Journal of the Korean Data & Information Science Society
2009, **20**(5), 779–786

한국데이터정보과학회지

로또복권의 당첨번호에 대한 무작위성 검정

임수열[1] · 백장선[2]

[12]전남대학교 통계학과

접수 2009년 5월 15일, 수정 2009년 8월 25일, 게재확정 2009년 9월 21일

요 약

로또복권은 이미 전 세계적으로 인기가 가장 높은 복권중의 하나이다. 우리나라의 경우 전체 45개의 번호 중 6개의 번호를 선택하는 방법으로 로또복권이 발행하고 있으며, 로또복권의 발행을 통해 얻은 수익 중 일부분을 복지사업에 지원하는 등 다양한 분야에 대한 지원 사업으로 사회 환원에도 적극적으로 앞장서고 있다. 하지만 이런 장점에도 불구하고 로또복권의 1등 당첨번호들이 과연 무작위로 선택되었는가 하는 의혹은 발행초기부터 지금까지도 끊이지 않고 제기되어 오고 있다. 따라서 본 연구에서는 로또복권의 총 331회의 1등 당첨번호들에 대하여 주관 사업체별, 그리고 로또복권의 가격변화로 구분하여 다변량중심극한정리와 몬테카를로 모의실험을 이용한 검정 방법으로 당첨번호들에 대한 무작위성 검정을 실시하였다. 그 결과 모든 경우에서 당첨번호들이 무작위성을 만족하였다.

주요용어: 로또, 몬테카를로, 무작위성, 복권, 중심극한정리.

1. 서론

전 세계적으로 많은 국가에서 가장 인기 있게 발행되는 복권은 바로 로또복권이다. 각 나라마다 로또복권의 주당 발행회수나 추첨방식에 있어 조금씩 차이는 존재하지만 우리나라 역시 세계적인 흐름에 따라 2002년 12월에 처음으로 로또복권의 발매가 시작되어 지금까지 온 국민의 관심 속에서 지속적으로 성장해오고 있다. 로또복권의 탄생 배경은 각 부처가 개별법에 근거하여 경쟁적으로 복권발행에 참여함에 따라 복권을 발행하는 기관과 복권의 종류가 지나치게 많은 문제점과 함께 복권발행기관 사이의 과다경쟁으로 인하여 판매수수료 등 유통비용이 상승하게 되어 공공재원조달의 효율성을 떨어뜨리고 오히려 국민의 사행심만 조장시키는 부작용이 우려되었기 때문이다. 즉, 기존의 사행성·도박성 위주의 복권사업 대신 건전한 선진형 복권인 로또복권으로의 유도가 필요하다고 판단되었기 때문이다 (한국조세연구원, 2003). 또한 로또복권의 전체판매액 중 상당부분은 소외계층에 대한 복지사업, 문화예술 진흥 및 문화유산 보존사업, 임대주택 건설, 저소득층의 주거안정 지원 사업 등의 사회 곳곳으로 환원되어 사용되어지고 있지만 이런 로또복권의 진정한 의미보다는 대박, 인생역전 등의 사행적 이미지가 국민들에게 강하게 인식되고 있다.

우리나라의 경우 이미 로또복권을 시행하고 있는 다른 나라들에 비하여 로또복권이 발행된 시기가 매우 짧은 시간임에도 불구하고 다른 나라와는 달리 로또복권에 대한 불신이 매우 높은 것이 사실이다. 특히 최근에는 로또복권에 대한 강한 불신과 의혹들로 인하여 로또복권이 발행된 이래 처음으로 감사원에서 로또복권의 시행에 있어서 실제로 객관적이지 못한 부분들이 있는지에 대하여 감사를 실시하고 있

[1] (500-757) 광주광역시 북구 용봉동 300번지, 전남대학교 통계학과, 박사과정.
[2] 교신저자: (500-757) 광주광역시 북구 용봉동 300번지, 전남대학교 통계학과, 교수.
E-mail: jbaek@chonnam.ac.kr

는 상황이다 (매일경제, 2009). 요즘에는 이런 문제점들에 대하여 경제 및 통계학자뿐만 아니라 다양한 분야의 연구자들에 의하여 로또복권의 설계과정의 투명성과 모의실험을 통한 추첨의 객관성에 대한 연구 활동들이 활발하게 이루어지고 있다 (Farrell 등, 1999; Russell 등, 2005; Scott 등, 1995; Chen 등, 2008). 또한 통계학을 기반으로 하는 다양한 확률모형들이 지속적으로 연구·개발됨으로써 로또복권의 당첨확률뿐만 아니라 당첨된 번호의 무작위성 검정 등과 같이 복잡하고 어려운 확률문제들도 쉽게 해결할 수 있게 되었다 (Helman, 2005; Boland 등, 1999; Genest 등, 2002).

