Estimating the duration of random walks

Eduardo Yukio Garrafa Ishihara
Gustavo Silva Garone

Prof. Elisabeti Kira, Ph.D

Institute of Mathematics and Statistics, University of São Paulo

eduardoyukio.ishihara@usp.br

Objectives

The Gambler's Ruin is a random walk, one of the most classical models of a markovian stochastic process. In this model, a gambler wagers money in each round and, according to a probabilistic rule, either gains or loses varying amounts. Such models are widely applicable in statistical modeling and computational algorithms [5]. Their behavior has been thoroughly studied in the simplest form — the gambler wins $\mathrm{R}\$1.00$ with probability p and loses $\mathrm{R}\$1.00$ with probability $1-p$ — particularly concerning the probability of ruin or absorption [4]. The aim of this project is to investigate the expected time to ruin in more complex models by proposing methods to estimate this duration and evaluating their performance through computational approaches. With this, we hope to expand the understanding of these processes and some of their generalizations.

Materials and methods

We analyzed the calculation of the expectation and variance of the duration of the classical Gambler's Ruin problem [1], highlighting the use of systems of finite difference equations. We developed an initial heuristic estimator based on an analogy with average velocity in Physics. This estimator yielded accurate approximations for most random walks, both in the simple case and in models where the gambler may win or lose amounts different from $\mathrm{R}\$1.00$ per round. In the fair game case, that is, when the expected gain per round is zero, the estimator deviated significantly from the theoretical and simulated values.

For random walks with more general rules, where analytical solutions are not available, we evaluated this estimator using Monte Carlo simulations, based on methods proposed by Ritter [3], using the Python programming language. As we considered broader rules and rules that greatly extended the simulation times and required more computational power, we migrated to the Julia language, as recommended by Godoy [2]. We further refined the estimator to account for ties within a round.

Having identified the limitations of this estimator in one dimension, we extended our study to random walks in $\mathbb{Z}^d$. By solving systems of equations computationally with symbolic libraries, we obtained analytical results for unit-step walks with different probabilistic rules. New estimators were proposed and analyzed for both unit and non-unit step walks, which provided good estimates when the number of dimensions was sufficiently large or when the probabilistic rule of the model was non-pathological.

Results

The proposed estimator showed efficacy in predicting the duration of one-dimensional uniform walks, as depicted by Figure 1, and adapts successfully when the transition probability depends on the gambler's current fortune.

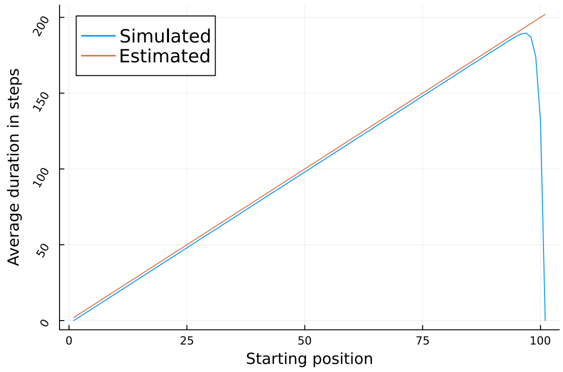

Figure 1: Comparing the estimator with 100.000 simulations per starting position

In higher dimensions the expected duration of the random walk must be obtained numerically due to the system of balance equations growing exponentially with the number of dimensions, rendering a closed form solution intractable. Still, a surprising result was revealed: once the dimensionality exceeds roughly the 100th dimension, the average time to game completion increases almost linearly with each additional dimension as shown in Figure 2.

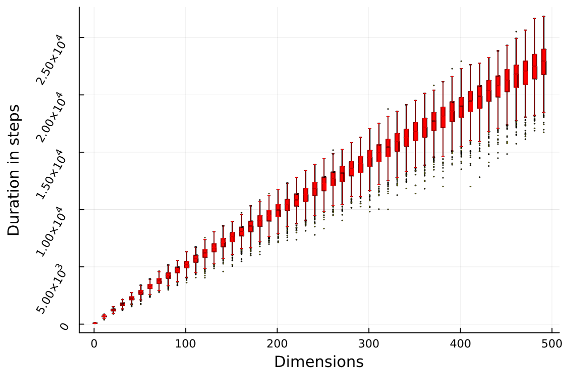

Figure 2: Simulated steps in 1.000 simulations em different dimensions.

In the particular model where the gambler cannot lose his money, we've obtained analytic expressions for the uniform, non-uniform, random and non-random variants. The complexity of the equations that describe the game's duration grows proportionally to the state space's size. In all cases, the estimated values show great proximity to simulated ones, although greater discrepancies show when the game is close to fair - that is, when the gambler's expected earnings per round are close to 0 - or when the starting position is close to a boundary value.

Conclusions

The Gambler's Ruin is a classical math problem proposed almost four centuries ago. Through intuition, undergraduates can formulate fresh questions and hypothesis, explore them computationally, and finally developing rigorous mathematics. It is our hope that this work encourages other undergraduates to explore and reexplore classical math problems in search of solutions to modern day challenges.

The authors declare no conflict of interests. Eduardo Yukio conceived the theoretical framework behind the estimators and derived the paper's analytical results. Gustavo Garone provided computational support and implemented the simulations used for improving the estimators.

Acknowledgements

We would like to thank our advisor, Elisabeti Kira, for her patience, trust, and mentorship that enabled this project to be completed. We would also like to thank the University of São Paulo for financing our research through the Programa Unificado de Bolsas (PUB).

References

[1] Jiří Anděl and Šárka Hudecová. Variance of the game duration in the gambler's ruin problem. 82(9):1750–1754.

[2] William F. Godoy et al. Evaluating performance and portability of high-level programming models.

[3] Frank E. Ritter et al. Determining the number of simulation runs. Springer.

[4] S.M. Ross. *Introduction to Probability Models*. Academic Press.

[5] S.M. Ross. *Markov Chain Monte Carlo Methods*. Academic Press.