

Probability Theory and Introductory Statistics

ALY 6010

Assignment 3

Title: Hypergeometric Distribution of Keno

Rahul Iyer

Professor Name: Amin Karimpour

**INTRODUCTION**

Before understanding the probability-distribution associated with the Keno game, it is vital to understand how this game works. The player can pick two to ten unique numbers, ranging from 1-80. Apart from these numbers, the computer shall pick three additional numbers (bonus numbers) randomly, excluding the numbers chosen by the player. These three numbers are symbolized with eggs and play a very important role in determining the winning amount (if any) of that player.

The computer then draws 20 unique numbers from the pool of eighty numbers, numbered 1 through 80. The player wins depending on his/her number of picks, and whether or not those picked numbers and/or bonus numbers match with these 20 winning numbers. For instance, a player that picked 4 numbers shall win 5$ if 3 of those 4 numbers match with the winning numbers (assuming that none match his/her bonus numbers). Additionally, if two or more of his bonus numbers are included among the 20-winning number draw, then the player shall receive 3 times and 6 times the winning amount for a match of 2 and 3 bonus numbers, respectively. As an example, if the player that wins 5$ from the previous example has two of the three bonus numbers matching the winning numbers, then the payer will receive 3 times the amount of 5$, which is 15$.

Now, we shall consider this game as a hypergeometric experiment to determine the probability of winning from the 20-number draw as well as matching the bonus numbers to claim the multiplier.

**ANALYSIS**

It is clear that the population of this game (N) is 80 since all numbers are drawn from this population of 1-80. The sample size would be 20 (r) as the computer picks 20 winning numbers from this population. The player picks ‘n’ numbers varying from 2 to 10 (also determining the number of possible successes), ranging from 1-80. The probability associated with hitting ‘k’ numbers from these ‘n’ numbers can be found out using the Excel’s function:

**Probability of hitting k numbers out of n numbers, matching r numbers from N numbers= hypgeom.dist(k,n,r,N,0)**

The last ‘0’ digit indicates that the resultant probability would calculate probability density function, and not the cumulative probability.

This logic holds true while calculating the bonus probability, probability of hitting the bonus by matching the eggs. However, some changes have to be made with the variables to increase the accuracy and reliability of the estimate. Since, we know that the computer picks those three bonus numbers uniquely from the population, except those numbers (p) chosen by the player. Therefore, the population would no longer be N, it would rather be N-p. In simple words, if a player picks 8 numbers, the computer shall randomly select 3 numbers from the remaining 80-8=72 numbers pool, to avoid duplicity.

**Probability of hitting k eggs out of n eggs, matching r numbers from N-p numbers= hypgeom.dist(k,n,r,N-p,0)**

Here, n=3 (the total number of eggs)

Therefore,

Probability of hitting 2 eggs out of 3 eggs, matching r numbers from N-p numbers= hypgeom.dist(2,3,r,N-p,0)

Probability of hitting 3 eggs out of 3 eggs, matching r numbers from N-p numbers= hypgeom.dist(3,3,r,N-p,0)

The value of ‘Return to player’ (or RTP) is calculated by the summation of the product obtained by multiplying the winning probability and the winning amount. It indicates the return on investment, RTP represents the monetary value a player can expect after each game.

Hit frequency of Keno is the number of times a player can win. It is the sum of probability associated with winning. A winning flag has been added to the Excel sheet that determines whether that instance is a winner or not, conditionally depending on the value in the ‘Pay’ column. A value of ‘0’ in the ‘Pay’ column would have a ‘0’ value in the winning flag too; any other positive value would result in a value of ‘1’ in this winning flag, reflecting that the particular instance is a winner.

Pulls per hit represent the average number of attempts, it would take a player to receive a winning number combination. Mathematically, it is the reciprocal of ‘Hit frequency’.

The RTP%, hit frequency and pulls per hit can be calculated in a similar way for the bonus numbers.

The volatility index is the standard deviation associated with the probability distribution of pay. In order to calculate standard deviation, we need to calculate mean and variance. Mean is given by the formula:

Where,

µ= Mean

X= the winning amount

P(X)= the relative probability of winning that amount

Variance is calculated using:

Variance= ∑(X-µ)2.P(X)

Where,

(X-µ) = difference between the mean and the winning amount.