로또복권의 1등 당첨번호들에 대하여 일정한 법칙이 존재하거나 당첨번호를 추첨하는 방식에 문제가 있었다면 지금까지 뽑힌 1등 당첨번호들에서는 무작위성이 나타나지 않을 것이다. 따라서 본 연구에서는 어느 특정 회차에서 임의로 조작된 1등 당첨번호를 탐지할 수는 없지만 당첨번호들이 무작위적인지를 판단할 수 있는 통계적 검정방법들을 적용하여 로또복권의 1등 당첨번호들이 과연 무작위적인지를 분석하고자 한다.

제2절에서는 로또복권이 당첨에 관한 이론적 근거들을 살펴보고, 제3절에서는 Coronel-Brizio 등 (2008)에서 제안된 로또복권의 1등 당첨번호에 대한 무작위성 (randomness) 검정에 대한 두 가지 검정 방법들에 대하여 살펴보겠다. 또한 제4절에서는 실제 우리나라에서 발행된 로또복권의 1등 당첨번호들에 대하여 제3절에서 살펴본 검정 방법들에 적용시킨 결과를 알아보고, 마지막 제5절에서는 결론 및 토의를 통해 본 연구를 정리한다.

2. 로또복권 당첨에 관한 이론

전 세계적으로 로또복권을 발행하고 있는 각 나라마다 조금씩 다른 특징을 가지고 있다. 멕시코의 경우 전체 선택할 수 있는 번호의 개수가 39→44→47→51로 증가하였으며 선택하는 번호의 개수는 모두 6개로 동일하였다. 또한 이탈리아의 경우 선택할 수 있는 번호의 개수는 90개이고 5개를 선택하는 방식의 로또복권 구조를 가지고 있다 (Coronel-Brizio 등, 2008). 우리나라에서 발행되고 있는 로또복권은 주 1회 발행되며 한 번의 게임 당 1,000원이며 당첨은 1∼45번까지의 숫자 중에서 순서에 상관없이 비복원 추출로 6개의 1등 당첨 숫자와 1개의 보너스 숫자를 추첨하는 방식으로 이루어진다. 로또복권에 당첨이 되는 경우는 표 2.1과 같다. 표 2.1 이외의 로또복권 당첨에 관한 보다 자세한 내용은 '나눔로또 6/45 (www.645lotto.net)' 홈페이지에서 확인가능하다.

Coronel-Brizio 등 (2008)에서 언급한 바와 같이 Chaitin (1975)은 임의의 수열이 무작위적인지 완벽하게 증명하는 방법은 없다고 밝혔다. 따라서 난수에 대한 무작위성 검정방법은 불완전하지만 무한히 많이 만들 수 있다. 보통 통계적으로 난수의 무작위성을 검정하는 방법은 대부분 무작위성을 감지할 수 있는 검정통계량을 선정하고 무작위성 (귀무가설)이 성립할 때의 값을 계산하여 그 것이 과연 관측된 값과 얼마나 떨어져 있는지를 살펴 결론을 내리고 있다. Frequence test, runs test 등과 같이 실제 난수의 무작위성을 검정하는데 가장 널리 쓰이는 National Institute of Standards and Technology (NIST)의 16가지 통계적 무작위성 검정방법들 (Andrew Rukhin 등, 2001) 모두 위와 같이 만들어진 것이다. 그러므로 당첨번호의 무작위성을 검정하는 방법은 여러 가지가 있을 수 있다. 그 중 한 가지 방법으로는 지금까지 1등번호로 뽑힌 숫자들이 1-45의 숫자들로부터 균등하게 출현했는지를 판단하여 검정하는 것이다. 그러나 본 논문에서 적용한 방법은 Coronel-Brizio 등 (2008)에서 채택한대로 45개의 숫자 중에서 무작위로 6개를 뽑았을 때 뽑힌 숫자들을 작은 것부터 큰 것 순으로 정리한 순서통계량을 고려하였다. 즉, 6개의 숫자들이 무작위로 뽑혔다면 그 뽑힌 숫자들의 순서통계량의 기대값이 무슨 값인지를 이론적으로 유도하여 그 값이 과연 실제로 출현한 순서통계량의 표본평균과 얼마나 차이가 나는지를 측정하는 검정통계량을 사용하여 검정하고자 한다.

