Section E1

501/E1

A study on equiprobability of outcomes of two lotteries in Sri Lanka

B M S G Banneheka*and K D Wijesinghe

Department of Statistics, Faculty of Applied Sciences, University of Sri Jayewardenepura, Nugegoda

With many different lotteries being available and being popular the lottery regulators need to ensure the public that the games are played fair. In a lottery game where one character out of *M* different non-numeric characters and *n* numbers out of *N* different numbers are drawn, the fairness would mean the independence and equiprobability of outcomes. Usually, there is no reason to doubt the independence of outcomes on different days. However, the equiprobability of outcomes could be violated in many ways. This study presents data on a study on testing the equiprobability of outcomes of two lotteries, labelled as lottery1and lottery2, where drawing the numbers resembles sampling with-replacement and without-replacement respectively. The non-numeric character was ignored and only the n-tuple of numbersconsidered as an outcome, for convenience. The values of (*N, n*) were (10, 6) and (67, 5)forthe lotteries 1 and 2 respectively. The number of outcomes used in this study were *k*= 487for lottery 1 and *k* = 468 for lottery2.

Theoretically, if the numbers are drawn with replacement, the equiprobability of n-tuples of numbers can be tested using the standard chi-square test on a$N^n \times 1$ contingency table. If the numbers are drawn without-replacement, that test is not valid. A computationally intensive, modified chi-square test on a$^NC_n \times 1$ contingency table is available for this scenario. However, these tests were not applicable as the observed frequencies in most of the cells were zero due to the fact that numbers of available outcomes were less than the numbers of cells in the two contingency tables. Therefore, a new and computationally simpler test was derived and applied. This test has the test statistic $T = \sum_{i=1}^{k} W_i/\sqrt{k}$, where$W_i = (\bar{X} - \mu_{\bar{X}})/\sigma_{\bar{X}}$, with $\bar{X}_i =$ the mean of the *n* numbers of the *i*th outcome, *k* = the number of outcomes used in the study, $\mu_{\bar{X}}$ and $\sigma_{\bar{X}}$ are respectively the mean and standard deviation of the sampling distribution of $\bar{X}$. Since the pool of numbers from which the draws were made was known, the exact values of $\mu_{\bar{X}}$ and $\sigma_{\bar{X}}$ for the two lotteries can be calculated separately depending on the sampling method. The decision rule is to reject the null hypothesis of equiprobability of outcomes at the α level of significance, if$|T| > Z_{(\alpha/2)}$. While there was not sufficient evidence to reject the null hypothesis of equiprobability of outcomes of the lottery 1, there was strong enough evidence to reject the same hypothesis for lottery 2 at 0.05 level of significance. By repeating the test for lottery 2 with partitioned data, it was found that the equiprobability of outcomes could have been doubtful during the early part of the study period.

sarath.banneheka@gmail.com Tel:+94716865801