] FARRELL L, HARTLEY R, LANOT G, et al. The demand for lotto: the role of conscious selection[J]. Journal of Business and Economic Statistics,2000,18(2):228-241.
- [15] WANG J S,LIN M Y. Number selection strategy of lottery players:an empirical study of the Taiwan lottery[J]. Taiwan Econ Forecast Poly,2006,37:49-67.
- [16] BAKER R D, MCHALE I G. Modelling the probability distribution of prize winnings in the UK national lottery: consequences of conscious selection[J]. Journal of Royal Statistical Society Series A (Statistical in Society), 2009, 172(4):813-834.
- [17] BAKER R D, MCHALE I G. Investigating the behavioural characteristics of lottery players by using a combination preference model for conscious selection[J]. Journal of Royal Statistical Society Series A (Statistical in Society), 2011,174(4):1071-1086.
- [18] 李珂. 基于空间面板模型的投资者控制幻觉对彩票销量的影响研究:以数字型彩票为例[D]. 上海:上海师范大学,2016.
LI K. The influence of investor control illusion on lottery sales based on spatial panel model: a case study of digital lottery[D]. Shanghai:Shanghai Normal University,2016.

Modeling on Three Moments of Random Return

MENG Kaiwen, FANG Kewei

(School of Economic and Management, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: [Purposes] Game rules of the so-called Double Color Ball Lottery have been examined via the study of relations of its sales and the prize pool accumulated, and it intends to explain why large prize pool cannot promote sales. [Methods] The study improved the models in the moments of the prize under conscious selection (lottery players choose numbers non-randomly), and to study the effect of lottery prize faction rules design variables on sales. [Findings] Lottery sales have fallen on decreasing tendency even pools have accumulated in large numbers, partly because the top prize of China's welfare lottery has been capped. [Conclusions] It makes the current game rules of Double Color Ball Lottery proper adjustments, such as increasing the amount of the bonus top and to regulation the proportion of the internal first prize award amount and among high prizes bonus not only will reduce the prize pool as the idle funds, but also to enhance sales and the lottery's welfare function.

Keywords: Double Color Ball Lottery; sales; the prize pool; prize rules; time series

(责任编辑 黄颖)