표 2.1 당첨구조 및 당첨금 배분

등위	당첨내용	당첨확률	당첨금 배분 비율
1	6개 번호일치	1/8,145,060	총 당첨금 중 5등 금액을 제외한 60%
2	5개 번호일치+보너스 번호일치	1/1,357,510	총 당첨금 중 5등 금액을 제외한 10%
3	5개 번호일치	1/35,724	총 당첨금 중 5등 금액을 제외한 10%
4	4개 번호일치	1/733	총 당첨금 중 5등 금액을 제외한 20%
5	3개 번호일치	1/45	5,000원

2.1. 확률

N개의 번호 중에서 k개의 번호를 비복원 추출하는 경우의 가능한 조합의 수는

$$\binom{N}{k} = \frac{N!}{k!\,(N-k)!}$$

이다. 이때 확률변수 X를 k개의 당첨번호와 일치하는 개수라고 한다면 그 때의 확률은

$$P\,(X=i) = \binom{k}{i}\binom{N-k}{k-i}\binom{N}{k}^{-1}, \qquad i=0,1,\ldots,k \tag{2.1}$$

이며, 이것은 초기하분포를 따른다는 것을 알 수 있다. 또한 k개의 당첨번호를 크기 순서로 정렬한 순서통계량을 $\boldsymbol{Y}' = \left(Y_{(1)},\ldots,Y_{(k)}\right)$라고 하면, 이때 $Y_{(i)}$는 뽑힌 k개의 당첨번호들을 오름차순으로 정렬한 후 그 값들 중 i번째 자리에 있는 번호를 의미한다. i번째 자리의 번호가 r이 되는 경우, 즉 $Y_{(i)}=r$이 되는 필요충분조건은 $i-1$개의 번호들이 $(1,r-1)$에서, $k-i$개의 번호들은 $(r+1,N)$사이의 값이어야 한다. 따라서 $Y_{(i)}=r$이 되는 확률은

$$P\left(Y_{(i)}=r\right) = \binom{r-1}{i-1}\binom{N-r}{k-i}\binom{N}{k}^{-1} \tag{2.2}$$

이다. 또한 i번째의 번호가 r이고, j번째 번호가 s인 경우의 결합 확률은

$$P\left(Y_{(i)}=r,\ Y_{(j)}=s\right) = \binom{r-i}{i-1}\binom{s-r-1}{j-i-1}\binom{N-s}{k-j}\binom{N}{k}^{-1}, \quad i<j,\ r<s \tag{2.3}$$

이다.

2.2. 1차, 2차 적률 및 평균벡터 μ와 공분산행렬 V계산

Coronel-Brizio 등 (2008)에 의하면 뽑힌 k개의 당첨번호 중 크기순서로 오름차순 정렬한 후, i번째 번호 $Y_{(i)}$의 기대값은 다음과 같다.

$$E\left(Y_{(i)}\right) = \binom{N}{k}^{-1}\sum_{r=i}^{N-k+i} r\binom{r-1}{i-1}\binom{N-r}{k-i} \tag{2.4}$$

또한, $Y_{(i)}^2$의 기대값은

$$E\left(Y_{(i)}^2\right) = \binom{N}{k}^{-1}\sum_{r=i}^{N-k+i} r^2\binom{r-1}{i-1}\binom{N-r}{k-i} \tag{2.5}$$

이다. 따라서 이때의 분산은

$$Var(Y_{(i)}) = E(Y_{(i)}^2) - \big(E(Y_{(i)})\big)^2 \tag{2.6}$$

이 되고, i번째와 j번째 번호의 공분산행렬은

$$Cov\left(Y_{(i)}, Y_{(j)}\right) = E\left(Y_{(i)}Y_{(j)}\right) - E\left(Y_{(i)}\right)E\left(Y_{(j)}\right), \quad (i < j) \tag{2.7}$$

