

安徽体育彩票开奖号码的数理统计分析

本文对安徽体育彩票6 + 1 玩法各位号码出现的数字进行了统计,作出了折线图,并进行了卡方检验,得出了各位数字中0 ,1 ,2 ,3 ,4 ,5 ,6 ,7 ,8 ,9 出现的概率是相等的结论。

1、背景材料

安徽电脑体育彩票从2000 年9 月12 日开始发行以来,受到了安徽省内外广大彩民的关注。人们对彩票出现的号码进行研究,希望能找到它们的规律,以便在博彩活动中赢得大奖。现在已经有传统型6 + 1 玩法和乐透型33 选7 玩法、30 选7 玩法、15 选5 玩法、21 选5 玩法及足球彩票等。在以上几种数字的选择玩法中,我们不能否认开奖的公平、公正和公开性,应该认为开奖中各个号码的出现具有随机性,各种号码出现的可能性大小是典型的概率问题。其中6 + 1 玩法的前6 个数字是从0、1、2、3、4、5、6、7、8、9 这10 个数字中重复抽取6 次每次抽取1 个的重复抽样问题,每个数被抽到的概率都是0.1 ;特别号是从0、1、2、3、4 这5 个数字中随机抽取1 个的抽样问题,每个数被抽到的概率都是0.12 (本文按这种情况计算概率),后来又改为从0、1、2、3、4、5、6、7、8、9 这10 个数字随机抽取一个,使得中特等奖的概率降低为原来的一半。33 选7 玩法、30 选7 玩法、15 选5 玩法、21 选5 玩法等都是不放回抽样问题。

2、案例分析

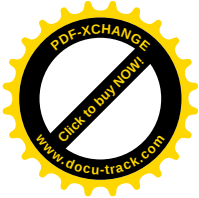
笔者统计了自从2000 年9 月12 日到2003 年9 月1 日为止384 期的实际资料(资料来源www.1ah1631.net) ,列出了统计表,希望为读者提供比较完整的信息,并从中发现其统计规律性。通过分析各位号码中各个数字出现的概率,以供彩民选号时作参考。

(1) 数字出现频数统计

根据安徽风采体育彩票6 + 1 玩法2000 年9 月12 日到2003 年9 月1 日为止384 期的实际资料,统计了0 - 9 这10 个数字在各位号码中出现的频数(见表1) ,以便于进行数理统计分析。

表 1 各位号码中各数字出现频率的统计表

数 字	0	1	2	3	4	5	6	7	8	9	合计
在第一位号码中的出现次数	37	46	31	32	39	40	30	42	33	54	384
在第二位号码中的出现次数	34	29	43	35	38	50	40	32	49	34	384
在第三位号码中的出现次数	38	47	36	34	32	39	41	50	33	34	384
在第四位号码中的出现次数	36	39	33	38	52	33	42	30	36	45	384
在第五位号码中的出现次数	39	34	42	38	36	34	51	36	40	34	384
在第六位号码中的出现次数	34	33	40	39	48	45	36	34	39	36	384
在六位号码中的出现的总次数	218	229	227	219	249	246	246	231	238	246	2304



(2) 数字出现均匀性的数理统计检验

为了检验各位号码中各个数字出现的概率是否相等,对第一、二、三、四、五、六位数字和特别号进行卡方检验,同时也对前六位号码总体进行卡方检验,计算列表和检验结果如下:

第一位号码	0	1	2	3	4	5	6	7	8	9	合计	临界值
实际频数 O_i	37	46	31	32	39	40	30	42	33	54	384	161919(0105)
理论频数 E_i	3814	3814	3814	3814	3814	3814	3814	3814	3814	3814	384	211666(0101)

$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i} = 131395.8 < 161919$$
, 因此, 在显著性水平 $\alpha = 0.105$ 之下, 认为这 10 个数字在第一位号码中出现是等可能的。

第二位号码	0	1	2	3	4	5	6	7	8	9	合计	临界值
实际频数 O_i	34	29	43	35	38	50	40	32	49	34	384	161919(0105)
理论频数 E_i	3814	3814	3814	3814	3814	3814	3814	3814	3814	3814	384	211666(0101)

$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i} = 111729.2 < 161919$$
, 因此, 在显著性水平 $\alpha = 0.105$ 之下, 认为这 10 个数字在第二位号码中的出现是等可能的。

第三位号码	0	1	2	3	4	5	6	7	8	9	合计	临界值
实际频数 O_i	38	47	36	34	32	39	41	50	33	34	384	161919(0105)
理论频数 E_i	3814	3814	3814	3814	3814	3814	3814	3814	3814	3814	384	211666(0101)

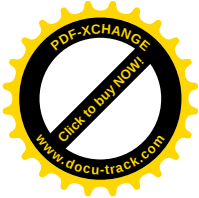
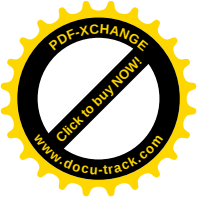
$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i} = 81604.2 < 161919$$
, 因此, 在显著性水平 $\alpha = 0.105$ 之下, 认为这 10 个数字在第三位号码中的出现是等可能的。

第四位号码	0	1	2	3	4	5	6	7	8	9	合计	临界值
实际频数 O_i	36	39	33	38	52	33	42	30	36	45	384	161919(0105)
理论频数 E_i	3814	3814	3814	3814	3814	3814	3814	3814	3814	3814	384	211666(0101)