**CALCULATION:**

Let us take a small example- we pick 2 numbers and hit two as well as obtain 3 matches on the bonus eggs. The different parameters associated with this instance are calculated below:

The probability of hitting 2 numbers if we pick only 2 is calculated by dividing the combinations possible of 2 numbers from 80 numbers and combinations possible of obtaining those 2 numbers from 20 numbers. There are 80 and 79 choices for the player’s first number and second number respectively. Further, the numbers itself may be interchangeable twice, like 12, 20 or 20,12 is the same. Therefore, the combinations for 2 out of 80 are 80\*79/2=3160. Similarly, the combinations for 2 out of 20 are 20\*19/2=190.

Probability of picking and hitting two numbers = 190/3160

= 0.060127

The RTP% in such a case is the product of winning amount and the probability. The winning amount (without the multiplier) is 10$. Therefore, the RTP% is **60.127%**, which means that for every dollar spent, an average 0.60127$ shall be won/returned to the player.

In this case, the hit frequency can be calculated very easily:

Hit frequency = ∑P(Pay)

=***0.060127*** (Since no other instance has an associated pay)

Further,

***=16.6316***

Similarly,

Bonus RTP%, hit frequency and pulls per hit for hitting 3 out of 3 eggs is **5.622%, 0.000937038** and **1067.192982** respectively.

Games Total RTP% is calculated by adding the Keno RTP% and Bonus RTP%:

=60.127 + 26.71 (Bonus- hitting 2 of 3 eggs) + 5.622

***= 92.454%***

**CONCLUSION:**

The table below contains the different parameters for the different choices an individual makes while selecting his/her numbers.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | PICK 2 | PICK 3 | PICK 4 | PICK 5 | PICK 6 | PICK 7 | PICK 8 | PICK 9 | PICK 10 |
| KENO RTP% | 60.127% | 63.827% | 65.557% | 64.71% | 66.691% | 67.457% | 66.668% | 68.667% | 68.034% |
| KENO HIT FREQ. | 0.06013 | 0.15263 | 0.25895 | 0.36713 | 0.161582 | 0.236579 | 0.317124 | 0.153051 | 0.211979 |
| KENO PULLS/HIT | 16.6316 | 6.55183 | 3.8618 | 2.723833 | 6.188805 | 4.226914 | 3.15334 | 6.533761 | 4.717453 |
| BONUS RTP% | 32.328% | 34.317% | 36.145% | 36.599% | 38.703% | 40.184% | 40.779% | 43.143% | 43.922% |
| BONUS HIT FREQ. | 0.00984 | 0.02498 | 0.04339 | 0.063014 | 0.028416 | 0.04264 | 0.058597 | 0.029001 | 0.041203 |
| BONUS PULLS/HIT | 101.637 | 40.039 | 23.0462 | 15.86944 | 35.1915 | 256.7185 | 184.0044 | 366.1148 | 253.703 |
| GAMES TOTAL RTP% | 92.454% | 98.144% | 101.701% | 101.311% | 105.394% | 107.641% | 107.447% | 111.810% | 111.956% |
| VOLATILITY INDEX | 7.77373 | 6.88409 | 13.7728 | 8.397946 | 14.6217 | 30.00412 | 16.05297 | 19.79174 | 16.20405 |

*Table 1*

It can be observed that the Bonus RTP% increases gradually with the number of picks, ranging from 32.328% for a two-number pick to 43.922% for a 10-number pick. Games Total RTP% follows this trend varying from 92.454% to 111.956% with some exceptions/outliers (Pick 5 and Pick 8) consisting of minor deviations.

On the other hand, the Keno RTP% is low for a 2-number pick and highest for a 10-number pick with fluctuations between those ranges, where values gradually increase and decrease repeatedly.

From a neutral standpoint, the values in the ‘Pick 5’ column are most favorable for winning since the volatility is low, while Keno pulls per hit and Bonus pulls per hit values are high. This directly means that the probability of hitting the winning numbers are the most optimum, compared to others if we pick 5 numbers for the Keno game.

**REFERENCE**

Hypergeometric Distribution. (2018). Retrieved November 18, 2018 from <https://stattrek.com/probability-distributions/hypergeometric.aspx>

Shackleford, M. (2018, August 27). Keno Odds. Retrieved November 16, 2018 from <https://wizardofodds.com/games/keno/appendix/3/>

The Excel HYPGEOM.DIST Function. (n.d.). Retrieved November 17, 2018 from <https://www.excelfunctions.net/excel-hypgeom-dist-function.html>