이다. 여기서

$$E\left(Y_{(i)}Y_{(j)}\right) = \binom{N}{k}^{-1} \sum_{r=i}^{N-k+i} \sum_{s=r+1}^{N-k+j} rs \binom{r-1}{i-1} \binom{s-r-1}{j-i-1} \binom{N-s}{k-j} \tag{2.8}$$

이다.
위 결과들로부터, $\boldsymbol{\mu} = E(\boldsymbol{Y})$의 i번째 성분 (component)의 값은

$$\mu_i = E(Y_{(i)}) = \frac{(N+1)i}{(k+1)}, \quad i = 1, 2, \ldots, k. \tag{2.9}$$

이며, 공분산행렬 $V = Var(\boldsymbol{Y})$의 각 원소 (element)들의 값은

$$v_{ij} = v_{ji} = Cov\left(Y_{(i)}, Y_{(j)}\right) = \frac{i(k-j+1)(N+1)(N-k)}{(k+1)^2(k+2)}, \quad 1 \le i \le j \le k. \tag{2.10}$$

이다.

2.3. 로또 6/45의 평균벡터 μ, 공분산행렬 V구하기

우리나라에서의 로또복권은 다음의 과정에 의하여 발행되고 있다. 전체 $1 \sim 45$의 번호 중에서 순서에 상관없이 6개의 번호를 선택하여 당첨번호 6개와의 일치된 번호의 개수를 비교한다. 따라서 우리나라의 로또복권의 경우 $N = 45$이고, $k = 6$이다. 따라서 로또 6/45의 평균벡터는 식 (2.9)로부터

$$\begin{aligned}\boldsymbol{\mu}_0' &= \begin{pmatrix} \frac{46}{7} & \frac{92}{7} & \frac{138}{7} & \frac{184}{7} & \frac{230}{7} & \frac{276}{7} \end{pmatrix} \\ &= (6.5714286 \quad 13.142857 \quad 19.714286 \quad 26.285714 \quad 32.857143 \quad 39.428571)\end{aligned}$$

이다. 또한 분산공분산행렬은 식 (2.10)으로부터

$$V = \begin{pmatrix} 27.459184 & 22.882653 & 18.306122 & 13.729592 & 9.1530612 & 4.5765306 \\ 22.882653 & 45.765306 & 36.612245 & 27.459184 & 18.306122 & 9.1530612 \\ 18.306122 & 36.612245 & 54.918367 & 41.188776 & 27.459184 & 13.729592 \\ 13.729592 & 27.459184 & 41.188776 & 54.918367 & 36.612245 & 18.306122 \\ 9.1530612 & 18.306122 & 27.459184 & 36.612245 & 45.765306 & 22.882653 \\ 4.5765306 & 9.1530612 & 13.729592 & 18.306122 & 22.882653 & 27.459184 \end{pmatrix}$$

이고, 분산공분산행렬의 역행렬은 3중대각행렬 (tri-diagonal matrix)로서

$$V^{-1}=\begin{pmatrix} 0.062430 & -0.031215 & 0 & 0 & 0 & 0 \\ -0.031215 & 0.062430 & -0.031215 & 0 & 0 & 0 \\ 0 & -0.031215 & 0.062430 & -0.031215 & 0 & 0 \\ 0 & 0 & -0.031215 & 0.062430 & -0.031215 & 0 \\ 0 & 0 & 0 & -0.031215 & 0.062430 & -0.031215 \\ 0 & 0 & 0 & 0 & -0.031215 & 0.062430 \end{pmatrix}$$

이다. 여기서 분산공분산행렬의 역행렬을 이용하는 이유는 제 3절에서 논의되는 다변량중심극한정리를 사용한 검정방법에서 사용되기 때문이다.