$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i} = 91958.3 < 161919$$
, 因此, 在显著性水平 $\alpha = 0.105$ 之下, 认为这 10 个数字在第四位号码中的出现是等可能的。

第五位号码	0	1	2	3	4	5	6	7	8	9	合计	临界值
实际频数 O_i	39	34	42	38	36	34	51	36	40	34	384	161919(0105)
理论频数 E_i	3814	3814	3814	3814	3814	3814	3814	3814	3814	3814	384	211666(0101)

$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i} = 61364.6 < 161919$$
, 因此, 在显著性水平 $\alpha = 0.105$ 之下, 认为这 10 个数字在第五位号码中的出现是等可能的。



第六位号码	0	1	2	3	4	5	6	7	8	9	合计	临界值
实际频数 O_i	34	33	40	39	48	45	36	34	39	36	384	161919(0105)
理论频数 E_i	3814	3814	3814	3814	3814	3814	3814	3814	3814	3814	384	211666(0101)

$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i} = 51687.5 < 161919$$
，因此，在显著性水平 $\alpha = 0.105$ 之下，认为这 10 个数字在第六位号码中的出现是等可能的。

1 至 6 位号码	0	1	2	3	4	5	6	7	8	9	合计	临界值
实际频数 O_i	218	229	227	219	249	246	246	231	238	246	2304	161919(0105)
理论频数 E_i	23014	23014	23014	23014	23014	23014	23014	23014	23014	23014	2304	211666(0101)

$$\chi^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i} = 61212.7 < 161919$$
，因此，在显著性水平 $\alpha = 0.105$ 之下，认为这 10 个数字在第 1 到第 6 位号码中的出现是等可能的。

由于特别号码中,2002 年 8 月 5 日以前采用从 0,1,2,3,4 这 5 个数字中任意抽取一个,取到每个数字的概率均为 0.12 ;但是从 2002 年 8 月 7 日以后,特别号也采取从 0,1,2,3,4,5,6,7,8,9 这 10 个数字中任意选择,取到每个数字的概率均为 0.11 ,这使得中特等奖的概率降低为原来的一半。尽管在此我们没有对历史数据的随机性进行检验,但可能推断各个数字的出现仍是均匀的,具有几乎相同的概率。

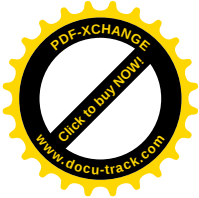
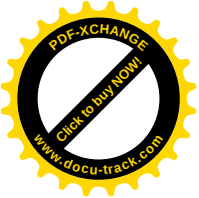
3、结论分析

通过对样本的检验,可以得出各个数字出现的概率基本上是相等的结论。所以,不存在固定的选号技巧。由于各个数字的出现具有一定的随机性,这就为广大彩民选号提供了广阔的范围。但是,不能钻进买彩票为了发财的牛角尖里。应该以一颗平常心,为了测试运气、娱乐,当然也为体育事业或教育事业或其他公益事业奉献自己的一份爱心。发行彩票,是我国政府在市场经济的条件下,为筹集社会资金,发展福利事业和体育事业而采取的一项重要举措,吸引了社会各界人士的积极参与,取得了显著的经济效益和社会效益。主要表现在促进消费、拉动经济增长、支持公益事业等各个方面。

综上所述,彩票发行的目的是为社会公众事业募集资金,是政府一种进行资金的再分配的手段,彩民可以从购买彩票中享受一定乐趣,而不应以“暴富”为目的影响日常生活。所以,彩民不应相信个别媒体和报刊错误诱导而试图寻找某种规律,拒绝充斥唯心色彩的博彩书籍和信息资料,更不要相信拆字、算命、占卜之类的迷信活动,以保证我国彩票事业的健康发展。

笔者看到许多关于所谓选号“技巧”的文章,基本上都是一家之言,没有公认最有效的方法。在报纸、杂志上有许多关于彩票选号的文章,对选号“技巧”的宣传也有不同观点。笔者认为,如果对于彩票号码出现规律持不同的意见,就会有不同的选号方法。以 6 + 1 玩法为例,不同的认识就有不同的观点。对此可以提出三种不同的观点和选号方法:

(1) 若认为数字的出现具有随机性,但大量次数的抽奖必然导致各个数字出现



的可能性大致相等。那么就应该根据过去各个数字的次数, 哪个数字以前出现的少, 以后就应该出现的多, 就选择那个数字。

(2) 若认为各个数字的出现的几率不等, 概率有大有小。以历史数据作为样本, 对各个数字出现的频率进行估计, 由于频率是概率的无偏、有效估计。那么就应该根据过去各个数字出现的次数, 哪个数字以前出现的多, 它出现的概率就大, 因此就选择那个数字。

(3) 如果认为各个数字的出现没有什么规律性, 不论以前的结果如何, 以后开奖时抽取的数字不受以前开奖结果的影响, 即开奖时抽取的数字与以前的结果具有独立性。那么选什么号都一样。

这三种观点都有道理, 并且都有理论根据。笔者认为, 从10个数字在20期中出现次数的统计是很难发现其规律的, 如果出现绝对平均的话, 每个数字平均只能出现2次, 这个标本未免太小了。根据数理统计学非参数检验的要求, 每个数字出现的次数不能小于5。