3. 무작위성에 대한 검정 방법

로또복권의 발행횟수가 n회일 때, 각 회차별 뽑힌 1등 당첨번호들을 오름차순으로 정렬하면 $\boldsymbol{y}'_n = (\boldsymbol{y}_1 \quad \boldsymbol{y}_2 \quad \cdots \quad \boldsymbol{y}_6)$이다. 이때 각 열들의 평균은 $\bar{\boldsymbol{y}}' = \left(\bar{y}_{(1)}, \ldots, \bar{y}_{(6)}\right)$가 된다. 본 연구의 로또복권의 1등 당첨번호에 대한 무작위성의 검정방법은 귀무가설 $E(\boldsymbol{Y}) = \boldsymbol{\mu}_0$를 검정하는 것이다. 검정통계량으로서는 Hotelling의 T^2통계량을 사용하였으며, 검정통계량의 분포는 대표본의 성질을 이용한 다변량중심극한정리 (multivariate central limit theorem)와 몬테카를로 (Monte carlo)를 이용한 모의실험 (simulation)에 의하여 각각 결정되었다. 다변량중심극한정리에 의하면 $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_n$은 평균벡터가 $\boldsymbol{\mu}$이고 공분산행렬이 $\boldsymbol{V}$인 집단에서 추출된 p차원 확률표본이라 했을 때, $n-p$가 상당히 크면 식 (3.1)의 T^2통계량의 분포는 χ^2_p에 근사하게 된다 (Park, 2004; Rahman 등, 2004).

$$M = n(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}_0)'\boldsymbol{V}^{-1}(\bar{\boldsymbol{Y}} - \boldsymbol{\mu}_0) \tag{3.1}$$

따라서 식 (3.1)의 $\bar{\boldsymbol{Y}}$대신에 실제자료인 로또복권 1등 당첨번호들의 평균 $\bar{\boldsymbol{y}}$를 이용하여 M값을 구한 후, $\chi^2_{(6)}(M)$의 계산을 통해 유의확률 (p-value)을 계산하였다. 또한 몬테카를로를 이용한 모의실험의 경우 1 ∼ 45까지의 번호 중 6개의 번호를 임의적으로 선택하는 과정을 발행횟수 (n)만큼 반복시행한 후 $\bar{\boldsymbol{Y}}^*$를 계산하고, 식 (3.1)의 $\bar{\boldsymbol{Y}}$대신에 $\bar{\boldsymbol{Y}}^*$를 사용, 식 (3.2)의 M^*를 구하는 과정을 각각 $R = 5,000$번씩 반복하였다. 그리고 그때의 유의확률값을 계산하였으며, 이때 몬테카를로 모의실험에서의 유의확률은 식 (3.3)와 같다.

$$M^* = n(\bar{\boldsymbol{Y}}^* - \boldsymbol{\mu})'\boldsymbol{V}^{-1}(\bar{\boldsymbol{Y}}^* - \boldsymbol{\mu}) \tag{3.2}$$

$$\widehat{p} - value = \frac{number\ of\ (M_i^* \geq M)}{R}, \quad i = 1, \cdots, R \tag{3.3}$$

4. 로또복권 6/45의 1등 당첨번호에 대한 무작위성 검정결과

우리나라 로또복권의 경우, 당첨번호의 수는 6개이며 공중파 TV를 이용하여 매주 토요일 저녁에 생중계로 당첨번호를 추첨하고 있다. 선택한 6개 번호의 순서에 상관없이 당첨번호 6개와 정확히 일치하게 되면 로또복권 1등에 당첨이 된다. 또한 2등을 위하여 1개의 당첨번호 (bonus number)를 더 추첨하게 된다. 본 연구에서는 45개의 숫자 중 6개 당첨 번호를 모두 맞힌 1등 당첨의 경우만을 연구하였다. 우리나라의 로또복권의 경우 1회∼261회까지의 사업자는 시중의 K은행에서 주관하였으며, 262회부터 N은행의 주관으로 로또복권이 발행되고 있다. 또한 87회까지는 로또복권의 1게임당 가격이 2,000원이었지만, 88회부터의 로또복권 1게임당 가격이 1,000원으로 인하되었다. 따라서 본 연구에서는 K은행

과 N은행의 로또복권 주관사 변경, 그리고 로또복권의 1게임당 가격이 변경되면서 당첨번호들에 어떤 변화가 있었는지를 알아보기 위하여 각각 추첨된 로또복권의 1등 당첨번호들에 대하여 무작위성을 검정하였다.

표 4.1은 실제 당첨된 6개의 1등 당첨번호에 대하여 오름차순으로 정렬 한 후 계산된 각 성분의 평균 $\bar{\boldsymbol{y}}' = \left(\bar{y}_{(1)}, \ldots, \bar{y}_{(6)}\right)$의 결과이다. 이 결과는 표현의 용이성을 위해 소수점 세 번째 자리에서 반올림하여 나타낸 값으로서, 실제로 SAS (default option) 프로그래밍 과정에서는 식 (3.1)의 $\bar{\boldsymbol{y}}'$의 값을 반올림하지 않고 계산된 결과를 그대로 사용하였다. 또한 제3절에서 제시되었던 다변량중심극한정리와 몬테카를로 모의실험에 의한 검정 방법별 유의확률의 값은 표 4.2와 같다. 표 4.2의 결과를 이용하여 식 (4.1)의 로또복권의 1등 당첨번호들에 대한 가설검정을 유의수준 0.05하에서 시행하였다.

표 4.1 로또복권 발행횟수별 순서통계량의 평균

		$\bar{y}_{(1)}$	$\bar{y}_{(2)}$	$\bar{y}_{(3)}$	$\bar{y}_{(4)}$	$\bar{y}_{(5)}$	$\bar{y}_{(6)}$	발행횟수
주관사	K은행	6.67	13.18	20.31	26.28	32.77	39.23	261
	N은행	5.94	11.89	19.34	26.49	32.71	39.64	70
가격	2,000원	6.40	12.92	20.43	26.34	32.87	39.28	87
	1,000원	6.56	12.90	19.99	26.31	32.72	39.34	244
전체자료		6.52	12.90	20.10	26.32	32.76	39.32	331

로또복권의 과거 주관사업체였던 K은행과 현제 주관사업체인 N은행에서 시행된 로또복권 1등 당첨번호들의 무작위성 검정 결과, 두 사업체 각각 두 방법의 유의확률의 값이 유의수준 $\alpha = 0.05$보다 매우 크기 때문에 식 (4.1)의 귀무가설을 기각할 수 없음을 알 수 있다. 또한 가격의 변화에 따른 로또복권의 1등 당첨번호들에 대한 두 가지 실험 모두 유의확률의 값이 유의수준 $\alpha = 0.05$보다 크기 때문에 귀무가설이 기각되어 당첨번호들은 무작위적임을 알 수 있다.

표 4.2 로또복권에 대한 검정 방법별 유의확률의 값

		M	중심극한정리 유의확률	몬테카를로 유의확률
주관사	K은행	6.0676	0.5843	0.3348
	N은행	4.7902	0.4290	0.5272
가격	2,000원	3.7563	0.2904	0.7468
	1,000원	3.2038	0.2171	0.7694
전체자료		6.0505	0.5824	0.3608

$$\begin{cases} \text{귀무가설 } (H_0)\text{: 로또복권의 1등 당첨번호는 무작위적이다.} \\ \text{대립가설 } (H_1)\text{: 로또복권의 1등 당첨번호는 무작위적이 아니다.} \end{cases} \tag{4.1}$$

마지막으로 주관사업체와 가격의 변화를 고려하지 않은 총 331회 발행된 로또복권의 1등 당첨번호에 대한 유의확률의 값 역시 0.5824와 0.3608로서 유의수준 $\alpha = 0.05$보다 크기 때문에 식 (4.1)의 귀무가설을 기각할 수 없다. 따라서 지금까지 발행된 로또복권의 1등 당첨번호들에 대한 무작위성을 검정한 결과 1등 당첨번호들은 무작위성을 따른다고 결론지을 수 있다.

5. 결론 및 토의

로또복권은 이미 전 세계 여러 나라에서 발행되고 있으며 많은 인기를 누리고 있다. 각 나라마다 로또복권의 발행 방식은 다르지만 공통적인 특성은 N개의 번호 중에서 k개의 번호를 선택하여, 추첨된 당첨

번호 k개와 선택한 k개의 번호 중 순서에 상관없이 정확하게 일치하는 번호의 개수에 따라 당첨금액이 차등 지급되는 방식으로 시행되고 있다는 것이다. 우리나라 역시 로또복권이 처음 발행된 이후 다른 어떤 복권보다도 가장 큰 인기를 얻어 오고 있지만, 로또복권의 당첨확률을 높일 수 있는 방법들이 존재한다는 주장들은 지금까지도 많은 사람들의 입에 오르내리고 있는 가장 큰 의혹이다. 따라서 본 연구에서는 로또복권의 주관 사업체별, 그리고 로또복권의 가격 변화를 고려하여 다변량중심극한정리와 몬테카를로 모의실험의 방법으로 1등 당첨번호들 무작위성에 대한 검정을 실시하였다. 그 결과 추첨된 1등 당첨번호들은 무작위성을 만족하였으며, 주관 사업체와 가격의 변화를 고려하지 않은 전체 자료의 경우에도 무작위성을 만족하는 것으로 판명되었다.

참고문헌

매일경제 (2009). <로또 당첨조작 의혹 규명한다>, 2009년 5월 18일 기사, 서울.

한국조세연구원 (2003). <로또복권 수익금 활용 및 관리방안>, 서울.

Boland, P. J. and Pawitan, Y. (1999). Trying to be random in selecting numbers for lotto. *Journal of Statistics Education*, **7**.

Chaitin, G. J. (1975). Randomness and mathematical proof. *Scientific American,* **232**, 47-52.

Chen, S. H. and Chie, B. T. (2008). Lottery markets design, micro structure and macro behavior: An ACE approach. *Journal of Economic Behavior & Organization*, **67**, 463-480.

Coronel-Brizio, H. G., Hernández-Montoya, A. R., Rapallo, F. and Scalas, E. (2008). Statistical auditing and randomness test of lotto k/N-type games. *Phisica A: Statistical Mechanics and its Applications*, **387**, 6385-6390.

Farrell, L., Morgenroth, E. and Walker, I. (1999). A time series analysis of UK lottery sales: Long and short-run price elasticities. *Oxford Bulletin of Economics and Statistics*, **61**, 513-526.

Genest, C., Lockhart, R. A. and Stephens, M. A. (2002). χ^2 and the lottery. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **51**, 243-257.

Helman, D. (2005). Combinatorial interdependence in lottery. *Teaching Mathematics and its Applications*, **24**, 203-207.

Park, C. Y. (2004). A note on the simple chi-squared test of multivariate normality. *Journal of Korean Data & Information Science Society*, **15**, 423-430.

Rahman, M. and Chakrobartty, S. (2004). Testing for uniformity. *Journal of Korean Data & Information Science Society*, **15**, 211-218.

Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A., Dray, J. and Vo, S. (2001). *A Statistical test suite for random and pseudorandom number generators for cryptographic applications*, NIST Special Publication 800-22 (with revisions dated May 15, 2001), NIST, Gaithersburg.

Russell, K. G. and Griffiths, D. (2005). A lotto system problem. *Australian & New Zealand Journal of Statistics*, **47**, 259-267.

Scott, J. F. A. and Gulley, O. D. (1995). Testing for efficiency in lotto games. *Economic Inquiry*, **33**, 175-188.

Journal of the Korean Data & Information Science Society
2009, **20**(5), 779–786

한국데이터정보과학회지

Statistical randomness test for Korean lotto game

Suyeol Lim[1] · Jangsun Baek[2]

[12]Department of Statistics, Chonnam University

Received 15 May 2009, revised 25 August 2009, accepted 21 September 2009

Abstract

Lotto is one of the most popular lottery games in the world. In korea the lotto considers numbers 1, 2, $\cdots$, 45 from which 6 numbers are drawn randomly, without replacement. The profits from the lotto supports social welfare. However, there has been a suspicion that the choice of the winning numbers might not be random. In this study, we applied the randomness test developed by Coronel-Brizio *et al.* (2008) to the historical korean lotto data to see if the drawing process is random. The result of our study shows that the process was random during two periods under the management of different business companies and of price changes, respectively.

Keywords: Central limit theorem, lottery, lotto, Monte carlo, randomness.

[1] Ph. D. Candidate, Department of Statistics, Chonnam National University, Gwangju 500-757, Korea.
[2] Corresponding author: Professor, Department of Statistics, Chonnam National University, Gwangju 500-757, Korea. E-mail: jbaek@chonnam.ac.kr